Typographic emphasis and contrastive focus:
An eye tracking study

Christopher William Norton

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy

The University of Leeds
School of Languages, Cultures and Societies

January 2018

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Christopher William Norton to be identified as Author of this work has been asserted by Christopher William Norton in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgements

Abstract

Two eye tracking experiments investigated whether italicising contrastively focused words facilitates processing. In speech, pitch accents can mark focus, signalling that there exist alternatives to the marked element which are relevant to interpreting the expression. In writing, typographic emphasis (italics, bold, etc.) can be used in a similar manner, particularly to mark contrastive focus: 'We expected Mark to bring *Mary* to dinner. However, when he arrived, he was with *Ellen*.' Previous studies have shown that processing is facilitated by congruent marking of information structure with pitch accents, and impeded by incongruent marking. This study sought similar effects for typographic emphasis.

Eye movements of participants were tracked as they read short texts (dialogues in Experiment 1, narratives in Experiment 2) in which a contrastively focused target word was italicised, or not. Experiment 2 also manipulated contrast on the target word. To the author's knowledge, this is the first eye tracking study that investigates the interaction of these factors within continuous text. Incongruence, where a non-contrastive word was typographically emphasised, or a contrastive word was not emphasised, was predicted to cause processing difficulty manifesting as re-reading.

Results did not suggest that incongruence had a processing cost, although some effects were found, including (amongst more subtle effects) longer overall fixations on target words, and higher likelihoods of them being fixated at all. These results are discussed from several perspectives, including a 'zero-impact' account, which holds there to be no linguistic effect of typographic emphasis, and a more probable 'effects elsewhere' account, which suggests the type of re-reading examined here may not have been the correct place to look. The possibility that the visual contrast inherent in typographic emphasis may simply 'catch the eye' rather than have a linguistic effect is also investigated, and assessed as being unlikely.

Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Term
AOI	Area of interest
CF	Contrastive focus
EEG	Electroencephalography
ERP	Event-related potential
fMRI	Functional magnetic resonance imaging
GLM	Generalised linear mixed-effects model
IS	Information structure
LMM	Linear mixed-effects model
PA	Pitch accent
PF	Presentational focus
RSVP	Rapid serial visual presentation
SPR	Self-paced reading
TE	Typographic emphasis

Typographic conventions

Italics indicate new terms, words qua words, phrases qua phrases etc., letters forming part of statistical notation, and titles of published work. On occasion italics are also used to mark the very type of contrastive focus that this thesis is concerned with.

SMALL CAPS indicate words that have pitch accents.

[Square brackets] indicate the extent of focus domains and similar information structure features (when appended with subscripts), binary experimental conditions, and interpolations within quotations.

{Curly brackets} indicate sets, in particular semantic sets of alternative elements and propositions.

⟨Angle brackets⟩ indicate graphemes.

/Slashes/ indicate phonemes and phonemic transcriptions.

One often hears of writers that rise
and swell with their subject,
though it may seem but an
ordinary one. How, then, with me,
writing of this Leviathan?
Unconsciously my chirography
expands into placard capitals.

Herman Melville

It doesn't just transport us, it
transports us—which I know is the
same word twice, but the second
time should be said louder and
slower.

Alan Partridge

Chapter 1

Introduction

This study investigates whether written emphasis can function in a similar way to intonational emphasis when it marks contrastive focus. This research question springs from the simple insight that spoken emphasis and visual emphasis in writing often seem to be recruited for the same purpose.

In speech, speakers add intonational information to their words which instructs hearers on how to connect the meanings of those words to what they already know. These instructions form part of the area of linguistics known as *information structure*. One type of information structure phenomenon is *contrastive focus*, which speakers use to indicate that an element of a sentence is unexpected to the hearer, and thus contrasts with a limited set of alternatives.

(1) We expected Mark to bring Mary to the party. However, when we opened the door, he was there with ELLEN.

In (1), the small capitals indicate a *pitch accent*, an intonational prominence marking the noun *Ellen* as being contrastively focused: the relevant alternative set with which it contrasts is made explicit in the prior sentence with the noun *Mary*. The speaker gives contrastive focus to *Ellen* because they judge that the hearer does not expect Ellen to be mentioned.

If (1) were to be written in a text unconcerned with metalinguistic analysis, the contrastive focus might be typeset in italics. This switching of font styles in order to emphasise parts of written text will be termed *typographic emphasis* here. Intonation-marked focus has been studied extensively in speech, and has been shown to be attended to, to facilitate processing when intonation and focus are congruent, and to impede processing when they are not. Less has been done on how *contrastive* focus is processed, and very little at all has been done to establish whether contrastive focus marked by typographic emphasis has similar effects:

Crystal might claim that 'the range of effects conveyed by switching between roman and bold and italic are quite well understood' (1998, p.14) but this is somewhat of an exaggeration, to say the least.

In fact, this issue can be seen as falling within a sub-discipline of linguistics that was originally proposed by Crystal himself, 'typographical linguistics'. If typographic emphasis can perform a similar function to intonational emphasis, if it actually helps the reader resolve the informational status of the emphasised elements in the same way that spoken emphasis helps the hearer, then this is persuasive support for the validity of investigating similar aspects of typography. If it does not, this may suggest that some typographic conventions serve relatively little linguistic purpose. Even such a disappointing eventuality would raise interesting questions about how and why such conventions have arisen.

1.1 Organisation of chapter

§1.2 introduces the putative linguistic sub-discipline **typographical linguistics**, to which this study is intended to make a contribution. It discusses the reasons why this is a relatively neglected area (§1.2.1).

§1.3 introduces **typographic emphasis**, the central topic of this study.

§1.4 discusses this study's adoption of a **psycholinguistic** approach to investigating typographic emphasis.

§1.5 outlines the **experimental approach** used here.

§1.6 details the **organisation of the thesis**.

1.2 Typographical linguistics

In 1997 the linguist David Crystal, addressing an audience of typographers, called for the establishment of a new sub-discipline he called 'typographical linguistics', which he proposed would address the central research question 'How do the various features of typography relate to the need to communicate meaning?' This suggests a huge, cross-disciplinary swathe of interest: semiotics, communication studies, stylistics and so on could all be brought to bear upon this research area. Indeed, all of these disciplines have already taken an interest in typography beyond the aesthetic.

However, in the twenty years since Crystal's suggestion, linguistics itself has

rarely investigated typography. This is despite Crystal following up his initial wide-ranging question with a targeted, explicitly linguistic remit for typographical linguistics, suggesting that it should investigate 'to what extent … the various features of typography convey linguistic meaning' (Crystal, 1998, p.9). Moreover, it is despite the inter-disciplinary field of psycholinguistics having put a considerable amount of effort into examining many other aspects of how written text represents language. It is also despite the evident richness of the sign system of written text, which contains many devices and conventions that would seem to go beyond being purely aesthetic, and instead appear to be used in a systematic, communicative manner. Despite all this, the call for a typographical linguistics still remains largely unanswered.

The present study is an attempt to address this outstanding issue. It stands with Crystal and a handful of other voices (e.g. Stöckl, 2005; Chafe, 1988) in reasoning that, if the connections between sound waves and meaning in spoken language warrant study by linguists, then the connections between visual marks and meanings in written language are also likely to reward linguistic investigation. This is true regardless of whether one adopts a position that writing encodes speech, which in turn encodes language;[1] or whether one takes speech and writing as being at an equal remove from whatever constitutes language itself. Either way, the intuition here is that certain aspects of typography perform a similar function to certain aspects of speech sounds.

1.2.1   Why the lack of typographical linguistics?

If the present study intends to address the lack of work on typographical linguistics, it is worth first asking why this area has been neglected.

For many language scholars throughout history, writing has seemed to be little more than a pale reflection of speech. To Plato, writing was 'a mere image' of 'ensouled speech'; to Saussure, studying language by looking at writing was like 'thinking more can be learned about someone by looking at his photograph than by viewing him directly' (Saussure, 1959, p.24); Bloomfield (1987, p.85) called writing 'merely a record of speech'. This perspective on writing has become 'axiomatic in the Western tradition' (Coulmas, 2003, p.4), although this axiom has certainly not gone without debate over the centuries.[2]

---

[1]This would make writing what Derrida calls 'the signifier of the signifier' (1967, p.7; see also Stöckl, 2005).

[2]Biber points out an irony that holds true across many areas of linguistic analysis: 'In theory,

As a theme within the history of science, the relationship between writing and speech is an interesting and complex issue of no small philosophical import, and as such, is rather too large a subject to treat of in detail here. However, part of the dismissal of writing as an object of linguistic inquiry is no doubt down to a sense that writing is speech denuded of much of its information. Whilst the sounds making up the atomic parts of speech, *phonemes*, are encoded in writing with a discrete set of visual marks or *graphemes* (Henderson, 1985; Rogers, 2005), writing might initially appear to lack analogues for the continuous elements of the speech stream, those suprasegmental elements that constitute *prosody*: 'a level of linguistic representation at which the acoustic–phonetic properties of an utterance vary independently of its lexical items' (Wagner and Watson, 2010, p.905).

And yet, writing clearly consists of more than a set of simple correspondences between graphemes and phonemes; otherwise, it would be an unbroken string of letters and nothing else. Instead, it encompasses devices which may be seen as functioning in similar ways to aspects of prosody. Chafe (1988) argues that one example of this is punctuation. Punctuation consists of a set of characters which, unlike most graphemes, do not correspond to phonemes: Chafe contends that the principal function of these punctuation graphemes is to indicate rhythm and intonation. This point can be contested, and it is undeniable that punctuation does not *only* encode prosodic information—but simply reading a punctuated stretch of text aloud provides compelling evidence for Chafe's point.

The present study is concerned with a different device, which will be referred to here as *typographic emphasis* (TE). Instead of supplementing the set of phoneme-encoding graphemes with an additional group of symbols, TE adds a layer of modulation onto the graphemes themselves.

1.3   Typographic emphasis

Sequences of speech sound units can be varied over the course of an utterance by alteration of the characteristics of the sound waves with which they are transmitted: different sounds can be made longer or shorter, louder or quieter, and can be given different pitches (Cruttenden, 1997). Written characters can be varied in a huge number of ways by alteration of the visual marks with which they are transmitted, arguably across a wider range of parameters: letters can be made thicker,

---

writing is disregarded as secondary and derivative from speech. In practice, however, speech is also disregarded as unsystematic and not representative of the true linguistic structure of a language.' (1988, p.7)

slanted, underlined, etc. in potentially infinite combinations (Stöckl, 2005). Of course, language requires conventionalised meanings, and if the suprasegmental aspects of writing are to perform similar functions to the suprasegmental aspects of speech, then the number of variations employed must be a small subset of all possible variations, and must be used in a more-or-less consistent manner. This appears to be the case in writing, which has developed a set of different variants, or *styles* of type, which exist within given typefaces.

One such variant is *italic* style. Italics are characterised by slanting letterforms which can convey a slightly more 'handwritten' air. They have come to be used as a visually marked alternative to the 'standard' *regular* or *roman* type style. Italics have a variety of functions, which on the whole tend to be categorised as 'emphasis'. As such, they will typically appear on isolated sections of language within larger surrounding stretches of regular-style type. This can be seen in the preceding paragraph, in which the term *styles*, new to the text in its technical usage, was italicised. This is an instance of TE.

Italic TE is arguably the most common form of TE to appear within same-level sections of continuous text in this manner: whereas bold or small cap styles may be used for headings, which stand at a distinct level, italics are frequently embedded within sentences. It is for this reason that italic TE has been chosen as the form under investigation here, and not due to any assumption that different types of TE have clearly defined uses. In fact, it is expected that the relatively weak conventions related to TE (arising from technical limitations and variation between written genres) will mean that the contrast caused by modulating between one form of TE and another is what drives any possible interpretation of meaning, rather than each form of TE being reliably tied to particular meanings.

Indeed, each form of TE (e.g. italic TE, bold TE, etc.) seems to be deployed for a variety of purposes, with considerable overlap. Restricting the sphere of interest to intra-sentential TE, some of the uses italics are put to include the indication of titles of published work, words under discussion as words, foreign terms, and contrastive focus (CF). It is the latter use which is the primary interest here, but the ways in which these apparently disparate uses are connected will also be discussed: briefly, TE usage can be divided into those uses which interpolate sections of language which are not analysed according to the surrounding grammar of the sentence, such as words, foreign terms and titles; and uses which appear more connected to sound features, including intonation features. The contention here is that this 'phonological' use of TE can indicate the same type of intonation as is associated in speech with pitch accents, and is used to mark CF.

1.4 Psycholinguistics

The central component of this study consists of two psycholinguistic reading experiments. A psycholinguistic approach has been adopted here for several reasons. Firstly, it is used as a way to sidestep a certain danger when conducting a linguistic analysis that involves typography. Any such linguistic analysis may be confused by aesthetic considerations: printed text is an artificial medium that requires conscious design, and typefaces are for the most part created by those trained in art and design. Due to this, much has been written on typography from the perspective of those who are interested in tradition, convention and aesthetics. Such analysis typically takes the kind of prescriptive approach that is anathema to linguistic inquiry (Walker, 2001): style guides may speak of 'appropriate' times to deploy bold or italic emphasis, but the reasoning behind this is often left opaque, and may well be rather arbitrary. For example, in the popular prescriptivist book *Eats, Shoots and Leaves*, Truss (2003) calls italics 'distracting', and some style guides go so far as to label TE as a feature of 'poor' writing (Saldanha, 2011, p.426).

To take a linguistic view of TE, then, a line in the sand must be drawn between such prescriptive/subjective opinions, and how this aspect of the written word is used and processed. Psycholinguistics stands firmly within an empirical, scientific tradition of research into language, and as such should be a suitable tool with which to develop a dispassionate and unbiased picture of this phenomenon. Moreover, the use of psycholinguistic experimentation rather than qualitative methods should allow for the isolation of those parts of the typographic stream that communicate *meaning* from those parts that are purely aesthetic, or those that are motivated more by convention than communication.[3]

Secondly, psycholinguistics has been perhaps the only branch of modern linguistic inquiry that has taken a consistently deep interest in the written form. As such, a large number of insights into how readers process the written word has already been gathered, and powerful, validity-tested frameworks and metrics to gather more data have been established. These offer an opportunity to investigate subtle effects, which is important here, as it may be that the effects sought here are rather subtle. As will be discussed later, TE-marking of CF is evidently an optional part of written language; whole texts containing no instances of it can easily be written and read. This suggests that if TE-marking of CF does have

---

[3]The relationship between convention and meaning is, however, complex, and will be discussed further below.

an effect on processing, it may not be as strong as those associated with pitch accents, which are mandatory in speech.

It is important to note that the psycholinguistic approach adopted here is slightly distinct from a large amount of work in that field. Psycholinguistic inquiry often takes advantage of the fact that reading is a complex mental operation that happens to be both conveniently manipulable and relatively widespread (at least in terms of the participant pools available to the typical academic researcher). It is also significant that reading requires concerted instruction and effort, unlike spoken language acquisition. Reading thus may be used as a point of access to a range of underlying cognitive processes, which may be the more general building blocks that are recruited to achieve reading. The following from Huey is often quoted in the literature:

> And so to completely analyze what we do when we read would almost be the acme of a psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind, as well as to unravel the tangled story of the most remarkable specific performance that civilization has learned in all of its history (Huey, 1908, p.6; quoted in Rayner and Reichle, 2010).

Similarly, Henderson et al. (1995) state that 'reading and language processing … [is] an area that has been central to the study of human cognition since the inception of modern cognitive psychology.'

In *this* study, however, psycholinguistic techniques and theories are not being used to access language processing in order to gain insights into the inner workings of cognition per se. Rather, the interest here centres on the relationship between written and spoken emphasis, and the effects which the particular characteristics of written emphasis may cause in terms of reading and processing strategies. The investigations into processing of text conducted here aim to interrogate this relationship, and as such they are underscored by a relatively traditional linguistic stance. This is in keeping with Crystal's call for typographical linguistics, which suggests a discipline more linguistic than psychological in flavour.

1.5 Experimental approach

The present study centres on experiments which test the hypothesis that TE-marking of CF will facilitate sentence processing, occasioning fewer refixations

to a contrastively focused word, and to an antecedent focused word with which it contrasts. This hypothesis is based on the following assumptions:

1. TE-marking of CF is processed;
2. CF *without* any explicit marking is harder to process;
3. Processing of focus occurs at a higher level than syntactic and semantic processing, and will thus manifest in later measures which capture refixations, as mental models of information structure are constructed and then re-assessed.

The experiments involve the use of an eye tracker to measure processing. The use of eye tracking to investigate cognitive processes exploits the reliable connections between cognitive load and eye movements; broadly, the duration of fixations on words reflects the level of processing difficulty they occasion, in real time. Eye tracking is attractive here as it allows for a higher level of ecological validity than other paradigms used for reading studies, such as self-paced reading tasks: during eye tracking, participants can read normally, and stimuli can be presented as whole, simultaneously available sections of text. This is particularly useful in the case of TE, which requires a visual contrast that would be lost if sections of text or words were presented in isolation. Eye tracking also allows for the analysis of both initial and subsequent fixations on given words: this should allow for a certain degree of insight into which stages of processing are affected.

This study is not entirely deductive and experimental in nature; given the lack of work on typographical linguistics, the research area is wide open, and any methodologies devised to test it will be relatively unproven and tentative. As such it would be a wasted opportunity to *only* run and present tightly-controlled psycholinguistic experiments. For one thing, the nature of eye tracking is such that huge amounts of data are available from any given experiment, often beyond that which is required for testing a given hypothesis: although the raw output of an eye tracker principally consists of a string of x–y coordinates (and perhaps pupil diameters), a large number of metrics can be extrapolated by combining these data streams with constructions included in the experimental design.

Whilst it would be problematic from a hypothesis-testing perspective to cherry-pick interesting results and retroactively construct theories around them, it would be equally unwise to ignore 'incidental' data. Next to nothing is known about how TE affects eye movements: this study represents an opportunity to begin building a body of knowledge regarding this. For example, it seems intuitive that there may be 'eye-catching' effects of TE, which may prompt non-linear reading pat-

terns, and perhaps confound the results of the experiments. Eventualities such as this will be considered alongside the central hypothesis-driven approach.

Furthermore, the experiments conducted here can only hope to address a minuscule area of the discipline of typographical linguistics, and progress in this area will require that broader issues are raised and discussed. As such, aside from presenting tests of the hypothesis presented above, this study aims to discuss TE and typographical linguistics from wider perspectives, in order to create a jumping-off point for future work in this area.

Note on perspectives This study concentrates almost entirely on reception rather than production. It is the reader, rather than the writer, that is at the heart of this investigation. Although completely divorcing reception from production is clearly impossible (and as a consequence, any investigation into either one will reveal a lot about the other), it does tend to be far easier to investigate reception, and a lot more work has been done on reception in the areas relevant to this study.

The sphere of interest here is also almost entirely situated within the English language, and does not attempt to make claims about other languages. Plenty of research on focus and information structure is based on other languages, not least because concerted efforts have been made to establish the existence of cross-linguistic information structure phenomena. No such wide-ranging claims will be attempted here, but the text will acknowledge when the literature it rests on pertains to languages other than English.

Finally, it is important to restrict this study to the English writing system. Writing systems vary considerably across languages. It would undoubtedly be interesting to compare TE in an alphabetic system such as that used for English with syllabic or logographic systems such as those used for Korean and Chinese; in fact, typographic conventions can vary even between such close cousins as the English and French writing systems.[4] However, just as cross-linguistic analysis is out of scope here, so is cross-writing system analysis.

---

[4]Writing systems should not be seen as distinct from languages (see §3.4.1). They will not be referred to here by the names of the languages they are used to write, but instead will be referred to as *the English writing system*, *the Chinese writing system* etc., with the assumption that the particular system referred to is the most prevalent associated with that language.

1.6   Organisation of thesis

Chapter 2: Contrastive focus reviews the literature on information structure, concentrating on the phenomena of presentational and contrastive focus. It examines how they can be marked and manipulated in speech, and then looks at research on how focus is processed.

Chapter 3: Typographic emphasis discusses the device upon which this study is based. The term *typographic emphasis* is defined and discussed from typographic and linguistic stances. Prior research on different uses of TE are reviewed, and a classification of the device into phonological and interpolational forms are proposed. Differences between TE and intonational emphasis are then discussed. The chapter ends looking at research relevant to the processing of TE: firstly general work on reading is reviewed, with emphasis placed on the role of sound in reading; secondly previous psycholinguistic studies on TE itself are be examined.

Chapter 4: Methodological approach & Experiment 1 first reviews the most suitable psycholinguistic techniques for investigating how TE is processed, and justifies the choice in this study of eye tracking. The connection between eye movements and processing is explored. The second part of the chapter is dedicated to a full report on the first experiment conducted for this study, which was run as a pilot.

Chapter 5: Experiment 2 reports the full-scale successor to the pilot Experiment 1. It discusses those elements of the design that were altered, placing particular emphasis on the addition of a contrastive focus as a factor. It reports and discusses the results, which were the product of a more in-depth and rigorous analysis than that conducted Experiment 1.

Chapter 6: Discussion provides summaries and comparisons of the results of both experiments, before considering the results of the weightier, more powerful Experiment 2 in detail. The limitations of the experimental approach adopted here are discussed, and suggestions are made for the future direction of research into TE processing.

Chapter 7: Conclusion gives a final summary of this study.

Chapter 2

Contrastive focus

2.1 Introduction

To address the research question of whether written emphasis can function in a similar way to intonational emphasis in marking contrastive focus, the concepts of *focus* and *contrastive focus* must be defined and discussed. *Focus* is an *information structure* phenomenon (Halliday, 1967) which indicates the presence of alternatives to a linguistic element (a word, a phrase, or even a complete expression) that has been made prominent (Krifka and Musan, 2012; Rooth, 1992). In speech this prominence is typically achieved by marking a word with a pitch accent; in writing it is typically unmarked (Stolterfoht et al., 2007).

(2) Q: Where did you go?
     A: We went to FRANCE.

In (2) the pitch accent on *France* indicates that this word has focus, and that a set of contextually relevant alternatives to its referent (e.g. a list of other holiday destinations) are necessary for the hearer to interpret the utterance.

Contrastive focus is a subtype of focus, in which a focused element is judged by the speaker to be unexpected by the hearer, and made correspondingly more prominent to indicate this (Coulson et al., 1998; Zimmermann, 2007). Whilst this higher prominence is often associated in the literature with it-clefts (e.g. É. Kiss, 1998), in speech this higher prominence can also be achieved with pitch accents alone (Selkirk, 2002; Watson et al., 2008), and in writing it can be achieved with typographic emphasis (McAteer, 1992; Sanford et al., 2006; Saldanha, 2011; Ashby, 2017).

(3) A: You went to Italy.
     B: We went to FRANCE.

In (3), *France* bears a stronger pitch accent than that which marks the non-contrastive focus in (2); furthermore, some would argue it might bear a different type of accent (e.g. Selkirk, 2011; Katz and Selkirk, 2002). This marks *France* as having contrastive focus, indicating to the hearer that the speaker thinks the hearer will not anticipate this word being uttered; to put it another way, they judge that the hearer does not expect them to make the proposition which results from uttering this word in this context.

This chapter reviews the literature on the concepts of focus and contrastive focus, which are, as the above outlines suggest, rather complex phenomena. They are discussed here principally in relation to spoken English. In part, the emphasis on speech is due to the fact that the majority of work in this area has related to speech rather than writing (although the final section of this chapter examines reading research). It is also due to the intent for this chapter to establish a background with which to contrast the subsequent chapter, *Typographic emphasis*. In this way there is a progression across the two literature review chapters, moving from the intensively researched area of information structure towards more preliminary, exploratory work on typographic emphasis.

2.1.1   Organisation of chapter

§2.2 introduces the overarching area of **information structure**, of which focus forms a central component.

§2.3 examines **focus** itself. Two central theoretical approaches are identified; alternative semantic theories (§2.3.1), and newness and assertion-based theories (§2.3.2). The alternative semantic approach is adopted from this point on, and used to discuss what can be focused (§2.3.4), before a discussion of how focus is marked (§2.3.5).

§2.4 deals with **contrastive focus**. Semantic (§2.4.1) and pragmatic (§2.4.2) approaches to contrastive focus are examined; the latter is preferred in this study. The corrective use of contrastive focus is also discussed (§2.4.3).

§2.5 touches on the issue of **focus sensitive particles**.

§2.6 reviews work relevant to the issue of **focus processing**. It asks, and attempts to answer, two questions which are pertinent to this study's aim of experimentally testing aspects of how focus is processed.

Note on terminology It is important to note that the information structure field has been 'notoriously variable' (Arnold et al., 2013, p.403) for some time, and is thus at a stage of 'unsettled terminology' (Matić and Wedgwood, 2012, p.3) and subject to 'terminological profusion' (Levinson, 1983, p.x). The review presented here takes a practical approach in selecting its preferred theoretical and terminological positions, as it supports an experimental study which must operationalise *from* theory rather than subsume itself *within* theory. However, it does seem likely that the terminological complexity of the information structure field reflects an intrinsic complexity, which should not be ignored or rejected outright.

2.2 Information structure

In speech, the phenomenon known as *focus* involves the emphasis of elements of a sentence, principally via the placement of *pitch accents* (PAs), which can be defined as 'an intonationally-signalled perceptual prominence of a word' (Birch and Clifton, 1995, p.366). Two sentences with different foci are shown in (4); small capitals represent PAs.

(4)   a. Mary gave John LESSONS.
       b. Mary gave JOHN lessons.

Intuitively, there seems to be a difference in meaning between the two sentences, triggered by the differing location of the single PA present in each: in (4a) one appears on *lessons,* and in (4b) one appears on *John.* At a semantic level, however, both sentences appear to mean the same thing. They would both have the same truth value for a given real-world situation: if Mary did give lessons to John, both sentences would be true, and if Mary did not, both would be false.

However, these sentences would have different felicities in different discourse contexts. Wh-questions are frequently used as a diagnostic for these phenomena (Partee, 1991; Zimmermann and Onea, 2011): whilst (4a) would be a felicitous response to *What did Mary give to John?,* (4b) would be infelicitous as a response to the same question (though it would be perfectly acceptable as a reply to *Did Mary give Mark lessons?*). What focus is doing here must be connected to discourse. To define this relationship more precisely, and thus to define focus itself, requires consideration of the wider area into which focus falls, *information structure* (Halliday, 1967).

Communication demands that the parties interacting must maintain models of each other's mental states. A sentence will only be informative if the data it contains adds to or updates the hearer's mental models. Therefore, speakers must first have an idea of what hearers do and do not know. If a speaker does not know what a hearer knows, that speaker risks either pointlessly giving the hearer what they already know, or giving them something that they are unable to integrate into their mental models, on account of it being unconnected to what they already know (Lambrecht, 1994; Clark and Brennan, 1991; Stalnaker, 2002).

The mutually-updated knowledge space which arises between language users when they communicate is known as the *common ground* (Stalnaker, 2002), and can be modelled as a set of mutually accepted propositions which are connected to a set of discourse referents (Krifka and Musan, 2012). Over the course of communication, knowledge must be continually updated as new propositions are made and new referents are introduced.

It appears that natural languages facilitate the updating of the common ground by allowing the grammatical encoding of the informational status of sentence constituents. It is this encoding which defines the *information structure* (IS) of the sentence (Halliday, 1967):

> The central function of information structure lies in the optimization of the processing of information coded in a linguistic utterance in light of the specific discourse needs of the interlocutors at the time of utterance (Zimmermann and Onea, 2011, p.1652).

IS encompasses a number of phenomena which have effects that cannot be accounted for purely in terms of truth-conditional semantics, as has already been seen in the alternation in (4). These phenomena result in the existence of sentences that may have identical truth conditions, but differ in *how* they convey their meanings:

> In every language there is an array of sentences that … differ only in the way they say what they say about the world. These alternatives cannot be used interchangeably in context. Use of particular alternatives is constrained by the current context (Vallduví and Engdahl, 1996, p.461).

IS can be seen as adding something to sentences that sits alongside their meanings. Vallduví and Engdahl (1996) suggest this 'something' takes the form of instructions, directing the hearer in how to interpret the status of parts of the

sentence; that is, they facilitate the integration of referents into the hearer's mental models. Similarly, Chafe (1976) offered the term *information packaging,* with the implication that these devices let speakers 'package' up sentences in order to assist the integration of information by the hearer.

The devices responsible for this packaging take the form of *informational articulations* (Vallduví, 1993; c.f. *information structure contrasts* Lambrecht, 1994). These articulations divide sentences into different notional functions. Theories of IS differ in terms of which articulations they recognise; the *focus–background* articulation is of principal interest here (*background* refers to anything not focused), but the *topic–comment* articulation (Lambrecht, 1994; Krifka, 2007) should also be briefly mentioned at this point. Broadly, this divides sentences into that which is being talked about, the *topic*, and that which is being said about the topic, the *comment* (some theories refer to the articulation as *theme–rheme*, e.g. Halliday, 1967). In English, topics are closely connected to grammatical subjects (Chafe, 1976), although this is not obligatory (Lambrecht, 1994).

The IS field diverges significantly here, with many different articulations having been proposed. For example, some theories (such as the Praguian approach of Hajičová, 1991; Hajičová and Sgall, 2004) treat topic and focus as a single informational articulation, rather than recognising focus–background and topic–comment separately; others have tried to unify those two bipartite articulations using a tripartite articulation, such as Vallduvi's proposal of sentences being divided into *focus* and *ground,* with the latter itself being divided into a *link* and *tail* (1993). In this study, which is concerned particularly with the sub-type of focus known as contrastive focus, a full overview of the many, many theoretical distinctions that have been proposed for this area would be extraneous: theories of how the entirety of IS works are less relevant than theories of how focus works. Before turning to the latter it is, however, useful to provide a broad overview of some of the tendencies and strands within IS research; this will be done with reference to the debate over the correct position of IS within linguistics.

2.2.1   The position of information structure within linguistics

The introduction given above has discussed the relationship between IS and discourse, and has thus implied that IS is closely associated with the domain of pragmatics. To place IS completely *within* pragmatics might be overly simplistic, however, given that certain aspects of it are also connected to syntax and semantics. Some of these connections will be touched upon now, despite them being covered in more detail later on; doing so should help to explain why the litera-

ture is so variable, and why it encompasses such multifarious views regarding the correct position of IS (Arnold et al., 2013; Matić, 2015). This will also establish a background against which to discuss the approach towards IS adopted here.

In terms of connections to syntax, the link between topic and grammatical subject has been mentioned, but there can also be a syntactic dimension to focus-marking. Focus marked by intonation was shown in (4), repeated here as (5a, b), but it can also be indicated by the syntactic device of *clefts* (É. Kiss, 1998). Clefts are syntactic constructions which divide what would otherwise be a single clause into one superordinate and one relative clause (Collins, 1991): in (5c), the former is *John* and the latter is *Mary gave lessons to*.

(5)   a. Mary gave John LESSONS.
      b. Mary gave JOHN lessons.
      c. It was JOHN Mary gave lessons to.

Furthermore, (5a), in which the PA is on the final noun, appears to be less marked than (5b), in which the PA appears earlier: focus is to some extent associated with word order (Birner, 1994; Matić, 2015).

IS and semantics also intersect. Whilst (5a, b) show a focus alternation in which there is no effect on truth-conditional semantics, there are, however, elements of IS that do have semantic effects, most notably *focus-sensitive particles* (Aloni et al., 1999).

(6)   a. John only showed Mary the PICTURES.
      b. John only showed MARY the pictures.

The presence of the particle *only* means that focus alternation between (6a) and (6b) could result in the two sentences bearing different truth conditions: if John showed Mary the pictures as well as something else, then (6a) would be false, but (6b) would be true (focus-sensitive particles are returned to below in §2.5).

For researchers working within formal linguistic traditions, the fact that IS intersects with pragmatics, syntax, and semantics makes its position within linguistics ripe for debate. Matić (2015) identifies two areas of contention here, the first being the question of whether IS belongs to semantics or pragmatics: he suggests that the effects of focus-sensitive sensitive particles, as in (6), have been taken by some, such as Rooth (1992), to demonstrate the semantic nature of IS. Other approaches here are to take IS as existing at the semantics–pragmatics interface, as Glanzberg (2005) does, or to view IS cognitively, as a domain which connects the core grammatical modules (syntax, phonology etc.) with the general

cognitive faculties that handle pragmatics and inference (Zimmermann and Onea, 2011). The second question Matić (2015) identifies is that of whether IS structures such as focus are in fact dedicated to encoding IS, meaning that they are a part of the grammar (Lambrecht, 1994), or whether instead they are pragmatic effects (Matić and Wedgwood, 2012).

Fully addressing these issues is a substantial task, and one clearly out of the scope of the present study. It is instead important to be clear about the approach towards IS and focus that will be taken here. Erteschik-Shir (2007) identifies two broad strands in the IS literature: research that concentrates on the structural side of IS, and research that concentrates on the pragmatic side of IS. Work concentrating on structural aspects of IS phenomena often takes a syntactic approach, as might be expected. An example of this would be work on *focus projection*, which seeks to explain how PAs on single words can mark larger foci by invoking deep structure and discussing propagation up trees (e.g. Selkirk, 1995; Gussenhoven, 1999; see §2.3.6). Meanwhile, studies that concentrate on the pragmatic aspects of IS phenomena are more concerned with how their semantic properties create different pragmatic effects; an example is the pragmatic account of *contrastive focus* from Zimmermann (2007). Note within these strands, the positions various authors adopt on the questions outlined by Matić (2015) do not necessarily neatly line up; this categorisation is based broadly on perspectives adopted and interests pursued, rather than on specific theoretical positions.

The present study will align itself more closely with the pragmatic strand of research. This study must control, rather than investigate, some of the more complex structural aspects of focus: in particular, focus projection, clefts, and focus-sensitive particles. The aim here is to investigate whether the marking of contrastive focus with typographic emphasis (TE) has an effect, meaning that all other structural factors must be controlled. The necessity for control, however, means that these factors must be attended to: they will be discussed below.

2.3   Focus

Theories of focus must account for two particular qualities of the phenomenon that seem intuitively apparent: wh-question congruence, and newness. The first has been mentioned briefly above: focused elements can be seen to correspond to the wh- part of a wh-question (Selkirk, 1995; Büring, 2007).[1] (7b) would be

---

[1]Observations of the connection between focus and wh-questions go back a long way: Krifka (2007) cites the 19th century German linguist Hermann Paul on this issue.

an appropriate response to (7a); the focused noun *snowboarding* corresponds to *what*. Making the focus incongruent with the question results in infelicity, as in (7c).

(7) a. What sort of thing is Alexa into?
     b. She's into SNOWBOARDING.
     c. # SHE'S into snowboarding.

The second quality of 'newness' is rather more difficult to pin down. The intuitive sense is that focus corresponds to something which might variously be considered 'unactivated', 'unrecoverable', 'informative', or 'new' in relation to the present discourse. If (7a, b) are considered together as a dialogue, then in (7b) the pronoun *she* can be seen as 'old' information, as its referent is already in the common ground. The focused noun *snowboarding* is 'new', as it does not exist in the common ground. Even cursory consideration of this brief summary of 'newness' will reveal several caveats and issues; these will be returned to below.

Matić (2015) identifies two theories of focus that are currently dominant. The *alternative semantics* approach to focus is particularly influenced by the quality of wh-question congruence. It holds that focus indicates alternatives relevant to the interpretation of the sentence (Rooth, 1985, 1992, 2016; Krifka, 2007). The *assertion-based* approach of Lambrecht (1994) is arguably more heavily influenced by the quality of newness—or at least, the research tradition that has concentrated on this quality—although it also covers wh-congruence. It sees focus as being that part of a sentence in which what is being asserted differs from what is presupposed.

An alternative semantics view of focus will be preferred here: it sees rather wider support in the literature than Lambrecht's theory, but has also influenced psycholinguistic work which is relevant to the present study (Fraundorf et al., 2010; Fraundorf et al., 2013; but c.f. Cowles, 2003, who adopts Lambrecht's theory to investigate IS processing).

2.3.1 Alternative semantics accounts of focus

A usable definition of focus must a) be precise enough to isolate its characteristics without encompassing other phenomena, and b) pertain to the property itself; that is, it must not confuse the property with how it is marked, or how it is interpreted. The view of focus adopted here is based on two definitions provided by Krifka and Musan (2012; based largely on Krifka, 2007), which are terse, compelling, and, in combination, compliant with these criteria. The first is:

> Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions (Krifka and Musan, 2012, p.7).

This is subsequently elaborated upon to create the following complementary definition of focus:

> The property of a string … to point out the existence of alternatives either to the expression or to its denotation (Krifka and Musan, 2012, p.8).

The central notion here is that of alternatives being indicated. This is a widely accepted account of focus (e.g. Krifka, 2007; Molnár, 2002; Zimmermann and Onea, 2011) which is based largely on the alternative semantics of Rooth (1985, 1992, 2016), a semantic theory which attempts to formally model various phenomena using sets of alternatives. Rooth describes it as 'a semantic framework that finds application in the analysis of questions, focus, disjunction, negative polarity, presupposition triggering, and implicature' (2016, p.19).

Alternative semantics draws on Hamblin's (1973) account of the semantics of wh-questions, and as such is particularly pertinent to the wh-congruence quality of focus. Broadly, Hamblin proposed that the semantics of a wh-question can be taken as corresponding to a set of propositions. Therefore the question *who killed him?* corresponds to the set {*Dave killed him, Ian killed him, Jenny killed him, …* } The purpose of a question is thus to prompt the listener to specify one member of that set in response: 'Pragmatically speaking a question sets up a choice-situation between a set of propositions, namely, those propositions that count as answers to it' (Hamblin, 1973, p.48).

Alternative semantics applies this idea of a set of propositions to the semantics of focus. It posits that, for a focused element, a set of propositions exists as a second type of semantic value alongside the ordinary semantic value. This set can be derived by substituting the focused element with other elements of the same type; the set thus contains *alternative* propositions (Rooth, 1992). For a phrase such as (8a), the alternative set would be the set of all possible propositions with the form *I'd like the X* as in (8b).

(8)   a. I'd like the WATER.
      b. {I'd like the COFFEE, I'd like the BEER, I'd like the JUICE, … }

Substituting the focused element for an element of the same type means that many alternative sets would contain astronomically large, if not infinite numbers of alternative propositions: (8b) would thus not only contain propositions

in which every possible drink is requested, but also every possible object of *like*; {*I'd like the Large Hadron Collider, I'd like the very large blue dog, …* }. Clearly there must be constraints on these sets; otherwise, from a pragmatic perspective they would consist largely of irrelevant propositions, and from a cognitive perspective they would be near-impossible to compute. Both of these concerns are external to a formal semantic theory such as this, but alternative semantics does allow for a contextual constraint on alternative sets (motivated by the contextual properties of focus-sensitive particles): Rooth assumes that this constraint is 'fixed pragmatically' (1992, p.15). How this is achieved is an open question (Umbach, 2004); it will suffice here to assume that alternative sets consist of alternative propositions which are relevant to the present discourse.

How might alternative sets be useful to the hearer? It has been stated here that IS is the encoding of the informational status of parts of sentences, in order to facilitate updating of the common ground (Krifka, 2007; Zimmermann and Onea, 2011). If the process of updating the common ground is conceptualised as a series of questions (Roberts, 1996), many of which are not explicitly posed, then alternative sets connect a sentence to a question that is currently open in the discourse:

> This set of focus alternatives serves as a preliminary evaluation context, for example, by relating the utterance containing the focused constituent to a particular relevant question in need of resolution at the current state of discourse (Zimmermann and Onea, 2011, p.1652).

This, then, can account for the connection between focus and wh-question congruence. To be clear, however, the questions referred to here are not necessarily 'questions' at all, but instead are the same underlying things which drive the creation of actual explicit questions: they might be better described as 'informational needs' (Krifka and Musan, 2012).

Returning now to the definitions of focus proposed by Krifka and Musan (2012): as they note, it is essential to avoid confusing focus as a property with those devices which are used to mark the property. Their definition is thus agnostic regarding how focus is marked. Aside from this confusion having complicated the study of focus across the literature, avoiding the inclusion of how focus is marked also allows the definition to work cross-linguistically. Focus marking will be discussed further in §2.3.5.

Similarly, the definition does not detail what effects focus might have. Krifka and Musan (2012) argue that there are different interpretational effects that can arise from the use of focus, as with other linguistic categories such as time, case, or gender. On this basis sub-types of focus can be identified: these will be discussed in §2.4.

Two final points regarding this definition must be noted: Firstly, Krifka and Musan (2012) define focus as the *indication* of an alternative set, rather than the set itself. This distinction must be made because, as they point out, alternatives exist for almost every element of every sentence, not just focused elements.

Secondly, where they say that focus indicates alternatives either to the string's 'expression or to its denotation', they are referring to a distinction first noted by Krifka (2007), who observed that the focus semantic value can be a set of alternatives either to the denotation of the focused element, or to its form. He defines *expression focus* as occurring when focus marks alternative forms. These forms are on the surface level of linguistic representation, meaning that they can range from comprising sentence constituents to involving sub-word units (as in corrections of pronunciation). *Denotation focus* is that type hitherto discussed here, in which alternative meanings are being marked. To illustrate expression focus, Krifka gives the example *Grandpa didn't [kick the bucket]$_F$, he [passed away]$_F$*, and argues that the two expressions are not alternative to each other in terms of their meanings, as they both have the same property DIE; instead, he suggests that the alternatives are the *connotations* of the expressions.

2.3.2 Newness and assertion-based theories of focus

The insight that focus is in some way connected to 'newness' is arguably an older one than that which draws a connection to alternatives (Jackendoff, 1972 makes one of the earliest mentions of alternatives). Halliday's pioneering paper *Notes on Transitivity and Theme in English* (Halliday, 1967) introduced several ideas that have endured throughout much of the subsequent work on focus, including the concepts of newness and informativity:[2]

> Information focus is one kind of emphasis, that whereby the speaker
> marks out a part (which may be the whole) of a message block as

---

[2]Whilst inquiry into what is now principally referred to as *focus* ranges back to at least the mid-19th century (see González, 2001), it was Michael Halliday who, drawing on ideas developed by the Prague School in the early 20th century, spearheaded Western research into the area (Brown, 1983; Krifka and Musan, 2012).

that which he wishes to be interpreted as informative. What is focal is 'new' information; not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse. If we use the … term 'given' to label what is not 'new', we can say that the system of information focus assigns to the information unit a structure in terms of the two functions 'given' and 'new' (Halliday, 1967, p.204).

In writing this, Halliday also introduced a terminological headache for the field that has continued to this day: 'newness' has become a somewhat controversial choice of word to indicate what focus marks (Steedman, 2000). A tension in the terminology is already apparent in the above quotation, in which Halliday has to flag his use of *new* as not meaning that focused words have not yet been mentioned—i.e. that *new* is not being used with its typical sense at all.[3]

(9)   a. I think I last saw you in SCHOOL.
      b. #I think I last saw him in SCHOOL.

Take (9) as being said in greeting by one person directly to another person standing alone. In (9a), the personal pronoun *you* does not need to have focus, despite its referent not having been mentioned; the hearer is, of course, immediately able to resolve the referent—it is the hearer themselves. Compare this with (9b), which also contains a backgrounded pronoun, *him*. This case is infelicitous, not because the referent of *him* must be 'new' despite being presented as given, but because it is impossible to resolve if there is no relevant third person in the common ground.

The literature aligned with the 'newness view' thus tends to concur that it is not the status of referents themselves that determines whether they receive focus; rather, focus is assigned based on the speaker's assessment of the hearer's ability to resolve referents. This means that temporal terms such as *new* are problematic, and there have been many attempts to define replacements; a large list could be provided here, but this would not be particularly useful. The term to be preferred here is one Halliday (1967) himself briefly uses when clarifying his own terminology: *recoverable*, which is also used by Lambrecht (1994).

Prince (1981a) notes that non-focused, background components can be inferred or accommodated, and suggests that both set–element and element–set re-

---

[3]To give Halliday his due, a) he draws his axis between *given* and *new* rather than between *old* and *new*, and b) he does admit that *given*, his chosen word for what is being termed *background* here, is 'rather inappropriate' (Halliday, 1967, p.204)

lations can serve as a route for the hearer to resolve referents. In Prince (1981b) she proposes a taxonomy of different focus–background elements, in which 'inferable' is one category. This means that elements can be left as background when the speaker judges them to be inferable by the hearer.

(10) Darren's looking tough these days.
      His neck is the size of a TREE TRUNK!

In (10), *his neck* has not been previously mentioned and thus is not in the common ground, but is presumed to be inferable from the context: the person previously referred to with *Darren* is expected to have a neck—and the speaker may also assume the size of one's neck is commonly associated with the quality of 'looking tough'.

Halliday's definition of focus implies a second way of looking at the distinction between focused and backgrounded parts of a sentence, alongside the issue of recoverability; he suggests that the unrecoverable part is that which is *informative* (focus is 'that which [the speaker] wishes to be interpreted as informative' Halliday, 1967, p.204). A problem with this definition is that a focused phrase or referent cannot convincingly be argued to be 'informative' on its own, as any useful definition of 'information' in linguistic terms must surely involve relations between entities: 'Information … is a propositional notion: A sentence, or the proposition it expresses, can be informative (w.r.t. a given stock of knowledge), but parts of sentences cannot' (Büring, 2007, p.449). Thus in the sentence *I like MARY*, it is the relationship between the two referents that is informative, rather than the focused word *Mary*. As Lambrecht (1994) notes, even single word sentences are propositional: if one sees some money on the floor and exclaims *Money!*, the proposition is that the speaker has noticed the money, or that the money is on the floor, etc.

On this basis, Lambrecht (1994) argues against what he calls the 'segmentational view' of information, claiming that sentence constituents cannot correspond directly to such notions as recoverability: information is conveyed by relations between the denotata of sentences, and propositions. Within this relational, propositional view of IS, Lambrecht advances a different approach to focus: he sees it as being the part of a sentence in which the assertion differs from the presupposition.

The view of sentences as consisting of these two types of proposition is widely accepted: 'If I say that the Queen of England is bald, I presuppose that England has a unique queen, and assert that she is bald' (Stalnaker, 1977, p.149). In keeping with his 'non-segmentational' approach, Lambrecht states that it is frequently

impossible to directly connect the presupposition or the assertion to particular sentence constituents. In his view assertion involves relating the presupposition (or set of presuppositions) to a new proposition: this proposition is thus 'superimposed' on the presupposition. Focus is the part of this proposition which is *not* the presupposition; in other words, the subject of the assertion is the presupposition, and its predicate is the focus.

(11) Q: Where did you go last night?
A: I went to the MOVIES.

Lambrecht notes that most accounts of focus would classify the NP *the movies* in (11) as being focused, but that in his account, however, the 'newness' (and thus the focus) is neither the constituent, nor its designatum. Instead, what is new and focused is the way in which the designatum of *the movies* functions as the predicate of the assertion, which has the presupposed open proposition 'speaker went *somewhere*' as its subject (Lambrecht, 1994).

2.3.3 Arguments against 'focus' as a category

An alternative position on focus has been advanced by Matić and Wedgwood (2012), who warn against what they see as a logically suspect trend within IS research:

> First, identify linguistic phenomena that, in use, are systemically accompanied by a particular interpretive effect (say, the evocation of alternatives or an indication of how an assertion relates to some implicit or explicit question). Then label each of these phenomena as realising/being dedicated to that interpretive effect (i.e. 'a focusing strategy' or 'focus construction'), thus unifying cross-linguistic phenomena (Matić and Wedgwood, 2012, pp.5–6).

Matić and Wedgwood thus suggest that what is analysed as 'focus' is actually a set of different pragmatic effects, and that even if these all arise from some underlying universal cognitive principles, it does not follow from this that focus is a single *linguistic* category.

This position is not well supported by other sources, although aspects of it have an intuitive appeal: certainly, their criticism of circular reasoning seems to be pertinent to a wide variety of research, inside and outside of linguistics. The aim of this study is not to establish linguistic categories, however: it can safely leave the status of focus as a category as an open question.

2.3.4   What can be focused?

The literature often uses examples in which the focused element is an NP—as have the examples given thus far. However, many grammatical categories can have focus, as illustrated in (12).

(12)   a. Q: How does Louise get to work?
           A: She CYCLES.
       b. Q: Which box is it?
           A: The CARDBOARD one.
       c. Q: Did he take a gun?
           A: He took THE gun.

Lambrecht (1994) argues that informational articulations such as focus can occur anywhere that predicate–argument relationships occur.[4] Focus can therefore be placed on phrases and expressions as well as on single words. The fact that focus can range from single words up to entire expressions (Krifka, 2007) is what has motivated the choice here of the term *element* (short for *linguistic element*) over *constituent,* which is also frequently used.

This study will restrict most of its discussion and all of its experimentation to NPs. This is motivated both by the tendency in the literature to use them, and by the fact that NPs constitute a large, well-studied category which lends itself to use in experimentation.

Another tendency in the literature is for frequent reference to be made to '*the* focus' of a sentence. Büring (2007) cautions against this phrasing for one compelling reason: multiple foci can be present in one sentence (this is also noted by Dik et al., 1981; Krifka, 2007; Rooth, 2016), as illustrated by (13).

(13) (What happened to my harp?)

       a. SOMEONE STOLE it.
       b. SOMEONE SENT it to NORWAY.

For the same reason, it is also unwise to refer to the background as being the 'complement' of focus. Where this thesis refers to 'the focus', this does not imply a view whereby a sentence can only contain one focus.

---

[4]For this reason it is unlikely that interjections could have focus, unless it were the *expression focus* posited by Krifka (2007), as in *he didn't say 'GOSH', he said 'WOW!'* (see §2.3.1).

2.3.5 How focus is marked: pitch accents

Three structural phenomena have been mentioned in connection with focus: pitch accents (PAs), clefts, and focus-sensitive particles. Of these, only PAs and clefts can be said to mark focus; focus-sensitive particles are rather 'associated with focus', in that they interact with a focused element to create semantic effects (Aloni et al., 1999; Ippolito, 2007).[5] Clefts will be discussed below in relation to contrastive focus (§2.4.1.2).

The literature tends to assume all focus in English *must* be marked with a PA (this is rarely stated explicitly, although Zimmermann, 2007 is an exception). This thesis will follow suit: it is hard to conceive of a focus without a PA, although empirical support for this would be useful. A PA was loosely defined above as 'an intonationally-signalled perceptual prominence of a word' (Birch and Clifton, 1995, p.266). Such a definition is sufficient for the majority of theoretical accounts of focus, which tend to abstract PAs into a single binary feature which either appears on a word or does not. However, abstracting PAs in this way obscures some subtleties in the grammar of intonation, and in the way it might indicate different types of focus. Furthermore, as the present study aims to compare PAs with TE, differences in their forms may thus be important. Comparisons between PAs and TE must wait until TE has been fully introduced, in Chapter 3, but some detail regarding PAs can be given now.

PAs are particular tones, or ordered pairs of tones, which form part of the contour of an intonation phrase (Pierrehumbert, 1980).[6] The intonation contour is also used for purposes such as marking questions (Hadding-Koch and Studdert-Kennedy, 1964), but this is performed by tones distinct from PAs (Pierrehumbert, 1980). PAs serve to make a particular syllable of a word perceptually prominent (Bolinger, 1958), although in the case of two-tone PAs the second tone may fall on the subsequent syllable (Pierrehumbert, 1980).[7] Typically the syllable which is made prominent is the lexically stressed syllable of the word (Hirschberg, 1993), although PAs may be placed on other syllables in instances of expression focus (e.g. *I said it's IMpossible, not POSsible*).

---

[5]This phrase was first coined by Jackendoff (1972), and then notably taken up by Rooth in his PhD thesis (1985).

[6]Following much of the focus literature, as a convenient simplification which does not appear to obscure many subtleties of focus, PAs will be treated here as operating on the level of sentences rather than intonation phrases.

[7]PAs are indicated in this thesis by small capitals on whole words, rather than on the graphemes corresponding to the stressed syllable as is done in some of the literature (e.g. Selkirk, 1995).

Despite the phonological literature recognising several types of PAs, the relationship between these and focus structure is a long-standing point of contention (Repp, 2010). In particular, there has been debate over whether a 'contrastive accent' exists within English. This will be discussed below in the section on contrastive focus (§2.4.1.3).

2.3.6 Focus projection

As has been said, focus can appear at many levels from the word up to the sentence, yet it is principally marked on a single syllable within a single word by a PA. It thus follows that the IS processor must have a method to compute the scope of a focus based on this low-level marking. Consider (14).

(14) a. Q: What happened?
          A: [He took POISON.]$_F$
      b. Q: What did you say he took?
          A: He took [POISON.]$_F$

The square brackets in (14) demarcate the *focus domain*, 'any structure that can be marked as being entirely [+focus] with only a single accent' (Gussenhoven, 1983, p.62). (14a) has *broad focus*, and (14b) has *narrow focus* (Ladd, 1980; cited in Weber et al., 2006). The answer to (14a) would require the entire sentence *he took poison* to have focus, as the question indicates an informational need for a full proposition. The answer to (14b) would only require the noun *poison* to have focus, as the question indicates an informational need for a single referent.

(15) a. Q: Who bought a book about bats?
          A: [MARY]$_F$ bought a book about bats.
      b. Q: Who did something with bats?
          A: *[MARY bought a book]$_F$ about bats.

Chomsky (1971) called the mechanism by which focus domains are calculated *focus projection*. Syntactic accounts of this assume that focus can propagate in some way up trees (e.g. Selkirk, 1995; Gussenhoven, 1999): particularly influential in this area is the Selkirk's theory of focus projection (1995), in which focus a) can project from the pitch-accented head of a phrase to the whole phrase, and b) can project from an internal argument of a head to the head itself. This allows focus to spread leftwards as in (14), but prevents it from spreading rightwards as

in (15). Whether focus projects to a constituent in line with these rules is dependent on discourse context; due to this, Selkirk defines these rules as regarding the *licensing* of focus projection.

Within this area there has been some debate over whether focus can actually be projected to whole sentences, and whether focus is always connected to a constituent (Gussenhoven, 1999). Alternative non-syntactic theories of focus projection have also been presented (e.g. the optimality theoretic account of Féry and Samek-Lodovici, 2006). The important issues in terms of this thesis are that a) focus domains can vary in size with no explicit marking and b) the size of the focus domain is dependent on discourse context. This study aims to investigate focus processing, but little psycholinguistic work has been done on how focus projection affects processing; there might, for instance, be a correlation between cognitive load/processing time and the distance across which a focus spreads, but this is not known. It is important therefore to tightly control focus breadth, ideally to the point of only examining focus on individual words.

2.4   Contrastive focus

Up to this point, focus has been discussed here largely as an undivided, monolithic concept. Some sources have, however, proposed sub-dividing focus into various categories. Although this is a contentious issue in focus research, those sources which do recognise a division of focus tend to coalesce around two particular types, *presentational* and *contrastive* focus (Zimmermann and Onea, 2011).[8] Presentational focus (PF) (Halliday, 1967; É. Kiss, 1998; Gussenhoven, 1999; Selkirk, 2002; Selkirk, 2007) can be seen as the less marked form, which simply marks elements which correspond to informational needs. Aside from appearing less marked in linguistic terms, PF is less marked in metalinguistic terms: the literature often uses the term *focus* alone to refer to presentational focus (É. Kiss, 1998; Gussenhoven, 2006).

Contrastive focus (CF) is typically used to refer to focus for which the alternative set is already particularly salient in the discourse (Zimmermann and Onea, 2011). Intuitively, instances of CF seem to involve the explicit exclusion of one or more of these alternatives, and this seems to be connected to the size of the alternative set.

---

[8]Gussenhoven (1999) goes as far as describing seven sub-types.

(16) a. A: What did Laura buy yesterday?
      b. A: I heard Laura bought hair straighteners yesterday.
      c. B: She bought a HAIRDRYER.

The focus in (16c) is presentational in response to (16a). Simplified somewhat, the alternative set for the focused word *hairdryer* can be taken to be the set of things which Laura would be likely to buy. The focused word is not placed in in opposition to any particular member of this large, open set. If the question preceding the focus is (16b), however, then the focus is contrastive, as it has a closed alternative set, limited to *hair straighteners*. Speaker B's response in (16c) thus seems to place *hairdryer* in opposition to *hair straighteners*.

CF is crucially important to this study, which aims to investigate whether placing TE on a minimally-marked focused word will result in that word being correctly analysed as having CF, instead of incorrectly analysed as having PF. Therefore, a position is required on how CF differs from PF in terms of the pragmatic effects it creates. The position taken here will be one influenced by the pragmatic accounts of Zimmermann (2007) and Calhoun (2009). Such higher-level accounts cannot dispense with the semantic side of CF, however: this must be reviewed first.

A prominent strand of research has attempted to identify grammatical distinctions between PF and CF, with some sources presenting prosodic evidence for this (Selkirk, 2002; Katz and Selkirk, 2011), whilst others have found syntactic evidence (É. Kiss, 1998, 2017). This study need not choose between the two positions in this debate, i.e. whether there is a grammatical distinction between the two sub-types of focus, or whether they only differ in terms of their pragmatic uses (Zimmermann, 2007; Zimmermann and Onea, 2011): those arguing for a grammatical distinction do not dispute the existence of the pragmatic effects with which this study is concerned.

2.4.1 Semantic approaches to contrastive focus

It should first be noted that care must be taken not to confuse *contrast* as it relates to contrastive focus with *linguistic* contrasts. Language is fundamentally contrastive: it consists of a system of contrasts which are employed to communicate meaning (Saussure, 1959). As Bolinger said about linguists, 'our main business is establishing contrasts at all levels' (1961, p.84). In this respect, every level of language involves items that contrast with each other, from phonemes, to morphemes, to syntactic and information structural constructions. If CF is to be dis-

tinguished from PF, the contrast it involves should be distinguished from this type of *linguistic* contrast.

Contrast involves opposition, such that an item that contrasts with other items is opposed to them (Molnár, 2002; Repp, 2010). Bolinger (1961) suggested that, at one level, this means that every focus, PF and CF alike, is intrinsically contrastive: putting it in alternative semantic terms, the focused element always contrasts with every item in its alternative set.

The idea of every focus being contrastive does not fit, however, with a second quality which some sources associate with contrast, that of *dominance* (Molnár, 2002) or *preference* (Bolinger, 1961). This means that the contrasting item is not only different to all others, but is in some way on a different level in terms of a given property (Molnár, 2002): in other words, 'this thing is not like the other things, *and* it's better/worse/etc.'

Various semantic qualities have been proposed to account for the differences between CF and PF, all arguably arising out of contrast's two basic qualities of opposition and dominance. The two qualities which tend to be most discussed pertain to a) the bounds of the alternative set, and b) the exclusion of alternatives (Repp, 2010).

2.4.1.1 Contrastive focus and the bounds of the alternative set

The bounds of the alternative set can be seen to vary for the focus in (16c). Recall that *hairdryer* has been proposed to have PF when it is preceded by (16a): the alternative set here is extremely large, comprising anything that Laura might have bought yesterday. This set can thus be said to be *open* (É. Kiss, 1998—although it should be recalled that alternative sets are always assumed to be restricted to some extent by pragmatic factors; Rooth, 1992). The alternative set for the focus on *hairdryer* when it is preceded by (16b) is far smaller in comparison, and can thus be said to be *closed*: it is limited to the focused *hair straighteners*.

É. Kiss (1998) argues that CF does not require a closed alternative set, but that the alternative set must have members which are easily *identifiable* by the participants in the discourse.

(17) a. Somebody who was in that room at 5pm stole the laptop, but I don't know who was there.
   b. Somebody who was in that room at 5pm stole the laptop; the only people who ever go in there at that time are Dave and Sally.
   c. It was DAVE who stole it.

Both (17a) and (17b) result in closed alternative sets for the focus *man* in (17c). However, in (17a) the members of the set are not identifiable, meaning that the focus on *Dave* can not be contrastive, according to É. Kiss. (17b) makes the members identifiable, and thus the focus is contrastive.

However, intuitively it is difficult to accept that *Dave* is more contrastive in (17b, c) than it it is in (17a, c). This is despite the fact that the focus in (17b, c) has both of the properties of CF as specified by É. Kiss; it has a closed alternative set, and the members of that set are explicitly identified. Krifka (2007) points this issue out using a similar example: he suggests there seems to be the same level of contrast on (18c) regardless of whether it is preceded by the open alternative set of (18a), or the closed set of (18b).

(18) a. Q: What do you want to drink?
      b. Q: What do you want to drink, tea or coffee?
      c. A: I want TEA.

It thus seems that, although restriction is a necessary quality of CF, 'a restricted alternative set is no guarantee for a contrastive interpretation' (Repp, 2010, p.1337).

Krifka's approach to this is to claim that CF should be restricted to instances in which the common ground contains the alternative set for the CF. According to his argument, the common ground itself does not contain the proposition in (18b); rather, this proposition, which defines the alternative set for the focus in (18c), exists within common ground *management*. This is a term Krifka introduces to refer to what is effectively a second mutually shared knowledge space, containing information about the 'communicative interests and goals of the participants' (Krifka, 2007, p.17), and existing alongside the *common ground content*. Questions indicate informational needs, and thus place propositions within common ground management.

2.4.1.2 Contrastive focus and exclusion

In the restricted set versions of both (17) and (18), the CF involves exclusion: the former implies the thief was *Dave* and not *Sally*; the latter that *tea* is desired and not *coffee*. Repp gives the follow exclusion requirement:

> Contrast marking on an element α indicates that there is a salient alternative α in the immediate context for which what is said about α does not hold (Repp, 2010, p.1338).

This type of exclusion is not a property of all types of focus. Whilst any focus might appear to exclude alternatives, in PF this exclusion is an implicature rather an entailment, as it can be cancelled. Umbach (2004) argues that this implicature of exclusion is created due to the maxim of quantity, according to which a cooperative speaker will not withhold information (Grice, 1989). Umbach shows that it can be cancelled, using the example (19).

(19)   a. Ronald made the HAMBURGERS.
          b. Ronald made the HAMBURGERS, and he made the SALAD, too.

Assuming an alternative set which includes various dishes, (19a) can generate the implicature that Ronald *only* made the hamburgers, and did not make the salad, the chips, etc. However, (19b) cancels this, suggesting that *hamburgers* did not truly contrast with every alternative.

The focus in (17c), repeated here as (20a), is not cancellable in the same way: (20b) is infelicitous.

(20)   a. It was DAVE who stole it.
          b. # It was DAVE who stole it, and SALLY.

What makes the difference here is the presence of a cleft. It appears that the cleft prohibits cancellation in (20). There is an intuitive sense that clefts, unlike PF, do involve exclusion. More precisely they appear to be *exhaustive*, in that they select a single alternative and exclude all others. There has been debate over whether this exhaustivity is entailed, conversationally implicated, or conventionally implicated (Horn, 1981; Vallduví, 1993; Velleman et al., 2012). Such semantic distinctions are out of scope here; it suffices to take clefts as a syntactic device which add an exhaustive quality to a focus, and thus make it a CF. Moreover, the present study has reason to avoid clefts altogether in its experimental items: whilst several similar studies have used clefts to manipulate focus (see §2.6.1.3), the resulting syntactic differences could introduce confounds to the design planned here, as will be discussed in Chapter 4.

2.4.1.3   Distinctive marking of contrastive focus

É. Kiss (1998) actually argues that CF *must* be marked in English with clefts, in addition to having the other properties of CF discussed above, i.e. a limited alternative set of identifiable members. She argues that CF is a cross-linguistic, syntactically distinct type of focus which is marked in Hungarian by a particular

syntactic position, and in English by clefts. This *identificational focus* is claimed to either be exhaustive or contrastive, or both exhaustive and contrastive, depending on the language: in Hungarian and English it is exhaustive, but can either be contrastive or not.

The claim that exhaustivity is not always connected to contrastiveness is controversial; it has been criticised with respect to Hungarian by Kenesei (2006). That CF in English must be marked with a cleft is also somewhat of an extreme claim, which is not well-supported by other sources, which tend to accept instances of CF with no clefts (e.g. Gussenhoven, 2006; Krifka, 2007). Subsequent work on focus in Hungarian has moved away from É. Kiss's proposal of a cross-linguistic, CF-like identificational focus (including her own work: see É. Kiss, 2017).

Some studies have argued that CF is marked by intonation features distinct from those marking PF. This goes back at least as far as the work of Daniel Jones, who stated 'contrast-emphasis is expressed mainly by intonation' (Jones, 1956; cited in Bolinger, 1961). Bolinger argued that there is no 'uniquely contrastive pitch' (1961, p.87), but instead that the degree of stress on words varies gradually according to how much contrast they carry. Empirical results here have been mixed: for example, Selkirk (2002) presented evidence suggesting that CFs are marked not only with a different type of accent to that marking PFs, but are followed by a phrase break which does not occur for PFs. Katz and Selkirk (2011), however, found no evidence for different types of PAs for CF, and instead found evidence that those marking CF are simply more phonetically prominent.

2.4.2 Pragmatic approaches to contrastive focus

Zimmermann (2007) adopts the position that CF differs from PF only gradually, and suggests that this is generally the case in intonation languages such as English. He then argues that there is no one-to-one relationship between grammatical marking, or emphasis, and particular focus uses (such as presentational or corrective focus); he claims this is the case not only in English but across 'languages as diverse as Finnish and Hausa' (2007, p.150). Because of this, it is not possible to predict whether a focused element will be marked for contrastive focus or not, if one only attends to its inherent properties.

Zimmerman defines his *Contrastive Focus Hypothesis* thus:

> Contrastive marking on a focus constituent α expresses the speaker's assumption that the hearer *will not consider* the content of α or the speech act containing α *likely to be(come) common ground.*

Zimmermann therefore claims that the contrast in CF is between what the speaker asserts, and what the speaker thinks is expected by the hearer. This is a different perspective than that of more semantically-orientated accounts: it does not wholly deny that CF involves restricted, explicit alternative sets, or that it involves the exclusion of alternatives; rather, these are side effects from the proposed contrast between mental states.

(21) Q: Who (all) did you invite?

A: PETER, I invited (but nobody else).

According to Zimmerman, in (21) the CF functions by rejecting the presupposition that multiple people were invited; a semantic approach would instead take it to function by excluding the alternative set of *Peter,* which would consist of the other relevant invitees. Note that this account does not discount alternative semantic theories of focus itself; it instead takes contrast to be a pragmatically manipulated effect which makes use of focus.

Zimmermann bases his theory on a judgement of likelihood: 'The less expected a given content is judged to be for the hearer … the more likely a speaker is to mark this content' (2007, p.147) However, he does not argue that the degree of marking is dependent on the degree of expectation, but rather that when a particular focus reaches some threshold of unexpectedness, the speaker will be motivated to mark it. There is no particular device used in this case; Zimmerman simply posits the use of a 'non-canonical' (and thus more marked) form.

Calhoun (2009) proposes a theory of CF which also involves a role for expectations, although she accepts a more 'traditional' semantic account of CF arising from a highly salient alternative set. She argues that speakers create CF by increasing this salience through the manipulation of hearer expectations. If a word is more marked than the hearer expects, then the bigger this difference, the more salient the alternative set—and thus the more contrastive the focus:

> If a word is more prominent than expected given its properties and position in the metrical structure, then it is more likely to be contrastive; conversely, if a word is less prominent than expected, it is more likely to be backgrounded. Expectations are built on the prosodic, syntactic and discourse properties of the words in the utterance (Calhoun, 2009, p.9).

Although Calhoun principally discusses prosodic marking, this theory implies that essentially any kind of marking can create CF, as long as it results in increased

prominence. Under this theory, then, in writing TE would presumably be as likely to prompt the interpretation of emphasised elements as other types of marking such as clefts (Fraundorf et al., 2013).

2.4.3   Corrective contrastive focus

Gussenhoven (2006) and Krifka (2007) both suggest that CF is typically used for correction.

> When the focus marks a constituent that is a direct rejection of an alternative, either spoken by the speaker himself ('Not A, but B') or by the hearer, the focus is *corrective* (Gussenhoven, 2006).

Correction involves an antecedent proposition being rejected by the assertion of a contradictory proposition. In the sentence conveying the contradictory proposition, what is focused is that which makes this proposition differ from the antecedent (van Leusen, 2004; Krifka, 2007). Some type of parallelism must exist between these two propositions: van Leusen (2004) suggests this is principally parallelism of thematic structure, and points out that parallel surface syntax is not required. This means that corrective CF can occur in a rephrasing of an antecedent proposition as in (22b), or with repetition as in (22c).

(22)   a. A: Matthew said he would visit on MONDAY.
      b. B: No, it was TUESDAY that he said he'd visit on.
      c. B: Matthew said he would visit on TUESDAY.

The mechanics of correction mean that the focus within a corrective sentence automatically has a restricted, explicit alternative set, limited to the referent of the antecedent focus: the focus on *Tuesday* in (22b, c) has a restricted alternative set containing only *Monday* from the antecedent proposition conveyed by (22a).

In terms of the pragmatic accounts of CF, under Zimmermann (2007) corrective sentences would clearly be judged by the speaker to be unexpected for the hearer, as they directly contradict a proposition already within the common ground. The issue of marking is interesting here; Calhoun's theory (2009) of CF being marked by any kind of relative increase in prominence might imply heavy accenting on the *Tuesday* in both (22b) and (22c), but the near word-for-word repetition in the latter might itself constitute a form of marking.

Some sources, such as Zimmermann (2007), use the term *corrective focus* in place of *contrastive focus*. However this would obscure a distinction Krifka (2007)

draws between corrective focus and *confirmatory focus*, which he proposes to occur when a proposition already made is included in a focus alternative. For example, a response to (22a) with confirmatory focus would be *Matthew said he would come on Monday*. Krifka suggests that the common ground must at this point include other alternatives under consideration: in (22) this would be the idea that Matthew might come on a different day.

2.5 Focus-sensitive particles

Aloni et al. (1999) list the notable *focus-sensitive particles* in English as being *only, even, too, also, always, usually, never* and *because*. The semantics of these words (c.f. *focusing adverb* in Rooth, 1992) depend on the IS of their containing sentences (Aloni et al., 1999).

(23) What did John show Mary?

    a. John showed Mary the PICTURES.
    b. John showed MARY the pictures.
    c. John only showed Mary the PICTURES.
    d. John only showed MARY the pictures.

The manipulation of focus in (23a) and (23b) has no effect on the truth condition of the sentence. The presence of *only* in (23c) and (23c), however, means that the same focus manipulation could now result in the sentences having different truth conditions: if in fact John showed Mary the pictures and something else, then (23d) is true but (23c) is false. *Only* excludes all of the alternatives to the focus with which it is associated (Rooth, 1992; Ippolito, 2007). How this is done is subject to a debate similar to that surrounding how CF excludes alternatives, regarding whether this it is a matter of entailment, implicature, etc. (Ippolito, 2007).

Focus-sensitive particles are evidently a complicating factor that should be avoided when experimentally manipulating TE in order to find out whether it can mark CF: if a particle such as *only* precedes a target word, it may prompt a contrastive reading regardless of the presence of TE, for example. Furthermore, if focus-sensitive particles were to be used in the experiments here, the potential interaction with truth conditions could create difficulty in constructing experimental items.

2.6 Focus processing

In order for comprehension to be achieved during language reception, multiple linguistic processes must occur (Wolf and Katzir-Cohen, 2001; Radach and Kennedy, 2004). At a low level, lexical processing must enable words to be identified from their orthographic representation. At a higher level, words must be placed into syntactic, semantic and pragmatic structures (Rayner et al., 1983). Above, focus has been argued to have a pragmatic component, which would suggest that it requires higher-level processing. Indeed, even if it is not viewed as being principally pragmatic, and regardless of the debate over the correct position of IS within linguistics, it seems clear that focus itself must involve some higher-level processing.

The central hypothesis of this study is that TE-marking of CF will facilitate sentence processing. The most direct experimental approach to investigate this would clearly involve the manipulation of the presence of TE on CF target elements within written sentences. A method of measuring how participants processed these sentences would be required, and any differences in processing between [+TE] and [−TE] sentences could then be analysed.

If this basic experimental approach is to be taken, then two questions regarding focus processing arise:

1. **To what extent is focus processed online?** This determines the choice of a method by which to measure processing, as well as the design of items; measuring online processing is arguably a more complex task than measuring offline processing. In addressing this question, it is worth also examining the underlying assumption that focus is processed at all.

2. **Are contrastively focused elements processed differently to presentationally focused elements?** The hypothesis assumes that TE-marking of CF will facilitate processing because it will make the CF status of the target elements easier to resolve: that is, it will enable them to be distinguished from PF elements. If different focus types have different effects on processing, this may help determine whether the participant has analysed the target as PF or CF.

The experimental work relevant to these questions does not fall neatly within a particular area of psychological research, and literature that refers to 'information structure processing' or 'focus processing' is relatively rare. This is undoubtedly due to the complex relationship between IS and several different linguistic

domains. Fortunately, however, many studies do relate to aspects of IS and focus processing; it is just that they fall within a range of different research areas.

Given the clear connection between focus and prosody, several studies have investigated how PAs affect the comprehension of speech. Other work on focus processing falls within *discourse processing,* which covers the ways in which language users relate different pieces of information within a text to each other; processes such as anaphora resolution thus fall under this umbrella, as well as focus and other IS phenomena (Sanders and Canestrelli, 2012).

There is also reason to look towards the syntactic parsing literature here, despite the fact that parsing is not directly of interest to this study (which does not aim to manipulate syntax in its experimental items, and indeed must take care to avoid the inclusion of syntactic constructions associated with focus, such as clefts). This is because research on IS phenomena has often been conducted with a view to how it affects syntactic parsing; to put it another way, the issue of discourse processing has been looked at through a syntactic lens.

Before continuing, a caveat: Cowles (2012) warns that some of the psychological literature uses *focus* to refer to the focus of attention, or the most salient referent within a sentence—both definitions which do not coincide exactly with linguistic definitions of *focus*. This issue is difficult to avoid completely, and is compounded by the fact that different experimental studies are based on different conceptions of focus: in particular, some are framed in terms of the newness account of focus rather than the alternative semantics/pragmatic approach adopted here. As will be seen below, Benatar and Clifton (2014) suggest that the variety of positions on focus which have been adopted across different studies may be what accounts for some inconsistent findings.

2.6.1 To what extent is focus processed online?

2.6.1.1 Connections between pitch accent and processing

The question of whether focus is processed online must be preceded by the question of whether it is processed at all. Some of the clearest evidence that focus does indeed have an effect on language processing comes from work in the field of speech comprehension, which has demonstrated clear links between the prosody of sentences and how they are comprehended.

Early general findings in this area showed that the prosodic structure of speech is attended to, and that structures coherent with the semantic content of sentences facilitate processing (see Cutler et al., 1997 for a summary). A more recent exam-

ple of this type of work is the study by Braun et al. (2011), in which participants completed word-monitoring, lexical decision and semantic categorisation tasks whilst listening to sentences which had either normal or aberrant intonation contours. Participants were slower in all tasks for the aberrant intonation contours.

In terms of PAs in particular, accented words appear to be processed faster than unaccented words. Shields et al. (1974) and Cutler (1976) tested this using phoneme monitoring tasks. In this type of task subjects listen to spoken stimuli and press a button as soon as they hear a target phoneme: the dependent variable is the latency of phoneme detection, which is taken to increase in line with processing cost (Foss, 1969). Shields et al. (1974) presented participants with sentences in which the target phoneme appeared word-initially either in a word with a PA, or in an unaccented word; they found that reaction times were faster for the words with PAs. Cutler (1976) used tape splicing to place neutrally-accented target words in two contexts: one where the intonation contour of the preceding words was consistent with that preceding a pitch-accented word, and one where the contour 'predicted' an unaccented word. The predicted pitch-accented words were recognised faster than the unaccented words, even though the target words themselves had no accent.

In light of this, the question then arises, does such a processing benefit result from PAs being taken as marking something linguistic, or just because an accented word is more salient? Cutler and Foss (1977) noted the latter would be a reasonable assumption, and that the benefit might result from accented words being easier to perceive unambiguously, due to factors such as increased volume, and vowels becoming closer to their 'citation form'. Terken and Nooteboom (1987) posed an obvious question in response to this: why would *any* words be left unaccented, if there is this processing benefit for accented words? They hypothesised that accented and unaccented words are processed in different ways: accentuation marks unrecoverable words and prompts greater attention to the properties of the word, to enable recognition of a previously unactivated word/referent; deaccentuation marks recoverable words and prompts the processor to restrict its search space to activated, recoverable words/referents.

Terken and Nooteboom (1987) tested this hypothesis using a task involving the verification of descriptions of changes to an on-screen set of alphabetic characters, arranged in a particular configuration. A letter would change position, and a spoken description would be presented (e.g. 'the p is on the right of the k' Terken and Nooteboom, 1987, p.149). Participants pressed buttons to indicate whether the description was correct or not. The recoverability of target words

(the grammatical subject of each description) was manipulated by a series of 'context' descriptions, which would either mention the subject of the upcoming target description, or not, thus making it either more or less recoverable. Target words were either accented or not. They found a significant interaction between recoverability and accentuation: unrecoverable words were processed faster when they were accented, and recoverable words were processed faster when they were unaccented. This finding supported their hypothesis of a qualitative difference in how accented and unaccented words are processed.

Further support for the basic hypothesis that PA-marked focus structure is processed came from Bock and Mazzella (1983). They tested the relationship between accents and focus by preceding target sentences with a sentence which either established a congruent or incongruent focus context, as illustrated in (24).

(24)   a. (Congruent) ARNOLD didn't fix the radio.
       b. (Incongruent) Arnold didn't FIX the radio.
       c. DORIS fixed the radio.

In the congruent condition (24a), the focus on *Arnold* creates a restricted alternative set of one member for the focus *Doris* in the target sentence (24c). The effect is of (24c) having CF, with an apparent corrective purpose. In the incongruent condition (24b), the informational need established by the focus on *fix* is not fulfilled by the focus on *Doris* (24c); this renders the target sentence infelicitous.

Comprehension times for the congruent condition were faster than those for the incongruent condition. Again, this is support for the idea that the processor is not simply aided by the increased salience and clarity of focused words, but that the processing benefit is related to the congruence of accentuation.

The above results suggest that PAs are attended to and that incongruence between PAs and focus results in slower processing. They answer the question posed here with a strong 'yes': focus is processed, and in particular the intonational marking of focus has measurable effects. This knowledge underscores the present study, which aims to discover whether typographic marking of focus in writing is similar.

Such studies do not demonstrate that focus structure is processed online, however, as they do not involve fine-grained (i.e. word-by-word) measurements. To seek evidence of this requires looking at experiments using different paradigms— but before doing this, models of how online language processing occurs should be considered.

2.6.1.2   The role of higher-level information during parsing

It is uncontroversial to state that much of language processing is extremely rapid, and occurs on a word-by-word, online basis. Work on syntactic parsing in particular has provided abundant evidence that language users construct interpretations of the linguistic input incrementally (Rayner and Clifton, 2009). It does not necessarily follow that focus is processed online, however: indeed, serial parsing models hold that such high-level information is not immediately accessible online (Ferreira and Clifton, 1986). Whilst consensus has been moving from serial models towards parallel models (Traxler, 2014), even within these models there are still some aspects of parsing which are more serial than parallel.

Broadly, in serial models one analysis of the syntax of a sentence is created at a time, and is then revised if necessary. Higher-level information (such as pragmatic information) is not available during parsing at all (Fodor, 1983; Ferreira and Clifton, 1986). In parallel models multiple analyses can exist simultaneously, and higher-level information is available during parsing (Lewis, 2000).

Serial models reflect the modular account of language processing, in which specialised cognitive systems or *modules* encapsulate the information they operate on, such that information being processed within a module is inaccessible to other modules (Fodor, 1983). In serial models an initial parse is constructed using only syntactic information, i.e. the grammatical categories of the words. During comprehension, listeners/readers tend to make similar initial parses of ambiguous sentences, suggesting that serial parsers would need certain strategies or algorithms; various strategies have been proposed for this, such as *minimal attachment* (Frazier and Fodor, 1978).

Models of parsing must also be able to account for higher-level information as well, as in some cases a correct parse must depend on contextual information. Serial models do this by positing two processing stages, in which a syntactic parsing stage is followed by a higher-level processing stage. This is why the influential *garden-path* model of Frazier and Rayner (1982) and similar models (e.g. Frazier and Fodor, 1978) are often referred to as *two-stage models*.

Parallel models differ from serial models in that they do not posit wholly encapsulated processing modules, and allow for interactive processing. They allow for multiple analyses to be constructed simultaneously as words are processed, and then involve some type of selection process in order to arrive at the correct analysis (Lewis, 2000).

Fully parallel models, which allow higher-level information to be available continually throughout parsing, are not supported because of the contrary evi-

dence that comes from *garden-path* effects. Garden-path sentences are syntactically ambiguous up to a certain disambiguating point: for example, in Bever's classic garden-path sentence *the horse raced past the barn fell* (1970, p.316), when the processor reaches *fell*, an analysis in which *raced* was the main verb can no longer be accepted, and instead the processor must arrive at the correct analysis in which *the horse raced past the barn* is taken as a reduced relative clause (Hale, 2003). A serial model in which the former analysis was being constructed would need to reanalyse the previously-processed input at this point in order to arrive at the correct analysis, and this would have a processing cost. A fully parallel model would allow the processor to be maintaining both possible analyses when it reached 'fell', and would be able to select the correct one at no cost (Lewis, 2000). However, evidence of strong garden-path effects, i.e. processing costs on arrival at the point of disambiguation, suggests that parsing must be at least partially serial (Gibson and Pearlmutter, 2000; Frazier and Rayner, 1982).

As was stated above, consensus has been moving from serial models towards parallel models (Traxler, 2014). If the parallel approach is valid, then this is fortunate for this study, inasmuch as these models predict behaviour more amenable to the detailed measurement of higher-level processing. The more interactive the processing, the more clearly the effects of focus should manifest during word-by-word processing. These differences in online processing may manifest as different patterns of activation and inhibition within the brain, which can be measured with neuroimaging, or as different external behaviours, which can be measured with techniques such as eye tracking.

2.6.1.3   Evidence from eye movements and brain imaging

Eye tracking and electroencephalography (EEG) can measure online processing at a high temporal resolution (Rayner, 1998; Brennan, 2016). Both can use spoken or written stimuli; eye tracking is, for various reasons, particularly suited to conducting reading studies. A fuller discussion of these technologies will be presented in Chapter 4, along with justification for why eye tracking has been selected for use in the present study. For now it should be stated that much of the use of eye tracking assumes that the length of time written words are looked at indexes their processing 'cost' (Just and Carpenter, 1980; Rayner, 1998; Kliegl et al., 2006).

Reading studies on focus present a problem, however, in that written stimuli contain no prosodic information, yet focus is typically marked with PAs. This difference is, of course, an issue at the heart of the present thesis, and will be

discussed in detail in the following chapter; however, this is brought up here, in terms of reading research concerned primarily with focus, because focus is frequently manipulated in written stimuli using clefts.

Clefts offer a convenient method of placing focus on words that would otherwise be backgrounded (Cowles, 2003). However, they are strongly associated with CF (see §2.4.1.2). This is not only the case in the extreme claims of É. Kiss (1998), who, as was mentioned in §2.4.1.3, argues for mandatory marking of CF with clefts in English; whilst most theories allow for CF to be marked in a variety of ways other than with clefts, they do not tend to claim anything other than a close association between CF and clefts. An alternation between unfocused, background elements and elements in the superordinate clauses of it-clefts is thus an alternation between background and CF—yet this is often treated as an alternation between background and focus in general, with little consideration as to whether PF and CF are processed differently.

The eye tracking studies of Morris and Folk (1998) and Birch and Rayner (1997) are examples of this issue; stimuli from these studies are given in (25).

(25) a. It was the gardener who watched as the barber trimmed the mustache.
      b. While the gardener watched, it was the barber who trimmed the mustache.
      c. It was the suburb that received the most damage from the ice storm.
      d. Workers in the suburb hurried to restore power after the ice storm.

In (25a) and (25b), which come from Morris and Folk, the nouns *barber* and *gardener* are alternately focused with an it-cleft, or left as background. In (25c) and (25d), which come from Birch and Rayner (who took their stimuli from Birch and Garnsey, 1995), the noun *suburb* is focused with an it-cleft in (25c) and is in the background as an object of a PP in (25d).

These stimuli do establish alternations between focus and background, but it is arguably *contrastive* focus, and treating this as equivalent to PF (as these studies do) may be problematic. Bearing in mind no context sentences were presented before these stimuli, it might be argued that the lack of an explicit, restricted alternative set to the focus on *suburb* in (25c) precludes it from being CF; however, *the most damage* in the relative clause implies that there is a relevant alternative to *suburb,* and it seems reasonable to take *suburb* as typically being a member of a small set such as {*city, country, suburb*}: that is to say, the restricted alternative set can be accommodated here.

Neither of these studies explicitly discuss CF in comparison to PF. This is an example of what Benatar and Clifton (2014) criticise as a general tendency in

experimental focus research, whereby various different manipulations of focus are used but then lumped together. They suggest this is why Birch and Rayner (1997) found results which conflicted with a study they conducted subsequently (2010): the earlier study found that focused elements (as Birch and Rayner defined *focused*) were looked at for longer than background, but the later study found the opposite. The Morris and Folk (1998) study discussed above also found that focused elements were looked at for less time.

Attempting to address this issue, in their own eye tracking study, Benatar and Clifton (2014) used written dialogues in which a target phrase was either backgrounded or had PF, using a manipulation that did not rely on clefts. Each stimulus started with a wh-question, but the target phrase in the reply was not congruent with this: instead, in the background condition the question included a hypernym/synonym of the target phrase, as in (26a), in which the target phrase is *the doctor*. In the PF condition, the question included the words *something* or *someone* in order to request information in addition to the referent of the wh-word, as in (26b).

(26)   a. (Background condition)
            A: Tell me, when did Caitlin leave to go to the cardiologist?
            B: I believe she left to go to the doctor a little before 11 this morning.
        b. (Presentational focus condition)
            A: Tell me, when did Caitlin leave to go somewhere?
            B: I believe she left to go to the doctor just a little while before 11 this morning.

As Benatar and Clifton note, the target phrase does not correspond to the wh-question in either condition, and thus might not be seen as having focus in either condition; they actually refer to their conditions as *given* and *new*. These labels have been changed to *background* and *presentational focus* here because the target phrase in (26b) should indeed have a focus: it addresses the informational need established by A's use of *somewhere*; furthermore, recall that there is no reason why a sentence cannot have multiple foci (see §2.3.4). Benatar and Clifton found that the PF target words were looked at for longer than the background words. This stands with the similar results of Birch and Rayner (1997), but against Birch and Rayner (2010) and Morris and Folk (1998).

Clearly there are issues in this literature, some of which may have been the cause for some of the conflicting results. However, given that all the studies dis-

cussed above did find that their various manipulations of focus and background had significant effects on eye movements, it is at least safe to conclude that they constitute evidence for focus being processed online.

2.6.2 How are contrastively focused elements processed in comparison to presentationally focused elements?

As has been discussed, instead of comparing CF and PF, experimental work looking at focus has tended to compare CF words with backgrounded words. There are two studies which have attempted to specifically compare these types of focus: those of Benatar and Clifton (2014) and of Cowles (2003). Benatar and Clifton have been highlighted above as having been critical of the imprecision regarding what type of focus has been compared to background in previous studies; in keeping with their concerns over this, they conducted an eye tracking experiment that attempted to compare CF with PF. This experiment used a similar format to that of its predecessor (also discussed above), with the addition of background information used to more clearly establish discourse referents, as shown in (27).

(27)   a. (PF condition) *John and Mary are working today.*
            A: Did you tell someone to go home early?
            B: I told John, but I don't know if it was a good idea.
        b. (CF condition)
            A: Did you tell Mary to go home early?
            B: I told John, but I don't know if it was a good idea.

In (27b), the target word *John* is intended to have corrective CF, as speaker B rejects the antecedent proposition that they have told *Mary* to go home early. Under the approach to CF advanced by Krifka (2007), the use of questions here might raise concerns about whether the alternative sets to the target foci exist within the common ground, or within common ground management, as the antecedent proposition is part of a question (see the discussion above in §2.4.1.1). Under the Zimmermann (2007) approach preferred here, however, speaker B's direct contradiction of the antecedent proposition is enough to suggest that speaker B judges that speaker A will not expect them to utter the word *John* at this point; this qualifies *John* as having corrective CF.

Benatar and Clifton found a significant effect of contrast, with participants looking at CF words longer than at PF words. They suggest that this processing cost is related to the complexity of updating mental models: invoking the 'file

change semantics' of Heim (1983), they suggest that PF requires a new 'entry' in a reader's mental model of the discourse (i.e. their representation of the common ground). CF requires an existing entry to be amended; although Benatar and Clifton do not speculate further, perhaps this is more costly because that entry must first be accessed. They do, however, extend this argument to their finding (discussed above in §2.6.1.3) that PF words are looked at longer than background words: they suggest that it is less costly to access an entry in the mental model, which is what happens for background words, than it is to create a new entry, which is what happens for PF words.

Cowles (2003) conducted an EEG study in which the presence of CF within a target sentence was manipulated by altering preceding sentences. Cowles used either a *wh*-question or a yes/no question to achieve this: both conditions for one of her items are presented as (28).

(28) a. (Contrastive focus context) Q: A butcher, a chef and a specialist were in the kitchen of a posh restaurant. They had started up the business together. It was successful, but they were very busy. All of them wanted everything to be perfect, but only one had time to stop and check the soup. Which one tasted the soup?

  b. (Presentational focus context) Q: The kitchen of a posh restaurant was filled with people trying to get orders filled. Near the door was a butcher and another person. A group of cooks was clustered around a stove, including a chef and a specialist. There was a pot of soup in the corner that was almost ready to be served. Did anyone taste the soup?

  c. (Target sentence) A: After a moment, the butcher tasted the soup.

The manipulation here relies on creating alternative sets of different sizes: in the contrastive focus condition (28a), the alternative set to the target phrase *the butcher* is restricted to three people (a butcher, a chef and a specialist). In the presentational focus context, the alternative set includes an indeterminate number of people.

In the CF condition a P600 component was found to be elicited at the target phrase, *butcher*. The P600 is a positive event-related potential (ERP) deflection which, whilst not having a clearly defined peak, typically has a midpoint at 600 ms (Osterhout and Holcomb, 1992). P600s have been found to be elicited both by syntactic anomalies (Osterhout and Holcomb, 1992; Coulson et al., 1998), and semantic anomalies as well (van Herten et al., 2005).

However, as (28) suggests, there were no syntactic or semantic violations in Cowles' stimuli. As potential explanations for this result, she points to theories regarding the P3b component, 'a domain-general brain response elicited by rare and/or informative events' (1998, p.22). The P3b takes the form of a positive deflection peaking between 300 to 500 ms after stimulus onset (Otterbein et al., 2012), and it has been suggested that it is part of the same family as the P600: Coulson et al. (1998) argued for this relationship using experimental results which showed that P600 amplitude had an interaction between the probability of ungrammaticality within blocks of sentences presented to participants, and the grammaticality of the sentences themselves. They took this to mean that the P600 is elicited not by ungrammaticality in particular, but by the relative *rarity* of ungrammaticality. This, in combination with the similar scalp distribution of both ERPs, suggests their close connection (Coulson et al., 1998; see also Gunter et al., 1997, who added support to this with a similar probability effect).

Cowles, assuming on this basis that the P600 is related to the P3b, suggests the finding of P600s elicited by CF may reflect updates to a mental model. This follows a theory of the P3b from Donchin and Coles (1988), which proposes that the sensitivity which the P3b shows to the subjective probability of events within the environment is due to it reflecting the updating of a mental model of the environment. The thinking here is very similar to that of Benatar and Clifton (2014): Cowles suggests that the CF condition (28a) would require an update to the mental model, and this is what prompts the P600.

Cowles also suggests that CF might be seen as conveying more information than PF, because it not only selects a single alternative as playing a role in a proposition, but also indicates that all other contextually relevant alternatives do *not* play a role in that proposition (see §2.4.1.2). If this is the case, returning to the file change metaphor of Heim (1983), multiple mental entries might have to be updated when CF is encountered, and this might account for the increased processing load.

The lack of studies in this area is unfortunate, but the results found by both Benatar and Clifton (2014) and Cowles (2003) suggest that there are differences in how PF and CF are processed, and that these differences relate to an increased processing cost for CF.

The basic experimental design for this study outlined at the start of this section features CF on the target word for every item, and only alternates the presence of

TE marking this word. This was the approach used in Experiment 1, which will be reported in Chapter 4. Given that the chosen method of measuring processing here is eye tracking, an increased processing cost manifesting as longer fixations on the target word (as found by Benatar and Clifton, 2014) may indicate that it has been analysed as having CF. However the intention is, of course, for all items to be controlled sufficiently that they are all processed in a similar way: this limits the utility of predicting longer fixations for CF, as there will be no shorter PF 'baseline' with which to compare them. It would have been more useful had there been evidence of a qualitative difference in eye movements, rather than this quantitative difference.

This prediction is, however, more useful for Experiment 2 (reported in Chapter 5), which is an expanded version of Experiment 1 that adds manipulation of CF as a factor. Given that this design is intended to alternate items between PF and CF, if participants spend longer looking at the [+CF] target words, this should provide evidence that they are indeed analysing them as having CF.

Chapter 3

Typographic emphasis

3.1 Introduction

This chapter discusses the concept of *typographic emphasis* (TE), the central subject of this study. This is a phenomenon in printed text which involves a temporary typographic change that serves to emphasise part of the text (Fraundorf et al., 2013; Saldanha, 2011; Ashby, 2017).

(29)   a. As it turned out, the package wasn't delivered to America. It seems it was sent to *Norway*, which was a bit of a surprise.

        b. The entire story was recounted in *What We Did Next*, which became an international bestseller.

As (29a) shows, italic TE can be used to mark contrastive focus (CF), the information structure (IS) phenomenon which was discussed in the preceding chapter. Intuitively, the emphasis of *Norway* seems to correspond to what would be a particularly prominent pitch accent (PA) in speech. However, in (29b), the TE does not appear to mark a phrase which is any way contrastive in the senses defined in Chapter 2. Instead, the italicisation marks the phrase as the title of a book. This is not a function of information structure; indeed, this appears to convey information which is not normally encoded by speech at all.

This study is concerned with the cognitive effects that may be engendered by TE-marking of CF. To investigate this a definition of the form of TE is required, along with an understanding of how it appears to be used. This is particularly necessary because it appears that TE has disparate functions, as (29) suggests. To test the processing of TE-marking of CF, the experimental stimuli must isolate this function to avoid ambiguity.

To investigate how TE-marking of CF is processed also requires an understanding of processing during reading; given the lack of specific research on TE, general principles must be extended from related psycholinguistic reading research. Particularly relevant is the role of phonological processing.

3.1.1 Organisation of chapter

§3.2 presents a definition of **typographic emphasis**. It discusses the motivation for this choice of term, bringing together details from the typographic tradition (§3.2.1) and a definition of the concept of *emphasis* itself (§3.2.2).

§3.3 reviews the small number of published studies which have investigated the **uses of typographic emphasis**. A division between *phonological* (§3.3.1) and *interpolational* (§3.3.2) typographic emphasis is proposed.

§3.4 considers the **differences between typographic emphasis and pitch accents** in terms of both form and function. The role of typographic emphasis within the writing system is considered, as is a series of differences between the two devices.

The perspective of this chapter then shifts from production to reception, from writing to reading: §3.5 examines how typographic emphasis is **processed**, taking account of reading research (§3.5.1) as well as previous psycholinguistic studies into TE (§3.5.2).

3.2 Defining typographic emphasis

From this point on, the following definition of the term *typographic emphasis* will be used to refer to the phenomenon that was exemplified in (29):

(30) **Typographic emphasis is a form of written emphasis created by the temporary alteration of the shape of letter forms within a section of printed text.**

This definition is broad enough to encompass an array of typographic manipulations, some of which are shown in (31). It includes the use of italics, bold text, capitals, small capitals (or caps), and switching between font sizes and typefaces (although the latter two are rarely used within sentences in professionally typeset text).

(31) a. This had *never* been seen before.
      b. He was THE biggest idiot she'd ever met.
      c. The word **engraved** is used in a technical sense here.

The definition only includes devices which alter the shape of letters, meaning that it excludes changes to the colour of letters, and devices which add elements to the letters, such as underlining and striking-out. In principal, these types of changes could be used for similar purposes. Colour switching of individual words and phrases is sometimes seen in textbooks and in tabloid newspapers (McAteer, 1992; Ashby, 2017), although such devices appear to be less frequent than shape alteration.

The definition in (30) is deliberately restricted to the form of TE, and does not make claims about its functions. It will be argued here that TE has several functions. This study is concerned with TE when it operates on the level of the sentence. Its functions at this level will here be divided into the categories of *phonological TE* and *interpolational TE*. Phonological TE appears to serve a similar purpose to that of pitch accents (PAs) in speech: in particular, it is used to mark contrastive focus (CF) as in (29a). Interpolational TE is used to mark sections of language which are embedded within others, yet are not analysed as part of the surrounding grammatical structure: examples of this include titles of published works as in (29b), and words qua words as in (31c).

Phonological TE is of principal interest here, but the fact that TE is used for two very distinct purposes is important: this may result in ambiguity that has an impact on how it is interpreted/processed. Both uses will therefore be examined here.

The term *typographic(al) emphasis* has rarely been used in the psycholinguistic literature; of the relevant studies that will be examined here, Ashby (2017) and Saldanha (2011) both use the term briefly, but do so without discussing it extensively. *Typographic emphasis* has been chosen here as being superior to terms for the same device such as *typeface emphasis* (McAteer, 1992) and *font emphasis* (Fraundorf et al., 2013; see also *emphatic italics*, Saldanha, 2011). This choice is motivated by a desire to be accurate and clear: McAteer's term *typeface emphasis* implies that the effect is achieved through the alternation of typefaces, and this, strictly speaking, is not the case. Fraundorf et al.'s use of *font emphasis* is more accurate in terms of typographic terminology, yet it risks confusion because *font* and *typeface* themselves are frequently confused.

3.2.1 Typographic terminology

Type is text produced mechanically or digitally, as opposed to text produced by hand (Baines and Haslam, 2002). *Typography* refers both to the compositional, aesthetic process of arranging this writing, and to the qualities of the result of this process. The modern typographic tradition owes its origins largely to the movable type system of Johannes Guttenberg (although it was predated by similar East Asian systems). Guttenberg's system, developed in Germany in the fifteenth century, involved arranging reusable metal components which corresponded to individual letters, in order to make templates with which to print pages of text (Lupton, 2010). For the majority of its existence, until the digital revolution of the late 20th century, typography was a mechanical, physical business (Jury, 2007); accordingly, much of its associated terminology pertains to the physical qualities of arranging metal, ink and paper.

The terms *typeface* and *font* are frequently confused in everyday language. This is in part a consequence of the explosion of what might be called 'amateur typography' in the wake of affordable home/office computing (Connors, 1993). To clarify using terms from professional typography, a *typeface* is a particular visual design for a set of letters. Times New Roman, Arial and Helvetica are all typefaces. *Fonts* can be seen as particular instances of typefaces: Times New Roman contains the fonts Times New Roman Italic 12 point, Times New Roman Regular 9 point, etc. Fonts vary in terms of the physical size of the letters, and their *type style*, which refers to different visual variations in the shapes of letters. Type styles tend to follow broadly conventionalised themes: examples are bold, italic and roman (or regular) styles (Baines and Haslam, 2002).[1] When typesetting was a physical practice, a font would consist of a set of metal blocks of a particular size and style, with one block for each character. Computer typesetting does not require the use of physical fonts, and thus has largely removed the need to distinguish between the terms *font* and *typeface* (Baines and Haslam, 2002).

Nonetheless, if systematic research is to be done on how typography and linguistics interact, then precision and clarity are crucial. As was mentioned above, the term *typographic emphasis* has been chosen here to sidestep any potential confusion between *font* and *typeface*: it is an attempt to strike a balance between being accessible to those not versed in typography, and being accurate enough to not contradict or muddy the typographic terminology. Although the term has been defined here in such a way that it could encompass emphasis involving switches

---

[1]*Roman* will be used here rather than *regular*.

of typeface, size or style, the present study is concerned primarily with type style alone. *Style emphasis* would have been an unhelpful choice given the technical use of *style* within some linguistic theories (e.g. sociolinguistics, stylistics).

Several different type styles have been mentioned thus far, including bold and small caps. This study focuses on TE that uses an italic style, appearing within sentences set in roman style. The *roman* font style is typically used as the default form of most typefaces (Ritter, 2002).[2] In linguistic terms, it might be seen as being unmarked. It consists of upright letters, such as those in this sentence. *Italic* is a style which typically features letters slanted left-to-right, which often have more cursive qualities than the letters in the roman style of the same typeface.[3] Italics originated in fifteenth-century Italy, and were originally used less as a style and more as distinct typefaces in their own right. In the sixteenth century they started being integrated with roman-styled text into typefaces (Lupton, 2010), thus offering typographers the option of creating TE:

> [The use of italics] for special material within roman text was a French innovation of the late sixteenth century and received a boost in England through their use for interpolated material in the Authorised ('King James') Bible of 1611; by the mid-1800s, authorities were complaining of overuse of italics in the previous century (Daniels, 2013, pp.68–69).

The central motivation for looking at italic TE within a roman setting is simply that the italic–roman contrast is common. It will be familiar to the vast majority of skilled readers, and is common in academic and narrative writing. A more nuanced motivation for this choice is not really necessary, as it is really only the contrast between styles that is theorised as being important here, rather than which styles are used:

> If we write in black on a white background, the black 'stands out' and is 'marked'. That of which there is less, that which is less usual, will be experienced as 'marked'. If we normally read roman type, italics are marked. In a text printed in italics, a word in roman type will stand out (Kuipers, 1975, p.43).

---

[2]Roman *style* should not be confused with Roman *script*. A *script* refers to an inventory of graphemes; *Roman script* refers to the script used by many writing systems, including the English writing system (Cook et al., 2009).

[3]Gelderman (1999) notes that it is possible to set italics upright; 'genuine', manually designed italic fonts should still appear distinctive without their typical slant.

An interesting illustration of this is the convention of TE within italic sections of text involving a switch back to roman style. An example of this is given in (32), in which the word *no* has 'inverse' roman TE, which is embedded within a longer stretch of italics.

(32) The notice, which read *Danger: Under* no *circumstances should you open this door*, was partially obscured by foliage.

What this suggests is that a relatively arbitrary choice of font styles for the experimental examination of TE can be made.

Nonetheless, it would be interesting to investigate whether TE using different font styles has different interpretive effects; whether bold TE or small caps TE has a stronger effect on processing contrastive focus, for example. Studies such as those conducted by McAteer (1992) and Fraundorf et al. (2013) have touched on this (see §3.5.2). However, the decision to restrict the present study to a single style contrast is prompted by the desire for tighter experimental control: as will be seen below, the existing work done in this area has not established a strong theoretical motivation to examine multiple styles, and the increased risk of confounding introduced by doing so would outweigh the potential benefits. Moreover, given that the connection between TE in general and interpretive effects is expected to be relatively weak here (for various reasons discussed in §3.4), it would seem intuitively likely that any connections between particular varieties of TE and certain interpretations would be weaker still, and thus difficult to find evidence for.

Whilst 'true' italic styles are manually designed by typographers, digital typesetting allows the automatic creation of italic-like styles in typefaces which lack a human-designed italic font. This digitally-slanted style is properly termed *oblique*, but is frequently simply referred to as italic (Bigelow and Holmes, 1993; Gelderman, 1999; Lupton, 2010). Such distinctions might seem pedantic, but may actually be relevant to the psycholinguistic investigation of typography: the level of visual contrast between a roman and italic style may be higher than that between a roman and oblique style, and this may serve to make TE more or less salient.

This is particularly noteworthy because eye tracking and self-paced reading research commonly uses the Courier typeface, which lacks a manually designed italic font (unlike its successor, Courier New). The popularity of typefaces such as Courier in this type of research is partly due to them being monospaced: in such

typefaces each character is of a uniform width, providing a degree of control over issues such as the perceptual span (Rayner et al., 2010); proportional type-faces (such as Times New Roman, or Charis SIL, in which this thesis is set) have characters of varying widths. The differences between roman and italic/oblique styles may also be smaller in monospaced typefaces than in some proportional typefaces, in which italic characters are narrower than roman ones.

These are concerns which are rarely brought up in the psycholinguistic litera-ture: greater attention to typographic detail may benefit this type of research. Fraundorf et al. (2013) mention that, during piloting of a self-paced reading experiment, they found participants struggled to distinguish italic and roman Courier, which they remedied by switching to Arial, which is a proportional typeface. Whether the lack of differentiation they found was related to charac-ter width, the minimal difference between the roman and oblique styles in that typeface, or to another property, is unclear.

The Times New Roman typeface was used in both of the experiments con-ducted for the present study, motivated by the concerns outlined here: it is a pro-portional typeface with a distinct italic font. Selection of this typeface was also influenced by the fact that it is extremely widely used, and should have therefore been familiar to all participants.

3.2.2 Emphasis

The concept of *emphasis* is frequently evoked in style guides when discussing what is referred to as TE here, yet they tend not to specifically define what *em-phasis* means (Saldanha, 2011). In the linguistic literature, there is a similar air of vagueness surrounding the concept: 'The term *emphasis* is in current use for multivaried phenomena that do not constitute a single unitary category' (Kohler, 2006, p.748).

In formal linguistics, emphasis is strongly associated with focus: recall Halli-day's definition of focus, discussed in the preceding chapter: 'information focus is one kind of emphasis' (Halliday, 1967, p.204). Furthermore, Trask provides the following definition of *emphasis*:

> A very general term for any phenomenon which serves to draw par-ticular attention to some element in a sentence or utterance, either to place that element in focus or to contrast it with some other element (Trask, 1992, p.89).

In fact, the two functions of emphasis Trask suggests are essentially both instances of focus; to contrast an element with another requires contrastive focus.

To link emphasis with focus too strongly here would risk implying that *typographic emphasis* itself only indicates linguistic focus, whereas it actually has a range of functions (as will be discussed here). A more theory-neutral definition is thus required:

(33) **Emphasis is the motivated, delimited creation of contrast through modulation.**

This definition can be split into the following elements:

1. Motivation. Emphasis is placed in order to convey something; it is not a result or side-effect. A word suddenly spoken with a drastically different pitch because the speaker was in pain is not emphasised.
2. Delimitation. Emphasis must have a beginning and an end. If a person spoke quietly for half their life, and then switched to speaking loudly for the rest of their life, neither the quiet nor the loud sections is emphasised.
3. Contrast. This has been discussed in the preceding chapter; although this discussion was framed in terms of the semantics of focus, the central points hold: contrast involves opposition and dominance.
4. Modulation. This is a process whereby a signal is overlaid in some way onto another signal. Modulation warrants further discussion.

3.2.2.1   Modulation and emphasis

The proposed role of modulation in emphasis can be further explained by modelling language output as a stream of data. This approach is influenced by the basic elements of *information theory* (or *communication theory*; Pierce, 1979), as originally advanced by Shannon (1948). To introduce this, spoken language will first be used as an example, and then written language will be discussed in these terms.

Shannon's model involves a *transmitter* and a *receiver*: to model speech, these can be seen as a speaker and a hearer. *Channels* constitute the medium across which *signals* are transmitted between these two entities. The varying sequence or *stream* of phonemic speech sounds can be seen as being the primary signal.[4]

---

[4]This is not to suggest that the real channel of speech is not actually the air between the speaker's vocal tract and the hearer's eardrum; it is simply a useful abstraction.

Additional signals can be transmitted simultaneously by altering or *modulating* an aspect of the primary signal which does not impact its meaningful contrasts. In speech, non-phonemic aspects of the phoneme stream are modulated, such as pitch and rhythm. This does not alter the phonemic content of the primary signal, but instead conveys a secondary *prosodic* signal. For example, in English, changing the pitch of the vowel /e/ from 200 Hz to 100 Hz does not remove or obscure its contrast with the vowel /o/, or any indeed with any other phoneme in the language. This alteration might, however, be part of the placement of a PA on the syllable that contains this phoneme; this prosodic signal can have its own system of contrasts.

Applying this model now to the written modality, the equivalent to the primary signal of speech is the sequence of letter forms or *graphemes*. A *grapheme* can be defined as the minimally distinctive unit in a writing system (Henderson, 1985; Sampson, 1985; Coulmas, 2003; see §3.4). The TE signal is carried by a layer of modulation on top of this, which alters graphemes without obscuring the contrasts between them. For simplicity's sake, this signal can be seen as having two levels, either [−TE] or [+TE]. This is roughly illustrated in (34) for an example sentence containing TE.

(34)   a. (Full output)   We didn't expect *disaster* on such a scale.
         b. (Grapheme signal)   We didn't expect disaster on such a scale.
         c. (TE signal)   − −   −   +   − − − −

The word *channel* has been used here in a slightly different way to its typical use in the literature on multimodal communication. For example, Engle (1998) argues that speech and gesture have been seen as separately interpretable channels in this field, but should instead be seen as part of a 'composite signal'. No claims are being made here regarding 'multimodal' communication in this sense; the conception of *signal* and *channel* here are not directly equivalent to the ideas which Engle argues against.

––––––––––––––––––––––

According to the definition of emphasis given in (33), the phenomenon referred to as TE is a type of emphasis, as are PAs. Some of the formal, structural parallels between the two are particularly apparent when considered under the framework of the model that has been presented here. This does not, however, imply that they are equivalent in terms of functions, or in terms of how they are interpreted; differences here will be discussed subsequently.

The view of emphasis presented here suggests some obvious parallels to the notion of *markedness*. This concept originates within structuralist linguistics; it has received particular attention in the field of phonology, but can be applied across linguistics, and beyond (Battistella, 1990; Haspelmath, 2006). It refers to 'an asymmetric relationship between two or more elements' (Janda, 1996, p.209) or 'an evaluative nonequivalence' (Battistella, 1990, p.1), with the hierarchical relation placing a *marked* form that possesses a feature over an *unmarked* form that does not. It must follow then that the emphasis of a linguistic element makes it marked. Emphasis involves contrast, and contrast imparts a quality of dominance (Molnár, 2002); in linguistic emphasis, this dominance can be seen as the possession of some additional feature (or features). The dominant view of PAs expressed in the preceding chapter conceives of them adding the feature of focus to the phrases in which they appear; at one level then, a focused element can be considered marked.

Note that the argument here is that all linguistic emphasis results in markedness, not that all markedness is caused by emphasis: in the question *How young are you?* (Battistella, 1990, p.3), the markedness of *young* is not a result of modulation, and thus is not an instance of emphasis as it has been defined here.

3.3   Uses of typographic emphasis

To examine TE usage from a linguistic perspective requires selecting a narrower subset of the phenomenon than that implied by the definition given in (30), which is general enough to allow for a range of uses, some of which are irrelevant to the present study. For example, TE is used in a stylistic manner to mark different sections of documents: bold text may be used (along with spacing devices) to indicate titles and headings. TE can also be used as more of an aesthetic device. Consider the stereotypical 'wanted' posters associated with the Wild West; many feature a wide variety of typographic manipulations which are arguably more aesthetically than linguistically motivated.

These uses are of little interest here, as their distribution is distinct from that of the *phonological TE* on which this study concentrates. They operate on the document level rather than on the sentence level, and as such are of more concern to stylistics than to linguistics. It is the sentence level that is of interest here; this is where phonological TE occurs. A second type of TE also occurs at this level; this will be referred to as *interpolational TE*. Because these two types of TE have similar distributions, ambiguity can arise: therefore, both must be examined here.

Research into different uses of TE has been rare. A large-scale corpus approach investigating TE usage has yet to be taken; in its absence, there exist a handful of studies which have used small corpora to research TE. These are the studies of Ashby (2017), Sanford et al. (2006) and Douglas (2009), all of which present surveys of italic usage in novels. Of these, the Sanford et al. (2006) study is perhaps the closest in spirit to the present study, approaching TE from a psycholinguistic perspective (although it does not provide as in-depth a survey of usage as the others). Ashby (2017) takes a more phonetic view of TE, whilst Douglas (2009) approaches the issue from within the field of translation studies.

All three studies manage to find a range of use patterns within very small samples; they each categorise the uses slightly differently, but there are many points of agreement between them. All three are concerned with narrative text, and thus examine uses of TE within sentences rather than the more stylistic uses associated with headings and aesthetics; as such, they align with the level of interest here.

It is interesting to note that these studies focus on TE in novels, despite each coming from different research disciplines. Sourcing linguistic examples from fiction might raise warning flags in some circles: however, in terms of TE there is as yet a complete absence of larger corpus studies into less 'artistic' genres of text. Furthermore, the studies provide support for each other's classification of different types of TE usage. The analysis of these studies will continue here with the proviso that these are nascent steps into TE analysis (Ashby, 2017 in particular is very clear about this fact).

Sanford et al. (2006) looked at the largest sample, using twenty 'full narrative' texts. They do not cite them all, but all the examples they give come from novels. Despite the larger sample, of the three studies here this gives the briefest of discussions about usage. Sanford et al. identify the following uses of italics (2006, pp.115–116):[5]

- **Changes of voice quality** (e.g. shouting, whispering): Whispering in my ear, *"Don't make a sound!"* (Dahl, 1979, p.319)
- **Contrast**: " … is it possible … that the deceased banged his head sometime *after* his death?" (Guterson, 1995, p.22)
- **Emphasising plot-crucial details**: A small saw and some nails. *Woodwork tools.* (Rankin, 2001, p.448)

---

[5]The following three lists are all taken almost directly from the three articles; in some places examples have been removed.

- **Reflecting surprise**: "Did I hear somebody say mother *exploded*?" (Banks, 1992)

The quotation for 'contrast' is a clear case of CF: semantically, a highly restricted, easily accessible alternative set can be accommodated for *after*, containing only *before*. Pragmatically, the suggestion that the dead person hit their head after death can be presumed to be unexpected from the perspective of the hearer (as the quotation is taken from reported speech, it may be unexpected by the character(s) hearing the utterance, and possibly by the reader themselves).

The quotation for 'surprise' could also be argued to be a case of CF; the quality of surprise is, after all, intrinsic to Zimmerman's view of CF (2007). It is difficult however to infer from the quotation alone what the alternative set would be.

'Emphasising plot-crucial details', suggests that this type of TE use is specifically connected to the text-level element of the plot, and thus operates in a stylistic manner outside the purview here. This TE may instead be intended to mark a thought of the author/narrator, distinguishing it from the rest of the ostensibly objective narrative; Ashby (2017) proposes such a use of TE.

Ashby conducted a survey of italic usage in three novels (Dickens, 1839; Storey, 1960; Theroux, 1967) and presents the following categories of usage:

- **Intonational**: It's picking *on* us, not picking *at* us. (Theroux, 1967, p.41)
- **Foreign words**: *lignum vitae* (Dickens, 1839, p.161)
- **Reported speech**: 'I told you six times, *don't look*!' (Theroux, 1967, p.219)
- **Representing thoughts**: *What am I doing?* Waldo thought. (Theroux, 1967, p.212)
- **Quoted texts** (songs, letters, etc.): *Thank you, Margaret, for coming. You've made it all worthwhile, 'Howarth'.* (Storey, 1960, p.133)
- **Pronunciation features** (e.g. h-insertion): *h*onours. (Dickens, 1839, p.183)
- **Imitating sounds**: *ploop*. (Theroux, 1967, p.238)
- **Longer phrases**: only *it was the wrong lady*. (Dickens, 1839, p.500)
  (Ashby, 2017, p.7)

Ashby's primary interest is in what she terms 'intonational' emphasis, which essentially corresponds to PAs. It should be noted that her example for this does not represent a PA marking presentational focus, but rather contrastive focus: 'It's picking *on* us, not picking *at* us' clearly contrasts the two prepositions. This is the only point where one of Ashby's proposed TE uses obviously coincides with one of Sanford et al.'s ('contrastive'), although, as was mentioned above, Ashby's 'representing thoughts' does seem close to the example Sanford et al. use for

'emphasising plot-crucial details'. Her 'reported speech' and 'quoted texts' also appear to pertain to text-level functions of TE, which can be ignored here.

Ashby's 'pronunciation features' and 'imitating sounds' can be seen as being related to 'intonation emphasis', in that all three are connected to phonology. The example she cites from Dickens of h-insertion might be seen as an instance of *expression focus* per Krifka (2007), in that the TE on the first grapheme indicates alternatives to the pronunciation of the word, i.e. alternatives to the form of the word rather than to its denotation. However, Douglas (2009) does suggest the similar 'representations of dialect' as a use of TE.

Ashby's 'longer phrases' is confusing: italics clearly aren't usually used simply to mark 'longer phrases', otherwise many texts would italicise all long phrases. Furthermore, the example she gives from this, from *Nicholas Nickleby* by Dickens, would actually appear to be a use of italics to represent a pitch accent, and thus to constitute 'intonational emphasis' in her terms. At this point Newman, the character speaking, is telling Nicholas that he has mistaken someone's identity; spoken, the contrastive focus might work thus: *it was the WRONG lady*. It is however interesting that Dickens puts TE on the whole phrase, rather than on the adjective alone; the domains at which TE occurs will be further discussed below (§3.4.4).

Douglas conducted a survey of italic usage within a small corpus of four novels and their translations: he used two Italian originals and their English translations, and two English originals with Italian translations. The cross-writing system aspect of this study is not particularly of interest here, but the uses of italics that Douglas proposes are pertinent (albeit lacking examples from the corpus, which is unfortunate):

- **Titles**
- **Quotations**
- **Foreign borrowings**
- **Onomatopoeia**
- **Neologisms**
- **Representations of dialect**
- **Narrative prominence**
- **Embedded quotations**
- **Tonic prominence**
  (Douglas, 2009, p.6)

This list clearly coincides with several of the proposed uses reviewed above. It also includes uses of TE for 'titles' and 'neologisms'. That the other studies did

not find these uses in their corpora is unsurprising; mention of titles of published work and use of neologisms is more likely in other forms of text (although science fiction and fantasy may well feature many neologisms). Indeed, use of TE to mark titles of published work has already been mentioned in this chapter, and is common in academic writing in particular. Similarly common in this genre is the use of TE to mark neologisms and newly introduced terms, which can be seen throughout this thesis, as well as throughout the literature it cites.

Considering the findings of these three studies together, it is interesting to note that many of the examples picked to illustrate italic usage are from sections of reported speech; more specifically, from *direct speech*, which implies the exact quotation of another speaker or writer (Coulmas, 1986). Naturally, direct speech is more common in novels than in many other types of writing, but even allowing for this, the number of uses these studies exemplify which come from direct speech still seems relatively high. It would certainly be interesting to run a corpus analysis investigating whether italics are relatively more common within stretches of direct speech. With that said, it is undeniable that italics and TE can also be used within texts that contain little to no direct speech; this issue will not be further investigated here.

Douglas' survey is unique amongst these studies in actually proposing a way of classifying these uses according to more general principles: his list is arranged to represent a cline of italic usage, which ranges from what he terms 'punctuation features' (at the top of the list), to 'prosodic features' (at the bottom). He does not define what he means by 'punctuation', but the fact that he places it in opposition to 'tonic prominence' seems to suggest that he sees it as less a matter of encoding sound, and perhaps more related to the encoding of non-linguistic information.

A similar, but more categorical approach approach to classifying TE usage will be adopted here. Table 3.1 collects the proposed uses from all three surveys together, and classes each as being either *phonological* or *interpolational*.

Table 3.1: New categorisation of italic usage

Phonological	Interpolational
Intonation/contrast[A, D, S]	Reported speech, thoughts & quotations[A, D]
Onomatopoeia[A, D]	Foreign words[A, D]
Pronunciation/dialect features[A, D]	Narrative prominence[D, S]
Changes of voice quality[S]	Neologisms[D]
Reflecting surprise[S]	Titles[D]
	Words qua words

[A]Ashby (2017) [D]Douglas (2009) [S]Sanford et al. (2006).

The 'words qua words' exemplified in (31c) has been added to the uses from the studies, and has been classified as interpolational TE. This is, of course, a usage fairly specific to specialised literature such as linguistic research; there are doubtless many more such specialised uses in different genres. All of the uses discussed here are presented as examples, rather than as an attempt to construct a full taxonomy. This holds for the proposed division between phonological and interpolational TE; no claim is made here as to whether there are more general types.

The phonological uses of TE all appear to have the indication of sound in common. The term *phonological* is used because it is unclear whether they work on the basis of actually evoking *phonetic* qualities, or if they are more abstract. The interpolational uses of TE all have the quality of *embedding* sections of language within other sections. These embedded sections are not required to conform with the grammar of the language that surrounds them.

3.3.1   Phonological typographic emphasis

Of the various uses of phonological TE, clearly the most important to the present study is its indication of PAs. It may also be a particularly common use: Ashby (2017) provides descriptive statistics for her categorisations of italic use across the corpus she used, and classes 86% of all instances of italics as 'intonational'.

When TE appears on a focused word, it seems to mark *contrastive* focus rather than presentational focus. Aside from the examples given thus far, further evidence comes from the fact that TE is rarely, if ever, used on every single word that would bear a PA in speech:[6]

(35) # He smiled at *Jessica*. She took his *hand*. They walked towards the *door*.

There thus appears to be a categorical distinction here, and one which is much easier to identify than the contentious difference between 'normal' PAs and 'contrastive' PAs (see §2.4.1.3).

It is certain that the vast majority of written sentences in English do not feature any TE, or any other apparent way of marking focus (aside from those syntactic/lexical devices outlined in Chapter 2). This suggests that readers are able to infer focus without PAs (assuming that focus structure is an essential part of language). How this may be achieved is unimportant here; what is interesting is

---

[6]However, something like this may be seen in the dialogue in comic books, which frequently appear to mark all foci with bold text (Khordoc, 2007).

that the writing system does not mark presentational focus (PF), and yet it can mark CF.

According to Calhoun (2009), CF marking is a probabilistic process, occurring when a word is caused to have a higher salience than the hearer expects. Her account of this concentrates on speech, but if this view is applied to writing it is compatible with TE-marking of CF. Furthermore, it is also compatible with a potentially limitless number of ways to increase salience. A more traditional linguistic approach might posit a systematic association between the written device and CF; this is closer to the approach of Zimmermann (2007), which is similar to that of Calhoun in that it is pragmatic and probabilistic, but implies that only certain types of CF marking exist. The question of how systematically TE-marking of CF is used across writing can only be addressed by a large-scale corpus approach; in the absence of this, it is sufficient to accept the evidence of the surveys presented above.

Whilst TE-marking of CF does occur, it does not *always* occur: it is optional, not obligatory. This again fits with Calhoun's probabilistic approach to CF marking, as it does with the majority of other theories, with the exception of the strong claims of É. Kiss (1998) about obligatory clefting.

It should also be noted that TE can indicate contrastive *topics*, as in (36).

(36) A: Do you know what Brian has been up to?

B: I don't care about Brian, but *Angie* has just done something incredible.

The preceding chapter only made brief mention of topic: the approach this study has adopted to focus does not require engaging with other informational articulations. Furthermore, eschewing topic, and in particular contrastive topics, means avoiding additional information structure debates; some sources such as Molnár (2002) and Repp (2010) have attempted to argue that *contrast* itself should be taken as an IS primitive alongside topic and focus, and part of such arguments is based on the existence of contrastive topics. Other sources (e.g. Partee, 1991) analyse contrastive topics as topics which contain a focus.

Viewing contrastive topics as topics which contain a focus is sufficient for this study's purposes, but this is not to make a strong claim on the status of contrast as a primitive in IS. At any rate, the apparent ability of TE to mark contrastive topics as well as contrastive foci provides little evidence for either side of the debate; more interesting questions here would relate to the fact that the same type of TE can mark a contrastive topic and CF, whereas in speech contrastive topics are associated with a different type of PA to those used for CFs (Hedberg and Sosa, 2008).

The present study will not engage explicitly with contrastive topics. The work necessary to create experimental stimuli in which topic is varied would appear to be even more challenging than that involved in manipulating focus alone. As such, the experiments conducted here will not be able to provide any explicit evidence that TE has any effect on the processing of contrastive topics. Despite this, in intuitive terms it seems extremely likely that if TE has an effect on the processing of contrastive foci, it will have a similar effect on the processing of contrastive topics.

3.3.2   Interpolational typographic emphasis

Titles, quotations, foreign borrowings and words qua words are all examples of what will be termed *interpolation* here: they are sections of language embedded into other sections of language.[7] They can be seen as chunks of language which are not required to conform to the grammar of the language which surrounds them (*setting* from hereon, for convenience), as can be seen in (37).

(37)   a. *As he was talking about he wrote, he smiled.
      b. As he was talking about *He Wrote*, he smiled.

(37b) is grammatical, despite containing a sequence of words that would be ungrammatical if they were taken as having no special status. A title can be grammatical or ungrammatical in itself; changing the title in the example to the internally ungrammatical *Wrote He* would not affect the setting. Effectively a title is treated as a noun, despite being able to consist of anything from non-words to relatively long stretches of language with complex syntactic structures.

In a similar way, the internal content of quotations, as representations of other language use, does not need to connect to or conform to the grammar of their setting.

(38) The scrawled note on the piece of paper read *Is the truth?*

Indeed, this disjunction can go as far a complete switch into another language, as occurs during code-switching in the speech of multilingual speakers (Auer, 1999):

(39) When they told me *hemos hecho todo*, I believed them.

---

[7]In the same vein as 'words as words' are phrases as phrases: 'She then spoke at length about the whole *to be or not to be* issue.' Whilst such phrases and quotations may be more likely to be delimited with quotation marks rather than typographic emphasis, as Douglas' survey finds, occasionally typographic emphasis is used.

The use of *interpolation* to describe this function of TE is inspired by Waller (1980), who puts forward a set of functions of punctuation. He defines *interpolation* as 'the insertion or juxtaposition of a short segment into a longer one in such a way that the continuity of the sentence, paragraph, page, chapter or book is not destroyed' (1980, p.248). As is often the case in such work, Waller does not include TE in his analysis; he suggests interpolation as a function of parentheses, dashes and commas. This definition is not fully compatible with the way *interpolation* has been presented here: whilst both imply that interpolated sections are pieces of language which are located within other sections of language, Waller's sense does not entail that those sections are permitted to be unanalysed according to the grammar of their setting.

The classification of TE uses into *phonological* and *interpolational* here is not intended to be absolutely categorical. Reading the examples in this section aloud may well suggest prosodic/phonological aspects to what have been classed as interpolational uses here. Indeed, for some theoretical purposes, it is likely that there would be benefits to taking a more subtle approach, such as the cline of italic usage suggested by Douglas (2009). For the purposes of this study, however, the division proposed here will be sufficient.

3.4 Differences between typographic emphasis and pitch accents

Whilst TE and PAs may sometimes perform a similar function, they are significantly different in terms of form. Furthermore, they both perform other, differing functions: in particular, some instances of TE mark interpolation, whereas PAs are not used for this.

3.4.1 Differences between writing and speech

At a basic level, it is obvious that the difference in form results from TE being a written phenomenon and PAs being spoken phenomena. However, the precise manner in which the forms differ should be examined, as this may impact how they are perceived and processed. It will also be argued here that the difference in functions is related to the fact that writing and speech are distinct semiological systems: writing is not purely a system of encoding speech, but can encode other information—which is what happens when TE marks interpolation.

Writing and speech are often referred to as *modalities*, although as Bernsen (2002) points out, the meaning of this term can vary across the literature. It is often used in such a way that implies writing and speaking are simply different methods of transmitting language. This view is sufficient for many research purposes, but it ignores the paralinguistic and non-linguistic features of writing (Nunberg, 1990). Whilst this thesis makes use of the term *modality*, this is done with no intent to imply equivalence between writing and speech; instead, a more nuanced perspective on writing will be taken here, drawing on work done on the analysis of writing systems.

Writing systems are typically seen as individual, unique methods of encoding a particular language through the use of *graphemes* (Mattingly, 1992; Coulmas, 2003). A grapheme is a minimally distinctive unit in a writing system (Henderson, 1985; Sampson, 1985; Coulmas, 2003); the term deliberately echoes *phoneme* (Henderson, 1985). In semiotic terms, graphemes are signs (Coulmas, 2003) which are typically realised as visual marks (although they may also be realised as tactile signs, such as those in braille; Roberts, 1996).

Nunberg (1990) argues that there has been a tendency for the linguistic analysis of writing systems to adopt what he terms a 'transcriptional' view, casting them merely as imperfect methods used to transcribe spoken sounds. Nunberg advocates a position less typically seen in the linguistic literature, which views writing systems as semiological systems in their own right. Such a view holds that writing systems have complex correspondences to the systems of their 'host' spoken languages, and also proposes that writing systems can encode other types of information not directly encoded by speech (Catach, 1986; Sproat, 2002):

> [Writing is] a semiological system among others, closely linked with the most important among them all, the spoken language. This relationship does not mean that writing has no attributes and functions of its own (Catach, 1986, p.1).

Evidence for the asymmetric relationship between writing and speech can be found by examining how graphemes correspond to linguistic units in the English writing system. Looking at this will also provide a background against which to clarify the position of TE within writing.

As Table 3.2 illustrates, graphemes do not always correspond to phonemes in a simple one-to-one manner: the correspondence can be one-to-many, as when ⟨x⟩ corresponds to /ks/, or many-to-one, as when ⟨ng⟩ corresponds to /ŋ/. In different contexts, graphemes can represent different phonemes or sequences of

Table 3.2: Grapheme to linguistic unit correspondences in the English writing system

Grapheme	Correspondence
⟨c⟩	/k/
⟨x⟩	/ks/
⟨ng⟩	/ŋ/
⟨i⟩	/ɪ/ or /aɪ/ or /ə/, etc.
⟨&⟩	*and*
⟨9⟩	*nine*

phonemes; in the English writing system, grapheme–vowel correspondence is particularly irregular: the grapheme ⟨i⟩ corresponds to /ɪ/ in *fish*, /aɪ/ in *fight*, and /ə/ in *terrible*. Furthermore, some graphemes correspond to morphemes, such as ⟨&⟩ and ⟨0⟩–⟨9⟩.

However, TE does not appear to directly correspond to linguistic units at all, in either its phonological or interpolational guises. In this respect, it has something in common with punctuation, which similarly fails to correspond directly to particular linguistic units, and also appears to have phonological functions:

> Punctuation must be considered together with a variety of other graphical features of text, including font- and face-alterations, capitalization and spacing, all of which can be used to the same sorts of purposes (Nunberg, 1990, p.17).

The similarity is also noted by McAteer (1992), and is implied by Douglas's proposal of a cline of italic usage (2009), one pole of which is 'punctuation features'.

Punctuation consists of a set of graphemes, or *punctuation marks* (Nunberg, 1990) such as ⟨.⟩, ⟨,⟩, and ⟨?⟩. None of these correspond to linguistic units, but have been suggested to encode rhythm and intonation:

> Although punctuation certainly fails to represent the total range of prosodic phenomena a writer or reader may assign to a piece of written language, it does capture some major aspects of a writer's prosodic intent (Chafe, 1988, p.397).

Chafe claims that punctuation graphemes tend to delimit units of written speech which are similar to his proposed *intonation units* (Chafe, 1987), suggesting that the 'inner voice' of the reader is guided by punctuation; intonation patterns are constructed from what Chafe terms the 'covert prosody' of a given piece of writing. His analysis is based on marks such as ⟨.⟩ and ⟨,⟩, which are associated with

the boundaries of sentences and clauses. The phonological TE discussed here is distributed differently, but essentially this function seems compatible with a prosodic view of punctuation.

Nunberg (1990), in keeping with his 'anti-transcriptional' stance, argues that there are many functions of punctuation which are not prosodic in nature. For example, he mentions the fact that commas are sometimes required where there would be no intonational break in speech, such as the comma in *John, as I said, will not be here* (1990, p.13). Many similar phenomena can be identified. For example, if one views the 'blank grapheme' of the space as part of punctuation, its indication of word boundaries arguably has no equivalent in speech (and no equivalent in some systems, such as the Chinese writing system);[8] the use of uppercase letters to mark proper nouns is a similar case. Interpolational TE appears to lack a clear relationship to a particular prosodic pattern, and is thus similar to these uses of punctuation.

An issue in grouping TE with punctuation is that TE does not exist as a unique set of graphemes, unlike punctuation. Whilst it might be argued that the roman ⟨a⟩ and italic ⟨*a*⟩ should be analysed as two different graphemes (similar questions arise for upper and lowercase letters), here it is preferable to recall the model that was outlined above (§3.2.2.1), and take TE as being conveyed by a modulation achieved by the alteration of type style. Analysing TE as a modulation of the grapheme stream in this way casts it as a *suprasegmental* feature, in that it can range across the segments of individual graphemes (less importantly, this implies that ⟨a⟩ and ⟨*a*⟩ are allographs of an underlying ⟨a⟩ grapheme). In this view, the form of TE is at one level similar to the form of PAs; they are both suprasegmental.

3.4.2   Differences in information quantity

Thus far, it would appear that TE-marking of CF is very similar to PA-marking of CF. However, there are two noteworthy differences between the two types of marking which may impact how they are interpreted and processed.

Firstly, TE tends to vary in an extremely limited manner: a given text is unlikely to contain more than roman, bold and italic styles, and of these only the latter two would be likely to be used for sentence-level TE. Framing this within the modulation model outlined above, the TE signal in this respect is limited

---

[8]Indeed, much has been made in the literature on reading acquisition about the fact that learning to read involves metalinguistic awareness to a level that is not required for learning speech (Nagy and Anderson, 1995).

to only two or three levels. In contrast, the ToBI prosodic transcription system recognises five types of PA: H*, L*, L*+H, L+H* and H+!H (Silverman et al., 1992).

Secondly, PAs are intonational devices, and aspects of intonation vary continuously. The five types of PA are categorised on the basis of relative differences between tones, but PAs can have variable pitch: 'it is customary to think of an intonation contour as having a linguistically distinctive shape or pattern and an independently variable pitch range' (Ladd and Morton, 1997, p.313). Whilst TE could itself theoretically take any visual form, and could be made to vary continuously (e.g. with increasingly thicker strokes across the graphemes in a word), in practice it always involves a discrete and sudden shift from one type of letter form to another, such as in the roman–italic examples given here. PAs thus involve what might be seen as analogue (or continuous) modulation, in comparison to the digital (or categorical) modulation of TE.

The amount of information involved in the *deployment* of these devices does not necessarily correspond to the amount of information they *convey*, however; it may be that perception of PAs is categorical rather than continuous, as Ladd and Morton (1997) suggest. With that said, very fine distinctions in pitch and rhythm are obviously perceptible: if CF itself is taken to become more contrastive the more the alternative set is made salient, per Calhoun (2009; see §2.4.1.3), then varying continuous properties such as pitch and volume may result in different interpretations.

How might the present study be impacted by these differences? If a given piece of written discourse is likely to contain fewer different types of TE than the number of types of PA that would occur in an equivalent spoken discourse, then this suggests that ambiguity may arise. This will be discussed below in §3.4.3.

In terms of the categorical nature of TE and the continuous nature of intonation, the following question arises: if CF is marked principally by degree of prominence, as various voices in the literature have implied (e.g. Bolinger, 1961; Calhoun, 2009; Katz and Selkirk, 2011), then just how prominent might the italic TE used in this study's experiments be? It might be that it is not prominent enough to prompt a CF interpretation (or, if CF itself is taken as gradient, a sufficiently contrastive interpretation to cause any measurable difference). The choice of italic TE was justified above (§3.2.1), but it might be that other forms, such as bold or all-caps, are interpreted as having greater prominence (this is discussed further in the review of previous psycholinguistic work on TE below; §3.5.2). It may also be possible that any type of TE alone is not prominent enough to effectively mark

CF, in which case the design of the experiments used here would preclude the target words from being interpreted as CF, because they were deliberately designed to have no form of CF marking other than TE (i.e. in the [+TE] condition).

3.4.3 Potential for ambiguity

A variety of functions of TE have been discussed here. Multiplicity of function is of course common to many linguistic devices, including aspects of intonation. However, as has been proposed here, writing is a semiological system distinct from speech, and its devices such as TE are able to encode information which is not linguistic: the main example here has been that of interpolation TE, which has no direct equivalent in speech.

When the exact same device has multiple functions, ambiguity may arise. A potential consequence of this may be that TE has less of a chance than intonation to 'hit its mark' and be interpreted as it was intended to be; an instance of TE intended to indicate a title might be misinterpreted as representing contrastive focus, for example. The likelihood of a skilled reader being thoroughly confused by this is, of course, low; what *is* possible is that a processing cost may be incurred by ambiguity. Consider (40):

(40) Robert said that *Jurassic World* was the best dinosaur film.
Ken said that *Jurassic Park* was the best.

If all titles are routinely italicised in this particular text, then the level to which *Jurassic Park* is intended to contrast with *Jurassic World* may be unclear.

Potential ambiguities such as this are an interesting area: the MA thesis (Norton, 2012) which preceded this study attempted to compare phonological and interpolational TE using self-paced reading. This proved methodologically difficult, and results were inconclusive (see §3.5.2); from the perspective of the present study, it seems that at this stage a wiser and potentially more powerful experimental approach is to limit the area of interest to phonological TE.

3.4.4 Differing domains

Whilst focus is a property that can be associated with words or phrases, prosodic marking of focus manifests on single syllables. Typically TE is used on whole words or phrases, rather than on graphemes or syllables. The spoken sentence

(41a) would most likely be rendered with TE as (41b).[9]

(41)   a. He smiled at JESSica.
        b. He smiled at *Jessica.*

Note that placing a PA on a given syllable may have an effect on surrounding syllables: modulation of frequency by the vocal folds works gradually rather than suddenly switching between higher and lower frequencies—hence the term *intonation contour* (Pierrehumbert, 1980; Ladd and Morton, 1997; this was touched on in §3.4.2). The shape of the pitch stream may thus be altered by an upcoming accented syllable, providing the listener with a 'preview' of the upcoming emphasis.

This is in contrast to TE, which can be immediately switched on and off on a grapheme-by-grapheme basis. However, TE itself may be 'previewed', although by a different mechanism: it may fall within the *perceptual span* of characters which are accessible to the right of the point at which the eye is looking (around 14–15 characters; Rayner, 1998).

Any differences caused by the way in which TE marks whole words and PAs mark individual syllables are likely to be insignificant, given that there is a focus projection mechanism that enables focus to propagate from the syllable level to the word and phrase level (Selkirk, 1995; Gussenhoven, 1999; see §2.3.6). Furthermore, although 'preview' effects may exist to different extents for both devices, at this early stage of investigation into TE, attention to such minor potential differences can probably be safely deferred to future studies.

3.4.5   Technical and conventional constraints

The ability of speech to convey information may sometimes be limited by external factors, such as the degradation caused by frequency information being lost when speaking on the telephone (Lawrence et al., 2008). Returning to the concept of *channels* from Shannon (1948), this can be seen as a constraint imposed by the nature of the channel, rather than a deficiency of the transmitter or receiver. It seems that it is far more common for the channel to be constrained for writing, and specifically for type, than for speech.

A speaker is generally free to make use of any and every linguistic device they are aware of. If they are constrained in some way, it is more likely to be a permanent impediment specific to themselves, such as that resulting from a disability;

---

[9]This is not always the case; for example, the instance of TE marking h-insertion which Ashby (2017) gives from Dickens (see §3.3).

this is a limitation associated with the transmitter rather than the channel. A writer is not necessarily free to make use of every orthographic or typographic device they are aware of, particularly when producing type rather than hand-writing. They can be constrained by restrictions in the electronic system used to generate the type, which might be seen as limitations associated with the trans-mitter. They can also be constrained by properties of the channel itself. In terms of TE, certain services, such as Short Message Service (SMS) 'text' messages, 'tweets' on Twitter, or plain-text emails, are restricted to a single font style throughout the message, and as such are unable to modulate between roman/bold/italic etc.

Where such constraints exist, interesting alternative methods of indicating that which TE would normally indicate have developed; in instant message for-mats such as SMS, there are grapheme-based methods such as the use of delimit-ing characters (e.g. *I was *so* angry*; c.f. the linguistic literature conventions used in this thesis to delimit items such as phonemes and graphemes); emojis may also perform similar functions. Switching the case of letters, as in *They were AMAZED*, is a manipulation available in all but the most restricted methods of print gener-ation, and falls under the definition of TE used here. It should be noted however that professionally printed material rarely uses all-caps TE, and instead tends to use small caps. Nevertheless, in their studies of TE interpretation, both McAteer (1992) and Fraundorf et al. (2013) investigate all-caps TE (see §3.5.2); despite its unpopularity in traditional, formal printing, it may be that the status of all-caps as the most universally accessible form of TE has made it popular in some genres.

Genre is an important issue here: it represents a factor which constrains how TE is used, in addition to the technical constraints discussed above. Genre is de-fined by Swales as comprising:

> communicative purposes. These purposes are recognised by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic struc-ture of the discourse and influences and constrains choice of content and style (Swales, 1990, p.58; quoted in Sanz, 2010).

This definition implies that genre analysis typically sees genre as having a certain level of formality; it tends to discuss 'experts' and the communities they operate in (Bhatia, 1997); however, arguably many non-formal, less strictly regulated and delimited text genres arise out of diverse communication needs and communities.

The organisational and structural forms employed within a given genre are a matter of convention (Bhatia, 1997). The use of different types of TE for different

purposes appears to vary across genre: for example, as has already been mentioned, words qua words marked by italic TE is a convention frequently seen in linguistic texts. Whilst conventions may be motivated by certain needs within the genre, they will at some level be arbitrary—arbitrariness is, of course, a hallmark of language in general (Saussure, 1959). Arbitrary variation in writing is likely to be particularly great: as Mattingly suggests, in writing systems there can be 'no one, natural set of input symbols' (1992, p.15).

The variety of forms TE could take is enormous. Although the following quotation from Mattingly concerns graphemes, it can equally be applied to TE:

> The shapes of the signs in the writing systems of the world and the way they are arranged are extremely various, and such limitations as exist are to be accounted for not by cognitive or linguistic factors but by practical ones, such as the nature of the writing materials available and what patterns are easily written by hand, or by aesthetic ones, such as the beauty of particular stroke patterns (Mattingly, 1992, p.15).

This means that different genres have a vast array of potential conventions to choose from. However, the potential for variation in form does not mean that the form of TE varies completely freely across genres. Typography, which stands at a level outside of text genres, has its own conventions, resulting in a limited set of visual forms (e.g. bold, italic, small caps etc.).

It is clear that, even within a genre, conventions for TE use vary. For example, whilst the interpolational use of italic TE to indicate new terms is widespread in the linguistic literature, it is not the only approach adopted: taking examples from some of the IS literature cited in this thesis, italic TE marks new terms in the text of Calhoun (2009), Zimmermann (2007), and Bolinger (1961), yet, in other documents, new terms are a) not marked at all, such as in Bock and Mazzella (1983) and Halliday (1967); b) delimited with quotation marks, such as in Benatar and Clifton (2014) and Birch and Rayner (1997); or c) marked with small-caps TE, such as in Lambrecht (1994). Part of this variation may be due to a lack of *explicit* convention. This is discussed by Waller (1980), who suggests that 'there is no assurance that writers, or indeed typographers, use typographic options consistently; there is no well-argued and theoretically-sound source of guidance for the composition or evaluation of complex text' (p.242).

This complexity and variability may result in high individual variation in terms of how individuals interpret and process TE. Depending on their expo-

sure to, and possible participation in different genres, they may associate different types of TE with different uses. Providing a genre-situated view of TE is beyond the scope of this thesis; it would be difficult to produce artificial experimental stimuli that controlled all necessary variables and also conformed to a given genre. However, in its experimental material it will attempt as far as possible to model professionally printed text such as might appear in the form of a published article or novel.

Individual variation is, of course, an issue in all research on humans, and one that can be dealt with in part with appropriate sampling. The fact that a higher level of individual variation is expected for responses to TE than for responses to its (hypothesised) equivalent of PAs means that the strength of predictions made here must be tempered, and effect sizes are expected to be relatively small.

A different limiting factor on TE usage may arise out of the fact that reading and writing must be formally taught, unlike speech. This means that different people may have different levels of acquisition of writing dependent on their educational history; furthermore, conditions like dyslexia that affect people's ability to read and write are relatively common. Spoken language acquisition can, of course, be subject to similar developmental issues—but its precognitive nature means that, whilst it is dependent on social factors, it is not additionally dependent on the quality of education received. A relatively subtle feature of writing such as TE may well bypass many people, even those that are relatively highly educated, given that it is rarely discussed or studied. It may be that the way people use and interpret TE is dependent on how frequently they encounter it—namely, how much reading they do.

Even for those relatively well-acquainted with TE, a perception of the device as being more a matter of aesthetics and style than a linguistic tool (e.g. Truss, 2003) may affect both usage and interpretation. Consider that some uses of TE do appear to primarily aesthetic (as mentioned in §3.3 above); this is quite distinct from spoken forms of emphasis, which might be employed in aesthetic endeavours such as singing, but are rarely aesthetic in and of themselves within normal communication.

3.5   Reading typographic emphasis

Up to this point, this chapter has discussed the form and intended functions of TE, and has framed this discussion largely in terms of writing rather than reading. This study is concerned with investigating how TE is interpreted (more specifically,

how it is processed), and so discussion must turn to reading.

The relationship between writing and spoken language has been treated as complex here: the argument has been that writing encodes speech, but not all aspects of speech, and also that it encodes information which speech does not. A naïve view of reading would be that it is a decoding process driven entirely by what is written. Psycholinguistic research into reading in fact suggests a more complex picture, removed from what might appear intuitive.

3.5.1   Reading

Reading is an enormously popular area within psycholinguistics (see Chapter 1), which has given rise to a huge body of research (for reviews see Rastle, 2007; Rayner and Reichle, 2010). In the face of this, a simple, general definition is a useful place to start: 'Reading is the ability to extract visual information from the page and comprehend the meaning of the text' (Rayner et al., 2012, p.23). This might seem to suggest that reading is a process of linear decoding. Smith makes the point that reading is not a cognitively unique process in any single respect:

> There is nothing special about reading in terms of what a reader has to do. Reading does not make any exclusive or esoteric demands on the brain. There are no unique kinds of movements that the eyes must make in reading that they do not make when we examine a picture or glance around a room. No particular kind or degree of visual ability is required to discriminate among printed letters or words (Smith, 1994, p.1).

However, reading is a highly *complex* process; its complexity is a result of both the subtleties of the semiological systems of writing and speech, their relationship between each other, and the processes required to negotiate this relationship. The definition given above from Rayner et al. (who, to be clear, do not in any way imply that reading is simple) splits reading into two components: a perceptual component, in which information is extracted from the text, and a comprehension component, in which meaning is arrived at.[10]

The perceptual component of reading involves eye movements: the reader must direct their attention towards a sequence of graphemes. This process has

---

[10]The temptation may be to think of these as 'stages', but the term *component* is used to here to avoid implying a simple order.

been studied in minute detail, using techniques which precisely record and measure eye movements. These techniques, and some of the insights they have provided into how language is processed, will be further discussed in the following chapter (§4.3.1).

For now it should suffice to touch on three basic aspects of visual behaviours during reading. Firstly, readers do not look at every single word; for example, shorter words, and more predictable words are less likely to be fixated (Rayner, 1998; Drieghe et al., 2005). This suggests that reading is not an entirely bottom-up, data-driven process, but involves top-down processing. Secondly, readers do not proceed 'forwards' from word to word at all times, but frequently backtrack (Rayner, 1998; Rayner et al., 2006). The ability to do this is a luxury not afforded when hearing speech; that it occurs shows that reading is a less linear process. Finally, the length of time readers look at a word varies in proportion to various properties of that word; for example, the more common a word, the less time it is fixated, and if a word increases the grammatical complexity of the sentence, it receives longer fixations (Just and Carpenter, 1980; Rayner, 1998). These properties mean that eye movement measuring techniques are immensely useful to psycholinguistics.

The second component of reading, comprehension, can be seen as a process of decoding, although two caveats should be raised concerning the use of this term: firstly, 'decoding' implies that the aim of comprehension is for the reader to arrive at the same meaning the writer intended, yet there is no guarantee that this will happen. Secondly, as has been discussed above, what is encoded, and thus what is decoded, may not be entirely linguistic in nature.

To ask 'how is comprehension achieved' would be to open the floodgates to a huge amount of research and debate within cognitive psychology. It is more pertinent, and somewhat more manageable to question how (or whether) *phonological recoding* fits into the comprehension process. Writing might be seen as a process of encoding meanings into phonemes, and then encoding those phonemes into graphemes. Reading might thus be seen as the reverse of this: the route of the input would pass from the *orthographic* level (that of the grapheme), through the phonological level, to the semantic level. Such a view has been advanced in the form of the *obligatory phonological mediation hypothesis*: in Rapp et al's rather evocative phrasing, this hypothesis sees 'written language [as] entirely parasitic upon spoken language' (Rapp et al., 1997, p.72).

This issue is highly relevant to the present study. To summarise what has been discussed above: writing encodes speech, which encodes meaning; it has been

proposed here that TE encodes PAs, which in turn encode CF. The question is, does the *de*coding process require the intermediate sound-based step: can readers go directly from the visual phenomenon of TE, to access the semantic and pragmatic aspects of CF, or must they first 'hear' the intonation which TE encodes?

3.5.1.1 The 'inner voice' and phonological recoding

The literature tends to agree that the intuition of an 'inner voice' when reading is extremely common (e.g. Slowiaczek and Clifton, 1980; Chafe, 1988; Perrone-Bertolotti et al., 2012); this is the voice which can appear to be speaking 'in the head' during reading. This phenomenon has been studied under various names: Leinenger gives *subvocalization, inner speech, speech recoding* and *phonological recoding* (Leinenger, 2014, p.1535) as examples; the latter will be used here. An intuition is, of course, not enough to go on—but it is actually logical to assume a phonological stage of processing can be activated during reading. Without such a stage, readers would be unable to read nonwords such as *swelve* or *froized* aloud, or, by the same token, valid words that they have not encountered before (Price, 2012).[11] The fact that it is possible to read aloud in general must mean that there can be a route between orthography and phonology (Taft and van Graan, 1998). The importance of phonology in learning to read is also widely accepted (Perfetti, 2003; Hulme et al., 2005), although this does not mean that skilled reading must still use phonology (Leinenger, 2014).

There is also psycholinguistic evidence for phonological recoding during reading. Reading rate can be influenced by whether the reader thinks the text was written by a person who speaks quickly or slowly: Alexander and Nygaard (2008) demonstrated this both for reading aloud and reading silently, by measuring reading rates for texts which were identified with either a slow or fast speaker whose recorded speech was played to participants prior to the test. On a similar note, Filik and Barber (2011) used eye tracking to show that participants reading limericks show evidence of processing disruption when reading rhymes which depend on accents different to their own.

The tasks in both of the studies described above had explicit connections to sound, which may have been what prompted these phonological recoding effects. Perrone-Bertolotti et al. (2012) addressed this issue by conducting an experiment which involved no auditory stimulation, and found that temporal voice-selective areas (areas of the brain which respond to heard voices) can be activated during

---

[11]Nonwords taken from the ARC Nonword Database (Rastle and Coltheart, 2002).

reading. This study used intracranial electrodes located in the auditory cortices of epileptic patients. Rapid serial visual presentation (see §4.3.2) was used to present two stories word-by-word, intermixed: the task was to attend to one of the stories, set in white words, and ignore the other, set in grey words. Perrone-Bertolotti et al. found that temporal voice-selective areas responded during reading, and that this effect increased for attended words. They suggest that inner voice effects may be modulated by attention, and point to another finding of Alexander and Nygaard (2008), of an interaction between more difficult texts and the assumed speech rate of the author: the more difficult the text, the more reading speed was determined by speech rate. Perrone-Bertolotti et al. (2014), reviewing this and related work on phonological recoding, suggest that 'the little voice in our head during reading is not present all the time'.

Perhaps the greatest difficulty here lies not in uncovering evidence of phonological recoding, but divining the point (or points) in processing at which it occurs. Leinenger (2014) points out that the evidence found thus far for phonological recoding, including that outlined above, has not yet clarified this issue. She groups together three strains of thought within the literature regarding how phonological recoding fits into reading: The first is that there is an early stage of phonological recoding, possibly required in order for the meaning of words to be accessed; the second posits a later stage which arises out of lexical access, with recoding playing other roles (such as helping short-term memory and adding prosodic information); the third viewpoint is that any phonological recoding occurring during reading is a by-product of how reading is taught, and is thus not strictly necessary to the process of reading as a whole.

The lack of clarity regarding the role of phonological recoding in reading is disadvantageous to the present study. It would benefit from being able to make more specific predictions: does TE-marking of CF first prompt a PA 'heard' by the inner voice, as per Chafe's view of punctuation (1988)? Or does TE instead prompt a contrastive reading more directly, as might occur with clefting?

3.5.2 Studies on typographic emphasis

As has already been mentioned, linguistic work in general on typography has been limited (Waller, 1996). Psycholinguistically-motivated work on TE specifically is rarer still. Only two such studies appear to have been published (McAteer, 1992; Fraundorf et al., 2013). Both of them must, of course, be reviewed here.

The earlier paper is Erica McAteer's 1992 article *Typeface emphasis and information focus* (based on McAteer, 1989, an unpublished Ph.D. thesis). Ostensibly,

her subject matter is very close to that of the present study; however, there are some issues with the methodology she employed, which will be discussed here. Before proceeding, it should also be noted that this research was conducted at a time when accessible word-processing was still in its infancy, and the explosion of text-based internet communication into day-to-day life was yet to occur. As a consequence, the behaviour and metalinguistic intuitions of her participants may not reflect those of a similar sample of participants today. At the same time, it is unlikely that the typographic character of professionally printed material has changed significantly in the last twenty five years.

McAteer conducted a series of five studies into TE. The first was qualitative and metalinguistic in nature: 40 participants were asked to give brief statements. The following is an excerpt from the prompt that was used:

> Thinking back over our own experience of typeface change when reading textbooks, magazines, fiction or whatever, please give me a brief statement saying why capital letters may be better for some cases where typeface emphasis is required, and italic print for others. Do you think they may actually mean something different? (McAteer, 1992, p.349)

A set of attributes was distilled from the freeform responses. Italics were found to receive more associations with 'contrast' and 'connotation'; all-caps had more associations with 'prominence' and 'importance'. As has been discussed in the present thesis, terms such as these can be rather difficult to pin down. Participants naïve to these issues can hardly be blamed for this; indeed, vagueness is a danger of soliciting qualitative judgements about such subtle phenomena (see §4.3.3).

McAteer's analysis of this is somewhat problematic. She proposes two functions of emphasis: 'Modulatory' emphasis is what the participants identified with all-caps TE, and is held to operate on an emphasised constituent such that it 'means itself, only more so'. 'Contrastive' emphasis, identified with italics, is held to involve opposition between the emphasised element and some other element (whether implicit or explicit). The latter is clearly identifiable with CF as it has been defined here, but the former, 'modulatory' emphasis is confusing.

(42)   a. Did John buy a hat?
           John bought a *book*.
       b. Guess what John bought?
           John bought a *book*.

McAteer uses (42) to illustrate the distinction, claiming that (42a) is contrastive, and, if John is known to rarely buy books, (42b) is modulatory. Each case certainly has different semantics for *book*, including different sizes of its alternative sets—but under Zimmermann (2007) both constitute CF, in that both are judged by the speaker to be unexpected by the hearer.

McAteer's description of modulatory focus as making a word 'mean itself, only more so' seems to refer to degree, perhaps as in *it wasn't just black, it was BLACK*.[12] However, this emphasis is still contrastive in nature; the alternative set is implied to consist of differing degrees of the quality, with an unexpectedly extreme degree being selected.[13]

This confusion seems to arise from two issues: firstly, given a psycholinguistic framework, studying multiple forms of TE simultaneously is overreaching, as surely the previously unstudied primary contrast between $[+\text{TE}]$ and $[-\text{TE}]$ is more important than intra-TE contrasts. Secondly, the lack of a coherent position on IS means that some of the comparisons drawn are underspecified and unhelpful. With this said, McAteer's work is useful to the present study, as it does provide evidence that TE is processed differently.

This evidence is found within the final study McAteer reports: a self-paced reading experiment which tested whether placing TE on a pronoun would affect how it was resolved. The rationale was based on the idea that placing a PA on a pronoun in speech makes it contrastive (Bösch, 1983; Hoop, 2004).[14] Participants were presented with a series of short texts such as that shown in (43).

(43) The glass bowl lay shattered on the floor.
Simon said Fred did it.
He did.
Carol sighed and fetched a brush.
Who did it?

Unlike in the majority of self-paced reading tests, participants advanced through the texts at their own pace a sentence at a time, rather than word-by-word (see

---

[12] That this is the intended meaning of 'modulatory' emphasis may be supported by Sanford et al. (2006), who define modulatory emphasis as 'raising the force of an item on some scale' (p.115).

[13] Further confusion is caused by McAteer citing Dik et al. (1981) for the modulatory–contrastive opposition: this is a misreading or a misattribution, as Dik et al. (1981) does not propose or discuss this particular division.

[14] As Bosch points out, not all stressed pronouns are contrastive: non-anaphoric pronouns used in a deictic way are stressed, as in 'Did anybody leave that lecture yesterday? *He* left' (Bösch, 1983, p.58); if there is no recoverable referent for *he*, it is likely to be stressed, as well as accompanied by a pointing gesture.

§4.3.1.2 for a description of self-paced reading studies). Each test item contained a verb phrase such as *Simon said Fred did it*, followed by a verb phrase containing an ambiguous pronoun that may have referred to either proper noun; *he did*. Italicisation of the pronoun was manipulated between participants. After each text the participant was asked to choose the referent of the pronoun, being presented with the nouns from the preceding verb phrases: for (43), the choice would thus be between *Simon* or *Fred*). McAteer was primarily interested in whether italicisation of the pronoun would cause participants to choose an 'alternative' referent to that of the 'default' reading.

Indeed, 'alternative' referents were chosen more often in the [+ TE] condition. For each stimulus McAteer decided the 'default' referent was whichever noun the majority of participants chose when there was no TE, and then tested the difference between conditions based on this; this is a somewhat circular approach, but does not detract from the fact that the TE provoked a 'shift' in interpretation.

One problem here is the pragmatic strangeness that results if one reads the [+ TE] condition as having contrastive focus. In (43) the 'default' [− TE] condition reading of *he did* would be that it is a confirmation of the first sentence, with the pronoun referring to the subject of the subordinate clause (*Fred*). If the TE on the pronoun is to be interpreted as marking a contrastive focus, this would imply that the pronoun refers to the subject of the preceding sentence (*Simon*) instead. However, the result is infelicitous, as the repetition of *did* with no other cues (such as clefting, or markers like *in fact*) actually implies that *Fred* is still the referent. This might account for why, in the [+ TE] condition, participants spent longer looking at the sentence containing the TE, as they might have been trying to process this infelicity—although such a result may also be simply due to the increased salience of the emphasised words prompting increased attention. Reading times at the word level rather than the sentence level would be useful here, in particular.

In the [+ TE] condition, participants were quicker to choose the referent. This may indicate something about emphasis promoting the encoding of alternatives—a hypothesis which is investigated more fully by Fraundorf et al. (2013), discussed below. However, this may have been due to the [+ TE] pronoun sentence being interpreted as pragmatically odd, necessitating re-reading (and thus a more deliberate decision over the referent at that stage).

Given the strong association between focused pronouns and contrast, McAteer's stimulus design here might be seen as a good way to test whether TE can perform the same function as PAs. However, a major problem here is that pro-

nouns are often not fixated at all during reading: the effects here may be due to visual salience making the TE words more noticeable rather than emphasis prompting a different IS reading. Word-level reading times such as those provided by word-by-word self-paced reading, or eye tracking, would help considerably. However, even if this study fails to cast much light on *how* TE affects processing, it is clear that it supports the basic hypothesis that it *can* affect processing.

In a more recent study, Fraundorf et al. (2013) look at the relationship between TE and IS from a different angle, investigating how TE affects memory. Although this is a more up-to-date, and indeed more rigorous piece of research, in some ways it is similar to McAteer's work: the TE in the stimuli is in the form of italics and all-caps, and choices between referents are used as a metric.

This study is based to a large extent on a prior study by the same authors (2010), which is thus worth examining in some detail first. This study involved spoken rather than written stimuli, and tested how memory was affected by different PAs. Participants were played short discourses in which a context sentence established limited alternative sets for focused elements in a 'continuation' sentence, as in (44).

(44)　　a.　(Context) Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys.
　　　　b.　(Continuation) Finally, the British spotted one of the monkeys in Indonesia and planted a radio tag on it.

In (44b), *British* has the alternative set {*British, French*}, and *Malaysia* has the alternative set {*Malaysia, Indonesia*}. Two conditions were used: either the first focused word was given a 'contrastive' L+H* PA, and the second a 'presentational' H* PA; or vice versa. Recognition memory was tested after presentation of these discourses: the text of each discourse was displayed on a screen with the target words blanked out, and participants had to make a choice between the two possible missing words for each target. Fraundorf et al. (2010) found participants remembered words with the 'contrastive' accent better; in a subsequent two-session experiment, they showed that this effect persisted over 24 hours.

This later experiment involved a subset of the same discourses, but instead of a 'fill-in-the-blank' task, participants were asked to decide whether probe sentences were true or false. The probes either named the correct item, the contrasting item, or an unmentioned item, as shown in (45).

(45) a. (Correct) The British scientists spotted the endangered monkey and tagged it.

  b. (Contrast) The French scientists spotted the endangered monkey and tagged it.

  c. (Unmentioned) The Portuguese scientists spotted the endangered monkey and tagged it.

Participants were more accurate in rejecting the probes containing the contrasting item than the unmentioned item. When the target word had been presented with the contrastive accent, this facilitated the rejection of contrastive probes; it did not facilitate rejection of the unmentioned probe. Fraundorf et al. took this to show that CF prompted improved encoding of the alternative set.

The experiments reported in Fraundorf et al. (2013) tested whether the same encoding of alternatives occurs when TE is used in the written form, in place of PAs in the spoken form. The materials and procedure were a near-replication of those of the 2010 study, with the principal difference being the modality. Participants were presented with written versions of the stimuli, with the presence of TE independently manipulated on the critical words. The type of TE was also manipulated between participants; both all-caps and italic TE were used. After this the same true/false test phase as before was used, with the same types of probe that were illustrated in (45).

Fraundorf et al. (2013) found that the presence of TE on the target word improved rejection of the other explicit alternatives, and did not facilitate rejection of unmentioned items; this was the same result as found for the PAs in the earlier study. Importantly, Fraundorf et al. note that their results must not solely be related to the perceptual prominence of TE, because if it were, then the TE would have facilitated rejection of items that had not appeared in the discourse. This suggests that TE has a systematic effect that is not simply related to drawing attention to certain elements; rather, it appears to be able to trigger the same kind of behaviour that PAs on CF do within speech.

They also note a 'marginal' non-significant result ($p = .09$) which suggests that capitals may have a stronger effect than italics. With all due provisos about accepting the concept of 'near-significant' p values, this may be due to the larger visual salience of all-caps; alternating between a word set in lowercase (or initially capitalised) and one set in all-caps involves larger changes of form on a letter-by-letter basis than most italic fonts. It could, however, be due to a qualitative difference in how the two types of TE are interpreted. The qualitative component of McAteer's study suggested this might be the case, although Fraundorf et al.'s

results go in the opposite direction to the metalinguistic opinions of McAteer's participants: she found that italics were more associated with contrast than all-caps, but the Fraundorf et al. study suggests the opposite.

Before concluding this chapter, it is worth briefly considering the MA thesis which preceded this study (Norton, 2012). This represented an ambitious attempt to investigate how both phonological *and* interpolational italics are processed, though it was hampered by theoretical and methodological issues.

Two self-paced reading experiments were conducted in which participants read short sentences. The presence of TE on the fourth word, which was always a noun, was manipulated.

(46)   a. (Phonological TE) Often they ignored *Edmund* at lunch.
       b. (Interpolational TE) Yesterday we delayed *Ribbon* by weeks.

To test phonological TE processing, sentences such as (46a) were used, in which the target word was the name of a person. To test interpolational TE processing, nouns which were not names, but which were capitalised, were used as target words. The intent was to imply that these words were the titles of published work; films, books, theatre productions, etc.

Several problems with this design are apparent even from this brief summary: in particular, the sentences were very short, and very little contextual information was available. It is difficult to see where contrast could have been created with such minimal discourse context. Furthermore, as has been noted before, attempting to draw comparisons between phonological and interpolational TE is an overreach: the field of typographical linguistics is too new, and the knowledge base too undeveloped for such investigation.

It is perhaps unsurprising, therefore, that no significant differences were found between items read in the [+TE] and [−TE] conditions. This null result was also probably related to a lack of visual contrast between the [+TE] target words and the surrounding text, given that the self-paced reading paradigm used did not allow both the target and its surrounding words to be viewed simultaneously: this will be discussed further in §4.3.1.2.

Chapter 4

Methodological approach & Experiment 1

4.1 Introduction

As was established in the previous chapter, there has been a lack of psycholinguistic studies on typographic emphasis (TE). Of the two directly relevant studies that have been published (McAteer, 1992; Fraundorf et al., 2013), neither investigated online processing. The MA thesis which preceded the present study did attempt to measure processing, but was hampered by methodological issues (Norton, 2012; see §3.5.2 for discussions of all three studies). In order to investigate whether typographic emphasis marking of contrastive focus (CF) affects sentence processing, it is thus necessary to evaluate the techniques that have been developed to investigate processing during reading, and to construct novel materials. The former issue is discussed in the first part of this chapter. The latter is discussed in the second part, in the form of a report on the design and results of a pilot experiment (referred to from hereon as Experiment 1).

Experiment 1 was designed to investigate the primary hypothesis that CF will be processed more quickly when it is marked with typographic emphasis, and the secondary hypothesis that TE will have a non-linguistic 'eye-catching' effect due to its visual salience. Participants read a series of short dialogues on a screen whilst their eye movements were recorded. TE on a target CF word in each dialogue was manipulated between participants. After this experiment was run as a pilot, it was decided to considerably alter the design before conducting a full study. The creation of Experiment 2 (the design and results of which are fully reported in the next chapter) was prompted by concerns regarding certain aspects of the original design and materials. Chief of these concerns was the fact

that in the pilot, when TE appeared, it always appeared on CF. This meant that the experiment would have been unable to show whether TE has a different effect on processing when placed on non-CF words (although it could address the hypothesis outlined above).

Despite the decision to substantially alter the original pilot experiment before the full study was conducted, the original pilot experiment is reported here for the following reasons:

1. Despite the low statistical power due to the small pilot sample ($N = 10$), and a resulting low number of significant results, it did find two statistically significant results: when the target word had TE, participants spent longer overall looking at the word with which the target word contrasted; they also looked at the word following the target word more times in the same condition.

2. As a whole this experiment represented an early step towards psycholinguistic investigation of TE, amongst scant comparable literature: at such an early stage reporting of null results can only further the field.[1]

3. The materials used were markedly different to those used in the full study, and thus may bolster the results of the full experiment by providing perspective on how TE is read in different types of texts: the pilot used script-like dialogues, whereas the full experiment used narrative texts.

4.2   Organisation of chapter

§4.3 reviews the most prevalent experimental **techniques used to investigate processing** during reading, dividing these into eye movement (§4.3.1) and neuroimaging techniques (§4.3.2). Qualitative methods are also discussed (§4.3.3).

§4.4 provides background on the **connections between eye movements and processing**, and gives an overview of eye tracking metrics (§4.4.1).

§4.5 outlines the two hypotheses that Experiment 1 was designed to test.

§4.6 then reports the methodology used for **Experiment 1**.

§4.7 reports the **results** in full. This section includes analysis of inter-participant and inter-item variation.

§4.8 presents a preliminary **discussion** of the results of this pilot experiment.

---

[1]Arguably, this is important at *any* stage; see Rosenthal (1979) on the 'file drawer' problem.

4.3 Techniques to investigate processing during reading

The central problem when investigating language processing is the same that is faced by cognitive psychology in general: gaining access to the workings of the mind is extremely difficult. Whilst input and output are manipulable and measurable, directly observing the processes that occur between them is so challenging that the mind has often been regarded as a 'black box' (Sternberg and Sternberg, 2017). Experimental techniques including eye tracking and electroencephalography have made this box more transparent, but, when sources such as Osterhout et al. (2004) discuss the 'ideal method' of measuring processing during sentence comprehension, it is under the assumption that such a method does not yet exist.

Osterhout et al. (2004) identify three qualities for this 'ideal method'. It should:

1. 'Provide continuous measurement throughout the process of understanding a sentence;
2. Have a temporal resolution exceeding that of the relevant processes;
3. Be differentially sensitive to events occurring at different levels of analysis (phonological, syntactic, semantic, etc.)' (Osterhout et al., 2004, p.271)

If no existing technique fully possesses all three qualities, then it follows that an ideal method for the more specific task of measuring TE processing does not exist yet either. Nonetheless, weighing potential techniques against an ideal standard, however unattainable, is a useful method of evaluation. In addition, then, to the general qualities proposed by Osterhout et al., the ideal method for the current task should:

1. Allow for the simultaneous presentation of enough text that the necessary manipulations of information structure can be achieved. Manipulation of CF may require several entire sentences, as it requires the reader to have an expectation of a certain level of prominence or markedness for the CF word, and for that expectation to then be exceeded (Zimmermann, 2007; Calhoun, 2009): discourse context is thus required. Therefore, the ideal method should allow for several lines of text to be presented.
2. Give the reader sufficient time to make as many visual passes (see §4.4.1) as are required for full comprehension of the stimulus. It is hypothesised here that re-reading of material prior to (and including) the target word will occur when the CF is not marked with TE; this must be captured if it occurs.

3. Display the text in a form that allows for the presentation of different font styles simultaneously, to allow TE to be placed.

4. Allow for naturalistic reading of text presented in an ecologically valid manner. Ideally, the participant should be in a comfortable environment, reading text which does not appear to them to be particularly unusual, and reading it in a naturalistic way. For this study, the concern is for the stimuli to have the same typesetting characteristics as professionally printed, published text. This is motivated by the relationship between TE usage and genre/convention.

It will be argued here that eye tracking is able to satisfy these four specific qualities (although it cannot fully satisfy the three general 'ideal' qualities suggested by Osterhout et al., 2004).

Techniques to measure linguistic processing during reading can be divided into those based on **capturing eye movements**, which include eye tracking and self-paced reading, and **neuroimaging techniques**, which directly measure brain activity, and include encephalography and functional magnetic resonance imaging (Garrod, 2006). Aside from this study's preferred method of eye tracking, self-paced reading and neuroimaging constitute a large part of psycholinguistic work similar to that conducted here, and are thus worth evaluating. Given the focus of this study (and the resources available to it), self-paced reading is the closest competitor to eye tracking, and will thus receive greater attention than neuroimaging (an additional reason for this attention is that self-paced reading was used in two of the prior studies on TE, McAteer, 1992; Norton, 2012).

In addition, qualitative methods, such as the use of questionnaires and interviews regarding what participants think of TE, will be considered. Such techniques are not experimental, yet neither are they without relevance, having been employed to examine TE before (e.g. Chafe, 1988; McAteer, 1992; Ashby, 2017).

One limitation that might be common to all the techniques discussed below should be considered: they may fail to distinguish between normal reading and the *mind wandering* (Feng et al., 2013) which can occur when a participant's attention leaves the task. More specifically, during reading a familiar phenomenon can occur in which the eyes move across the text as if reading is occurring, whilst the mind is elsewhere (an understandably difficult phenomenon to study; see Schooler et al., 2004). The best defence against this is probably careful task design: in the experiments conducted here, comprehension questions were used to ascertain whether participants were paying attention, and effort was taken to make the material sufficiently interesting and motivating.

4.3.1 Eye movement techniques

Eye movements during reading consist of the same types of movements that occur when viewing any two or three-dimensional image or scene: the eye makes a continual series of movements, or *saccades*, to different parts of the visual field (Rayner, 1998), in order to project whichever parts of it are of interest onto the *fovea*, the part of the retina with the greatest visual acuity (Holmqvist et al., 2011). Saccades are extremely fast—fast enough that no new information is obtained during them, as only a blur would be perceived (Rayner, 1998). Rather, information is obtained between saccades when the eye is relatively still, during *fixations*, which typically range from 200–300 ms (Rayner, 1998).[2]

Psycholinguistic techniques that measure eye movements rely on the assumption of there being what Kliegl et al. (2006, p.13) call 'a fundamental link between language-related and oculomotor processes'. In reading, there are connections between online language processing and the duration of fixations, although the relationship between some types of processing, particularly higher-level processing, and eye movement behaviours itransparent less clear. Issues pertaining to this are discussed in more detail below (§4.4.1), but for now it suffices to state that the basic existence of a relationship between eye movements and processing has meant that techniques which measure fixation times on words and phrases have become very popular. This popularity has resulted in well-developed, well-tested methodologies being developed within these paradigms (Rayner, 1998; Holmqvist et al., 2011).

Before examining eye tracking and self-paced reading separately, it should be noted that, viewed in the context of Osterhout et al. (2004)'s overarching 'ideal' method of investigating processing, both the eye movement techniques discussed below share one central limitation: neither technique is differentially sensitive to different linguistic levels. No particular eye movements appear to be directly associated with particular linguistic levels. This is in contrast to neuroimaging techniques; for example, some of the event-related potentials captured by electroencephalography appear to be sensitive to semantic violations, and some to syntactic violations (Brennan, 2016).

At the same time, eye movement techniques share the advantage of a lower

---

[2]There are other types of eye movement, including *pursuit* movements when tracking moving objects and corrective movements that alter focus, and the constant small tremor-like movements known as *nystagmus* and *microsaccades*. Pursuit movements are unlikely to happen during reading in an experimental setting (both text and reader are likely to be stationary relative to each other), and nystagmus and microsaccades are typically treated simply as noise in reading research (Rayner, 1998).

cost than neuroimaging. In addition to their lower fiscal cost, they are cheaper in terms of the amount of time they demand to be spent on design, testing and analysis (Boland, 2004).

Eye movement techniques typically use computer displays in order to present stimuli to participants, and this raises certain issues. In terms of the present study, there may be some effect of using the computer screen rather than printed material, in that the digital medium may imply a different genre with different conventions of TE, but it seems unlikely that this would greatly affect results; differences in behaviour here are likely to be very subtle and small, if they exist at all. Eye tracking can actually be used to record participants looking at things other than computer displays, but it would be prohibitively difficult to conduct reading experiments such as those that were carried out for this study using physical, printed stimuli; calibration, trial order randomization, comprehension question display etc. would all be more complicated and time-consuming.

On a related note, there is a considerable body of literature, particularly within the education field, on how comprehension is affected in general when reading on paper versus reading on screens (e.g. Mangen et al., 2013; Margolin et al., 2013). This may have implications for the generalisability of the results from this study, although only inasmuch as it has implications for the generalisability of the vast majority of psycholinguistic research on reading, which overwhelmingly favours the use of computer displays in experiments for the reasons of convenience mentioned above. At the same time it should be acknowledged that reading from screens is now an everyday activity for many, if not most literate people, and certainly for the typical university population from which this study draws its participants (Walsh, 2016): psycholinguistics has been using computer displays for a long time, but whilst screen reading once might have contributed to the unnatural atmosphere of the testing experience, it is now prevalent enough to mitigate this particular issue.

4.3.1.1 Eye tracking

Eye tracking has been chosen as the method to be used in the present study. Boland (2004) succinctly summarises the two key advantages it holds over other techniques:

> The eye-movement record provides an on-line measure of processing difficulty with high temporal resolution, without relying on any secondary task to produce the dependent measures (Boland, 2004, p.52).

As such, eye tracking goes some way towards being the ideal method posited by Osterhout et al. (2004). It also fulfils all of the more specific qualities posited above: these will be examined following an overview of the technique.

The most prevalent technique used to measure eye movements and to capture where participants are looking involves video tracking with infra-red light sources and high-speed cameras. This is the method discussed here, but other techniques do exist, such as electrooculography, which measures eye movements by monitoring electrical impulses in the muscles used to move the eyes (see Holmqvist et al., 2011 for a thorough overview of different techniques, both past and present). Note that the following overview takes the EyeLink 1000 Plus system used in this study as a baseline example, although the details discussed should hold true for most video tracking systems.

A typical eye tracking set up involves an infra-red light source shining towards the participant's eyes as they look at a computer screen. The light source creates a reflection on the participant's cornea, and this is visible to an infra-red camera focused on the eye (Holmqvist et al., 2011). The point at which the eye is gazing on the computer screen can then be calculated from the image of the eye by an algorithmic comparison between the position of the corneal reflection and the position of the pupil (Holmqvist et al., 2011). One or both eyes can be tracked in this manner.

The initial output is typically a sequence of two-dimensional Cartesian coordinates representing pixels that were at the centre of the foveally-fixated area; this stream of coordinates can be processed in order to infer details of fixations and saccades (Holmqvist et al., 2011). Fixations are taken to occur when sequences of coordinates remain within a certain radius (given that the eye is never truly still), whereas sequences that show a rapid move from a start point to an end point indicate saccades (different eye tracking systems may use different algorithms to distinguish these). Each pair of coordinates represents a *sample*, calculated from a single image from the camera. The sample rate (in other words, the recording speed) of the eye tracker is thus of crucial importance: fixations that are shorter than the sample rate will not be captured. The EyeLink 1000 Plus records at a rate of 1000 Hz, giving the system a resolution higher than the generally-accepted minimum duration of fixations which reflect processing, which is around 100 ms (McConkie, 1983).

Eye tracking systems require calibration on a per-participant basis to establish baseline relationships between the location and size of the participant's eyes, the position of the camera, and the position and size of the display (Duchowski,

2017). At a bare minimum, calibration must occur before a recording begins, but may also need to occur throughout an experiment. The need for recalibration is dependent on factors such as the participant slightly shifting their position: a sufficient degree of movement in any direction can alter the angles between the corneal reflection, the pupil and the camera, thus reducing accuracy. Most calibration routines involve the participant being asked to fixate a series of fixation points displayed on the screen (Holmqvist et al., 2011).

Subsequent processing and statistical analysis depends on the research question at hand. For reading studies such as the present one, one or more *areas of interest* (AOIs) are defined for each stimulus. These are two-dimensional areas (typically rectangular in reading research) which are analysed as discrete items. A single AOI might cover a letter, a word (as in the experiments conducted here), a sentence, or any arbitrary section of the image. Some analyses can involve overlapping AOIs, or AOIs treated as groups (Holmqvist et al., 2011).

The division of stimuli images into AOIs allows for the calculation of a variety of statistics, such as the total time the area of interest is fixated, the number of times the AOI was fixated, the order in which a series of AOIs were fixated, etc. These are referred to as *metrics*, and will be discussed in more detail below (§4.4.1).

Advantages of eye tracking for the current study   Eye tracking fulfils each of the specific ideal characteristics outlined above, qualifying it as the most suitable method to investigate processing of TE:

1. It allows for the presentation of multiple sentences at once: whole pages of text can be presented if necessary. For the type of stationary camera tracking used in this study, the only limit in terms of quantity of text would be whatever is legible without the participant moving their head from the chin rest.
2. The stimuli can be presented for as long as it takes participants to read and re-read them; there is no time limit on the display;
3. Typographic manipulations can be displayed in any conceivable combination. Eye tracking lends itself to the presentation of images, meaning that, if necessary, images of professional-quality text typeset using external programs can be used. This is what was done for the experimental stimuli here, using the LaTeX typesetting engine;
4. Naturalistic stimuli can be used: it is even possible to use professionally printed material on paper in an eye tracking experiment—although a com-

puter display is far more convenient. Whilst there are limitations on the nature of the stimuli, these are more due to experimental design concerns than limitations of eye tracking itself.

Limitations of eye tracking for the current study   Unlike self-paced reading, which requires a relatively unusual mechanism for reading, eye tracking allows more naturalistic behaviour. At the same time, however, the presence of the camera and the necessity for repeated calibration may increase the likelihood of the participant becoming more aware of the process, and more self-conscious.

Depending on the eye tracking set up, an additional issue can be the presence of an investigator in the room (or at least in an adjoining room, controlling the system remotely). In this study, the investigator was present in the same room as the participant. The requirement for an investigator to be present is due to a need for them to monitor the experiment throughout, chiefly in order to administer recalibration when necessary. The presence of an investigator may give rise to observer effects, although not all such effects need be deleterious from the perspective of the success of the experiment: their presence could also have a motivating effect.

4.3.1.2   Self-paced reading

Self-paced reading (SPR) involves the presentation of successive sections of text on a screen, with control over the display of sections given to the participant. It is the closest competitor here to eye tracking for the task at hand, and warrants discussion here in part due to its use in Norton (2012) and McAteer (1992), both studies which have had a heavy influence on the present study (see §3.5.2). However, there is little that SPR offers which is not achieved by eye tracking, aside from a lower cost and shorter development time. At the same time, eye tracking offers substantial improvements over SPR in terms of accuracy and ecological validity.

The technique was originally developed by McConkie and Rayner (1975), and exists in various configurations, with the *non-cumulative moving window* configuration particularly prevalent in reading studies (Rayner, 1998). In this version of SPR, the participant is typically presented with a section of text (e.g. a sentence), with each character masked with a non-letter such as a hash or an underscore. When the participant presses a button, the first word is unmasked. Repeated button presses unmask the following words, whilst re-masking the previous word (the revealed segments could in fact be anything from letters to whole sentences,

as in McAteer, 1992). The effect is thus of a 'moving window' of visible text. (47) illustrates the consecutive stages of a participant reading the phrase *Anne is a girl* (each line of the example would be displayed in the same position on-screen).

(47)   a. `#### ## # ####.`
        b. `Anne ## # ####.`
        c. `#### is # ####.`
        d. `#### ## a ####.`
        e. `#### ## # girl.`

The dependent variable in an SPR study thus measures the amount of time looking at a word. Whilst this is not as precise as the fixation data captured by eye tracking, it has been shown to be a reliable indicator of processing (Mitchell, 2004).

Advantages of self-paced reading for the current study   A major advantage of SPR over eye tracking and neuroimaging is the simplicity of building, running and analysing experiments using the technique: provision for quickly building SPR experiments exists within several experimental presentation software packages, and to run this type of experiment typically only requires a display and a keyboard (or a response pad). Furthermore, there is a single dependent variable which does not require large amounts of pre-processing or cleaning (Mitchell, 2004).

A second advantage is that, in some respects, SPR is less intrusive a procedure than eye tracking, and is considerably less intrusive than neuroimaging. SPR does not require the participant's head to be stabilised or covered with electrodes, and does not restrict their movement. Both of these things can serve as additional reminders to the participant that their behaviour is under scrutiny, and in general seem intuitively likely to result in somewhat unnatural behaviour.

Limitations of self-paced reading for the current study   Whilst SPR does allow for the presentation of sufficiently long stretches of text, and thus satisfies the first quality required for this study, it fails to fully satisfy any of the remaining qualities:

In terms of the quality of allowing for sufficient visual passes, if the 'moving window' is set to display one word in an SPR experiment, then regressive fixations to previous words are not possible: the participant is forced to soldier on, with recourse to memory their only option if re-reading would otherwise be required. This restriction also has implications for the 'naturalistic reading' quality discussed below.

Considering the quality of allowing for the presentation of different font style raises several issues. A major problem with using the moving window paradigm to investigate TE is that the visual contrast of an emphasised word may be reduced: if an italicised word is presented alone, the visual contrast between it and the surrounding text only exists temporally, i.e. in memory, rather than in the stimulus itself. This may make the emphasis less likely to be perceived at all.

In the MA thesis project which preceded this study (Norton, 2012), two non-cumulative moving-window SPR experiments were conducted which manipulated TE between participants on target words within short sentences. The design and theoretical perspective was markedly different to that developed here, but still the single point in which the sentences differed was the presence or absence of TE. No significant results were found in terms of reading times on the target or post-target words. It is highly likely that the above issue of a lack of visual contrast was a major factor in this; this reduction in visual contrast was also compounded by the 'traditional' choice of a fixed-width typeface (Courier New), which does not feature a distinct italic font but rather uses an oblique slanted version of the regular font (see §3.2.1). Such typefaces are normally preferred in SPR and eye tracking studies, as the uniform character width allows for easier calculations and analysis pertaining to what is foveally fixated, visible parafoveally, etc. It is hard to say to what degree this issue might have affected the results for the MA experiments, although, in post-test discussion, several participants reported not having noticed any TE at all appearing in the experiments.

A possible solution to the issue of visual contrast would be to adopt a *cumulative* version of SPR, in which previously unmasked sections remain unmasked as the participant advances through the words. However, this would reduce the correlation between the time between button presses and the time spent looking at the last word revealed, because the participant would be free to make regressive fixations to previous words (Rayner, 1998). Another validity issue can occur if participants manually advance the display several times in quick succession, in order to 'line up' a few words to read in one group; this is not unknown (Just et al., 1982), but could be warded against by explicit instruction before testing.

The final quality for which SPR falls down is that of allowing naturalistic reading. Whilst involving fewer reminders to the participant that they are being monitored and recorded, SPR demands they adopt a very distinct reading style. The peculiarity of this may be a major contribution to what Rayner et al. (1989, p.22) call a 'grossly different' rate of reading: reading speed during SPR tends to be around half that of normal reading (Rayner, 1998). A component of this is the re-

action time needed to move the fingers, which is slower than that needed to move the eyes (Rayner, 1998). In terms of whether this is 'natural', however, Mitchell argues that nowadays SPR-style reading is not all that unusual: 'In computer displays, advancing *is* typically achieved by pressing a key, and in cell-phone text messages the display is routinely restricted to just a few words' (2004, p.23). This recalls the points made about ecological validity above (§4.3.1.1).

A slower reading speed is not necessarily problematic if it is the only difference. However, Mitchell (2004) points out SPR studies which have suggested that different segmentation (i.e. which sections of the material are presented each time the participant advances the display; words, phrases, etc.) of the same material can result in that material being processed differently.

4.3.2 Neuroimaging techniques

Neuroimaging techniques are arguably more powerful, and certainly more complex, than eye movement techniques. However, where SPR may come across as a budget form of eye tracking, eye tracking is not a budget form of neuroimaging. Eye tracking instead offers a considerably different window onto the language processor, and has clear advantages for reading studies in terms of its visual nature. With this said, cost has been mentioned several times above, and it is true to say that this was a factor in selection of eye tracking here over the considerably greater expense of neuroimaging techniques.

The most popular neuroimaging techniques in psycholinguistics are electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) (Brennan, 2016).[3] Both record neuronal activity: EEG records the polarity and amplitude of electrical impulses within the brain using electrodes on the participant's scalp, resulting in measures known as event-related brain potentials (ERPs); fMRI records changes in blood oxygenation within the brain (Brennan, 2016). Such techniques potentially offer a more transparent window onto processing than that provided by eye movement/reading speed-type measures, as they represent direct measures of brain activity (Osterhout et al., 2004). In order to investigate language processing, these techniques are often combined with the paradigm known as rapid serial visual presentation (RSVP), which involves words being presented in the centre of a screen at a fast rate, typically out of the control of participant (Ditman et al., 2007).

---

[3]Magnetoencephalography is also used to obtain event-related brain potentials, but is less common in linguistic research (Brennan, 2016).

Very broadly, the two techniques have inverse strengths and weaknesses. Where fMRI has high spatial resolution accurate to within millimetres within the three-dimensional space of the brain, EEG has low spatial resolution, and where functional magnetic resonance imaging has low temporal resolution, EEG is accurate to within milliseconds (Osterhout et al., 2004; Brennan, 2016). This is an accuracy that even 'exceed[s] that of the processes of language comprehension' (Osterhout et al., 2004, p.276).

4.3.2.1 Advantages of neuroimaging techniques for the current study

The temporal imprecision of fMRI is enough to dismiss its suitability here, but from a purely theoretical perspective several aspects of EEG are attractive for achieving the aims of the present study. Returning to the ideal qualities for measuring processing proposed by Osterhout et al. (2004), ERPs allow for continuous measurement, high temporal resolution, and differential sensitivity to linguistic levels. Furthermore, research into ERPs has been particularly productive in investigating how anomalous sentences are processed, with certain types of ERPs having been shown to reliably occur after syntactic and semantic anomalies: a central assumption being made in the experimental approach here is that written sentences that have CF but fail to mark it will be initially interpreted as being anomalous. A sizeable body of ERP research has already been conducted on information structure. Much of this is on the processing of information structure in speech rather than in writing—but see the analysis of the EEG study of Cowles (2003) given in §2.6.1.3.

4.3.2.2 Limitations of neuroimaging techniques for the current study

RSVP is unsuitable for investigating TE: it violates three of the four specific qualities outlined above: it does not allow for large amounts of text to be presented simultaneously, it does not give the reader sufficient time to re-read, and it is highly unnatural. One reason RSVP is used in some types of reading studies is because it allows the experimenter to know which word is being fixated at which time. An alternative to this might be to use eye tracking simultaneously with neuroimaging (Ditman et al., 2007); this can be difficult with EEG given that eye movements introduce artefacts into the data, although some techniques can be used which combine the eye movement and neuroimaging data to filter these out (Plöchl et al., 2012). Such a combination might eventually constitute the ideal method for TE processing research, but at this early stage of investigation into

this area, the huge amount of resulting data coupled with a lack of well-developed models of processing would risk an unfocused, inconclusive analysis. Regardless of this, such a triangulated approach was far out of scope and means of the present study.

4.3.3   Qualitative, metalinguistic methods

Relevant qualitative methods in the context of TE processing might include the use of interviews and surveys that investigate how participants think they interpret TE. As a window onto processing, the use of qualitative data may seem a dubious choice: metalinguistic reflection may be revealing in some ways, but cannot be expected to provide direct evidence of which processes are at play, or when (see Kretzschmar et al., 2013 for an illustration of this). However, given the general 'fuzziness' around TE in terms of the breadth of its usage and the lack of explicit conventions associated with it, there might be an argument for investigating people's conscious impressions of how TE works: it could be that there is considerable variation in how participants interpret TE, in which case their conscious beliefs about it might account for their pre-conscious processing behaviour.[4] Furthermore, it may turn out that TE only works on CF when it is being consciously attended to: it would be difficult to use purely experimental techniques to distinguish pre-conscious from conscious processing. The use of qualitative methods can complement an experimental approach, rather than standing in for one.

There are few instances of linguistically-focused research using qualitative methods to engage with typography. Much has been written on typography within prescriptive traditions; some of this could be argued to constitute self-reflective, metalinguistic analysis of how readers are affected by TE—yet this work does not reliably inform psycholinguistic questions. In his article about the connections between punctuation and prosody, Chafe (1988) in part uses an introspective approach (this paper was discussed previously in §3.4.1): he discusses his own intuitions (and those of others) regarding the 'inner voice', and makes various suggestions about what punctuation causes the reader to 'hear'. Chafe bolsters this with the results of two experiments: a production experiment in which he

---

[4]The issue of 'consciousness' in information processing is far from being well-agreed upon or well-understood. Use of 'pre-conscious' and 'conscious' here refers to those aspects of behaviour that participants are able to notice and discuss in a metacognitive manner: while a reader cannot usefully describe millisecond differences in their first fixation times on words, they are more likely to be able to describe returning to a problematic word or sentence.

analysed how punctuation affected participants reading written texts aloud, and a novel task in which participants had to repunctuate a text stripped of all punctuation marks.

Another study already discussed here (§3.5.2) which used a combination of qualitative and experimental data was that of McAteer (1992). Aside from her reading experiments, she solicited statements from participants regarding their opinions on the functions of italic and all-caps TE. As was discussed in Chapter 3, the results of this were somewhat vague, and less useful than the results of the experiments: there is a clear contrast between utilising metalinguistic insight from Chafe, an experienced (and prominent) linguist, and attempting to utilise metalinguistic judgements from naïve participants.

The current study does not make much use of qualitative methods, as it seeks to adopt a more traditional psycholinguistic approach. Nevertheless, informal debrief sessions were conducted after each participant was tested, and this qualitative approach was useful in establishing that participants did not tend to consciously notice aspects of the experimental manipulations.

4.4 Eye movements and processing

The ways in which eye movements reflect processing have been well-studied for over a hundred years (Rayner, 1998), with strong consensus in many areas: the physical characteristics of eye movements are understood to a level beyond that which linguists need concern themselves, and many aspects of how these movements reflect low-level linguistic phenomena such as word recognition are also quite firmly established. Connections between eye movements and higher-level phenomena are, however, less well understood, and more subject to debate.

As has been previously mentioned, psycholinguistic techniques that measure eye movements rely on the assumption of there being a link between eye movements and language processing. Early work in this area was based on two assumptions: the *immediacy* assumption held that processing of a word starts immediately when it is encountered by the eye, and the *eye–mind* assumption held that the eye remains on a word until that word has been processed (Just and Carpenter, 1980). As Kliegl et al. (2006) point out, this assumption implies that the metric known as *first-pass dwell time* (the total length of all fixations on the word before the eye leaves the word for the first time; Clifton et al., 2007) should provide the single necessary measure of processing.

Later work has revealed a more complicated picture: there may be a certain

amount of distributed processing in word identification, meaning that processing for a given word might occur when the eye is no longer looking at that word, but instead has moved to a subsequent word.[5] Furthermore, in the course of normal reading there are a relatively high number of fixations (10–15%; Rayner, 1998) that take the eye back to previously fixated words: whilst some of these *regressions* are due to errors, some are driven by a need to re-process that word (Rayner et al., 2006). This additional complexity has prompted the development and use of a wide variety of eye tracking metrics.

4.4.1   Eye tracking metrics

(Some eye tracking metrics are referred to with multiple terms, often depending on the context or field in which they are used. Where metric names are introduced here that will be referred to again, the first name given will be that used throughout the rest of this thesis.)

As stated above, the output of an eye tracker mainly consists of a series of x–y coordinates representing which pixels of a display were fixated. In conjunction with the definition of AOIs, this data can be used in order to calculate a large variety of metrics. Boland (2004) points out that there are really only two behavioural measures for each AOI: the duration of fixations on it, and the size and direction of saccades before and after these fixations. All other metrics are computed from this basic data, and thus do not constitute fully independent measures on their own. However, eye movements during reading and linguistic processing interact to produce complex patterns of the two primary behavioural measures, making the metrics computed from them useful.

The metrics used for psycholinguistic eye tracking research are often categorised according to whether they are 'early' or 'late' measures (Clifton et al., 2007; Staub and Rayner, 2007). As Clifton et al. are careful to note, these terms refer to the point in the eye movement record at which the effect captured by a metric occurs, relative to the given AOI; they do not refer to whether this effect corresponds to early or late stages of processing. However, whilst the relationship is not direct, and whilst this issue is still subject to debate, Clifton et al. concede that early measures are more likely to reflect early processing, and vice versa.

The ways in which eye movements reflect lower-level processing are better researched and more agreed-upon than the ways they reflect higher-level pro-

---

[5]Indeed, this question depends on what type of processing is being referred to—in terms of offline, non-linguistic processing, the fact that we can learn things from reading suggests that processing that started with a fixation can continue for a long time (Holmqvist et al., 2011).

cessing. Lexical processing is widely seen as having a relatively clear, direct effect on eye movements: essentially, the harder a word is to identify, the longer the eye remains on the word (Balota et al., 2006; Staub and Rayner, 2007). Word identification happens very rapidly on initial fixation (Rayner and Clifton, 2009), and factors such as word frequency and number of meanings have been shown to have reliable effects on fixation time (see Rastle, 2007 for an overview). Staub and Rayner (2007) argue that these low-level lexical variables constitute the 'engine' that drives eye movements, accounting for the bulk of behaviour; or, as Clifton et al. (2007) phrase it, 'much of the systematic variance in fixation time and location can be attributed to processes of recognizing individual words in the text'.

The connections between eye movements and syntactic processing have also been well-researched, although the relationships between the two are slightly less clear than they are for the lexical level. The bulk of work on syntactic processing has examined syntactic ambiguity. Ambiguity has been shown to prompt reanalysis, which manifests in the form of longer initial fixations on the disambiguating word, and regressions to points in the sentence from which a different analysis can be constructed (Staub and Rayner, 2007; von der Malsburg et al., 2015). Such findings have played a large part in uncovering the incremental, word-by-word nature of syntactic processing (Tanenhaus et al., 1995; Rayner and Clifton, 2009).

Lower-level processes are not of central importance to the present study, but their impact on the eye movement record means that it must attend to them. For example, if target words vary in word frequency this may affect how long they are fixated, due to differences in lexical access speed. This may then impact 'late' metrics such as total dwell time, and thus potentially obscure the effects sought from those metrics. Whilst target word frequency was not controlled in Experiment 1 (see §4.7.8), it was subsequently deemed advantageous to control it for Experiment 2—not simply to avoid lexical access speed confounding the results, but because there is an interest here in establishing whether italics have an early 'eye-catching' effect of visual salience, and this too could be affected by varying lexical access speeds.

Metrics typically used to investigate lower-level processes include the following (N.B. all metrics referred to here are measured in milliseconds, except where otherwise stated):

- **First fixation time:** the length of the first fixation on the word (Clifton et al., 2007);[6]
- **Single fixation time:** the length of a fixation on a word when only a single fixation was made to it (Clifton et al., 2007);
- **First-pass dwell time**: the total length of all fixations on the word before the eye leaves the word for the first time (Clifton et al., 2007). A *pass* can be thought of as a distinct 'visit' to a word, and consists of a sequence of contiguous fixations to it.[7]

The relationship between eye movements and the higher-level postlexical processing of semantic and pragmatic information is considerably less clear than that between eye movements and lower-level processing (Clifton et al., 2007; Staub and Rayner, 2007). There appears to be no convenient one-to-one correspondence between the basic units of these levels and quantifiable eye movements. In the face of conflicting results, there is far less consensus on how to measure the processing associated with higher levels (Clifton et al., 2007; Reichle et al., 2009).

That there is a relationship between eye movements and higher-level processing is not in dispute, however. In terms of fixation durations, Staub and Rayner (2007) suggest that, whilst low-level lexical processing drives the eye forward through a text, 'higher-level variables may primarily serve to slow down processing (and increase fixation times) when something doesn't compute well' (Staub and Rayner, 2007, pp.336–337). In terms of saccades, regressions are associated with higher-level processing: Vitu (2005) states that interword regressions may be triggered as responses to either visuomotor error, lexical processes, or to higher-level processes. However, Vitu suggests that it may be the former two lower-level factors which are predominant in prompting regressions.

There are two pre-eminent models of eye movement during reading, and their predictions about regressions and higher-level processing differ: the E-Z Reader model (Reichle et al., 1998, 1999; Rayner et al., 2004, 2007; Pollatsek et al., 2006) predicts the vast majority of regressions to be due to postlexical processing difficulty (i.e. higher-level processing), whilst the SWIFT model (Engbert et al., 2002, 2005; Kliegl and Engbert, 2003) predicts that *no* regressions are due to postlexical processing, and that the majority are due to lexical processing diffi-

---

[6]*Time* and *duration* are used interchangeably in the literature; *time* will be used here.

[7]*Run* and *pass* are also used interchangeably in the literature, making for alternate terms such as *first-run dwell time*, etc.; *pass* will be preferred here.

culties.[8]

One reason for the lack of clarity regarding the impact of higher-level processing on eye movements (which may also account for the differences in the models) may be the intrinsically high level of variability across different readers. Clifton et al. (2007) suggest that at higher levels of processing readers have more options in terms of how they handle processing difficulty: during word identification, readers can either stare at a problematic word or give up—but at the sentence level they can choose to go back to an earlier point in the text, pause at the problematic point, or continue onwards to seek further information in later text (for a detailed description of different types of eye movement pattern which are in keeping with these different categories of behaviour, see Frazier and Rayner, 1982).

This presents difficulties to studies which aim to investigate higher-level processing. This study is one of them: processing of the information structural phenomenon of CF must occur at higher levels, as it depends on discourse-level information. The typical approach to investigate higher-level processes is to employ multiple metrics: this is done under the assumption that the effects of this processing may manifest as complex patterns. Whilst an approach that requires using multiple statistical tests has a certain epistemological danger (Holmqvist et al., 2011), this is a necessary evil given the current state of knowledge in this area.

The following metrics are typically used in studies into higher-level processing:

- **First and second-pass dwell times:** see above for first-pass dwell time. Dwell times can be calculated for any number of repeated passes on a word, although analysing anything after the second pass is rare;
- **Regression path time** (also *go-past time*): the total time from the first fixation on a word until the first fixation on a subsequent word,[9] including any time spent in fixations to preceding words if there are first-pass regressions (Boland, 2004; Clifton et al., 2007). Logically this will include the time taken to make saccades (at a minimum, the time taken to make the

---

[8]Outside of the automatised reading with which these models deal, there is also the possibility that a regression to a previously fixated word could be the result of an even higher level of processing, that of conscious thought; the reader might consciously decide to look at a word again. Experimentally, there is little way to account for this other than to try to minimise participants' awareness of their own reading processes during the task.

[9]The literature on eye tracking metrics used during reading tends to assume single-line stimuli, and thus uses terms such as 'leftwards' and 'rightwards' to indicate regressive/non-regressive fixations. With multiline stimuli such as those used in the experiments here, the first word of a new line will appear leftwards of a preceding word; therefore here words will be referred to as 'previous' and 'subsequent' instead.

saccade to the subsequent word); indeed, the longer the regression path is, the higher the percentage of this metric that will reflect time spent in saccades rather than fixations;

- **Regressions out** (also *probability of a regression*): the percentage of regressive saccades away from the word, typically restricted to the first pass only (Boland, 2004);
- **Total dwell time** (also *total reading time*): a blanket measure of all fixations on a word, regardless of which pass they fall into (Boland, 2004).

Regression path time and regressions out are different to the other metrics discussed here, in that they reflect not only fixations made on the current AOI, but are rather measures relative to it, which can provide information regarding fixations on multiple other AOIs. In this way, an analysis of eye tracking data for a reading study can take an AOI-based approach, yet not be limited to looking at individual AOIs one-by-one.

It is important to explicitly note that current models of eye movements do not view saccades to other words as only occurring when the currently fixated word has been 'completely processed'; rather, processing of one word can continue whilst the next is fixated. This can result in 'spill-over' effects, whereby difficulty processing one word has an effect on fixations to the next (Reichle et al., 1999). In this way, in some situations the so-called 'early' metrics for a given word may in fact reflect late processing of the word that was fixated before it.

On a similar note regarding the linearity of eye movements and processing, some processing can occur *before* a word is fixated, if that word falls within parafoveal vision. It is crucial to consider such *preview effects* (Rayner et al., 1989; Schotter et al., 2012) when using AOIs: even if there are no direct fixations on a particular AOI at all, this does not mean that the word has not been at least partially processed.

4.5   Experiment 1 hypotheses and predictions

4.5.1   Primary hypothesis

A sentence containing a word that has CF and TE will be processed more quickly than an equivalent sentence in which that word has no TE.

To expand on this, this hypothesis held that a) if a word is presented in a sentence that is only felicitous if this word is taken as having corrective CF, and

b) this word is displayed without TE, then c) there will be a processing cost (from reanalysis) for this sentence in comparison to when the target word is presented with TE.

The predicted processing cost was expected to manifest as re-reading of the target word (i.e. that word which was either displayed with or without TE) and the words preceding it. This re-reading was expected to occur when the target word did not have TE, because that word would then be initially analysed as not having CF. This analysis would, however, be infelicitous on a discourse level, due to the way in which the stimuli texts were designed (see §4.6.2 for more details): thus, depending on the point at which contrastive focus structure is processed, a re-analysis would then have been triggered to resolve this infelicity either after the end of the sentence was reached, or after the first pass was made on the target word. For either case, re-analysis would prompt regressions to the target word and the words preceding it as the processor attempted to search for cues to the information structure of the target word.

4.5.2   Secondary hypothesis

Initial fixations will be longer on a word with TE than on one without, due to the increased visual salience.

This hypothesis anticipates an 'eye-catching' effect, whereby TE increases initial attention on a word. This increased attention would seem likely to manifest as longer initial fixations; this would not be linguistic processing of the sort that would more probably affect later fixations. Increased attention due to salience is an aspect of the *isolation effect*, which is defined by Ashcraft as 'improved memory for one piece of information that is made distinct or different from the information around it, such as underlining one word in a list in red ink' (2006, p.222).[10] Such a difference in memorability may result from increased attention on outstanding elements such as words with TE.

If there was strong support for this secondary hypothesis, but weak or no support for the primary hypothesis, this might indicate that some type of non-linguistic effect driven by visual salience had obscured what connections there might be between TE and CF.

---

[10]This is often called the *Von Restorff effect*, yet Hunt (1995) argues that von Restorff's discussion of such effects (1933) was more nuanced than modern usage of this term implies.

4.6 Experiment 1 methodology

An eye tracking experiment was conducted to test the two hypotheses outlined above. It manipulated the presence of TE on a target word within short on-screen dialogues, which were designed so that the target word had corrective CF.

4.6.1 Participants

10 participants were recruited from the undergraduate population of the University of Leeds. The mean age was 22.3 years (SD = 1.4 years). Participants were required to have completed two years of university-level study or more, in order that they had had sufficient exposure to the type of professionally printed material in which TE is more likely to be used. All participants were third-year undergraduates studying within the School of Languages, Cultures and Societies. All were native monolingual English speakers, with normal or corrected-to-normal vision and no diagnosed reading problems. Prior to the experiment participants were informed that this was a linguistic investigation into reading during which their eye movements would be recorded, but were not made aware of the experiment's specific purpose until after they had completed the procedure.

A within-subjects/between-items structure was used in which the first group of participants received half of the stimuli with the target word in italics ([+TE] condition), and half of them in the same roman style as the surrounding text ([−TE] condition); the second group received the same items but with the conditions reversed (i.e. Group A received Item 1 in the [+TE] condition, and Group B received it in the [−TE] condition).

4.6.2 Materials

40 experimental items were designed in the form of short written dialogues between two speakers (see Appendix A for a full list of all test items, fillers and check questions). In each item, two sentences were attributed to the first speaker (Speaker A), and a single-sentence reply was attributed to the second (Speaker B). The structure will now be explicated using Test item 1, shown in Figure 4.1. In this example, the discourse topic was *men Amanda met last year*, the antecedent focus word was *John*, and the target word was *Stephen*.

Discourse topic sentence (Speaker A): A discourse topic was introduced. This topic was a group or set of things (here, *men Amanda met last year*), typically in

> A: Amanda's getting married to this guy she met last year. She said his name was John.
>
> B: She said his name was *Stephen,* and I'm sure of that because I actually know him.

<div align="center">Figure 4.1: Test item 1</div>

the form of an NP, but conveyed with various different grammatical structures to create variety and reduce the likelihood of the test items being recognised as such. It was introduced with the proposition that one member of this set was somehow superior/different. The fact that this member was unspecified in this sentence was intended to establish an informational need, to be fulfilled in the following sentence.

> 'Amanda's getting married to this guy she met last year.'

Antecedent focus sentence (Speaker A): A member of the set was identified. This member took the form of one focused noun, which constituted the *antecedent focus word,* that word which would establish the alternative set for the contrastively focused target word in the following sentence.

> 'She said his name was John.'

Target sentence (Speaker B): A correction was made using a repetition of the previous sentence, with the member of the set under discussion changed to a previously unmentioned single-word item within the same set. This item thus had corrective CF. Some additional words appeared after the target to position it within the middle of the line, guarding against issues relating to oculomotor error that can occur at the end of lines (see §4.6.5.1).

> 'She said his name was *Stephen,* and I'm sure of that because I actually know him.'

This structure was designed in order to create a target sentence which could only be interpreted as felicitous if the target word was analysed as having CF, yet at the same time that target word would have as little CF-marking as possible. The combination of these two factors was intended to prompt a reanalysis after the target sentence was first read in the $[-\text{TE}]$ condition, as an initial analysis would take the target word to have focus but not CF, because it had comparatively little marking. The unacceptability of this analysis would be due to the infelicity of

Speaker B making this proposition *without* contrastive focus, as such a proposition would be irrelevant to the discourse (Grice, 1975).

The minimal marking of the target word was achieved by creating corrective CF via the repetition of the previous sentence with a single word changed; this was thus a 'direct rejection of an alternative', per Gussenhoven's definition of corrective CF (Gussenhoven, 2006; see §2.4.3). By using this explicit rejection, the target word has CF in Zimmermann's (2007) sense of it being unexpected, but without any particular device marking it: most importantly, the target word was not preceded by any devices associated with focus such as clefting or focus-sensitive particles, which could guide the processor to an initial correct analysis, regardless of the presence of TE, before the target word itself was fixated and processed for the first time.

Similarly, using word-for-word repetition alone allowed for Speaker B to contradict and correct Speaker A without using any words or constructions preceding the target word that might mark that the sentence would contain corrective CF: a reply to *She said his name was John* could be *No, she said his name was Stephen*, or *Actually, I think he's called Stephen*, or *That's wrong, Stephen is his name*, etc.

The use of the discourse topic sentence served two purposes: to add enough context that the successive items did not come across as jarringly sudden leaps into the middle of different conversations, and to delimit a semantic set from which both the antecedent focus and target words could be taken. The presence of this hypernym of the following focused words was intended to provide a level of semantic priming control for these words (Meyer and Schvaneveldt, 1971), by making it more likely that the processor would have a similar baseline level of activation for both.

4.6.2.1 Use of dialogues

The use of dialogues is not unusual in studies on focus processing. Dialogues as stimuli can take advantage of the relationship between wh-questions and focus (e.g. Bock and Mazzella, 1983; Birch and Clifton, 1995). Fewer studies have used written dialogues, although Birch and Rayner (1997) is a notable example. The choice here of dialogues instead of narrative text was intended to provide a level of control over the type of mental modelling necessary to deal with information structure: explicitly presenting two interlocutors in conversation should prompt the reader to assess focus from the speakers' perspectives. In order to understand the conversation, they must build a model of each speaker's knowledge, as well as the common ground between them. This is in contrast to the situation that occurs

in many texts (including this thesis), in which there is an implied 'interaction' between an active authorial voice and a passive, receptive audience. In such texts, rather than standing outside of the interaction, as when reading a dialogue, the reader is an implicit part of the discourse. The authorial voice assumes a mental model of the reader, and the reader mentally models the authorial voice in turn. Given that the reader is silent, certain manipulations of information structure then become harder: the dialogues adopted here allow for the inclusion of CF in the form of corrective CF, in a simpler way than if narrative texts were used.

The use of dialogues meant that the stimuli consisted almost entirely of direct speech (bar the *A* and *B* labels assigned to the two interlocutors). This has two implications that were judged to be helpful. Firstly, as discussed in §3.3, actual usage patterns of TE seems to show that it is used with a higher frequency in sections of direct speech. Therefore, presenting dialogues with TE was hoped to be somewhat ecologically valid.

Secondly, the use of dialogues allowed for the use of corrective CF, which is useful for the reasons discussed above. Corrective CF could be used in narrative texts, but a single authorial voice correcting itself across repeated texts could seem odd, given that such texts tend to be composed iteratively and edited to remove mistakes. 'Disagreement' and inconsistency in narrative texts might have made participants view it as being deliberately unreliable (like the proverbial 'unreliable narrator' in literature), perhaps increasing the likelihood they would question and second-guess the content. The use of a dialogue which represented two interlocutors disagreeing was intended to avoid this.

4.6.2.2 Check questions

In order to encourage participants to read for comprehension rather than to just skip through items, and to provide a rough way of checking how well they actually comprehended them, check questions were designed for 20% of the items (including fillers: $N = 32$). These took the form of 'true or false' statements relating to the content of the dialogue. The propositions in the check questions were designed to avoid reference to the target or antecedent focus words, in order to reduce the chance that participants might notice the experimental manipulation. Half of the statements were true.

4.6.2.3 Fillers / distractors

40 experimental items were designed in the form of short written dialogues be-tween two speakers.

120 filler items were designed using the same two-speaker dialogue format. These comprised three types of 40 items each. The first two types were designed specifically as distractors. In the test items the following two factors co-occurred, leading to a possible risk of participants becoming aware of the experimental manipulation:

1. The set/member structure established in Sentences 1 and 2;
2. The presence of a typographically emphasised word (in 50% of the test items seen by a given participant).

In order to mitigate this, two types of filler item were designed: the first type (Figure 4.2) had the same type of set/member structure as the test items, with the first sentence introducing a set and the second naming a member of the set. In these fillers the final sentence differed from those of the test items, with Speaker B exactly repeating Sentence 2, followed by some additional words expanding on this: the intended effect was for Speaker B to appear to be emphatically agreeing with Speaker A, rather than correcting them using contrastive focus. This is what Krifka (2007) calls *confirmatory focus* (see §2.4.3).

A: Rachel has gone to cast her vote in the election. She'll definitely vote for the Greens.
B: She'll definitely vote for the Greens, like she always does.

Figure 4.2: Confirmatory focus distractor (Filler item 5)

The second type of fillers had a wider-ranging structure that did not follow the set/member formula, but instead contained a typographically emphasised word in the final sentence (Figure 4.3). Unlike the CF words in the test items, this word was always an intensifying adverb, and bold TE was used instead of italic TE. The intended effect was an increased degree of intensity: this recalls the 'modulatory' emphasis referred to by McAteer (1992) and Sanford et al. (2006; see §3.5.2 and footnote 13, p.81).

A third type of filler item followed no particular pattern other than the di-alogue framework (Figure 4.4). These items were intended to provide enough variety to reduce boredom, and thus maintain motivation.

A: I'll be bringing the car when we go to the cinema. Are you sure you don't need a lift?

B: I'm **absolutely** sure I don't need one, I'm working next door to the cinema all day anyway.

Figure 4.3: Bold TE distractor (Filler item 42)

A: The results of the assessment were better than we expected. All departments have improved, as far as the inspectors are concerned.

B: We have done exceptionally well, considering the challenges we've been facing.

Figure 4.4: General filler (Filler item 61)

4.6.2.4 Item order

Each participant received the test and filler items in a different pseudorandomized order: sequences of four items were constructed, each containing three fillers and one test item, and the order of these sequences was randomized. Items followed by check questions were similarly pseudorandomized: in both cases, the aim of this was to reduce predictability and to prevent the consecutive appearance of test items or check questions. With this structure, there were a maximum of 2 consecutive test items (or consecutive items followed by check questions).

4.6.2.5 Semantic controls

The words used for the target words and antecedent focus words appeared only once, within their particular test items, across the entire experiment. This was done to avoid repetition priming effects (Sloman et al., 1988).

4.6.2.6 Visual/typographic characteristics

Items were typeset in a manner broadly consistent with the standard typesetting of English-language books, in an attempt to preserve a level of ecological validity, and to establish a convincing typographic setting in which TE would seem in keeping with the apparent genre of the stimuli text.

All text, including that used in the instructions, was presented in 30pt Times New Roman as black text on a white background. The text started at the top of the screen, with each sentence following the last on the same line: lines were broken at a margin on the right. The character count of a full line was roughly 66

A: Amanda's getting married to this guy she met last year. She said his

name was John.

B: She said his name was *Stephen*, and I'm sure of that because I actually

know him.
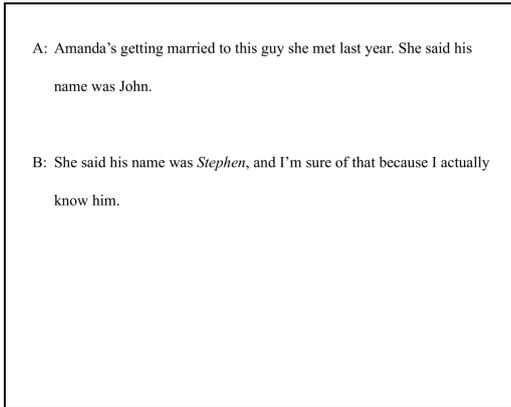
Figure 4.5: Screenshot of Test item 1

True or false: Speaker B knows Amanda's fiancé.

<True>   <False>
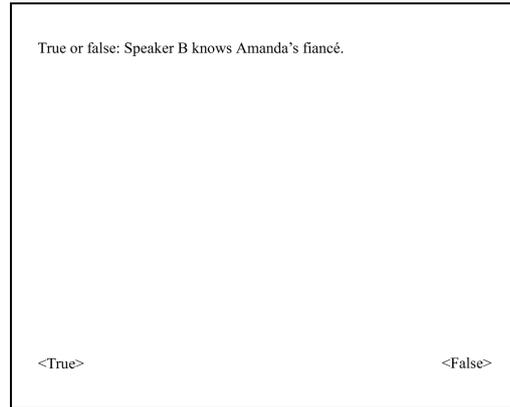
Figure 4.6: Screenshot of Check question 1

characters.[11] The text was not justified to both left and right margins (a feature of the typesetting of many, but not all books), but was rather left-aligned, which allowed for inter-word and sentence spacing to remain consistent without the need for hyphenation (which could complicate analysis). A second departure from typical professional book typesetting was the use of double spacing, in order for the analysis to more clearly distinguish which line of text was being fixated.

Each item used the strings 'A:' and 'B:' to indicate the different speakers, and contained a blank line separating the end of Speaker A's sentences and the start of Speaker B's sentence. The stimuli were aligned so that Speaker B's sentence, including (where applicable) the target words, appeared close to the vertical centre of the monitor, where initial testing suggested eye tracking recordings were slightly more accurate. Figure 4.5 provides a screenshot of Test item 1.

Check questions began with the text 'True or false:' followed by the statement itself. The strings '<True>' and '<False>' were also presented at the bottom left and right of the screen along with the check questions, representing the locations of the buttons on the response box, as an attempt to lessen a tendency that was noticed in pre-pilot testing, whereby participants would look away from the screen at the response box to select the appropriate button. Figure 4.6 provides a screenshot of Check question 1 (the question that followed Test item 1).

---

[11]This is the number of characters that sources such as Bringhurst (2002) recommend—whilst there may be little evidence that this prescription has any particular beneficiary effect on reading, this viewpoint is widely adopted in printing and thus should serve to make the stimuli fit slightly better in the 'book-like' genre aimed for here.

Table 4.1: Stimuli word count statistics

		Words			
Category	Item count	Count	% of total	Mean	*SD*
Practice items	10	356	6%	36	4
Test items	40	1333	21%	33	3
Filler items	120	4352	70%	36	5
Check questions	34	213	3%	6	2
All	204	6254	100%	31	12

4.6.2.7 Stimuli statistics

Each participant saw 170 items, consisting of 10 practice items, 40 test items and 120 fillers. They also saw 34 check questions. Table 4.1 gives statistics for words in the stimuli: in total participants were presented with 6,254 words over the course of the experiment (excepting text other than that in the stimuli, such as the instructions, break messages etc., and the *True or false:* prepended to each check question). The mean word count of the stimuli was 36 ($SD = 5$), and the mean word count of the check questions was 6 ($SD = 2$).

4.6.2.8 Apparatus

Eye movements were recorded with an SR Research EyeLink 1000 Plus eye tracker system, sampling at a rate of 1,000 Hz. A chin rest was used to restrict head movements. This was adjusted where necessary to allow for the participant to comfortably view the entire screen. Each participant's dominant eye was determined using a Miles test prior to the recording (Miles, 1930), and the movements of this eye alone were recorded, although both eyes were used to view the stimuli.

The stimuli were presented on a 17-inch Dell 1708FPf LCD monitor in its native resolution of 1,280 × 1,024 pixels. This was connected to a computer recording the eye movement data, and a second computer running the experiment display program (a Python program developed in SR Research Experiment Builder). The second computer also recorded participant input, which was captured with a Cedrus RB730 response pad. The pad was set up with a button on the left labelled '<True>', one in the centre labelled '<Next>', and one on the right labelled '<False>'.

4.6.3 Procedure

After having read an information sheet and having signed a consent form, participants were helped into a comfortable position with their head in the chin rest, and then presented with written instructions on the monitor. These explained to them that the task was to read dialogues between two speakers on the screen, and to press the '<Next>' button on the response pad when they understood the dialogue. Then the check questions were explained, and participants were told to click the relevant button on the response pad when they were displayed.

After the participant had read the instructions, the eye tracker was calibrated.[12] This calibration procedure was run again after breaks, and for some participants between trials where necessary; the investigator was present in the room throughout the experiment, sat at a separate desk and facing away at a 90 degree angle from the participant. Throughout, the investigator monitored the accuracy of the recording by watching real-time eye movement data on the display of a control computer connected to the eye tracker.

A ten-item practice block was then presented. The practice items included two items adhering to the same format as the test items, one with the target word italicised and one without; the other eight items were 'fillers', one of which included bold text, and two of the items were followed by check questions. In this way the 'test item' to 'filler' ratio and the check question ratio were the same as those in the main experiment. These practice items were presented in the same order to all participants.

On finishing the practice block, text on-screen re-summarised the task and explained that the main experiment would contain three equally-spaced breaks. The main experiment followed, with four blocks consisting of 40 stimuli. The first three blocks were followed by an automatic message instructing the participant to take a break; they were encouraged to move out of the chin rest. After one minute the on-screen text changed to inform the participant they could continue when they wished by pressing the '<Next>' button.

Each trial began with a small, circular fixation point being displayed in the top left of the screen at the point where the centre of the first character of the stimulus text would appear. Eye position was monitored by the investigator and, when the participant was fixating the point, the investigator accepted this and triggered

---

[12]Additional instructions were given verbally to the participant concerning fixation points and calibration—this was in order to separate out the mechanics of the task from the method of measurement, and to avoid overwhelming the participant with front-loaded information in the form of initial instruction screens.

the display of the stimulus (if recalibration was required, it was conducted at this point). The stimulus was then displayed until the participant pressed the '<Next>' button. At this point, when items were followed by a check question, that question would appear, remaining on screen until the participant pressed '<True>' or '<False>'. The next trial would then begin immediately.

Eye movements were only recorded during the display of stimulus text or check questions. The sections of recordings during which check questions were displayed were not included in the analysis; nor were the practice and filler items.

4.6.4 Selection of areas of interest

The analysis here was AOI-based, meaning that all the metrics were calculated relative to AOIs. Figure 4.7 shows the AOIs for Test item 1.

A: Amanda's getting married to this guy she met last year. She said his name was John*.

B: She said his name was *Stephen*[†], and[‡] I'm sure of that because I actually know him.

<div align="center">

Figure 4.7: AOIs in Test item 1
*antecedent focus word, [†]target word, [‡]post-target word

</div>

The principal AOI in the analysis was the target word, which appeared in the third sentence of the stimuli, and was presented in both italic font style ([+TE] condition) and roman font style ([−TE] condition).

A second AOI was considered for inclusion in the analysis, as this would be useful in capturing re-reading; the best candidate was that word which has been termed the 'antecedent focus word' here, which appeared in the second sentence of each test stimulus. The expectation here was that, in the [−TE] condition, the need for disambiguation would drive the participant to return to earlier sections of the text. The antecedent focus should have been the most relevant word to return to, given that it made explicit the alternative set for the CF of the target word, and was thus necessary for understanding why the target word was mentioned at all.

Finally, a third AOI was included in the analysis, the word immediately following the target word. This was intended to capture potential spill-over processing.

4.6.5 Selection of metrics

It is imperative to select metrics a priori, given the temptations afforded by the large quantity available: Holmqvist et al. (2011) warn against adopting a 'fish-

ing trip' approach of searching for statistical significance across too many metrics, in particular because the number of significant results occurring by chance (i.e. Type I errors) will increase in proportion to how many individual tests are conducted. With that said, the fact that most eye tracking metrics are correlated to an extent (Boland, 2004) should mean that the increase in the probability of false positives due to multiple tests should be lower than it might otherwise intuitively seem to be.

The analysis used here rested primarily on 'later' metrics. Such metrics were judged better able to reflect the higher levels of processing that should be associated with focus; in particular, re-reading was expected in this experiment. This was because the experimental design aimed at something somewhat similar to a garden path effect (Frazier and Rayner, 1982; Gibson and Pearlmutter, 2000), but with ambiguous information structure instead of syntax: reanalysis was expected, and it was assumed this would prompt re-reading.

4.6.5.1 Capturing re-reading (testing the primary hypothesis)

Many eye tracking studies into processing during reading predict that online effects will occur in a linear fashion, with the effects of the manipulation of a target word manifesting relatively rapidly in fixations to that word and to subsequent words. As Hyönä et al. (2002) note, this is partly due to such experiments being focused on low-level phenomena, with designs that seek to minimise higher-level processing (typically, less text is presented, and the stimuli are single sentences, phrases, or words). This study, however, anticipates that the absence of TE on the target word will have a complex effect that will include re-reading, and instead seeks to provide enough material that higher-level processing is *encouraged* (it does also anticipate some lower-level effects that will manifest in a more linear way, such as comparatively longer fixations on the target word and subsequent words—i.e. spill-over processing).

'Re-reading' as it is used here should be distinguished from certain regressive movements of the eye: it is intended to refer to the reader returning to sections of the text that have been previously fixated *and processed*. Rayner et al. (2006) suggest that the majority of regressions are very brief returns to the previous word, driven by oculomotor error or lexical access issues. Such oculomotor error can often be seen when the eye makes a large movement to a subsequent line of text and 'lands' in the wrong place (Rayner, 1998). This is not considered re-reading here. Similarly, returning to text that was fixated but that wasn't actually processed, due to mind wandering (see §4.3) is not of interest here. To an extent,

stimulus design can ward against the dependent variables being thrown off by oculomotor errors, by the placement of target words away from the start of lines: this was done here. The issue of 'unprocessed' words during mind wandering seems to be extremely difficult to avoid, however; indeed, it is doubtful whether existing eye tracking techniques are able to fully distinguish between the types of re-reading that are of interest here, and those types which are not.

The exact pattern of the re-reading that was expected to occur here could not be explicitly predicted, both because existing theories are not sophisticated enough to base predictions on, but also because the high-level processing that was assumed to be involved was also likely to result in high inter-participant variation. Indeed, von der Malsburg et al. (2015), strike a cautionary note after analysing re-reading behaviour in an eye tracking dataset (which had been collected previously by Meseguer et al., 2002): 'readers differ tremendously in how they orchestrate the various fixation strategies'. The possibility of different participants responding differently to the experiments here was considered a compelling argument for the use of linear mixed effects modelling in the analysis: random effects and slopes could be included for participants, which could account for them responding to the experimental manipulation at differing baseline levels and to differing degrees. Whilst Experiment 1 did not use this on account of its low N, linear mixed effects modelling was used to good effect in Experiment 2.

Dwell time measures that reflect passes after the first one were possible candidates for capturing re-reading, from second-pass dwell time, through third-pass dwell time etc., up to total dwell time. **Second-pass dwell time**[13] is frequently used to study higher-level effects (Clifton et al., 2007), but times for subsequent passes are rarely (if ever) used. Second-pass dwell time was judged capable of capturing some re-reading; although second passes may also reflect lower-level processing, it was judged that any significant differences between conditions found for this metric in the experiments here would only reflect higher-level processing differences caused by the experimental manipulation.

Total dwell time was also selected as a useful metric, as it was expected to be able to capture multiple re-readings: if participants were prompted to re-read the target word (and the other AOIs to be analysed, the antecedent focus word and the post-target word; see §4.6.4) more than once, this would not be captured in second-pass dwell time, but would affect the total dwell time. However, there are risks related to this metric: it is somewhat of a blunt instrument, as it lumps together all fixations on a word without distinguishing how many separate passes

---

[13]Metrics selected for this experiment are typeset in **bold** style.

were made on it. Moreover, these passes could be separated by passes on a single adjacent word, multiple adjacent words, distant words, etc., and the metric would not reveal any of this. Due to these concerns, the **pass count** was selected to be analysed in conjunction with total dwell time; it was hoped that this would go some way to solving this problem.

Regression path time was considered for use here, but ultimately was not employed. If this metric were analysed relative to the target word, it would only capture re-reading that occurred immediately after the first fixation on the target word; if the participant continued to the subsequent word, or finished the phrase/sentence and *then* returned to earlier parts, this metric would miss this. A similar problem applies to the use of the regressions out metric on the target word.

A potential solution to this would have been to instead analyse metrics such as regression path time or regressions out relative to an AOI larger than the target word—perhaps starting at the target word and containing all remaining words. There was, however, some variation across the stimuli in the number of words after the target word (due to a deliberate attempt to keep test items relatively varied): in Experiment 1, the number of words ranged from 4 to 14, mean = 8.4, $SD = 2.7$. An additional complication was the fact that the post-target words in some items appeared on a subsequent line of text. A separate issue is that, if AOIs larger than words were used, there would be an attendant loss of granularity: it would not be possible to know whether one or all of the words within the larger AOI had been fixated (Radach and Kennedy, 2004).

Scanpath analysis, whereby AOIs can be dispensed with, and instead the path the eyes take across a stimulus is examined, might be considered a valid method to interrogate re-reading behaviour. However, moving into the analysis of the complex geometric entities constituted by scanpaths seems a risky proposition, principally because there is very little precedent for this type of analysis in the current literature: contrasting early eye tracking research into reading with the current state of the field, von der Malsburg et al. write 'it is striking how little attention scanpaths received in later research' (2015, p.1676). von der Malsburg and Vasishth (2011) do present a technique for comparison of scanpaths in order to compare re-reading behaviour, but this is a relatively new method, which was eschewed here in favour of the more established AOI approach.

Another approach would be to look at the most all-encompassing metric of all, trial dwell time, which represents the combined total dwell times for every single AOI in the trial (although only three AOIs were included in the analysis

here, every other word in every item was also defined as an AOI for use with such metrics as this). This is a comparatively 'brute force' approach, which would be particularly vulnerable to inter-item variation; the items used here varied considerably across variables such as word number, word frequencies, etc., and so trial dwell time was not used.

4.6.5.2 Capturing eye-catching effects (testing the secondary hypothesis)

The 'late' metrics required to test the primary hypothesis have now been discussed and selected, but the secondary hypothesis required an 'early' metric in order to show whether there is an initial 'eye-catching' effect of the visual salience of TE. **First-pass dwell time** on target words was judged suitable for this, as it would capture everything which was *not* re-reading (as *re-reading* has been defined here).

This metric could also be used to validate the experimental design to some degree: if all was well, there should have been no effect of TE on first-pass dwell time on any words preceding the target. This is because the manipulation should not have been encountered at that point, as the participant should not have read the target word before first fixating the words before it. Effects on first-pass dwell time on preceding words would thus indicate that something had occurred; perhaps an unusual reading pattern, or perhaps even that the TE on [+TE] words was registered parafoveally before the word itself (or even its line of text) was fixated; if this were so, it could have an impact on the processing of the entire dialogue. Effects being prompted in this way by an as-yet-unfixated [+TE] target word seem extremely unlikely, however.

The four metrics that were used in the analysis of Experiment 1 were **first-pass dwell time**, **second-pass dwell time**, **total dwell time**, and **pass count**. Although Experiment 2 tested revised hypotheses, the same essential approach was taken: the same metrics were employed, and the above justification of their selection for Experiment 1 also stands for Experiment 2.

4.7   Experiment 1 results

Four metrics from the eye tracking data were analysed: first-pass dwell time, second-pass dwell time, total dwell time, and pass count. These metrics were

calculated for three AOIs per trial: the target word, the post-target word, i.e. that word immediately following the target (to investigate potential spill-over effects), and the antecedent focus word, the noun with which the CF on the target word contrasted.

The metrics were initially calculated using the SR Research Data Viewer software, before being processed by Python and R scripts written by the investigator. All the statistical analysis discussed below was conducted using R (R Core Team, 2017). For programming and data manipulation, the packages *dplyr* (Wickham et al., 2017), *purrr* (Henry and Wickham, 2017), *reshape* (Wickham, 2007), *tibble* (Müller and Wickham, 2017) and *tidyr* (Wickham and Henry, 2017) were used. For linear mixed effects modelling, *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2016) and *broom* (Robinson, 2017) were used.

4.7.1   Data preparation

Eye movement recordings were first checked on an item-by-item basis for all participants. Manual corrections were performed where it appeared that fixations were vertically offset, which happened in several places; it seems that there were some accuracy issues, likely caused by inexperience in use of the equipment. Corrections were performed in the following manner, to minimise any possible effects of experimenter bias: when several fixations within a trial were clearly offset relative to the AOIs (in almost all cases they appeared lower than the words), *all* fixations in that trial were moved by the same amount of pixels vertically to correct this. No fixations were moved horizontally, even when their position on this axis was clearly off, as this a) would risk bias (fixations could be moved into or out of one of the main AOIs being analysed) and b) would have been harder to do systematically (whilst it was clear when fixations were systematically offset from the generously separated lines of text, as they fell between the lines, the words were much more closely spaced horizontally).

Consecutive fixations shorter than 100 ms and located within 1 degree of a longer fixation were merged with that longer fixation, and then remaining fixations under 100 ms were removed. This reflects standard practice, prompted by consensus in the literature that sub-100 ms fixations do not reflect linguistic processing (McConkie, 1983; Inhoff and Radach, 1998). In some eye tracking studies fixations longer than around 800 ms are removed as they can reflect distraction and fatigue (e.g. Binder and Rayner, 1998; van Gompel and Majid, 2004), but this was not performed here. The hypothesis predicted processing difficulty and re-reading, but the exact pattern of this was unspecified, and extremely long fixa-

tions might still reflect some type of processing. Furthermore, fatigue effects were controlled for with pseudorandomized trial orders on a participant-by-participant basis.

Trials for which there were no fixations on the target word were removed from the analysis, as in this case it could not be confidently assumed that the participant had been subject to the experimental manipulation. 83 of the 400 trials (21%) were removed in this way. This was an unexpectedly large quantity, which may have been related to the accuracy issues mentioned above. A generalised linear mixed effects model was constructed to investigate whether the probability of a target word being skipped was affected by TE or by item presentation order (i.e. whether there was an effect of fatigue): the outcome was whether the target word was skipped or not. TE and item presentation order were included as fixed effects, and subject and item as random effects. The model was built using the binomial distribution (with a logit link function). It did not show significant effects of either TE or item presentation order. This suggests that the probability of a target word being skipped was affected neither by whether it was italicised, nor by effects of fatigue as the experiment went on.

4.7.2   Data distribution

Shapiro–Wilk tests indicated that none of the data were normally distributed ($p < 0.05$ for all metrics and AOI word types). Histograms suggested that the data had an ex-Gaussian distribution for first-pass dwell time and total dwell time on all word types (Figure 4.8). The ex-Gaussian distribution, which is characterised by right-skewed data with a short left tail and a long right one, is common for reaction-time data such as fixation data (Staub et al., 2010). The zero-inflated distribution of second pass dwell times shown in Figure 4.8 is notable, and was caused by the rarity of second passes being made on the words. Given that the distributions were non-normal, the principal measure of central tendency used here will be the median for all metrics except for pass count data, which is better represented by the mean, due to its extremely right skewed distribution and very low range. This can be seen in Table 4.2, which shows descriptive statistics for each AOI.

Second passes were extremely rare, as is suggested by Table 4.2: indeed, only 16% of target words and 17% of antecedent focus words across all items received a second pass. This is problematic for the primary hypothesis.

Table 4.2: Descriptive statistics for all metrics, all areas of interest by TE

	First-pass dwell time		Second-pass dwell time		Total dwell time		Pass count	
	[−TE]	[+TE]	[−TE]	[+TE]	[−TE]	[+TE]	[−TE]	[+TE]
Target words								
N	150	167	150	167	150	167	150	167
Minimum	105	106	0	0	111	106	1	1
Lower quartile	156	151	0	0	160	162	1	1
Median	184	185	0	0	198	210	1	1
Upper quartile	241	250	0	0	343	324	1	1
Maximum	910	830	1164	610	3341	1523	4	6
IQR	85	99	0	0	183	162	0	0
Mean	222	216	67	52	316	290	1.25	1.3
SD	117	107	172	121	339	218	0.579	0.707
Post-target words								
N	150	167	150	167	150	167	150	167
Minimum	0	0	0	0	0	0	0	0
Lower quartile	0	0	0	0	0	0	0	0
Median	0	0	0	0	0	0	0	0
Upper quartile	160	188	0	0	160	196	1	1
Maximum	514	592	214	512	578	818	3	3
IQR	160	188	0	0	160	196	1	1
Mean	80	90	5	18	86	113	0.413	0.503
SD	114	119	28	73	127	166	0.57	0.684
Antecedent focus words								
N	150	157	150	157	150	157	150	157
Minimum	0	0	0	0	0	0	0	0
Lower quartile	0	0	0	0	0	0	0	0
Median	152	178	0	0	160	188	1	1
Upper quartile	221	228	0	0	271	294	1	1
Maximum	878	463	792	756	1520	1109	6	4
IQR	221	228	0	0	271	294	1	1
Mean	165	161	32	41	208	220	0.92	0.968
SD	145	121	101	109	219	215	0.782	0.812

a) Target words

b) Post-target words

c) Antecedent focus words
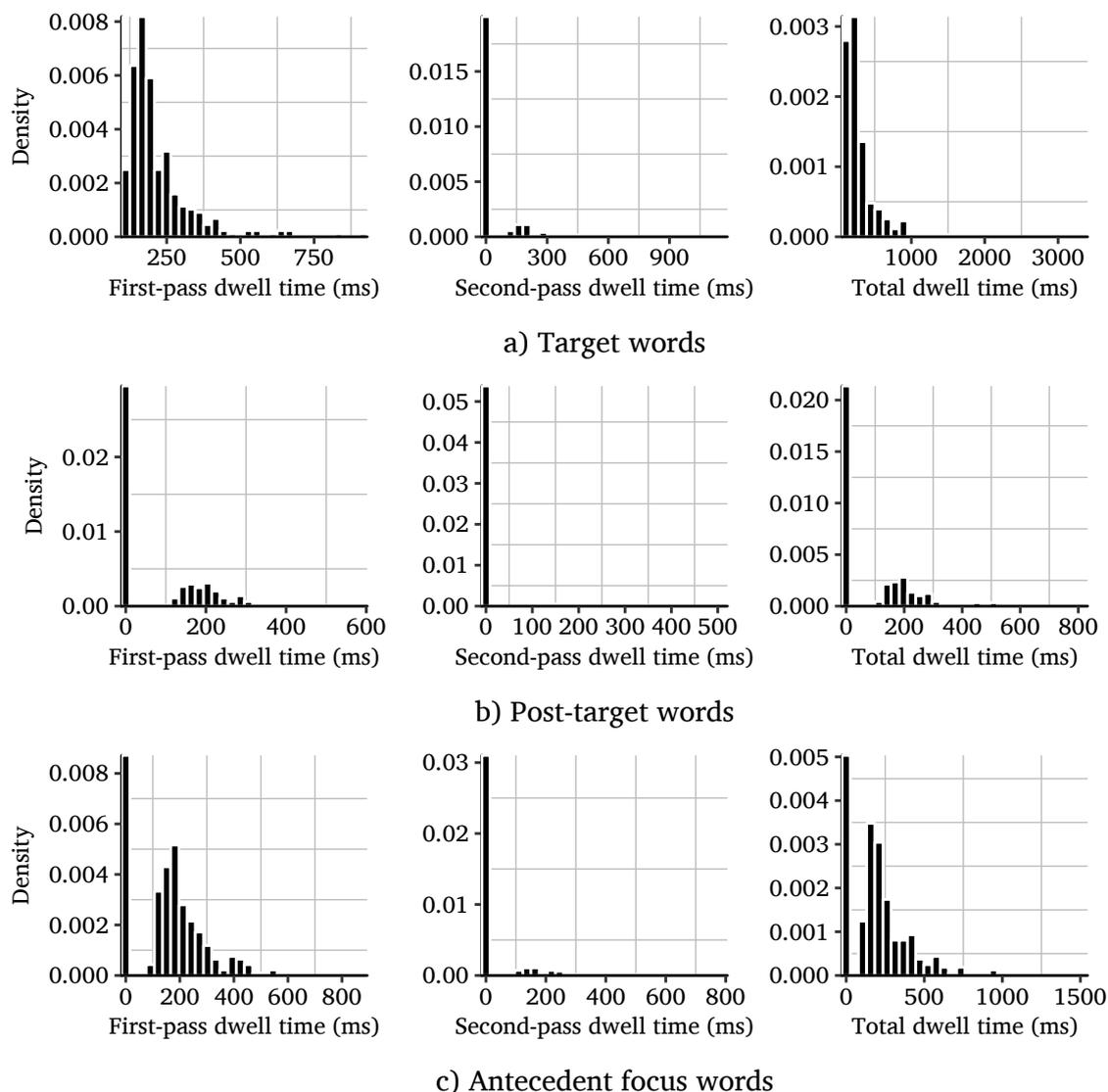
Figure 4.8: Densities for dwell times on AOIs

4.7.3 Analytical approach

The experiment was designed with the intention of using linear mixed effects models (LMMs) for analysis. Initial attempts to construct LMMs for this dataset quickly suggested that the statistical power afforded by the small sample was too low to obtain significant results (or even results 'approaching' significance).[14] For this dataset, Wilcoxon signed-rank tests were conducted instead; the use of these

---

[14]LMMs were subsequently used to analyse the full experiment that followed this, Experiment 2; these are reported in Chapter 5. The results of Experiment 1 suggested that the use of LMMs would be advantageous in Experiment 2, as they showed considerable variability for participants (§4.7.7) and items (§4.7.8): LMMs enable such variability to be accounted for by including participant and item as random effects.

Table 4.3: Wilcoxon signed-rank tests on typographic emphasis

Metric	Z	p	r	Change
Target				
First-pass dwell	-0.19	0.846	0.044	2
Second-pass dwell	0.00	1	0	0
Total dwell	-1.40	0.16	0.314	12
Pass count	-0.26	0.798	0.057	0.0527
Post-target words				
First-pass dwell	-1.34	0.181	0.299	0
Second-pass dwell	—	—	—	0
Total dwell	-1.34	0.181	0.299	0
Pass count	-2.21	0.0273*	0.493	0.0897
Antecedent				
First-pass dwell	-1.09	0.275	0.244	26
Second-pass dwell	0.00	1	0	0
Total dwell	-2.40	0.0165*	0.536	28
Pass count	-0.19	0.846	0.044	0.0482

non-parametric tests instead of their parametric equivalents was prompted by the non-normal distribution of the data. Table 4.3 shows results for all the Wilcoxon signed-rank tests conducted, with the *change* column showing the difference between the medians for [−TE] and [+TE] (or the difference between the means in the case of pass count).

4.7.4 Testing the primary hypothesis

If re-reading occurred as predicted in the [−TE] condition, the second-pass and total dwell times and the pass counts of the target and antecedent focus words would be higher than those in the [+TE] condition. This would result in the difference between the medians being negative, but as Table 4.3 shows, this was not the case: there was either a positive change or no change for all metrics, on all AOIs.

There were only two significant effects found, both in the opposite direction to that predicted. The first of these was an effect on total dwell time on antecedent focus words ($Z = -2.4$, $p = 0.0165$, $r = 0.536$). This result showed that the median total dwell time for the [+TE] condition was longer by 28ms. The second effect was found on post-target words, and will be discussed below (§4.7.6).

4.7.5 Testing the secondary hypothesis

If there was an effect related to the visual salience of the [+TE] target word, this was expected to be reflected in longer first-pass dwells to that word. This was not the case, as there was no significant difference here (furthermore, the non-significant difference between medians for the two conditions was a miniscule 1 ms).

4.7.6 Spill-over effects

A Wilcoxon test showed an effect of TE on post-target words, where it was associated with an increased pass count ($Z = -2.21$, $p = 0.0273$, $r = 0.493$). An increased pass count on post-target words is unlikely to reflect spill-over processing from the target word; this would be expected in the form of a longer first-pass dwell time. Wilcoxon tests for all other metrics on post-target words showed no significant differences.

4.7.7 Analysis of participants

The mean accuracy for comprehension check questions was 94% ($SD = 3\%$): all but one participant answered every check question for the test items correctly. All participants thus appeared to have been reading for comprehension throughout, meaning that no data needed to be discarded.

In informal debriefing sessions after each run, the participants were asked after the experiment whether they had become aware of the experimental manipulation. Some mentioned having noticed TE: they tended to notice both the italic TE of the test items, and the bold TE included as a distractor in some fillers. However, none articulated an awareness of its relation to contrast in the stimuli. The results were thus judged not to have been influenced by conscious awareness of the manipulation.

Figure 4.9 show that a variety of reading behaviours were exhibited across participants. Looking at fixations on target words as illustrated by Figure 4.9.a, Participant 7 clearly made longer second passes, and engaged in more re-reading than any other participant. Only three participants returned more than once to target words; Participants 9 and 10 did not make second passes at all.

Fixations on antecedent focus words were also variable. As Figure 4.9.b shows, all participants made second passes on this AOI, with the exception of Participant 4. Again, Participant 7 was unique in having longer reading and re-reading
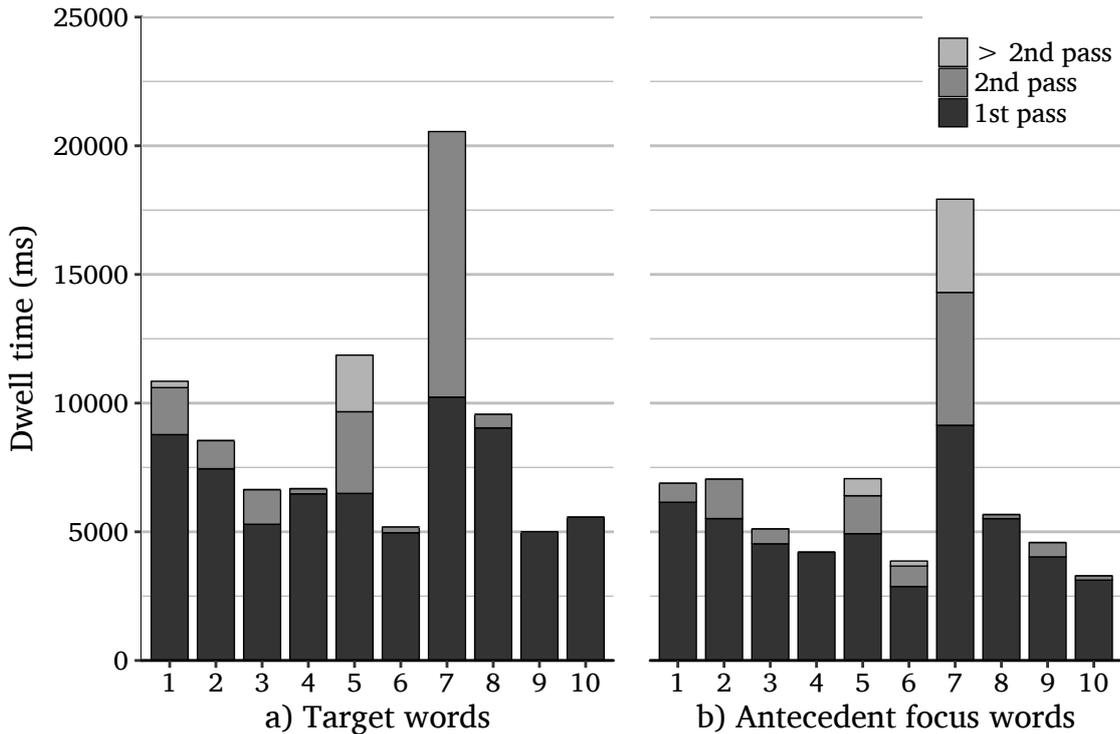
Figure 4.9: Summed n-pass dwell times by participant

times. As with target words, it was rare for anything more than a second pass to be made on the AOI, although Participants 9 and 10 did make second passes on these words, unlike on the target words.

4.7.8 Analysis of items

Variability across items was high, as can be seen in Figure 4.10. Items 21 and 28 are notable for being fixated for far less time; Item 38 was fixated and revisited for much longer than any other (the target word in this item was the low-frequency proper noun *Breitling*).

Whilst sentence and dialogue structure were carefully controlled, the individual target words used were relatively variable. Word frequency has been mentioned here (§4.4.1) as having an effect on fixation durations; other properties known to affect this include character count (Barton et al., 2014) and syllable count (Yap and Balota, 2009).

To check whether this variability may have confounded the results, Pearson correlation coefficients were calculated between the metrics and character count, syllable count and word frequency (measured on the Zipf scale; see van Heuven et al., 2014) of the target words: the results are shown in Table 4.4. No significant correlations were found ($p > 0.05$ for all tests).
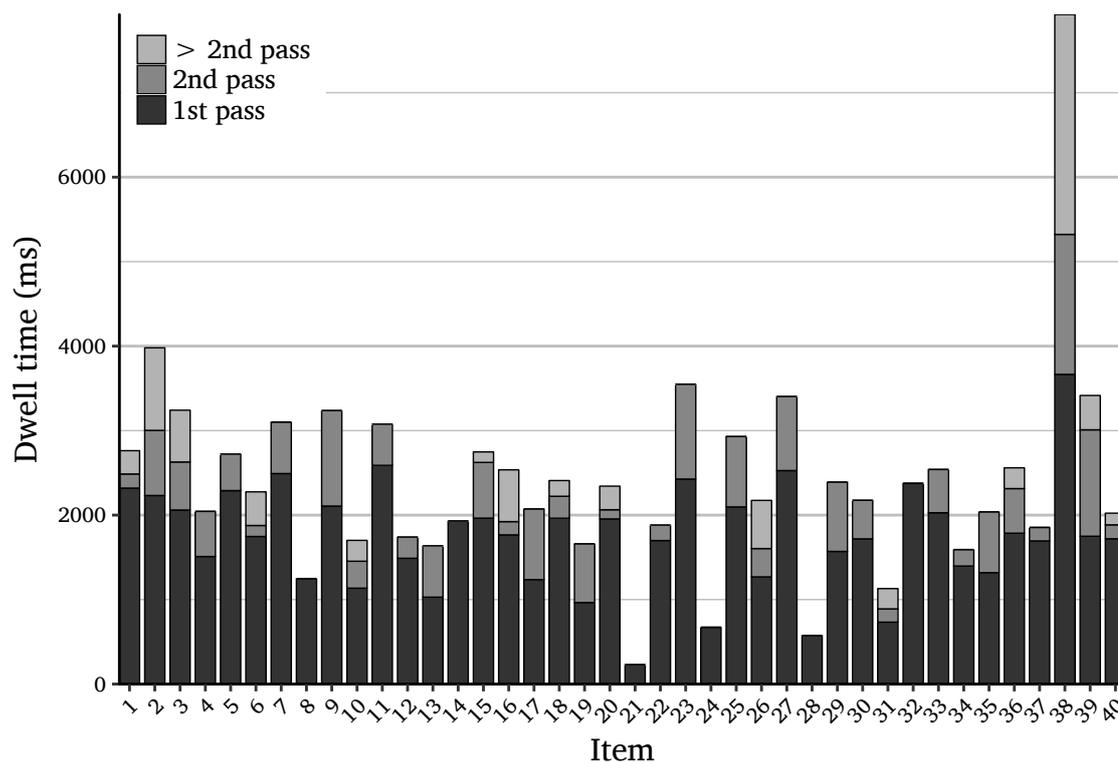
Figure 4.10: Summed n-pass dwell times on target words

4.8 Experiment 1 discussion

These results do not provide strong support for either the primary or secondary hypotheses. This is unsurprising given the low statistical power attained by a small *N*: the lack of significance here was not taken as discouraging. Indeed, given the low power, it was promising to find significant results for two of the twelve tests conducted (four metrics × three AOIs). With that said, neither of these effects was in the predicted direction.

The effect found on antecedent focus words, which were looked at for longer in total when the target word had TE (by 28 ms), suggests that the experimental manipulation was doing *something*. Although total dwell time is, as has been suggested above, something of a 'blunt instrument' which does not provide temporal information, it seems safe to assume that this effect on the antecedent focus word happened after the target word was fixated, and was thus related to the manipulation rather than a coincidence. This is supported by the lack of a significant effect on first-pass dwell time for the antecedent focus word. This word received longer fixations, rather than the shorter ones that had been predicted: the prediction was that less re-reading would occur in the [+ TE] condition, and this result may suggest the opposite. However, in the absence of significant results for the

Table 4.4: Pearson correlation coefficients between target word properties and dwell times.

	First-pass dwell time	Second-pass dwell time	Total dwell time
Word frequency (Zipf)	-0.16	0.06	-0.01
Character count	0.23	0.06	0.19
Syllable count	0.13	-0.1	0.06

$p > .05$ for all coefficients.

other metrics intended to capture re-reading, this result alone is inconclusive.

The other effect was found on post-target words, which were visited a very small number of times more (0.0897) when the target word before them had TE. Again, this was unexpected; as was mentioned above, this is unlikely to reflect spill-over processing from the target word. It may be that this result actually reflects a higher pass count on the target word itself: repeated passes on this the post-target word may have actually been to the adjacent target word, but the recording may have been inaccurate. It is especially surprising to see any significant 'later' effects on this the post-target word, as the majority of these words were grammatical/functional words, which typically receive fewer and shorter fixations during reading (Rayner, 1998); this might also point to issues with accuracy.[15]

Concern over the eye tracking accuracy that had been achieved during this pilot was one factor that led to it not being run as a full experiment. There were also concerns over the stimuli: as has been shown, there was considerable inter-item variability, and target words had not been strictly controlled. To do so would require rewriting stimuli more or less completely—which would also afford the opportunity of making the stimuli less repetitive and mechanistic-seeming, by altering the manipulation by which corrective CF was created.

With that said, this experiment was a successful test of the approach adopted here: by finding significant results, it demonstrated that the use of eye tracking to investigate how TE is processed is a viable approach. The essential elements of the experimental design, including the choice of italic TE and the single text-per item format, appeared to be working well. A full version of this pilot would have probably been worth conducting despite the issues mentioned above, but

---

[15]Of the post-target words, 50% were prepositions, 12.5% adverbs, 12.5% auxiliary verbs, 12.5% determiners. The remaining 12.5% were pronouns, adjectives and conjunctions. Average word frequency was 6.39 Zipf with a standard deviation of 0.84.

there was one central concern which prompted a redesign: in this experiment italic TE only occurred on CF words, and never on words that did not have CF. If this experiment was run with a full sample, it would thus not be able to show whether any effects of TE happened regardless of the information structure of the sentence in which it appeared. Consequently, a second experiment was designed and carried out: this will be discussed in the following chapter.

Chapter 5

Experiment 2

5.1 Introduction

As was discussed in the previous chapter, Experiment 1 was run as a pilot. The results suggested that typographic emphasis (TE) has an effect on processing, although clear patterns did not emerge. It appeared that when the target word had TE, antecedent focus words were fixated for longer in total, and that slightly more passes were made on post-target words.

Concerns over the design of Experiment 1 prompted the development of a second experiment: in particular, in Experiment 1 TE had always co-occurred with contrastive focus (CF), meaning that this design was unable to fully interrogate the relationship between the two. Experiment 2 was therefore designed to manipulate CF as a second factor.

The central difference between the two experiments was the addition of this factor. Whilst the approach taken in Experiment 1 should have been sufficient to test the hypothesis that TE-marked CF would be processed faster, the manipulation of CF as well allowed Experiment 2 to address the following question: does TE have a different effect on processing depending on whether it is placed on a word that has contrastive focus, or on a word that has presentational focus? If the central hypothesis holds that TE can facilitate processing of CF, because it is taken as marking it, then there should be an effect of 'congruence' when the two factors match: the congruent conditions $[+\,\text{CF}, +\,\text{TE}]$ and $[-\,\text{CF}, -\,\text{TE}]$ should facilitate processing in comparison to the incongruent conditions $[-\,\text{CF}, +\,\text{TE}]$ and $[+\,\text{CF}, -\,\text{TE}]$.

Experiment 2 was thus designed to test whether congruence between CF and TE facilitates processing. *Congruence* is used here only to refer to whether the presence of CF and TE matched, and should not be confused with any effects this

matching may have had on *felicity*: the term *felicity* is used here in relation to a secondary hypothesis, which predicted that sentences which had a TE-marked target word without CF would incur a processing cost, because it would be infelicitous (and thus confusing) for this word to be marked for contrast.

A second major difference between the experiments was a switch from using short written dialogues as stimuli to using short narrative-style texts. This change enabled the creation of considerably more varied stimuli without a great loss of control over the key words (i.e. the target and the antecedent focus word).

5.2 Organisation of chapter

§5.4 reports the revised **methodology** for Experiment 2. Of particular importance are the changes to the stimuli (§5.4.2); the switch to narrative-style stimuli is discussed here, as is the addition of the CF factor. Predictions are then given (§5.3).

§5.5 reports the **results** of this experiment, giving a full account of the analytical approach, which used linear mixed effects modelling (§5.5.4). The results of the models built to test effects of congruence (§5.5.5) and felicity (§5.5.6) are given, as are supplementary results including main effects of both factors (§5.5.8).

§5.6 provides a preliminary **discussion** of the results, ahead of the fuller discussion of both experiments in Chapter 6.

5.3 Hypotheses and predictions

5.3.1 Primary hypothesis

A sentence containing a word which has congruent CF and TE (i.e. both CF and TE, or neither) will be processed more quickly than an equivalent sentence in which that word which has incongruent CF and TE (i.e. CF but no TE, or TE but no CF).

As with Experiment 1, the intent was for this experiment to measure the processing cost in order to test this hypothesis: as before, processing cost was expected to manifest as re-reading of the target word and preceding words.

5.3.2   Secondary hypothesis

A sentence containing a word that is infelicitous (with no TE but with CF) will be processed more slowly than an equivalent sentence in which that word is felicitous (any other combination of present/absent CF and TE).

Again, this processing cost would manifest in the form of re-reading. The assumption here was that TE would prompt an initial analysis of the target word as having CF, but this would then conflict with the discourse context, in which it would be infelicitous for this word to have contrast. Again, this processing cost would manifest as re-reading of both the target word and preceding words.

5.3.3   Conditions

Testing these hypotheses required the use of two binary factors, CF and TE, which could be combined to create four conditions. Full details of the approach taken to manipulate these factors within stimulus frames are given below (§5.4.2); for now, Table 5.1 provides a simplified overview of the conditions.

The congruent conditions were $[-\,CF, -\,TE]$ and $[+\,CF, \,+\,TE]$ (rows 1 & 2 in Table 5.1), and the incongruent conditions were $[-\,CF, +\,TE]$ and $[+\,CF, -\,TE]$ (rows 3 & 4). To find support for the primary hypothesis would require finding a difference between these congruent and incongruent conditions: processing for the congruent condition would be facilitated in comparison to the incongruent condition. Each condition will now be examined in turn. Throughout, where processing difficulty is discussed, this was expected to manifest as re-reading which would be reflected by longer second-pass dwell times, longer total dwell times, and higher pass counts on all AOIs.

The condition $[-\,CF, -\,TE]$ (Table 5.1, row 1) was **congruent** in that there was

Table 5.1: Conditions

| Condition | Metafactors | | Hypothesised difficulty | Example |
	Congruent	Felicitous		
$[-\,CF, -\,TE]$	True	True	Low	I wanted fruit. I got oranges.
$[+\,CF, +\,TE]$	True	True	Low	I wanted apples. I got *oranges.*
$[+\,CF, -\,TE]$	False	True	Higher	I wanted apples. I got oranges.
$[-\,CF, +\,TE]$	False	False	Highest	I wanted fruit. I got *oranges.*

The example target word is *oranges.*

no dissonance between the information structure (IS) of the target word and how it was marked for contrast. Indeed, this is because it was *not* marked. Whilst it might be argued that the syntax of the sentence did mark the target word because, for example, it was an object, it appeared towards the end of the sentence, etc., this only constitutes focus marking and not contrast marking. The $[-CF, -TE]$ target word was a typical example of a non-contrastive, presentationally-focused word in the written modality. The containing target sentence was thus predicted to be interpreted as simply adding relevant information, and thus should have placed minimal load on the processor.

The $[+CF, +TE]$ condition (Table 5.1, row 2) was also **congruent**. The target word had contrast, and its TE marked it as such (if the central hypothesis of this study is assumed to be supported, and it is held that TE can be taken as marking contrast). Accordingly, this sentence was predicted to place minimal load on the processor, similar to $[-CF, -TE]$. Whilst it was predicted that the presence of contrast may have resulted in this condition being processed slower than the other congruent condition $[+CF, +TE]$, this was relatively unimportant to the primary hypothesis; what mattered was the difference in processing between these two congruent conditions and their incongruent equivalents.

The $[+CF, -TE]$ condition (Table 5.1, row 3) was **incongruent**. The target word had no TE to mark the CF, and this lack of marking was predicted to make the containing sentence harder to process. There was no pragmatic reason for this proposition to be made, unless it was being contrasted with the previous proposition, so the processor was predicted to eventually arrive at the contrastive analysis—yet the total lack of marking of the CF was expected to make this harder for the processor, thus delaying arrival at the correct analysis.

The $[-CF, +TE]$ condition was **incongruent**, and was also the sole **infelicitous** condition according to the theoretical framework established here. The discourse context did not create any reason for the target word to have contrast, and yet it had TE, which marked it as having contrast. The secondary hypothesis predicted that this infelicity should have resulted in an increased processing load as a workable analysis was (fruitlessly) sought.

Note that $[+CF, -TE]$ was not infelicitous despite being incongruent, as there should be nothing unusual about a CF appearing without TE. The primary hypothesis predicted it would be harder to process this CF than when it was marked with TE, but it should still have been possible—whereas in $[-CF, +TE]$, the secondary hypothesis predicted the spurious, unresolvable presence of TE was expected to hamper the processor even more severely.

Alternative explanations for why the primary hypothesis might *not* be supported were also entertained before running the experiment, including the following two possibilities: 1) either TE can 'force' a contrastive interpretation of a focus, perhaps even when the discourse context is infelicitous, or 2) TE is not (in this environment at least) reliably interpreted as marking CF. If there was an effect, but incongruence *facilitated* processing, it was assumed that this would indicate that there were problems with the theoretical position adopted here, and perhaps with the experimental design.

If there was no effect of felicity and thus the secondary hypothesis was not supported, it was predicted that this might suggest that the presence of TE is not accorded enough weight to make a sentence unacceptable.

First-pass dwell time was analysed in order to a) investigate whether there was an early, 'eye-catching' effect related to the visual salience of TE, and b) to check that there were no early effects of TE on antecedent focus words; if there were, this might have indicated that antecedent focus words were fixated after target words, which would be undesirable. The investigation of eye-catching effects recalls the secondary hypothesis of Experiment 1 (§4.5.2), although, in light of the lack of support found for that hypothesis, this was only a minor concern for Experiment 2.

5.4 Methodology

An eye tracking experiment was conducted which manipulated the presence of TE and CF on a target word within short on-screen narrative-style texts.

5.4.1 Participants

28 participants were recruited from the staff and student populations of the University of Leeds. Conditions of participation were the same as in Experiment 1. The mean age was 23 years ($SD = 2.8$). The mean years of completed university study was 3.3 ($SD = 0.97$).

Participants were randomly assigned into one of four groups for counterbalancing purposes: each group received 40 test items, ten of each of the four new conditions in a complementary pattern. For example, participants in Group A received Test item 1 in the $[+CF, +TE]$ condition, Group B received it in $[+CF, -TE]$, Group C in $[-CF, +TE]$, and Group D in $[-CF, -TE]$.

5.4.2   Materials

40 test items were designed in the form of short narrative-style texts (see Appendix B for a complete list). As with the stimuli in Experiment 1, each consisted of three sentences; this will be further discussed below, but (48) provides an example in the [+CF, +TE] condition.

(48)   Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure he couldn't eat peanuts. She spoke to a mutual friend, who said that if Nigel ate *almonds* he'd be very unwell.

5.4.2.1   Changing to narrative-style stimuli

Experiment 1 used script-like dialogues for its stimuli. Whilst several advantages of this form were discussed in §4.6.2.1, there were downsides, not least of which being the structural rigidity of the sentences containing the target words: to make the corrective nature of the sentence clear, the third sentence would be a repetition of the second sentence, with only the target word differing. This was likely to have increased the predictability of the test items, and to have had a negative effect on participants' motivation and interest levels as the experiment progressed. The repetitive nature of these items was compounded by the inclusion of several filler items which followed the same pattern, but repeated the antecedent focus word as well, meaning that they had confirmatory rather than corrective CF.

The dialogue format had been chosen for two reasons, the most important of which being the idea that corrections in dialogues would appear more natural than corrections in narrative text, as an inconsistent authorial voice is unusual in most written genres. However, it was decided subsequently that there would be nothing unnatural or unusual about third-person narration which involved characters having expectations, which were then contradicted. This contradiction could be expressed with corrective CF, as will be illustrated below.

The other motivating factor behind the adoption of the dialogue format had been the connection between direct speech and italics suggested within the literature on italic usage. At this stage of research into TE, this is only a tentative connection which has not yet seen strong support from corpus analysis; furthermore, many instances of TE usage outside of direct speech were brought up in Chapter 3: this issue was therefore not enough of a reason to discourage the adoption of narrative text in place of dialogues.

Moving away from the dialogue form allowed for the stimuli in Experiment 2 to adopt much more varied structures. In particular, it allowed for the creation

of corrective CF without exact word-for-word repetition of the second sentence in the third: this will be further discussed below. This change should have also resulted in a slight increase in ecological validity; when reading professionally typeset material, the average person is presumably more likely to be reading narrative prose than scripts.

5.4.2.2 Manipulation of factors

The presence of TE was manipulated on the target word in the third and final sentence. The presence of CF on the target word was manipulated by altering an antecedent focus word in the second sentence.

A note should be made on the use of the term *antecedent focus word* here. Recall that an *antecedent focus* has been defined here as a focus which selects the alternative set of a subsequent corrective CF (§2.4.3). However, this experiment manipulated the degree to which the referent of the word referred to here as the *antecedent focus word* contrasted with the target word; consequently, these words did not always technically follow the definition given, as their referents did not always contrast with the referent of the target word. With this having been noted, for convenience the term *antecedent focus word* will still be used here to refer to these words.

Antecedent focus words were chosen that had differing hyponymy in relation to the target words, so that the second sentence either made a proposition that contrasted with that of the third sentence, or did not contrast with it: the words alternated between either being a co-hyponym of the target word [+CF], or a hypernym of it [−CF]. This relationship is illustrated in Figure 5.1.

The following explication takes Test item 1 as an example. The discourse topic was *foods one can be allergic to*, the antecedent focus word either *nuts* or *peanuts*, and the target word was *almonds*.

1. Discourse topic sentence: A discourse topic was introduced; as in Experiment 1, this was a group or set of things, typically in the form of an NP, and

[−CF]: *Types of food*   [+CF]:   *Types of food*
            | ╱   ╲
         Nuts*   Peanuts* Almonds[†]
            |
        Almonds[†]
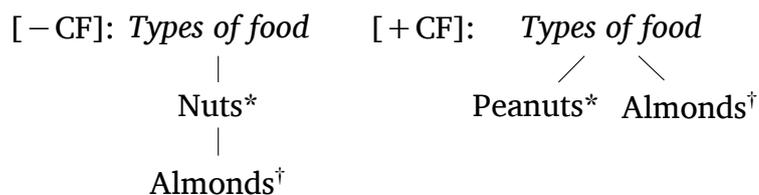
Figure 5.1: Hyponymy manipulation
*antecedent focus word, [†]target word

introduced using various different constructions in order to limit the predictability of test items.

   'Steph's friend Nigel was allergic to a certain kind of food'.

2. Antecedent focus word sentence: As in Experiment 1, a proposition was made about a referent that was a member of the discourse topic set. This referent was denoted by a focused noun, the **antecedent focus word**. Altering the hyponymy of this word either made the target sentence contradict this sentence, thus placing CF on the target word, or made the target sentence simply add more specific information to the discourse topic, placing presentational focus on the target word. In the $[-\text{CF}]$ condition, the antecedent focus word was a hyponym of the target word (*almonds* in this item) and either a co-hyponym or hypernym of the discourse topic (*foods one can be allergic to*):

   (Antecedent focus word: *nuts*)
   'Steph was fairly sure he couldn't eat nuts.'

In the $[+\text{CF}]$ condition, the antecedent focus word was a co-hyponym of the target word, and a hypernym of the discourse topic:

   (Antecedent focus word: *peanuts*)
   'Steph was fairly sure he couldn't eat peanuts.'

3. Target sentence: A proposition was made about a referent which was a hypernym of the discourse topic. This appeared as a focused noun, the **target word**, which was set in either italic or roman style, depending on the level of TE.

   (Target word: *almonds*)
   'She spoke to a mutual friend, who said that if Nigel ate *almonds* he'd
   be very unwell.'

(49) shows Test item 1 in all four conditions arising from the combination of the two factors.

(49)   a. $[-\text{CF}, -\text{TE}]$ Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure he couldn't eat nuts*. She spoke to a mutual friend, who said that if Nigel ate almonds[†] he'd be very unwell.
       b. $[+\text{CF}, -\text{TE}]$ Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure he couldn't eat peanuts*. She spoke to a mutual friend, who said that if Nigel ate almonds[†] he'd be very unwell.

   c. $[-CF, +TE]$ Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure he couldn't eat nuts*. She spoke to a mutual friend, who said that if Nigel ate almonds† he'd be very unwell.

   d. $[+CF, +TE]$ Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure he couldn't eat peanuts*. She spoke to a mutual friend, who said that if Nigel ate almonds† he'd be very unwell.

*antecedent focus word, †target word

This experiment did not use the technique used in its predecessor, in which the second sentence of each test item was repeated in the third sentence. This structure had been deemed necessary for the creation of CF given the dialogue format of that experiment, as it seemed to be the only natural way one speaker might correct another which did not involve linguistic devices which would 'tip the processor off' before the target word was encountered (such as clefts, outright negation such as *No, I think you'll find that…*, etc.). Freed from the constraints of two 'independent' characters obeying conversational maxims, the new narrative format was able to express its final proposition, whether it was corrective or not, using a greater variety of sentences.

The CF factor relied wholly on the manipulation of the hyponymy of the antecedent focus word. It is perhaps debatable whether the $[+CF]$ condition, where the antecedent focus was a co-hyponym of the target word, involved any actual 'marking' of the target word as being contrastive, as required by the definitions of CF provided by Zimmermann (2007) and Calhoun (2009). In Experiment 1, the target sentence's word-for-word repetition of the second sentence could be argued to be a form of marking. In Experiment 2, the target word in the $[+CF]$ condition was unmarked—unless the fact that it formed a proposition that directly contradicted the previous proposition were taken to constitute 'marking'. Whilst this contradiction or rejection still meant that the focus word had *corrective focus* in the sense of Gussenhoven (2006), it may be more accurate to see what is referred to as a [CF] feature here as something more subtle than the presence of corrective CF: $[+CF]$ target words can be seen as words that had to be interpreted contrastively for the sake of felicity, and $[-CF]$ target words were the converse of this.

A possible confound may have arisen here due to semantic priming: there may have been a different level of priming for different semantic levels. In $[-CF]$, when the antecedent focus word was a hyponym of the target word, this may

have primed the target word to a different degree than in [+CF] when it and the antecedent focus word were co-hyponyms.

A solution that was considered for this potential problem was making the antecedent focus word exactly the same word as the target word for [−CF], completely eliminating the contrast. This would have resulted in the target word having confirmatory focus.[1] However, it seems likely that the repetition priming effect of this would have been greater than the semantic priming that may have occurred with the chosen approach of manipulating hyponymy.

Note on use of names   A large number of the stimuli referred to named characters, in an effort to avoid a large quantity of sentences featuring indefinite nouns and anonymous roles, such as 'a woman/man was doing x', which might seem unusual. Names were chosen so as to maintain a roughly equal split between traditionally gendered names across items.

5.4.2.3   Check questions

Check questions followed the same format as those of Experiment 1. 20% of the items were followed by check questions (36 items). 50% of the questions were true. The check questions did not contain the target word, nor did they relate to its referent. See Appendix B for the check question text (each is shown after the item it followed).

5.4.2.4   Word selection

In Experiment 1 target words were not strictly controlled, meaning that properties such as word frequency and character length were highly variable. Whilst no significant effects of these properties (and syllable count) were found in the preliminary results, nevertheless it was decided that the risk of confounding effects in this area warranted a higher level of control.

New target words were selected from Subtlex-UK, a corpus built from television subtitles (van Heuven et al., 2014). Subtlex-UK was preferred over the commonly-used British National Corpus (BNC, 2007) due to it being more recent, and to it having been shown to account for 4% more variance in lexical decision times than the British National Corpus (van Heuven et al., 2014): this suggests that Subtlex-UK provides more accurate word frequencies. All target words were

---

[1]Some fillers in Experiment 1 were structured in this way to provide a distracting 'balance' from the repeated corrective focus constructions in the test items, but they were not analysed.

two syllable common nouns controlled for (in descending order of priority) word frequency, character count, and visual width.

Descriptive statistics for these and other word properties are given for each type of word in Table 5.2 (N.B. post-target and antecedent were not strictly controlled for these properties).

Table 5.2: Descriptive statistics for word properties

	Targets		Post-targets		Antecedents	
Property	Mean	*SD*	Mean	*SD*	Mean	*SD*
Word frequency (Zipf)	3.49	0.02	6.15	1.35	4.03	0.76
Character count	6.8	0.8	4.2	2.5	6.6	2
Visual width (pixels)	119	14	72	45	110	35
Syllable count	2	0	1.35	0.99	1.95	0.84
Phoneme count	5.8	0.8	3.5	2.5	5.5	1.7
Orthographic neighbourhood size	1.68	1.46	11.42	8.73	4.92	5.93
Phonological neighbourhood size	2.42	2.72	23.32	16.77	8.35	11.26

Frequency data taken from SUBTLEX-UK (van Heuven et al., 2014); orthographic and phonological neighbourhood sizes from CLEARPOND (Marian et al., 2012).

Because of the new stimulus frames, which were more varied in content than those of Experiment 1, target words no longer appeared in quite such predictable places on screen (in Experiment 1 they always appeared on the third line of text). Furthermore, the addition of the second factor, CF, resulted in slightly different visual widths of some lines across the condition, as different antecedent focus words were used. The new stimuli were designed to ensure that no target words appeared at the start or end of a line, or in appreciably different places on the x axis across the CF condition (however, different items had target words appear in different places on the y axis, i.e. in different lines of text).

It should be noted that the characteristics of the antecedent focus words were not controlled as tightly as those of the target words, as doing this whilst also manipulating their hyponymy in relation to the target words (and maintaining their relationship to the discourse topic) would have been prohibitively difficult. As can be seen in Table 5.2, whilst antecedent focus words were more variable than the target words, on the whole they were still fairly similar to each other. The slightly higher level of variability may have reduced the likelihood of finding significant differences between factors in first-pass dwell times on these words; however, these tests were of less importance than those testing the primary and secondary hypotheses.

5.4.2.5 Fillers

Experiment 1 had included a complicated array of fillers designed to present every possible combination of the basic features of the test stimulus frame. For example, 40 of the 120 fillers had the same set/member structure as the test items, but instead of Speaker B correcting Speaker A, they repeated Speaker A's second statement to create confirmatory focus (Krifka, 2007).

In designing Experiment 2 it was decided that the negatives of this approach had outweighed the positives: as a consequence of the filler design in Experiment 1, a full half of the items presented to each participant had almost the exact same structure. Whilst this may have obscured the subtleties of the experimental manipulation at a low level, at a higher level it resulted in extremely repetitive and uninteresting items. To reduce this repetition, the new fillers were designed to be highly varied. The majority consisted of three sentences, but the content and structure of these sentences varied considerably (see Appendix B for the full list of fillers).

The ratio of fillers to test items was also adjusted: A 2:1 ratio was used instead of the previous 3:1 ratio. This reduced the number of items each participant saw by 24% (from 170 to 130). The number of words each participant read across all items and check questions was 17% smaller (from 6,254 words to 5,217 words in Experiment 2: see Table §5.3 below). It was hoped that this slight reduction would reduce participant fatigue. The number of breaks was kept the same, with four blocks of 30 trials separated by a mandatory 30 second break (half as short as before), giving way to a 'press <Next> to continue' message.

Experiment 1's use of distractors featuring bold TE on intensifying adverbs was repeated: 20 items of this type were included, meaning participants saw the same number of bold TE as had italic TE.

5.4.2.6 Semantic controls

As with Experiment 1, word frequencies within the experiment were checked, and each target word and antecedent focus word only appeared once. Additionally, the frequencies of all words across the experiment were checked using the Subtlex-UK corpus (van Heuven et al., 2014), and unusually low-frequency words were substituted with higher-frequency equivalents to keep the text readable and undemanding. This was done in case a higher reading difficulty (other than that on items presented with the deliberately incongruent conditions) might demotivate or fatigue participants.

Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure that he couldn't eat peanuts. She spoke to a mutual friend, who said that if Nigel ate *almonds* he'd be very unwell.
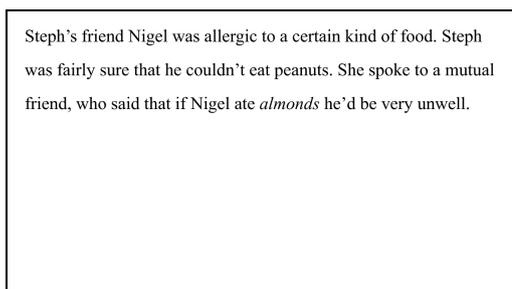
Figure 5.2: Screenshot of Test item 1

True or false: Poppy's hair was very slow to grow.

\<True\>   \<False\>
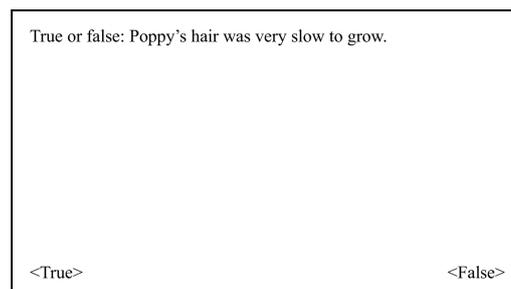
Figure 5.3: Screenshot of Check question 1

5.4.2.7 Visual/typographic characteristics

Items, instructions and check questions had the same visual characteristics as in Experiment 1. Figures 5.2 and 5.3 show screenshots of the first test item and first check question respectively.

5.4.2.8 Stimuli statistics

As Table 5.3 shows, each participant saw 130 items, consisting of 10 practice items, 40 test items and 80 fillers. Additionally they saw 26 check questions. In total they were presented with 5,217 words over the course of the experiment (excepting text other than that in the stimuli, such as the instructions, break messages etc.).

5.4.2.9 Item order

Each participant received the test and filler items in a different pseudorandomized order: sequences of three items were constructed, each containing two fillers and one test item, and the order of these sequences was randomized. The display

Table 5.3: Stimuli word statistics

Category	Item count	Words			
		Count	% of total	Mean	*SD*
Practice items	10	384	7%	38	3
Test items	40	1629	31%	41	5
Filler items	80	3029	58%	38	6
Check questions	26	175	3%	7	2
All	156	5217	100%	35	11

of those items that were followed by check questions was similarly pseudoran-domized.

5.4.3   Apparatus

The stimuli were presented on a 21.5inch Iiyaama Prolite B2280HS-B1 LCD monitor in its native resolution of $1{,}920 \times 1{,}080$ pixels. This was a larger, higher-resolution display than that used in the first experiment. It was hoped that the higher-quality display would improve reading conditions, increasing eye tracking accuracy and decreasing fatigue (one participant in Experiment 1 had reported their eyes feeling tired after finishing the experiment). Other than this, the same apparatus was used as in Experiment 1.

The position of the eye tracker, screen and head rest were brought more closely in line with recommendations in the eye tracker manual than they had previously been. This, along with the improved display, seemed to improve the accuracy of the recordings: fewer recalibrations were required during the experiment, and visual inspection of the recordings suggested they were more accurate than those attained previously (see §5.5.2).

5.4.4   Procedure

The same procedure followed in Experiment 1 was used. The ten-item practice block was slightly different in that it did not contain a test-like item for every condition, but instead had three test-like items with the conditions $[+\,\mathrm{CF}, -\,\mathrm{TE}]$, $[+\,\mathrm{CF}, +\,\mathrm{TE}]$ and $[-\,\mathrm{CF}, -\,\mathrm{TE}]$, as well as seven fillers: for this block to contain every condition and maintain the ratio of fillers to test items, it would have had to contain twelve items, which was judged to be too many for a practice.

5.4.5   Selection of metrics and areas of interest

The same assumptions were held about processing difficulty prompting re-reading: in Experiment 1 the $[-\,\mathrm{TE}]$ condition was incongruent with the universally $[+\,\mathrm{CF}]$ target words, and the $[+\,\mathrm{TE}]$ condition was congruent. Experiment 2 added a second incongruent condition and a second congruent one, to test a primary hypothesis which represented an evolution of the one tested in Experiment 1. As such, the same metrics and areas of interest (AOIs) could be used as before. To capture the re-reading predicted in the primary and secondary hypotheses, **second-pass dwell time**, **total dwell time** and **pass count** were measured. **First-pass dwell**

time was also measured to investigate whether the increased visual salience of [+TE] words had an effect, and to check that unusual reading behaviour was not occurring. The AOIs were the **target word**, the **antecedent focus word**, and the **post-target word** (to check for spillover processing).

5.5   Results

Four AOI-based metrics from the eye tracking data were analysed: first-pass dwell time, second-pass dwell time, total dwell time, and pass count. These metrics were calculated for three AOIs per trial: the target word, the post-target word, i.e. that word immediately following the target (to investigate potential spill-over effects), and the antecedent focus word, the noun with which the CF on the target word contrasts. No effects were found on post-target words, and they are largely excluded from the analysis that follows (but see §5.5.8.3).

The same programs were used to generate and analyse metrics as were used in Experiment 1 (see §4.7).

5.5.1   Issues with recordings

The software used to generate this experiment was a Windows 10 beta version of SR Research Builder. Assurances in private communication were obtained from SR Research that recording accuracy would not be endangered by the use of a pre-release version of the program; however, they warned that there might be slight timing issues, causing the point at which stimuli were displayed at the start of each trial to vary minutely. This was not judged to be an impediment for this experiment, because all the metrics being analysed would relate to individual AOIs rather than the trial as a whole. During testing, a handful of trials across participants did have a pause, ranging between roughly one to ten seconds, between presentation of the fixation point and the stimulus: when a pause over one or two seconds occurred, the participants were told to continue fixating the place where the point had been and await the appearance of the text.

During testing of Participant 16 the experiment display program crashed immediately after the second break (i.e. after two blocks had been presented). The experiment was re-started and the participant was asked to skip through the practice block (by hitting the 'Next' button immediately as each item was displayed), before being presented with the two remaining blocks. The participant was thus not presented with any of the items twice, and still received the standard number

of test and filler items in the correct order. During analysis, the results for this participant did not appear unusual and were not discarded.

5.5.2 Data preparation

Eye movement recordings were inspected and prepared following the same procedure as used in Experiment 1 (see §4.7.1).

In Experiment 1, trials in which the target word was not directly fixated at any point were removed from the analysis. However, such an approach was decided against for Experiment 2: whilst it was possible that a lack of direct fixation metrics on a given word indicated that it had not been processed, it was also possible that these words *had* been processed when they were in parafoveal vision (the so-called *preview effect*; Rayner et al., 1989; Schotter et al., 2012—see §4.4.1). At any rate, there was a lower rate of skipped target words in this dataset, which may have been due to improved tracking accuracy: 5% of trials had no fixations on the target word, in comparison to 21% in Experiment 1.

5.5.3 Data distribution

Distributions for dwell times on all AOIs are shown in Figure 5.4. As in Experiment 1, the dwell time metrics for all AOIs were non-normally distributed (Shapiro–Wilk tests returned $p < 0.05$ for all metrics and AOI word types). Distributions again conformed to the ex-Gaussian pattern of a high number of shorter durations, and a long tail of longer durations.

It can be seen that the data contained no zeroes, unlike that of Experiment 1 (see Figure 4.8 in §4.7.2): in this analysis, unlike that of Experiment 1, zeroes were treated as missing data. This was intended to improve validity, as comparing pass times of zero with ones that are greater than zero is not a particularly valid way of testing processing costs or advantages: a word that has not been fixated at all has not been processed infinitely more quickly than one that has been fixated for a certain time.

As the dwell times were not normally distributed, the principal measure of central tendency used here will again be the median for all metrics except pass count. Pass count again had a very low range (*Min* = 1, *Max* = 5) and was heavily right-skewed; the mean is a more useful measure for this.

Tables 5.4 and 5.5 show descriptive statistics for the data, grouped by CF and TE respectively. Target words were the most likely type of word to be fixated at least once (92% of AOIs presented). Antecedent focus words were also frequently

a) Target words

b) Post-target words

c) Antecedent focus words
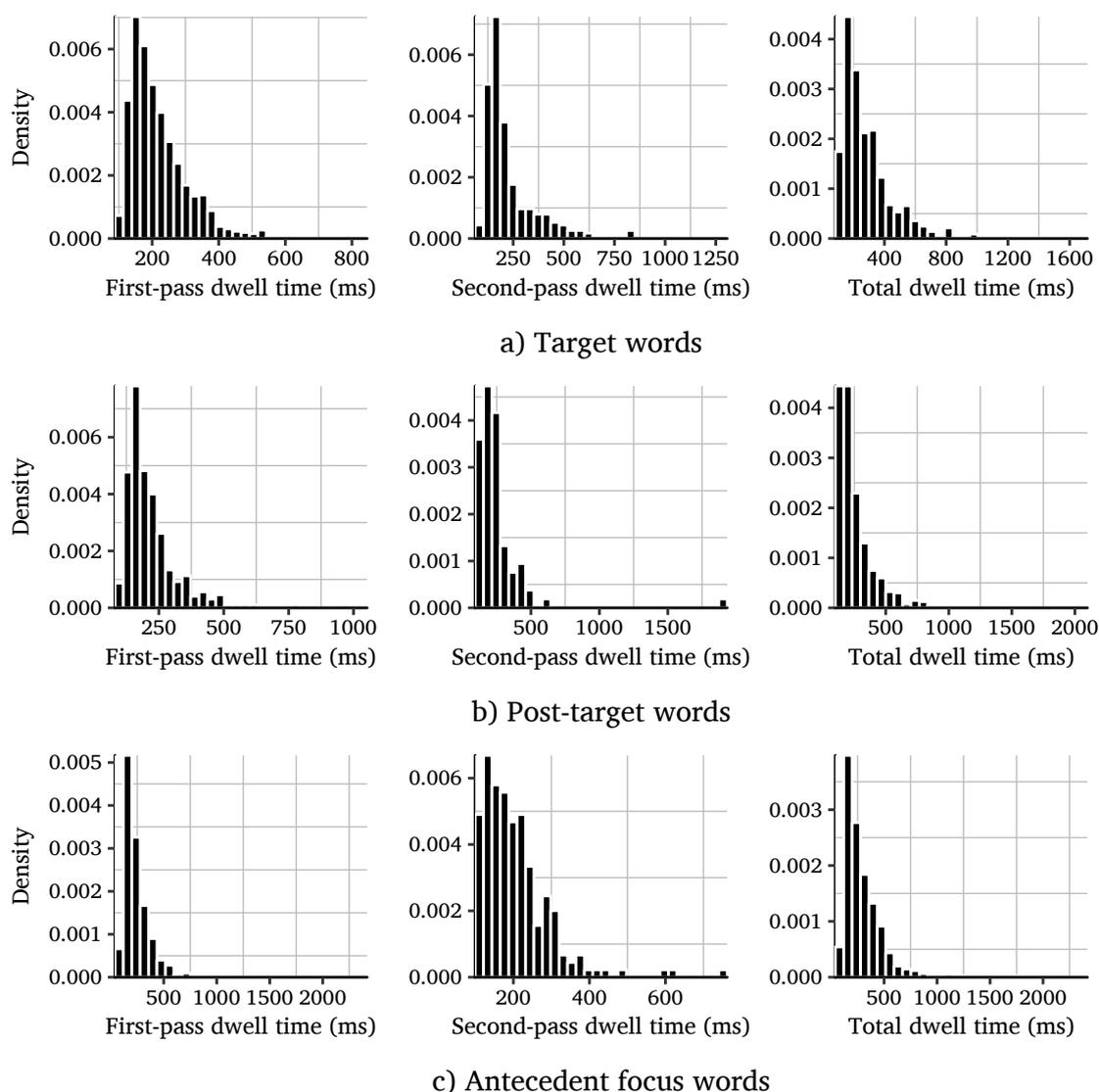
Figure 5.4: Densities for dwell times on AOIs

fixated once or more (85%). Post-target words, which were shorter than the targets (4.2 mean characters, in comparison to 6.8 for targets—see Table 5.2) and often functional in nature, were unsurprisingly fixated less frequently (54%).

Differences between median dwell times were very low for both factors, ranging from -23 ms (for total dwell time on post-targets by CF) to 38 ms (for second-pass dwell time on post-targets by CF). The mean difference was 4 ms. Similarly, differences between pass counts were low, ranging from -0.07 (for pass count on targets by CF) to 0.11 (for pass count on targets by TE). This suggests that any effects found here will be small.

Table 5.4: Descriptive statistics by CF for all metrics, all AOIs

	First-pass dwell time		Second-pass dwell time		Total dwell time		Pass count	
	[−CF]	[+CF]	[−CF]	[+CF]	[−CF]	[+CF]	[−CF]	[+CF]
Target words								
N	514	518	149	124	514	518	514	518
Minimum	101	101	101	103	101	101	1	1
Lower quartile	157	157	146	152	169	167	1	1
Median	196	199	176	191	240	240	1	1
Upper quartile	255	264	237	304	359	346	2	1
Maximum	833	618	964	1304	1378	1680	5	4
IQR	98	107	91	152	190	179	1	0
Mean	219	221	223	254	300	294	1.36	1.29
SD	91	88	137	172	190	190	0.64	0.56
Post-target words								
N	299	303	46	40	299	303	299	303
Minimum	102	101	101	103	103	101	1	1
Lower quartile	155	153	148	172	163	158	1	1
Median	195	188	181	219	222	199	1	1
Upper quartile	265	236	245	287	315	292	1	1
Maximum	1040	923	508	1882	1040	2034	4	5
IQR	110	83	97	115	152	134	0	0
Mean	231	220	215	284	268	261	1.17	1.15
SD	126	115	100	279	158	190	0.42	0.43
Antecedent focus words								
N	475	477	111	92	475	477	475	477
Minimum	101	101	103	103	101	101	1	1
Lower quartile	155	158	135	155	167	171	1	1
Median	207	211	182	192	243	243	1	1
Upper quartile	284	298	229	270	372	369	1	1
Maximum	1085	2401	608	744	1728	2401	8	5
IQR	129	140	94	115	205	198	0	0
Mean	242	254	196	217	306	309	1.32	1.25
SD	129	190	81	102	209	242	0.75	0.59

Table 5.5: Descriptive statistics by TE for all metrics, all AOIs

	First-pass dwell time		Second-pass dwell time		Total dwell time		Pass count	
	[−TE]	[+TE]	[−TE]	[+TE]	[−TE]	[+TE]	[−TE]	[+TE]
Target words								
N	500	532	112	161	500	532	500	532
Minimum	101	104	101	101	101	104	1	1
Lower quartile	154	160	143	152	160	176	1	1
Median	193	201	180	184	235	245	1	1
Upper quartile	259	262	226	298	338	381	1	2
Maximum	833	799	699	1304	1221	1680	5	5
IQR	105	102	83	146	178	205	0	1
Mean	217	223	214	254	275	318	1.27	1.38
SD	89	91	114	176	164	210	0.55	0.64
Post-target words								
N	309	293	47	39	309	293	309	293
Minimum	101	103	101	103	101	103	1	1
Lower quartile	157	152	153	169	162	159	1	1
Median	189	191	199	218	207	207	1.00	1.00
Upper quartile	255	248	247	306	306	307	1	1
Maximum	970	1040	508	1882	970	2034	4	5
IQR	98	96	95	138	144	148	0	0
Mean	224	228	216	284	260	270	1.17	1.15
SD	111	130	95	285	147	200	0.41	0.44
Antecedent focus words								
N	475	477	95	108	475	477	475	477
Minimum	101	102	105	103	101	102	1	1
Lower quartile	158	158	157	135	172	168	1	1
Median	209	207	194	180	242	244	1	1
Upper quartile	286	297	272	224	362	377	1	1
Maximum	1779	2401	744	608	2012	2401	8	7
IQR	128	139	115	89	190	209	0	0
Mean	243	253	220	193	303	312	1.27	1.30
SD	143	180	101	81	215	237	0.68	0.67
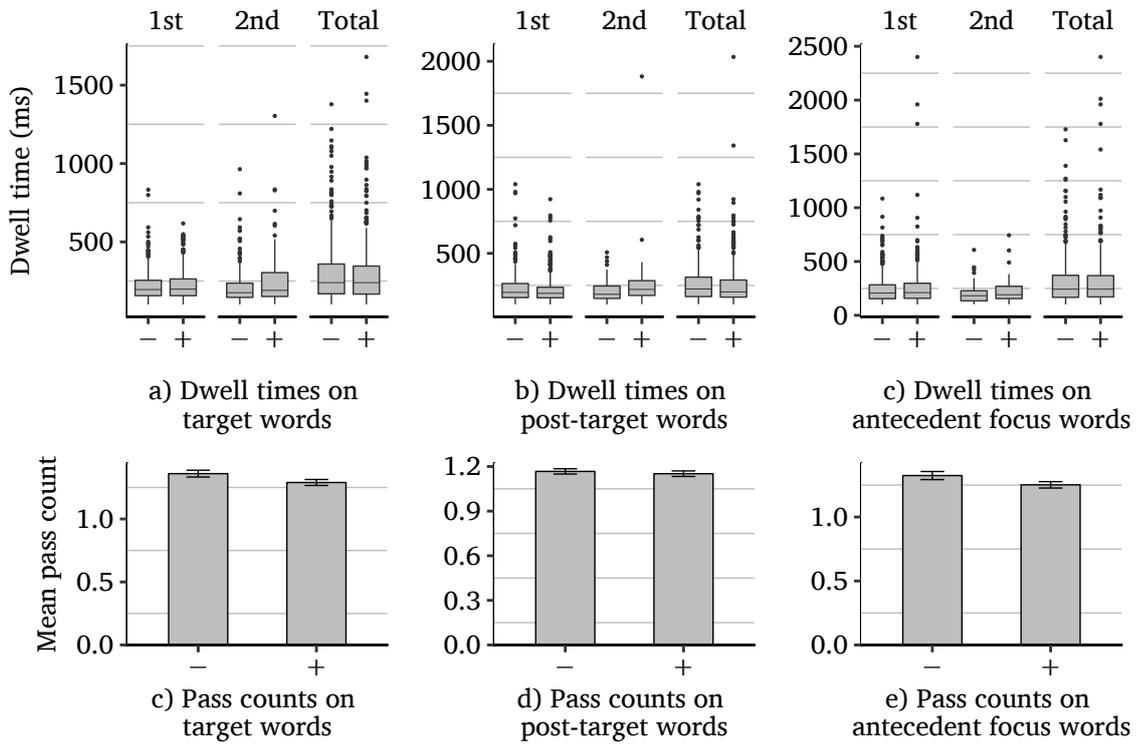
Figure 5.5: Charts by CF for all metrics, all AOIs
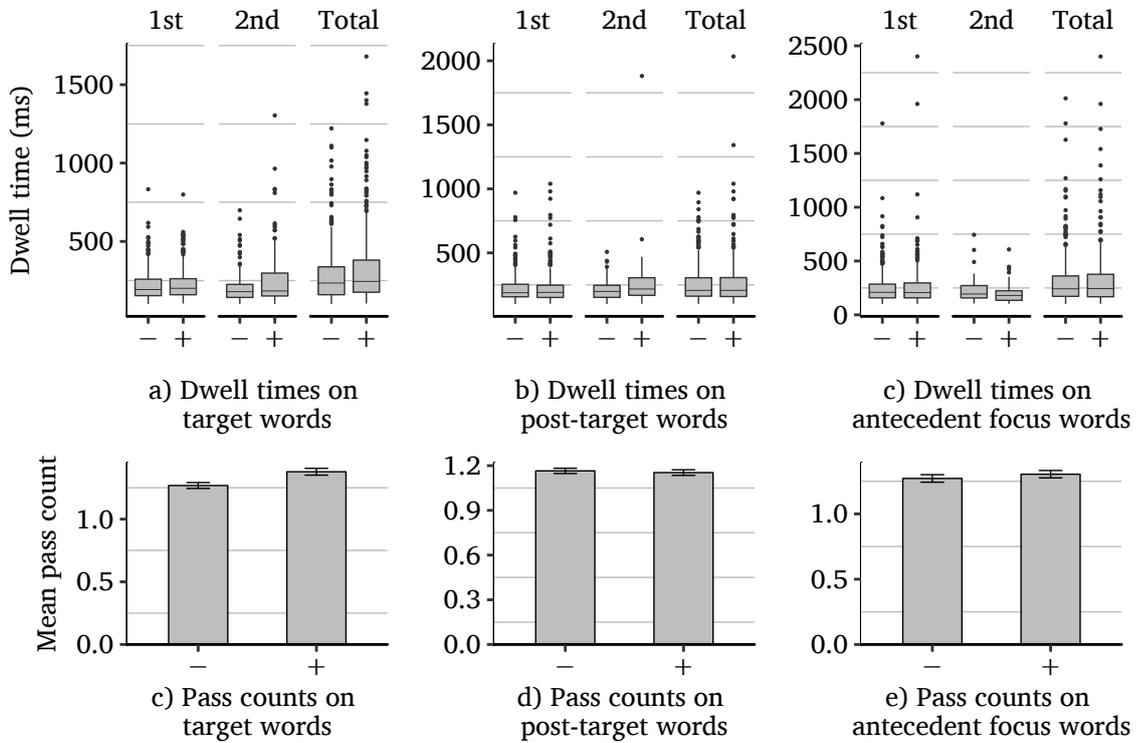
Figure 5.6: Charts by TE for all metrics, all AOIs

5.5.4   Analytical approach

a) Untransformed

b) Transformed (following Box−Cox procedure)
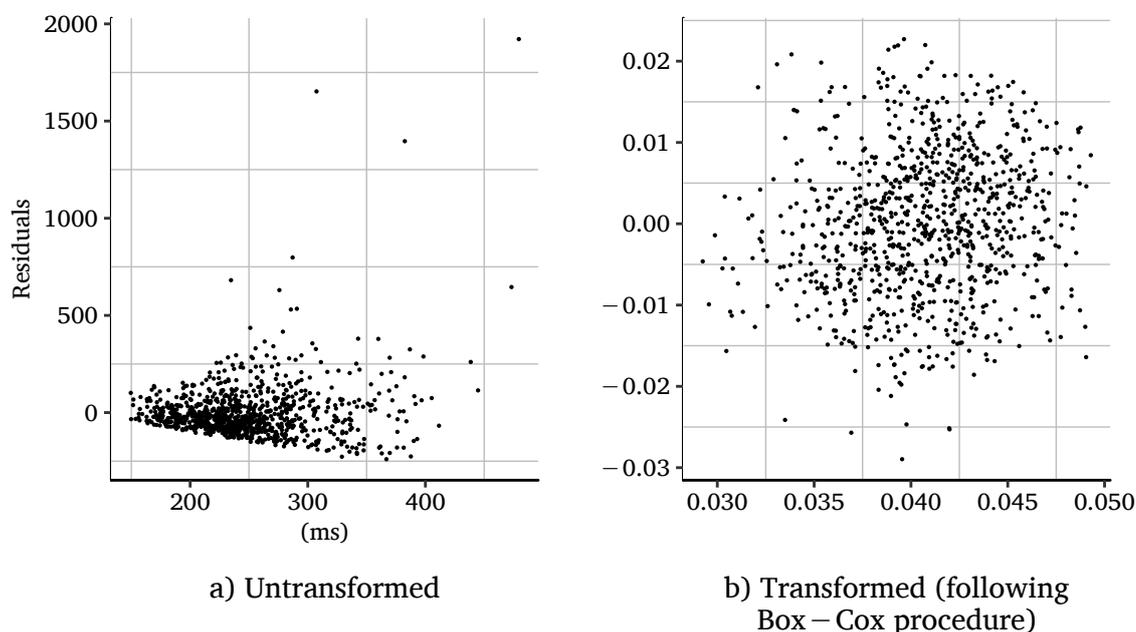
Figure 5.7: Residuals for linear mixed effects model for first-pass dwell time on antecedent focus words

Linear mixed effects models (LMMs) were constructed for each dwell time metric and word type. Visual inspection of residual plots for initial models suggested heteroscedasticity for all metrics, with variance increasing as the outcome increased (Figure 5.7.a gives an example). The Box–Cox procedure (Box and Cox, 1964) was used to arrive at exponents by which to transform the data; visual inspection of plots for the transformed data suggested appropriately normal distributions of residuals (e.g. Figure 5.7.b).

The *lme4* R package (Bates et al., 2015) was used to construct the models. The following approach was taken to construct each model. First, a minimal model was constructed with the metric as the outcome, CF and TE as interacting fixed factors, and subject as a random factor.[2] This model (fitted by maximum likelihood) was then compared with a model which added item as a random factor: if a log-likelihood ratio test found a significant difference between the two models, this newer model was adopted, and compared in the same way against successive models adding target word frequency, and target word length as random factors. For example, the code used to create the final model for first-pass dwell time on

---

[2] An LMM must have at least one random factor; of the two 'traditional' psycholinguistic random factors, subject was judged more likely to account for more variance than item, given that items here had been deliberately designed and controlled.

target words was `lmer(first.pass.dwell.time ~ CF * TE + (1 | item) + (1 | subject), data = target.words)`. Results from these models, including their random effects structures, are reported in Tables 5.6, 5.7 and 5.8. Correlations between random factors were checked and found to all be below 0.18.

Generalised linear mixed effects models (GLMs) were constructed for pass counts, with Poisson error distributions. The same procedure of model construction was followed.

5.5.5   Testing the primary hypothesis: effects of congruence

The primary hypothesis predicted that sentences containing a target word on which CF and TE were incongruent would incur a processing cost, relative to sentences that were identical except with target words that had congruent CF and TE. Thus (as was summarised in Table 5.1) $[-\text{CF}, -\text{TE}]$ and $[+\text{CF}, +\text{TE}]$ were congruent, and $[-\text{CF}, +\text{TE}]$ and $[+\text{CF}, -\text{TE}]$ were incongruent. Differences between these two pairs of factors would appear in the results of the LMM as an interaction between CF and TE. As can be seen in Tables 5.6–5.9, across the three AOIs and four metrics there was only one significant interaction between CF and TE, which was found on second-pass dwell time on antecedent focus words (estimate = 0.0076, $SE = 0.0036$, $t = 2.12$, $p = 0.036$). The fact that there were so few significant interactions means that these results provide little evidence to support the hypothesis: multiple metrics on multiple words were predicted to be affected by congruence.

The existence of an interaction alone is not sufficient to provide any support to the hypothesis. That would require that the interaction occurred between the incongruent and congruent conditions, and to have done so in the direction predicted. To investigate this, pairwise comparisons were conducted with a Tukey HSD test, the results of which are shown in Table 5.10. There were two significant results which involved one of the two types of incongruent condition, $[+\text{CF}, -\text{TE}]$: there was a significant difference between trials presented in this condition and those in the $[+\text{CF}, +\text{TE}]$ condition (estimate = -0.0079, $SE = 0.026$, $z = -3.01$, $p = 0.0138$), and a significant difference between trials presented in the same $[+\text{CF}, -\text{TE}]$ incongruent condition and the 'opposite' congruent condition $[-\text{CF}, -\text{TE}]$ (estimate = -0.0073, $SE = 0.026$, $z = -2.79$, $p = 0.0268$). In both, incongruence was associated with longer second pass times. However, support for the hypothesis would require significant results when comparing the second type of incongruent condition, $[-\text{CF}, +\text{TE}]$, with the congruent conditions, and none were found.

Table 5.6: Linear mixed effects model results for first-pass dwell times, all AOIs

Coefficient	Estimate	SE	t	p	Random effects	Variance
Target words						
(Intercept)	0.0718	0.0012	60.3	< 0.001*	Item	0.000007
CF	−0.0001	0.001	−0.11	0.914	Subject	0.000021
TE	−0.0011	0.001	−1.08	0.282	Residual	0.000123
CF:TE	−0.0004	0.0014	−0.290	0.77		
Post-target words						
(Intercept)	0.0151	0.0005	32.3	< 0.001*	Subject	0.000001
CF	0.0001	0.0004	0.16	0.87	Item	0.000003
TE	−0.0002	0.0004	−0.38	0.706	Residual	0.000013
CF:TE	0.0006	0.0006	0.92	0.357		
Antecedent focus words						
(Intercept)	0.0407	0.0012	34	< 0.001*	Item	0.000005
CF	−0.0008	0.0009	−0.98	0.326	Subject	0.000012
TE	−0.0011	0.0008	−1.26	0.206	Characters	0.000004
CF:TE	0.0015	0.0012	1.26	0.207	Residual	0.000084

Significance values estimated by the *lmerTest* R package (Kuznetsova et al., 2016). Box–Cox transformation coefficients: target words = −0.5, post-target words = −0.8, antecedent focus words = −0.6.

Table 5.7: Linear mixed effects model results for second-pass dwell times, all AOIs

Coefficient	Estimate	SE	t	p	Random effects	Variance
Target words						
(Intercept)	0.0098	0.0004	22.8	< 0.001*	Subject	0.0000003
CF	−0.0009	0.0006	−1.44	0.152	Residual	0.0000102
TE	−0.0008	0.0005	−1.46	0.144		
CF:TE	0.0002	0.0008	0.28	0.777		
Post-target words						
(Intercept)	0.0741	0.0028	26.9	< 0.001*	Subject	0
CF	−0.0047	0.0045	−1.05	0.298	Residual	0.0002204
TE	−0.0028	0.0045	−0.62	0.534		
CF:TE	−0.0021	0.0065	−0.32	0.75		
Antecedent focus words						
(Intercept)	0.0751	0.002	38.1	< 0.001*	Item	0.0000053
CF	−0.0073	0.0026	−2.79	< 0.01*	Subject	0.0000136
TE	0.0004	0.0024	0.16	0.874	Residual	0.0001
CF:TE	0.0076	0.0036	2.12	0.036*		

Box–Cox transformation coefficients: target words = −0.9, post-target words = −0.5, antecedent focus words = −0.5.

Table 5.8: Linear mixed effects model results for total dwell times, all AOIs

Coefficient	Estimate	SE	t	p	Random effects	Variance
Target words						
(Intercept)	0.0664	0.0015	45.2	< 0.001*	Item	0.000019
CF	0.0005	0.0013	0.39	0.693	Subject	0.000024
TE	−0.0036	0.0013	−2.85	< 0.01*	Residual	0.000201
CF:TE	0	0.0018	−0.01	0.995		
Post-target words						
(Intercept)	0.0236	0.0008	28.8	< 0.001*	Item	0.000012
CF	0.0009	0.0007	1.39	0.166	Subject	0.000004
TE	0.0005	0.0007	0.72	0.47	Residual	0.000033
CF:TE	−0.0005	0.001	−0.55	0.58		
Antecedent focus words						
(Intercept)	0.0641	0.002	32.8	< 0.001*	Item	0.000014
CF	−0.0008	0.0014	−0.570	0.566	Subject	0.000024
TE	−0.0018	0.0014	−1.31	0.189	Characters	0.000014
CF:TE	0.0028	0.0019	1.44	0.15	Residual	0.000220

Box–Cox transformation coefficients: target words = −0.5, post-target words = −0.7, antecedent focus words = −0.5.

Table 5.9: Generalised linear mixed effects models results for pass counts, all AOIs

Coefficient	Estimate	SE	z	p	Random effects	Variance
Target words						
(Intercept)	0.2735	0.0335	8.152	< 0.001*	Subject	0.0105
TE1	0.0408	0.0271	1.506	0.132		
CF1	−0.0254	0.0271	−0.937	0.349		
TE1:CF1	−0.0014	0.0271	−0.052	0.959		
Post-target words						
(Intercept)	0.1471	0.0379	3.881	< 0.01*	Subject	0
TE1	−0.0049	0.0379	−0.130	0.897		
CF1	−0.0056	0.0379	−0.149	0.882		
TE1:CF1	0.0300	0.0379	0.792	0.42		
Antecedent focus words						
(Intercept)	0.2514	0.0298	8.430	< 0.001*	Subject	0.00156
TE1	0.0124	0.0286	0.434	0.664		
CF1	−0.0280	0.0286	−0.979	0.328		
TE1:CF1	−0.0087	0.0286	−0.305	0.760		

Table 5.10: Pairwise comparisons between incongruent and incongruent conditions for second-pass dwell time on antecedent focus words

Comparison	Estimate	SE	z	p
$[+\mathrm{CF},-\mathrm{TE}]-[+\mathrm{CF},+\mathrm{TE}]=0$	-0.0079	0.0026	-3.01	0.0138*
$[+\mathrm{CF},-\mathrm{TE}]-[-\mathrm{CF},-\mathrm{TE}]=0$	-0.0073	0.0026	-2.79	0.0268*
$[-\mathrm{CF},+\mathrm{TE}]-[+\mathrm{CF},+\mathrm{TE}]=0$	-0.0003	0.0024	-0.12	0.9993
$[-\mathrm{CF},+\mathrm{TE}]-[-\mathrm{CF},-\mathrm{TE}]=0$	0.0004	0.0024	0.16	0.9986

Irrelevant comparisons (i.e. congruent–congruent, incongruent–incongruent) removed.

Table 5.11: Pairwise comparisons between infelicitous and felicitous conditions for second-pass dwell time on antecedent focus words

Comparison	Estimate	SE	z	p
$[-\mathrm{CF},+\mathrm{TE}]-[+\mathrm{CF},+\mathrm{TE}]=0$	-0.0003	0.0024	-0.123	0.9993
$[-\mathrm{CF},+\mathrm{TE}]-[+\mathrm{CF},-\mathrm{TE}]=0$	0.0076	0.0025	3.108	0.01*
$[-\mathrm{CF},+\mathrm{TE}]-[-\mathrm{CF},-\mathrm{TE}]=0$	0.0004	0.0024	0.159	0.9986

Irrelevant comparisons (i.e. felicitous – felicitous) removed.

5.5.6 Testing the secondary hypothesis: effects of felicity

The secondary hypothesis predicted that sentences containing an infelicitous target word would incur a processing cost in comparison to those containing felicitous target words. An infelicitous target word was defined as one that was not contrastive but was typographically emphasised, i.e. $[-\mathrm{CF},+\mathrm{TE}]$.

As with congruence, an effect of felicity would appear as an interaction between CF and TE—meaning that again the only significant result relevant to felicity was that discussed above, on second-pass dwell time on antecedent focus words. Table 5.11 presents the comparisons relevant to testing effects of felicity.

Only one of the three relevant comparisons returned a significant result: in trials in which the target word was infelicitous, second passes were longer on antecedent focus words than in trials in which the target word had CF but no TE (i.e. the incongruent, $[+\mathrm{CF},-\mathrm{TE}]$ condition). This is insufficient evidence to support the secondary hypothesis.

5.5.7 Investigating 'eye-catching' effects

There were no of significant effects on first-pass dwell times on any AOI. This suggests that the potential issue raised in §5.3.3, of antecedent focus words being fixated prior to target words, was not occurring. It also fails to support the hypothesis discussed in that section that there is a 'visual salience' effect of TE

Table 5.12: Words skipped by factor

	[−TE]	[+TE]	Total
[−CF]	3.04%	1.07%	4.11%
[+CF]	2.32%	1.43%	3.75%
Total	5.36%	2.5%	7.86%

which results in longer initial fixations: the results from Experiment 1 had also failed to support such a hypothesis.

However, further exploration of the data suggested there may have been a different type of visual salience effect at play. Table 5.12 shows how many of the total trials presented were skipped for each factor. Of the 7.86% of the total trials presented in which the target word was skipped, it was notable that participants were more than twice as likely to skip a [−TE] target word (5.36% of trials) than to skip a [+TE] target word (2.5%).

To investigate this (with the acknowledgement that this is exploratory analysis, supplemental to the hypothesis testing presented above), a logistic generalised linear mixed effects model using the binomial distribution (with a logit link function) was fitted to a dummy variable representing whether a target word was skipped: any trial with a missing first-pass dwell time for the target word was treated as having been skipped. The bottom-up model building approach described in §5.5.4 was used: TE and CF were included as fixed effects (with an interaction), and item and subject were included as random effects. The results of this model are shown in Table 5.13. A main effect of TE was found on the probability of target words being skipped (estimate $= -0.428$, $SE = 0.1189$, $z = -3.597$, $p < 0.001$).

Table 5.13: Generalised linear mixed effects model results for skipped target words

Coefficient	Estimate	SE	z	p	Random effects	Variance
(Intercept)	−2.841	0.2488	−11.421	< 0.001*	Item	0.403
CF	−0.003	0.1186	−0.024	0.98	Subject	0.492
TE	−0.428	0.1189	−3.597	< 0.001*		
CF:TE	0.152	0.1187	1.282	0.2		

5.5.8   Other effects

5.5.8.1   Typographic emphasis

A main effect of TE was found on target words, where it was associated with a longer total dwell time (estimate $= -0.0036$, $SE = 0.0013$, $t = -2.85$, $p < 0.01$).

5.5.8.2   Contrastive focus

Specific predictions were not made here regarding main effects of CF, but it is included here for the sake of completeness, and to facilitate further discussion regarding the manipulation of this factor in the following chapter.

A main effect of CF was found on antecedent focus words, where it was associated with a longer second-pass dwell time (estimate $= -0.0073$, $SE = 0.0026$, $t = -2.79$, $p < 0.01$).

5.5.8.3   Spill-over processing

No effects for any factor were found on the post-target words. As was the case in Experiment 1, most of these words were short, high-frequency, grammatical words: such words are rarely directly fixated (Rayner, 1998). This was a consequence of the constructions used to control the position and context of the target word. Because of this, the lack of effects found within the metrics on these words does not necessarily point towards a lack of spill-over processing. However, post-hoc analysis of the same metrics combined over two and three of the words following the target word did not reveal any effects either, which suggests that there was in fact no spill-over occurring.

5.5.9   Analysis of participants

The mean accuracy for comprehension check questions across all items was 95% ($SD = 4\%$), and the mean accuracy for questions on test trials alone was 96% ($SD = 8\%$). All participants thus appeared to have been reading for comprehension throughout, and no data were discarded on the basis of inattention.

As was done in the first experiment, informal debriefing sessions were conducted after each run, in which participants were asked if they had noticed repeating elements in the stimuli. Many mentioned bold text first, and when prompted recalled italic text, but as before none linked this to contrast. Furthermore, very few remarked on the contrast manipulation, although some mentioned that some of the sentences were odd, presumably in a pragmatic way. Additionally, several
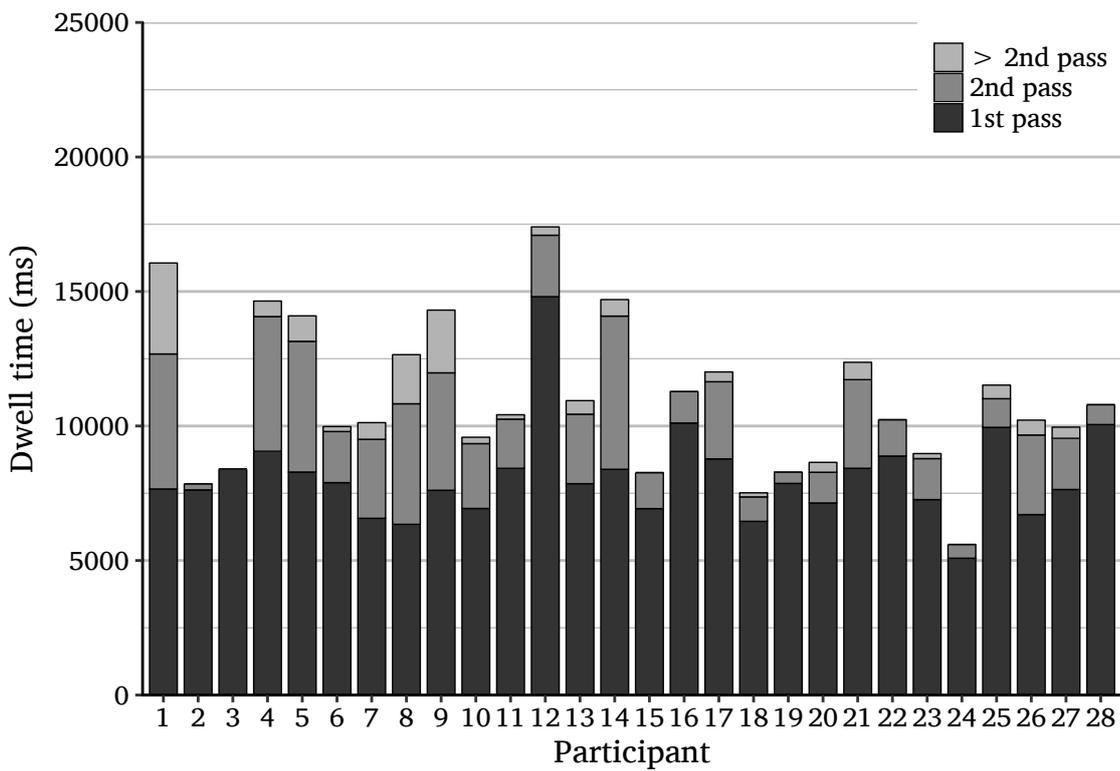
Figure 5.8: Summed dwell times on target words by participant
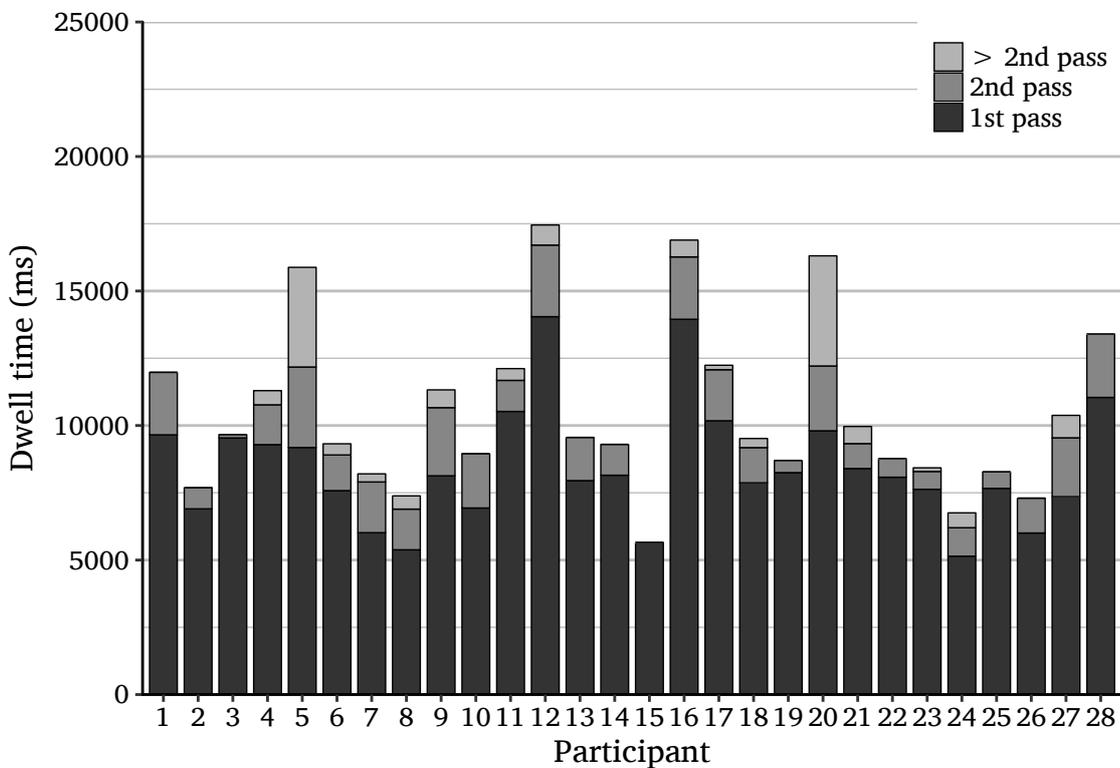
Figure 5.9: Summed dwell times on antecedent focus words by participant
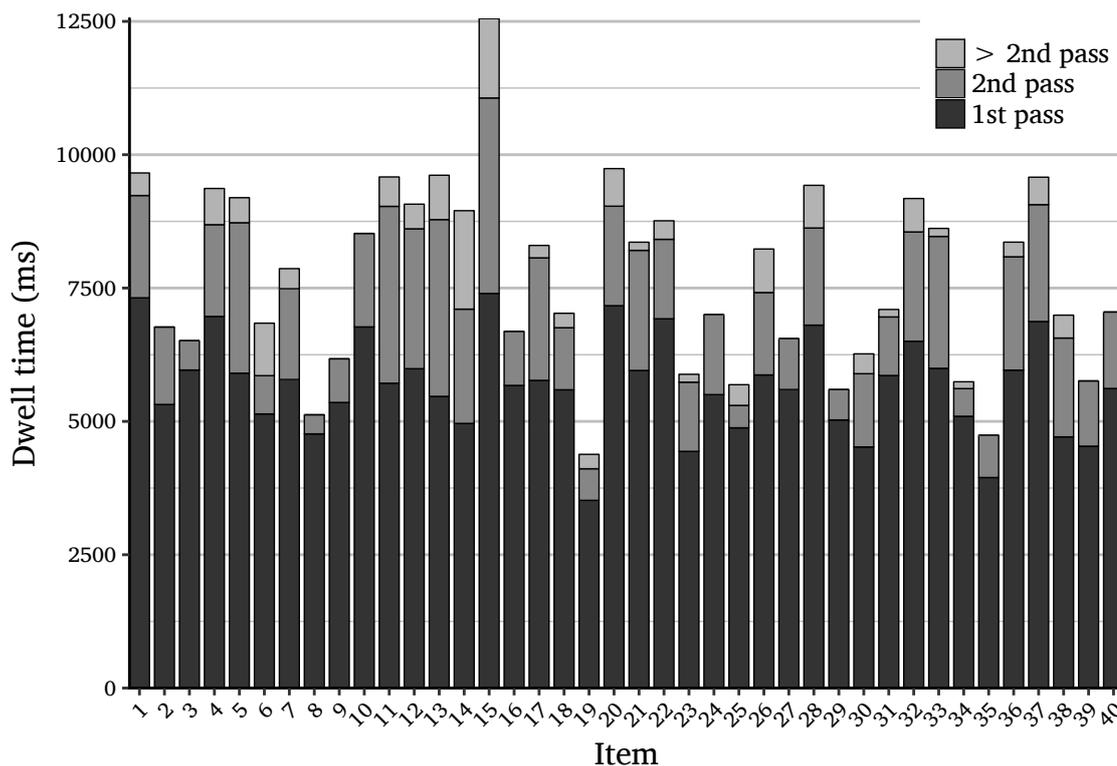
Figure 5.10: Summed n-pass dwell times on target words by item

participants remarked on the relatively high number of names used in the stimuli (see note on use of names, 5.4.2.2).

Reading behaviour was varied across participants, as was expected. Particularly notable in Figure 5.8 is Participant 12, who had considerably longer cumulative first-pass dwell times than any other participant. Similarly, Participant 24 was much more rapid overall than the others. Several participants barely made second passes; Participant 3 made none at all. On antecedent focus words (Figure 5.9), it is notable that Participant 12 again had a high first-pass time, as did Participant 16. Participant 24 was again very quick overall.

The LMMs included subjects as a random factor, and the results tended to show that they accounted for larger amounts of variance than item did in those models in which it was included. This was in keeping with expectations (see footnote 2, p.153).

5.5.10 Analysis of items

Target words were more strictly controlled than those in Experiment 1. It appears that the high level of control was successful in reducing variability of responses; although Figure 5.10 clearly shows that there were a variety of dwell times on

different target words, when compared to the corresponding figure for the items from Experiment 1 (Figure 4.10 in the previous chapter), the dwell times per item in Experiment 2 have a much tighter distribution. Indeed, during the construction of the LMMs, items in general appeared to account for less variance than participants, and item was frequently excluded as a random effect because it made no significant difference to the model. This suggests that the design of the items was sufficiently consistent.

5.5.11   Summary of results

- The primary hypothesis was not supported: incongruent target words did not prompt more re-reading, although one type of incongruence was associated with a longer second-pass dwell time on antecedent focus words.
- The secondary hypothesis was not supported: infelicitous target words did not prompt more re-reading.
- TE did not have an 'eye-catching' effect in terms of causing longer first-pass dwell times.
- TE was associated with a higher likelihood of fixations on target words.
- TE was associated with longer total dwell times on target words.
- CF was associated with a longer second-pass dwell time on antecedent focus words.
- No spill-over effects were found.

5.6   Discussion

This experiment manipulated CF and TE on target words within narrative texts, to investigate whether TE has an effect on processing, and whether this effect is caused by it marking CF, or by something else (such as its visual salience).

Congruence between CF and TE was investigated here in order to test the primary hypothesis, which predicted that incongruence would prompt more re-reading than congruence. Fewer effects were found for congruence than were expected. One type of incongruence, $[+\mathrm{CF}, -\mathrm{TE}]$, was found to be associated with a longer second-pass dwell time on antecedent focus words in comparison to when CF and TE were congruent. This was insufficient to support the hypothesis, which required a) the existence of effects on more than one metric across more than one AOI, and b) both incongruent conditions, $[+\mathrm{CF}, -\mathrm{TE}]$ *and* $[-\mathrm{CF}, +\mathrm{TE}]$, to differ significantly from the two congruent conditions.

The lack of significant effects of congruence may indicate that the primary hypothesis was too strong, and that the factors interact in a more subtle way than expected. Looking at the pairwise comparisons for the one place in the data where there was an interaction between CF and TE, second-pass dwell time on antecedent focus words (Table 5.10), the results seem to support this idea of a more complex interplay.

As was stated above (§5.3), a lack of support for this hypothesis might imply either that TE is able to 'force' a contrastive interpretation of a word despite its discourse context, or that TE is not reliably interpreted as marking contrastivity. No post-hoc testing of these results would be able to test either of these hypotheses, as there was no measure of whether or not the participant actually took the target word as contrastive or not.

Felicity was investigated here in order to test the secondary hypothesis, which predicted that infelicity would prompt more re-reading than felicity, where the single infelicitous condition is $[-\mathrm{CF}, +\mathrm{TE}]$. As with congruence, fewer effects were found than were expected: given that effects of this metafactor would also only manifest as an interaction between CF and TE, the single relevant effect was again on second-pass dwell time on antecedent focus words. Again, this was insufficient to support the hypothesis. The relevant pairwise comparisons (Table 5.11) showed that the only significant difference was between the infelicitous condition and the incongruent condition $[+\mathrm{CF}, -\mathrm{TE}]$; other than that, it appeared that the infelicity was not associated with a processing cost.

A lack of significant effects on first-pass dwell time on target words suggests a lack of support for an 'eye-catching' effect of TE; such an effect had been the subject of the secondary hypothesis in Experiment 1, which had predicted longer fixations on $[+\mathrm{TE}]$ words because of its visual salience. The results of Experiment 1 had not supported this, nor had those of Experiment 2. However, exploratory analysis of the Experiment 2 results did reveal an interesting effect of TE that could be seen as a different type of 'eye-catching' effect: the presence of TE made it more likely that participants would fixate a target word. If the TE prompted participants to fixate the target word, this suggests that the presence of TE is detectable without direct fixations (i.e. by the *preview effect*— Rayner et al., 1989; Schotter et al., 2012; see §4.4).

Regarding participants: the analysis above (§5.5.9) revealed high variability across participants, as was expected; this is likely to be a crucial factor in the lack of clear results. Whilst including participant as a random intercept enabled the LMMs to 'parcel out' variance due to different participant behaviour, highly

variable strategies for re-reading would still reflect across different metrics in different ways. The modelling approach used here, in which separate models were constructed for each metric, may have been unable to uncover such divergent behaviour.

Chapter 6

General discussion

6.1 Introduction

The present study set out to investigate whether typographic emphasis (TE) can be interpreted as marking contrastive focus (CF) during reading. Two eye tracking experiments were conducted to test the central hypothesis that **a sentence with TE on a CF word would be processed faster than an equivalent sentence without TE on that word**. Experiment 1 (Chapter 3) was a small pilot ($N = 10$), and Experiment 2 (Chapter 4) was a larger, revised version of this ($N = 28$). Both experiments relied on the following assumptions:

1. A written sentence containing a word that should be contrastively focused in accordance with the discourse context, yet has no marking, will be more difficult to process than an identical sentence with that word marked by TE.
2. This processing difficulty will manifest as increased re-reading across the text, with particular re-reading of the antecedent focus word with which the CF contrasts, and of the CF word itself.

Both experiments presented participants with 40 short texts that were manipulated between participants so as to either have a target word in italics for the [+ TE] condition, or in the same roman style as the surrounding words for the [− TE] condition. Experiment 2 expanded upon the pilot by additionally manipulating whether the target word had CF. This was achieved by the alteration of an antecedent focus word so that it had different hyponymy in relation to the target word: in the [− CF] condition it was a hyponym of the target word, and in the [+ CF] condition it was a co-hyponym of the target word. Combining the different levels of these factors created four conditions, which enabled Experiment 3 to test whether TE only had an effect when placed on CF words, or whether it had an

effect on words which had presentational focus (PF) rather than CF. The analysis of both experiments employed a combination of AOIs and eye tracking metrics intended to capture the predicted re-reading behaviour within the resulting eye movement recordings. Results from Experiment 1 were analysed principally with Wilcoxon signed-rank tests, and results from Experiment 2 were analysed with mixed effects models.

The results did not support the central hypothesis stated above, as differences were not found for all metrics between the sentences that contained TE and those that did not—and where differences were found, they did not sufficiently align with the specific predictions made. Instead, a complex picture was presented, which suggests that TE does have effects on the eye-movement record, and may interact with CF in some circumstances.

6.1.1 Organisation of chapter

§6.2 gives a **summary of both experiments**, including a comparison between them (§6.2.3). After this, discussion centres on Experiment 2, which was superior in several ways, including the statistical power afforded by the larger sample, the analytical approach adopted, and the deeper level of insight afforded by the additional manipulation of CF on the target word.

§6.3 presents an **interpretation of the results** in the form of a series of questions which pertain to the hypotheses tested by Experiment 2, and to issues that arose during analysis.

§6.4 examines the **limitations** of this study, and §6.5 suggests possible **future directions**.

6.2 Summary of results

6.2.1 Experiment 1

The first experiment tested two hypotheses: the primary hypothesis predicted that more re-reading would occur in the [−TE] condition than in the [+TE] condition. This assumed that [−TE] target words would first be incorrectly analysed as not having CF, and then a reanalysis would occur. This hypothesis was not supported. In the [+TE] condition participants spent longer in total fixating antecedent focus words (measured by the *total dwell time* metric), which was contrary to predictions: participants were expected to return to the antecedent focus

word for less time in this condition, which would have been reflected by a *shorter* total dwell time.

Similarly, in the [+TE] condition, participants returned to the word immediately following the target word more times (measured by the *pass count* metric); a change in pass count such as this was not predicted for this word, which had been analysed to capture 'spill-over' processing that may have been associated with the target word itself, and would have manifested in a different metric on this post-target word. As discussed in §4.8, this may actually have reflected increased passes on the target word itself, due to accuracy issues: either way, this increased attention in the [+TE] condition was the opposite of what was expected.

The secondary hypothesis predicted that participants would spend longer fixating [+TE] target words on their first visit to them (measured by the *first-pass dwell time* metric). This was based on the assumption that the visual salience of the [+TE] word would prompt early attention to it. This hypothesis was not supported: no significant effects were found here at all.

The results of Experiment 1 must be viewed with some caveats, the most important being that, as a pilot study run on 10 participants, the statistical power was low. Low power may have been compounded by the fact that the size of the effects was expected to be low: as a consequence, the probability of making a Type II error was relatively high. With eye tracking analyses such as this, which use metrics calculated relative to AOIs, some metrics may be calculated using far fewer samples than others: for instance, in Experiment 1 second passes on the target words only happened in 19% of trials, meaning that tests on the second-pass dwell time metric were considerably less powerful than those on first-pass dwell time, as first passes happened in 88% of trials.

Another issue arising from this being a pilot study was the presence of some accuracy problems, likely to have been caused due to inexperience in using the eye tracker system. This may have affected the quality of the eye movement recordings. Finally, it should be noted that the analytical approach adopted was different to that of Experiment 2: some trials were excluded from analysis where there were no fixations to the AOIs, but in the later experiment it was decided that the possibility of parafoveal processing of these AOIs meant that it was wiser to include those trials.

6.2.2 Experiment 2

The second experiment tested two hypotheses using the conditions [+CF, +TE], [+CF, −TE], [−CF, +TE] and [−CF, −TE]. The primary hypothesis was that 'in-

congruence' between CF and TE on a target word would result in more re-reading than when the two factors were 'congruent'. Congruent conditions were defined as those in which the two binary factors had the same polarity: $[+CF, -TE]$ and $[-CF, +TE]$ were incongruent, and $[+CF, +TE]$ and $[-CF, -TE]$ were congruent.

This hypothesis was not supported. Incongruent target words were not associated with increased re-reading, with the exception of significant differences between the incongruent $[+CF, -TE]$ condition and the two congruent conditions. These differences only occurred within second-pass dwell time on antecedent focus words; whilst these differences were in the expected direction (incongruence was associated with longer fixations), support for the hypothesis would require evidence of re-reading to manifest on more AOIs and across more metrics. Moreover, it would also require differences between *both* types of incongruent condition and the congruent conditions.

The secondary hypothesis was that the infelicity created by placing TE on a non-CF word would result in more re-reading in comparison to all other combinations of CF and TE, which were felicitous. This was not supported. The same metric on the same AOI, second-pass dwell time on antecedent focus words, was the only place in which the relevant interaction showed a significant difference, and this was only between the infelicitous condition and one type of felicitous condition, $[+CF, -TE]$.

This experiment also provided further evidence against the 'eye-catching' hypothesis tested in Experiment 1, as there was a similar lack of significant effects for first-pass dwell times on all AOIs. However, the presence of TE on target words was found to increase the likelihood of that word being fixated—this may constitute a different type of 'eye-catching'.

Two other effects were found which, whilst not directly pertinent to the hypotheses being tested, may be worthy of consideration: a main effect of TE was found on target words, where it was associated with an increase in total dwell time, and a main effect of CF was found for antecedent focus words, where it was associated with an increase in second-pass dwell time.

6.2.3   Comparisons between results

In attempting to draw comparisons between the two sets of results, it is important to acknowledge that the two experiments were very different. To recap some of the key differences in the designs:

- The text types of the stimuli: Experiment 1 used dialogues, Experiment 2 used short narrative-style texts;
- The level of contrast on the target word: all target words were contrastive in Experiment 1, whereas target words alternated between having presentational and corrective CF in Experiment 2;
- The filler/distractor structure: this was complex and frequently similar to test items in Experiment 1, but more varied and distinct from test items in Experiment 2.

The 10 participant sample tested by Experiment 1 was more homogeneous in terms of age and study experience than the 28 participant sample of Experiment 2. Experiment 2 benefited from a more rigorous and accurate approach to eye tracker setup and experiment programming (which was a direct result from experience gained during Experiment 1). In short, the results cannot be compared directly. Of course, there is no equivalent in Experiment 1 to those aspects of Experiment 2 that investigate CF as a separate factor; however, even comparing the results for TE alone needs to be done with the above differences kept in mind.

With all this said, the two experiments did both manipulate the presence of TE on a target word in the same manner, and both also had an antecedent focus word that was broadly similar in position and context. In terms of significant results, the one place in which they both closely coincided was in finding that TE was associated with a longer second-pass dwell time on antecedent focus words.

Despite the other individual tests conducted for each experiment returning different results, the only significant differences found across both experiments showed TE associated with *more* re-reading rather than *less* re-reading: i.e. increases in total dwell time (Experiment 2 on target words), pass counts (Experiment 1 on post-target words), and likelihood of fixation (Experiment 2 on target words).

This increase in re-reading is contrary to predictions, and may indicate that the presence of TE simply prompts greater attention. Despite this going against the predictions, these results are important, in that they provide support to the basic assumption here that TE has an effect on eye movements. The exact pattern in which this re-reading manifested may have been different in each experiment due to differences between the materials.

Neither experiment found any significant effects of TE on first-pass dwell times, on either the target word or antecedent focus word. A significant effect on the latter would have been highly unusual, as the antecedent focus word should have been first visited well before the target word, which was where the TE ma-

nipulation was realised. The lack of effects on first-pass dwell time on the target word is notable in that it may suggest the visual salience of [+TE] words is not what drives the longer total dwell times: this will be further discussed below (§6.3.3).

6.3 Interpretation of the results

This section will take the form of a series of questions. They encompass both those questions that this study set out to answer, and questions that arose from analysis of the results. Note that, from this point on, the focus will fall chiefly on the results of Experiment 2.

6.3.1 Did incongruence result in more re-reading?

Insufficient support was found for the primary hypothesis that incongruence between CF and TE would result in more re-reading than congruence between the factors. This was tested by constructing linear mixed effects models in which CF and TE were included as fixed effects with an interaction: significant interactions would point towards possible support for the hypothesis, and the presence of significant effects found within pairwise comparisons between the different conditions would confirm this support.

The congruent level consisted of [−CF, −TE] and [+CF, +TE], and the incongruent level consisted of [+CF, −TE] and [−CF, +TE]. In the incongruent condition, an interaction was found for second-pass dwell time on the antecedent focus word: pairwise comparisons showed that fixations were longer when comparing the incongruent condition [+CF, −TE] to both congruent conditions.

To recap the rationale behind this hypothesis: Incongruent stimuli were predicted to slow processing, because in the incongruent conditions the processor would be forced to reanalyse an initial, incorrect analysis of the IS of the target sentence. In the [+CF, −TE] condition the difficulty would be caused by the lack of explicit marking of the CF (recall that no other devices, such as clefts, were used to mark CF), which would result in an initial analysis of the focus as presentational. Such an analysis would be acceptable on the sentence level, but would be infelicitous on the discourse level, as the antecedent context meant that at this point there was no reason for a non-contrastive proposition to be made about the referent of the target word. The expected behaviour was for a reanalysis to be performed to resolve this. In the [−CF, +TE] condition, the difficulty would be

caused by the converse situation, in which the initial analysis would take the TE to mark CF, but context would prohibit this the presence of CF at that point. A reanalysis was thus expected here too.

The single significant difference between $[+CF, -TE]$ and the two congruent conditions on second-pass dwell time on antecedent focus words is insufficient support for the congruence hypothesis, which required evidence of a similar interaction on a greater number of metrics and AOIs, and for those interactions to show *both* incongruent conditions differing significantly from the congruent ones.

Looking at the antecedent focus words, it is slightly surprising that there was no significant effect on total dwell time on them in addition to the effect on second-pass dwell time, given that total dwell time must always be correlated with second-pass dwell time. The strength of the correlation between total dwell time and any given individual pass time logically depends on the number of passes and their durations; considering that there were rarely more than two passes, this might suggest that the impact of the significant effect on second-pass time was 'masked' by greater variability within the first pass. This seems to have been the case: first-pass dwell time here had a substantially larger standard deviation (163 ms) than the second-pass dwell time had (91 ms). Greater variability of first-pass dwell times may be due to differing lexical access times (Rastle, 2007; Staub and Rayner, 2007), caused by variability in the uncontrolled properties of the antecedent focus words, such as their frequencies and lengths (see Table 5.2 for some of these properties).

6.3.2 Did infelicity result in more re-reading?

Insufficient support was found for the hypothesis that infelicity created by placing TE on a target word that did not have CF would cause processing difficulty.

Support for this hypothesis was, as with the primary 'congruence' hypothesis, expected to manifest as an interaction between CF and TE, meaning that the single interaction found on second-pass dwell time on antecedent focus words in the course of investigating congruence effects was also relevant here. Of the three comparisons that could be made between the infelicitous condition and the other conditions, only the one between the infelicitous $[-CF, +TE]$ and the felicitous $[+CF, +TE]$ was found to have a significant difference: trials with infelicitous target words were associated with a longer second-pass dwell time on antecedent focus words in comparison to those with target words that had both CF and TF.

Infelicity appears not to have prompted more re-reading. However, the single significant interaction which was relevant to both the main hypotheses should not

be discounted totally. As an interaction between a visual variable and a linguistic one, it may still be a small piece of evidence for one of the central assumptions of this study: that TE can have a linguistic effect.

6.3.3 Does typographic emphasis have a visual salience effect?

The labelling of the different conditions as being either 'congruent' or 'incongruent', and 'felicitous' or 'infelicitous', was based on several assumptions regarding discourse and focus processing. If those assumptions are false, then the primary and secondary hypotheses are invalid as they have been stated here, as they rest upon the validity of these congruence and felicity metafactors. Moreover, the validity of the manipulation used for the CF factor itself may also be questioned. Both of these issues are discussed below; see §6.4. However, the TE factor was inarguably a valid manipulation of the presence of TE: either the target word was displayed in italics for [+ TE], or it was displayed in the same roman style as the rest of the stimulus for [− TE].

Alone, the TE factor represents a visual variable, as opposed to the linguistic variable of CF. Analysing any main effects of TE aside from interactions with CF must be less useful, then, in terms of finding support for the central hypothesis that TE-marking of CF can facilitate processing. This is why CF was added as a factor to Experiment 2. However, it is still worthwhile to look at the two main effects of TE that were found—especially because it is important to consider whether this visual variable had a purely visual, non-linguistic effect.

There was a main effect of TE on target words, where it was associated with a longer total dwell time. Additionally, during the exploratory analysis of target word skipping, it was found that that [+ TE] target words were more likely to be fixated than [− TE] words. This shows that, regardless of other factors, placing TE on a focused noun resulted in it receiving more attention.[1]

This increased attention might be related to TE triggering a different linguistic interpretation, but it could also simply arise from the higher visual salience of TE in comparison to the surrounding text. In other words, TE might have what will be referred to here as a *visual salience-only* effect, rather than a linguistic effect.

Support for TE being interpreted linguistically may come from the fact that it had an effect on total dwell time, but no effect on first-pass dwell time. If there

---

[1] It is worth recalling the results of Experiment 1 at this point; they also pointed towards TE being associated with greater attention (on both antecedent focus words and target words).

were a visual salience-only effect, that effect might be expected to happen early in the eye movement record, and thus be reflected by 'early' metrics such as first-pass dwell time. The visual salience of TE must be perceptible very early on when a [+TE] word is fixated, because the entire reading process relies on the rapid identification of visual detail in order for graphemes to be recognised (Balota et al., 2006), and TE is achieved through the visual alteration of letter forms. The effect of TE on the probability of fixating the target word also suggests that TE is identified early on, as it shows that the presence or absence of TE is detectable parafoveally (this effect will be returned to shortly). It thus seems likely that the processor is able to register when a word has TE before or during the first pass on that word. This may suggest that the association between TE and an increase in the late metric of total dwell time reflects processing unrelated to its visual salience.

On this point, it is also worth noting that the literature strongly associates longer fixations that occur later in the eye movement record with linguistic processes (Rayner, 1998; Drieghe et al., 2005; Clifton et al., 2007)—yet caution must be taken in looking to this previous work: the vast majority of the research that has established connections between eye movements and linguistic processing has done so using linguistic dependent variables alone, without the complicating aspect of a visually salient dependent variable like TE.

Whilst it is possible that the increased total dwell time constitutes evidence for TE having a linguistic processing effect, the increased fixation likelihood for [+TE] target words might seem to be a clear visual salience-only effect, whereby the TE made the word more 'eye-catching'. Word skipping is an interesting case, however: whilst saccades are generally accepted as being driven by low-level visual factors (Drieghe et al., 2005), skipping probability is also affected by linguistic factors. Although the best predictor of whether a word will be skipped is the visual factor of word length (i.e. shorter words are skipped more frequently), words are also more likely to be skipped when they are highly predictable, and when they have a high frequency; (Rayner, 1998; Brysbaert et al., 2005; Drieghe et al., 2005). This particular result thus does not cast much light on whether [+TE] add linguistic meaning to target words or not.

A different type of visual salience-only effect might be one whereby the gaze is attracted to a word out of sequence—an eye-catching effect extreme enough to be 'disruptive' to the normal course of reading, causing words to be fixated out of sequence to an unusual extent. Linguistic information from unattended lines of text has been shown to be unavailable (Inhoff and Topolski, 1992; Pollatsek et al.,

1993), but it is obviously possible for reading to be disrupted or interrupted by some form of distraction, or by deliberate conscious control. Reading can be disrupted by some types of visual salience; for example advertising images displayed adjacent to text have been shown to disrupt reading (e.g. Simola et al., 2011, although this study involved animated adverts). The analysis here did not include a measure capturing the order in which AOIs were first fixated, and thus cannot fully address this. However, if this type of disruptive effect occurred, causing the target word to be fixated out sequence, and if no other processing was prompted by [+TE], then the target word would be primed; when it was returned to, it would receive shorter fixations, rather than the longer fixations suggested by the increased total dwell time that were found here. The existence of a disruptive visual salience-only effect here thus seems unlikely.

Examining the main effects of TE in isolation would seems to provide inconclusive evidence for TE impacting linguistic processing. As has been suggested here, the increased total dwell time on target words may reflect linguistic processing—but as with the results as a whole, this is an isolated significant result standing against an array of null results from other tests.

6.3.4 Does contrastive focus affect eye movements differently to presentational focus?

Whilst effects of CF alone have not been foregrounded in this study, it is nonetheless interesting to note the lack of significant main effects found for this factor, particularly in light of gaps in the current literature. The presence of CF on the target word was manipulated by altering the hyponymy of the antecedent focus word (see §5.4.2.2). Crucially, though, the target word always had focus; the intended alternation was between CF and PF. A single main effect of CF was found on antecedent focus words, where it was associated with a shorter second-pass dwell time.

The lack of previous eye tracking research on the reading of CF is particularly unfortunate, as knowledge of eye movement differences between CF and PF would be useful in assessing whether the manipulation of CF used here was successful (see §6.4.3.1). However, there is some reason to believe that CF is associated with an increased processing load, which might manifest as longer fixations on those words: this comes from Cowles' ERP study on focus processing (2003); in particular, from her consideration of why her ERP experiment found a P600 being elicited by CF target words and not be PF words (this was discussed

in more detail in §2.6.2).

Cowles suggests the P600 component may have related to either the participant needing to match the target word to one of the alternative set, or to the extra information conveyed by the contrastive focus: i.e. it does not only specify the role of the target word, but unlike PF, it also specifies that the other members of the alternative set are to be excluded. Whilst Cowles' neuroimaging results suggest that CF has a different effect on the processor than PF, they do not provide a strong basis for making suggestions about how (or even whether) eye movements are affected by this difference. However, this view would seem to tally with the sole main effect found of CF in Experiment 2, which was an increase in second-pass dwell time on antecedent focus words.

Nonetheless, there was still a relative lack of effects of CF; as with the other results for TE and for the metafactors congruence and felicity, it is difficult to accept such isolated results as being convincing evidence for any given hypothesis. Assuming the CF manipulation utilised in the present study was successful, there are two possible explanations for the lack of main effects of CF:

1. CF affects eye movements differently to PF, but the effect it has was not fully captured by the metrics/AOI examined here;
2. CF and PF both have the same effects on eye movements as each other.

If the first possibility is true, and a difference between the two focus types *can* be captured by the eye movement record, it is unclear where this would manifest, if it was not captured by the approach taken in this study: aside from the target word and antecedent word, there were no significant effects of CF on post-target words (as with all other factors), so a spill-over effect can be discounted. Furthermore, in case the post-target AOIs represented words too small and functional in nature to attract many fixations anyway, post-hoc analysis of all metrics on three words after the target was conducted as well; no significant effects were found (see Chapter 5).

The second possibility, that CF and PF have the same effects on eye movements, seems plausible, especially if viewed from the perspective that has been adopted here which sees the form of CF as differing only gradually from PF, and the function of CF being more pragmatic than semantic (Zimmermann, 2007; Calhoun, 2009).

In summary, examining the main effects of CF alone is inconclusive, but may provide some support to the idea that eye movements do not differ significantly during reading of presentationally or contrastively focused words. This experiment was not designed to test this, however, and these results alone do not yet

lend strong support to such a hypothesis. Instead, this should be taken as a potentially interesting result, which might contribute to future research into this issue.

6.3.5   Why were there no effects on pass counts?

No effects on pass counts on either word were found for any factor. The number of passes tended to be lower than was anticipated, across the board; for example, only 26% of the target words fixated received more than one pass, and only 5% more than two passes.

Pass count was included as an attempt to bolster the validity of using second-pass and total dwell times as measures of re-reading, without requiring more complex scan-path analysis. This was an attempt at innovation: this metric is rarely if ever used in the literature in this manner. This may well be because large differences in pass counts are relatively rare, meaning that high statistical power is difficult to achieve.

6.3.6   Can typographic emphasis function in a similar way to intonational emphasis when marking contrastive focus?

This is, of course, the central research question that this study has investigated. As the above discussion suggests, the answer is 'it does not appear that TE functions in a similar way to PAs when marking CF.'

From the start, it has been evident that TE and PAs are considerably different beasts, and it has never been the contention here that TE would be processed in a similar way to PAs in general. The two devices are used very differently: the clearest example of this is that TE rarely, if ever, marks PF, whereas PAs always mark PF. However, both TE and PAs are associated with CF. The fact that PAs are associated with CF is uncontroversial (see Chapter 2), but the review of surveys of TE that was conducted here (see Chapter 3) shows that TE is also associated with CF. This study has set out, however, to investigate how the device is *received* rather than *produced*: this is what prompted its focus on processing.

In terms of processing, the results suggest that on a superficial level TE *is* similar to PAs in terms of them both being attended to during language reception (see Cutler et al., 1997 for attention to PAs): otherwise, there would have been no effects whatsoever of TE. However, this study has failed to find results as clear as those that came from such early studies into PA processing as Terken and Nooteboom (1987) and Bock and Mazzella (1983), where congruence between

information structure and PA marking had benefits, and incongruence caused problems. The interaction that was found between the TE and CF factors has not provided strong enough evidence to make any strong conclusions.

6.3.7   Why were the main hypotheses not supported?

Given the lack of support for it, it may be that the main hypothesis is false, and that TE simply does not result in facilitated (and thus faster) processing when it marks CF. Three possible explanations for why the hypothesis was not supported will now be considered: the 'zero impact' account is a strong interpretation, which holds that the underlying research question of this study is ill-founded, and that TE-marking of CF has no linguistic or cognitive effect whatsoever. The 'effects elsewhere' account is a more plausible approach, which suggests that the effects and benefits of the phenomenon exist in a form different to that expected here. Finally, an alternative view is presented which invokes 'good-enough' processing: this suggests that participants may have only processed the stimuli to an incomplete, 'good-enough' level, and were thus not affected by the experimental conditions in the same way as they would have been if they had been motivated to process them fully.

6.3.7.1   The 'zero impact' account

The 'zero impact' account holds that TE-marking of CF has no linguistic or cognitive effect. This is an extreme viewpoint which seems, at the very least, to be intuitively unlikely. Furthermore, it would be difficult to empirically support a hypothesis arising from this account, as it would be tantamount to proving a negative. However, it will be discussed here briefly, both for the sake of completeness, and as a background with which to contrast the alternative 'effects elsewhere' account that follows, as well as the suggestions in the subsequent *Limitations* section (§6.4).

There can be little doubt that TE *is* used to optionally mark CF in writing, as has been discussed and illustrated in Chapter 3. Given this, if the 'zero impact' account were true, the question would arise as to why TE-marking of CF occurs at all, and why it has been occurring in writing for such a long time (again see Chapter 3). A counter argument might be that the simple presence of phenomena in production does not logically imply that those phenomena serve a communicative purpose; some phenomena may be due to error, such as slips of the tongue or typos. However, it seems obvious that the vast majority of instances of TE-marking

of CF cannot be due to error, as they are too systematic. Moreover, writing, particularly the type of professional writing in which TE is more likely to appear, is planned and edited to a degree that is rare (if not impossible) for most instances of speech (Akinnaso, 1982).

Non-communicative phenomena in language production may also result from the limitations of the mechanisms of production. It is harder to find examples of this, however, as the by-products of limitations can be recruited for communicative purposes: for example, pauses for breath during speech are an inescapable consequence of the respiratory system, yet they can be used systematically to group phrases (Grosjean and Collins, 1979). The fact that it is hard to find other examples of a non-communicative by-product is itself a point against TE-marking of CF being merely a by-product itself. At any rate, it would seem particularly unlikely for such a complex phenomenon to exist without communicative purposes in a system as subtle (and artificial) as writing.

The strongest evidence the current study provides against the 'zero impact' account comes from the significant interaction found between CF and TE on second-pass dwell time on antecedent focus words. This is regardless of whether the hyponymy-manipulation approach use here succeeded in actually alternating 'contrastive focus' itself (see §6.4.3.1): the only appreciable aspects in which $[+CF]$ and $[-CF]$ differed were linguistic. Therefore, it seems unlikely that there is zero impact of TE, and much more likely that what impact it had was not fully predicted, and thus was not fully captured by the experimental design used here.

6.3.7.2 The 'effects elsewhere' account

If the effects of TE-marking of CF exist outside of where they were sought here, there are two possibilities. The first will be addressed in the *Limitations* section below (§6.4.2): it may be that the approach here of seeking re-reading effects was along the right lines, but the AOIs and metrics used were not sufficient to capture re-reading.

The second possibility will be addressed here: it may be that the effects caused by TE-marking of CF principally manifest either a) as other eye movement behaviours instead of re-reading effects, or b) outside of the eye movement record altogether.

The question has been raised above as to whether the results for felicity actually reflected re-reading: they may instead show a different type of eye movement effect. The longer 'late' fixations on the target word may indicate an increased processing load caused by the inappropriate/'spurious' presence of TE.

It had been assumed that processing cost here would be caused by a need to reanalyse the text, and it was anticipated that this reanalysis would be *overt*, meaning that it would leave a trace in the eye movement record: such reanalysis happens for syntactically ambiguous sentences, and can involve the eye returning accurately to alternative points of syntactic attachment (Lewis, 1998; Meseguer et al., 2002). In this design, the antecedent focus word was expected to be similarly useful in resolving infelicity or incongruence.

A process similar to Frazier and Rayner's *selective reanalysis* (1982) was anticipated (although it should be acknowledged that this theory relates to syntax and not higher-level processing); this theory predicts that syntactic reanalysis will prompt regressive eye movements as the parser tries to find alternative attachment points. Selective Reanalysis has found support from experiments showing that regressions are made as part of an 'intelligent repair process' (von der Malsburg and Vasishth, 2011, p.18), whereby the eye returns efficiently to the points in the prior text at which an alternative analysis could be made, instead of a systematic search serial search occurring back through the prior text.

However, it may be that reanalysis was required, but was achievable without need to return the eye to earlier parts of the text. When the processor relies on consultation of the short term memory, this type of reanalysis is referred to in the literature as 'covert' (Lewis, 1998). The processor must of course be able to perform covert reanalysis; for one thing, without it speech comprehension would be much more difficult. Lewis (1998) points out that both overt and covert reanalysis have a cost; either short-term memory storage for covert reanalysis, or time for overt reanalysis. Time tends to be far less of an issue during reading than it is during listening, and the experiment did not at any point encourage speed over accuracy. Accordingly, the expectation here was that the incongruent conditions would prompt overt reanalysis. Despite this, there may have been no need for it: if short-term memory storage was not sufficiently taxed, it may have been easy for the processor to conduct a covert reanalysis.

––––––––––––––––––––––

There are many ways in which the effects of TE-marking of CF could manifest outside of eye movements altogether—none of which, of course, could be revealed by the experimental approach that was taken here. Whilst eye tracking is an excellent tool with which to investigate online processing, it may be that TE has offline effects.

The only directly relevant psycholinguistic investigation is that of Fraundorf et al. (2013), which was reviewed in Chapter 2 (§3.5.2). They found that TE on

a CF target word improved recognition memory for explicit alternatives to that target word. Their experimental approach did not provide measures of online processing; it would be useful to know if this enhanced memory was associated with longer fixations on the target words.

6.3.7.3 No reanalysis: good-enough processing

An alternative conception of linguistic processing is the *good-enough* (or *shallow*, or *noncompositional*) approach advanced by Fernanda Ferreira et al. (Ferreira et al., 2002; Ferreira and Lowder, 2016), which differs sharply from traditional psycholinguistic approaches by dispensing with the assumption that the processor always generates 'representations of the linguistic input that are complete, detailed and accurate' (Ferreira et al., 2002, p.11).

Instead, a shallow processing approach may be adopted in which the processor does not fully process all linguistic input. A classic example frequently used in support of this is the fact that people accept the following 'semantic illusion' (Erickson and Mattson, 1981, p.540), answering 'two' to the question *How many animals of each kind did Moses take on the Ark?* The good-enough theory holds that the processor uses heuristics in order to analyse sentences efficiently; whilst the ways in which these heuristics might work are subject to debate, it is undeniable that under some circumstances sentences which are infelicitous can be accepted, and that this constitutes those sentences being misinterpreted as being felicitous (Ferreira and Patson, 2007).

Whilst the expectation here was for 'forced' reanalysis to occur so that the incongruent IS could be resolved, if good-enough processing was being used, participants may have simply accepted the sentence regardless. It is not possible to investigate whether this happened as there was no assessment of what participants comprehended (see §6.4.3.3).

On a related note, it is conceivable that, if participants had started off with a good-enough strategy which involved only superficial focus processing, but had then received comprehension checks which hinged upon whether the target word had been contrastive or not, then they might have learned to allocate more resources to processing the IS of the final sentence. Such checks were not used here in order to avoid drawing participants' attention specifically towards what was being investigated. More discussion of the limitations of the comprehension checks used will be provided below (§6.4.3.3).

6.4 Limitations

6.4.1 Level of statistical power

It may be that the statistical power achieved in Experiment 2 was not sufficient to uncover all the effects sought; this was clearly the case in Experiment 1, which used a small sample, and was possibly also affected by recording accuracy problems. Despite the larger sample in Experiment 2 and the better recording quality, it may be that effect sizes related to these phenomena are relatively small.

Throughout this study it has been predicted that the degree to which processing will be facilitated by TE marking of CF is likely to be smaller than the degree to which it is facilitated by pitch accent (PA) marking of focus and CF in speech. This position is prompted by the fact that many instances of CF in writing are not marked by TE, whereas instances of CF in speech are much more reliably marked by PAs. Due to this, and various related reasons (discussed in §3.4), the relationship between CF and TE has been predicted to be relatively weak.

The experimental design employed here relied upon the idea that unmarked CF would be particularly disruptive. Pains were taken to avoid including other methods of indicating CF such as clefting, and it was assumed that the [+CF, −TE] condition would be difficult to process due to the conflict between the lack of CF marking and the fact that the discourse context required the target word to have CF in order for its use to adhere to pragmatic maxims: in particular, if it was interpreted as non-contrastive, the use of the target word in the [+CF] condition would violate Grice's 'maxim of relation' [or relevance; Grice (1975); see §4.6.2, which holds for this aspect of the design of both experiments' materials]. However, it may be that the cost of violating this maxim was overestimated—perhaps because participants were able to accommodate the violation in some way, or perhaps because this type of infelicity does not have a large effect on online processing (recall also the discussion of 'good-enough processing' above; §6.3.7.3). Regardless of the reason, it may be that the minimally-marked CF did not pose much of a processing problem to the participants, thus reducing the intended contrasts and in turn reducing the power of the analysis.

If TE-marking of CF were to have effects, but they were extremely small, the question would arise as to whether it were an important phenomenon at all. If this were found to be the case for italic TE as has been used here, it would make sense to investigate whether other forms of TE might have greater effects. McAteer (1992) and Fraundorf et al. (2013) took initial steps in this area, looking at both italic and all-caps TE; however, far more work could be done in this area.

6.4.2   Appropriacy of metrics used

Given that the particular manifestation of processing difficulty caused by incongruence was predicted to be re-reading, accurately measuring and quantifying re-reading was crucial to the experimental design. As was mentioned above (§6.3.7.2), it is possible that re-reading behaviour was in fact prompted by incongruence but that the methods chosen to capture this were insufficient. There are two interrelated aspects to this; metric selection, and AOI selection.

To recap, the metrics used were first and second-pass dwell time, total dwell time, and pass count. Three AOIs were analysed: the target word, the post-target word, and the antecedent focus word (post-target words were included only to investigate spill-over processing on the target word, and were excluded from both analyses when no significant effects were found on them).

A major limitation of using dwell times and pass counts is that they lack temporal and ordinal information. Looking at them alone is not able to distinguish whether, for example, a second pass on the antecedent focus word occurred immediately after a single regressive fixation on the immediately preceding word, or if it happened after reading the entire rest of the stimuli up to and past the target word. As von der Malsburg et al. (2015, p.111) point out, 'it is possible that aggregated eye-tracking measures are misleading because they could in principle arise from a blend of several classes of fixation patterns'. In the literature it is relatively common for such measures to be allowed to stand alone, but the studies that use them do not always predict the type of re-reading pattern that had been expected here.

However, the tendency in these results has been for fewer significant effects than expected to be found on these metrics, rather than a high number of significant effects. Therefore the problem here may not be misinterpretation of dwell times as providing support for the hypothesis, i.e. a type I error, but a type II error caused by missing information. It might be that the re-reading patterns caused by incongruence were more complex than expected, and that only using two AOIs was insufficient to capture re-reading that may have involved two or more passes on many words.

This issue might be solved using scanpath analysis, mentioned in Chapter 4 (§4.6.5.1). As was discussed there, such an approach has relatively little precedent, which is unsurprising given the level of complexity involved. Von der Malsburg and Vasishth (2011) present a technique for the statistical comparison of scanpaths during re-reading, but as was mentioned before, this is a new method which has as yet gone untested by large amounts of studies, unlike the AOI ap-

proach that was adopted here. Nevertheless, even a qualitative approach involving the visual inspection of different scanpaths may be useful here (although this would require a considerable time investment in a study this size).

6.4.3 Issues with the materials

6.4.3.1 The contrastive focus manipulation

It should be acknowledged that the CF manipulation, intended to make the target word alternate between having corrective CF and PF, can be questioned. Chapter 5 discussed the fact that both Zimmermann (2007) and Calhoun (2009) take CF to only make sense in the context of a focus that is more marked than is anticipated. It was suggested (in §5.4.2.2) that this marking came in the form of the fact that the [+CF] target word formed part of a proposition which directly contradicted a previous proposition, and was therefore effectively 'marked' at the propositional level. The ice is thin here; in Chapter 5 it was acknowledged that it might be better to see what is referred to here as the [+CF] feature not as referring to the presence of CF itself, but as indicating a word that needed to be interpreted contrastively in order to make sense as part of the discourse.

Another, perhaps greater issue here is the possibility that [+CF] and [−CF] could possibly have been interpreted as both being corrective.

(50)   a. [+CF] Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure he couldn't eat peanuts. She spoke to a mutual friend, who said that if Nigel ate *almonds* he'd be very unwell.
        b. [−CF] Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure he couldn't eat nuts. She spoke to a mutual friend, who said that if Nigel ate *almonds* he'd be very unwell.

For example, in (50b), after Steph's belief that Nigel can't eat *nuts* is established, when the friend states that Nigel can't eat *almonds*, the target sentence might constitute the addition of information, which would mean the target word has PF, but it might instead be a correction itself. A different reading, however, could take the mutual friend as assuming that Steph thinks Nigel is allergic to *all* nuts, when in fact the friend believes Nigel is only allergic to one type of nut, almonds: the friend thus corrects Steph by rejecting the presupposition of an allergy to all nuts with the proposition that Nigel is allergic to a specific type of nut, *almonds*.

In some cases, then, the [−CF] condition may have been read as a correction regarding *specificity*, and [+CF] as a correction regarding *category*. This would

render what was treated as 'incongruent' as 'congruent', and what was treated as 'infelicitous' as felicitous.

Given this, it is unfortunate that the experiments as conducted were not able to measure whether participants had interpreted the target words as contrastive or not. Instead, the design rested on the assumption that participants would reanalyse the entire stimulus until they arrived at the 'correct' reading of the sentence containing the target word, or until they reached the conclusion that there was no resolution to the infelicity. As Ferreira and Lowder (2016) note, it is fairly common in psycholinguistic research for online processing studies such as this to rely on assumptions about participants' comprehension, rather than to attempt to measure it. That this is a common state of affairs is, of course, no reason to refrain from improving on it.

6.4.3.2 The infelicitous condition

An example item in the infelicitous condition is given here as (51).

(51) (Test item 13) A rehabilitation program was launched that would place ex-offenders in community service roles. Community leaders feared that criminals would be assigned to their neighbourhoods. They later found out that *burglars* made up the majority of the program.

The claim that $[-\text{CF}, +\text{TE}]$ was infelicitous was based on certain assumptions: chiefly that the appearance of TE on the target word was unmotivated by the context. Looking at (51), there is no apparent reason for the narrator to place emphasis on the target word *burglars* if TE is to be taken as marking CF. The word does not contrast with the antecedent focus word *criminals*; a burglar is quite clearly a type of criminal. The lack of motivation for the narrator to place TE on *burglars* arguably remains true even if many of the other possible uses of TE (discussed in §3.3) are considered: clearly *burglars* is not the title of a published piece of work, or a foreign word, for example.

However, in some other stimuli the target word appeared within sections of indirect reported speech (Coulmas, 1986), such as in Test item 11: 'Robert mentioned how *platinum* was the most attractive metal'. The TE on such target words could perhaps, at a stretch, have been interpreted as indicating a change in the speaker's voice quality. Despite this, even if a particular instance of TE in the stimuli could be taken as fulfilling a different function, within the wider context of the experiment it seems probable that the repeated use of italic TE on contrastive words elsewhere would cue participants that the sole use of italic TE within the

experimental stimuli was to indicate contrast. Moreover, the target words in less than half of the items appeared within reported speech (the precise figure is debatable, as in some items it is ambiguous whether the target word-containing clause constitutes reported speech or not).

6.4.3.3 The assessment of comprehension

Assessment of the comprehension of each test item would require an altered experimental design. Firstly, *all* test items would need to be followed by a comprehension check: 8 of the 40 test items had comprehension checks in this experiment. The downside to this would be a greatly increased chance of the participant being aware of the aims and manipulations of the experiment. An increased number of comprehension checks on filler/distractor items as well might have ameliorated this, at the cost of greatly increasing the duration of the task: these are the two main reasons why this was not done here.

Secondly, measuring comprehension regarding the contrastive status of the target word would require different types of question to those that were used here, which deliberately did not hinge on the participants 'correctly' interpreting the IS of the target words. Again, this was to prevent them gaining potentially problematic insight into the aims of the experiment.

It may be that the approach here was overly conservative: the attendant benefits of a design in which the comprehension checks did not function simply as motivating factors for participants, but instead evaluated whether they had read the target as being contrastive or not, may have outweighed the risks of learning effects. In any case, learning effects could be checked for during the analysis, and compensated for in the modelling process.

Another solution to this issue would be to conduct norming studies on the stimuli prior to running the eye tracking experiments: surveys could be used to establish how the stimuli were interpreted, and stimuli could then be tailored following the results in order to better achieve the desired CF manipulation.

6.4.3.4 Authorial voice

The selection of the 'genre' of the stimuli has been given considerable attention in this study. As mentioned above, major change was implemented in Experiment 2, in which the short dialogues used in Experiment 1 were substituted for short narrative-style texts. The reasons that were given for this change in Chapter 5 still stand (see §5.4.2.1), but some possible concerns can be raised over the narrative

format. Principally, there is the issue of authorial voice. Consider (52), which reproduces Test items 12 and 19 from Experiment 2. Both are in the $[+CF, +TE]$ condition.

(52)   a. (Test item 12) Victoria was helping her uncle move boxes of art equipment into his new studio. Victoria had presumed the majority of boxes would be full of paints. She ended up spending most of the day lugging boxes full of *brushes* up the stairs.

   b. (Test item 19) The pet shop assistant prided herself on figuring out what sort of product any given customer would want. That morning, she told a small boy approaching her where the hamsters were before he spoke. He then asked her if the *goldfish* were expensive.

In (52a) the target word appears within a third person description of a physical act; in (52b), it appears within reported speech. The question here is whether the TE is taken either as being a property of the narration, as if the narrator had included a parenthetical remark afterwards such as 'italics mine', or of the speech of the character themselves. Graesser et al. summarise this issue usefully:

> There are multiple levels of dialogue to worry about in narrative. Not only are there explicit speech acts between characters in the plot, but there are implicit acts of communication between characters, implicit dialogues between the narrator and audience, and implicit dialogues between the writer and reader. (Graesser et al., 2003, p.84)

Different readings are thus possible for (52b): the small boy may be taken to be deliberately placing a heavy pitch accent on *goldfish* to contrast it with the assistant's suggestion about *hamsters,* or, his utterance may be taken as having originally had only PF on *goldfish,* with the CF a property of the narrator's own IS (there may also be a blurring of the lines between the narrator and the assistant here).

There may be additional processing costs here associated with modelling the common ground between multiple characters, on top of the modelling done for the common ground between the narrator and the reader themselves; the stimuli were varied in this respect, meaning that this may be a confound.

All of the aforementioned potential readings involve the TE still being interpreted as CF. If, however, the TE is taken to be a property of the boy's utterance, this might also increase the likelihood of it being interpreted as marking something other than CF. Two of the surveys of italic usage discussed in Chapter 3 give

examples that suggest TE can be used in reported/direct speech, Ashby (2017) and Sanford et al. (2006), whilst Douglas (2009) does not provide explicit examples, but does mention similar phenomena.[2] Between the three TE is claimed to be able to mark dialect features, voice quality change, and intonation other than pitch accents. All of these would seem more likely to occur in reported speech: it may be that TE within reported speech is interpreted differently to TE outside it.

6.5 Future directions

6.5.1 Eye tracking studies

The most obvious future direction would be to revise the analytical approach taken here, and to conduct an exploratory scanpath analysis on the data collected for both experiments that were conducted in this study. Experience gained in doing this could then be leveraged in designing a new experiment (or series of experiments), which could be built from the ground up with the intention to capture re-reading through scanpath analysis.

The design of materials for any subsequent experiment would require the utmost attention: ideally they would be pre-tested extensively, in order to establish whether participants were interpreting target words that were intended to have CF as actually having CF. In the absence of data regarding this for the materials from Experiment 2, it is unclear whether they would require major revision or not—but it is certainly possible that some items could benefit from being adjusted.

6.5.2 EEG studies

If a revised eye tracking experiment (or series of experiments) were to return more significant results than those obtained here, then the next step could be to conduct an EEG experiment. This would be an expensive and complicated route to take, especially given that, to conduct an experiment in any way similar to this one, eye tracking or self-paced reading would have to be used in concert with the neuroimaging process: as was discussed in Chapter 4 (§4.3.2.2), presenting TE in context would require the simultaneous presentation of multiple words, and then the neuroimaging data would have to be analysed alongside the eye movement data in order to know when a given word was being fixated, returned to, etc.

---

[2]Ashby (2017) also gives an example showing TE as marking reported speech itself—see §3.3.

The issue of timing would imply that the high temporal precision of EEG would be preferable over fMRI (Osterhout et al., 2004; Brennan, 2016), although as was noted in Chapter 4 the combination of EEG and eye tracking requires complex analysis to filter out eye movement artefacts from the EEG data (Plöchl et al., 2012).

A combined neuroimaging and eye tracking experiment might sidestep some of the problems associated with interrogating the eye movement record alone: in particular, the results of this experiment may have been impacted by the issue of visual salience. Eye tracking experiments that investigate reading tend to rely on the eye–mind connection (Just and Carpenter, 1980), but are able to vary the linguistic properties of their stimuli without varying visual salience. TE by definition relies on visual salience, so this study was not able to do this. Whilst attempts have been made above (§6.3.3) to draw conclusions about whether there was a confounding visual salience-only effect at play here, eye movements are not (to invoke one of the 'ideal qualities' of a psycholinguistic technique as advanced by Osterhout et al., 2004) differentially sensitive to linguistic levels. The visual salience may obscure or alter what would otherwise be eye movements reflecting linguistic effects, and this interference may be impossible to factor out through eye tracking alone.

Neuroimaging is much more sensitive to different types of processing (although whether these correspond to 'linguistic levels' is another complex issue), and may help in separating out the visual aspects of TE processing from linguistic aspects. There have been no relevant neuroimaging studies on typographic emphasis; Cowles (2003, discussed in §2.6.2) represents a rare EEG experiment into information structure, dealing in particular with CF. However, this experiment used the more typical rapid serial visual presentation (RSVP) technique instead of concurrent eye tracking, and hence may not provide a suitable template: a neuroimaging approach to TE would most likely require a considerable amount of innovation.

6.5.3   Production experiments

Another angle of attack on the subject of TE and CF could come from investigating production. What might be considered 'production' experiments have been used before by studies which touch upon TE themselves. Chafe (1988) used production to investigate the relationship between punctuation and intonation (recall that TE has been argued here to be a close relative, if not a subset of punctuation; §3.4.1): Chafe both gave participants unpunctuated texts and asked them to

punctuate them, and recorded participants reading written texts aloud in order to analyse how their intonation patterns coincided with the punctuation marks (see §4.3.3). Similarly, one of McAteer's (1992) experiments involved participants being presented with text lacking TE, and being asked to indicate sections of that text which could be given TE in order to produce a 'non-default' effect, i.e. CF (see §3.5.2).

The present study is not chiefly concerned with how TE is deployed: it is more concerned with the effects of TE-marking of CF, regardless of how common this is. Therefore examining how participants would deploy TE when writing would not be particularly useful; information about usage, which is still useful on a secondary level to a study such as this (hence the examination of usage in §3.3) would be better gathered with a corpus approach (see below). Instead, what would benefit this research area is something closer to Chafe's comparisons between participants reading texts aloud and the texts themselves. A problem here was the inability to gain an accurate picture of whether participants had interpreted stimuli as having contrast or not (see §6.4.3.3): a supplementary experiment could be conducted with different participants, in which the stimuli were presented in all conditions and read aloud. CF in speech is marked with increased intonational prominence in comparison to PF (Bolinger, 1961; Katz and Selkirk, 2011), whether or not there are specific contrastive PAs (see Selkirk, 2002; Katz and Selkirk, 2011 and §2.4.1.3): the recordings could be analysed alongside eye tracking data to provide evidence of how different conditions were interpreted.

6.5.4   Corpus studies

All three studies (Sanford et al., 2006; Douglas, 2009; Ashby, 2017) that have been found in the literature investigating TE usage using corpora were small-scale (see §3.3), covering less than thirty individual texts between them. Furthermore, they all appear to have used novels exclusively. The need for a corpus study on a far larger scale is clear: as was said above, the present study itself was more concerned with reception than production, yet the two cannot be completely divorced from each other. More knowledge regarding how TE is used can only benefit studies such as this one.

A particular methodological challenge in creating a corpus of TE usage would be capturing TE itself. Typographic information is rarely encoded in existing corpora, meaning that new corpora would be required. It is rarely encoded as it is rarely studied, meaning that novel methods of analysis would also need to be developed, although existing corpus techniques are well-established for adding

and handling metadata to words: if words can be tagged for grammatical categories and that data can be analysed, the same should be possible for typographic properties such as type style.

Chapter 7

Conclusion

This study set out to investigate whether typographic emphasis can function in a similar way to intonational emphasis when it marks contrastive focus. In having done this, it represents a rare contribution towards the field of typographical linguistics (Crystal, 1998).

This study has reviewed the literature on both information structure, the area of linguistics into which contrastive focus falls, and typographic emphasis itself. It is hoped that the latter review might serve as a useful point of reference to future researchers in the field: there has been very little published in this area, and this may be the only review that both encompasses surveys of typographic emphasis usage, and psycholinguistic work on typographic emphasis.

Leading out of this review, a novel classification of sentence-level typographic emphasis usage has been presented here: it splits the uses of typographic emphasis into those that appear to be connected to phonological features (including the use similar to intonational emphasis for contrastive focus, which has been experimentally investigated here), and those that indicate interpolation, whereby sections of language are embedded within other sections of language.

Two eye tracking experiments were conducted for this study, both of which attempted to test how phonological typographic emphasis is processed during reading: the central hypothesis was that it would have a similar effect to that of pitch accents in speech, and facilitate processing when it was placed on a contrastively focused word.

To the author's knowledge, these experiments represent the most ambitious work on the online processing of typographic emphasis yet conducted; to date, no similar eye tracking study has been published that investigates the relationship between contrastive focus and typographic emphasis. The experiments (and in particular, the second experiment) involved complex, highly controlled de-

signs which, by typical psycholinguistic standards, presented a relatively unusual quantity of text to participants. All the material used for stimuli was created from scratch, and is available in the appendices of this thesis.

The results of these experiments were mixed, and did not provide clear support to the hypotheses that they were designed to test. However, they have demonstrated that a) typographic emphasis is attended to; b) it has an effect on the eye movement record; and c) that it can interact with contrastive focus, and thus can have some linguistic effect.

The field of typographical linguistics is wide open. This study has taken an early step into this neglected area, and has started to uncover evidence that there is something to be found within it.

References

Akinnaso, F. N. 1982. On the differences between spoken and written language. *Language and Speech*. **25**(2), pp.97–125.

Alexander, J. D. and Nygaard, L. C. 2008. Reading voices and hearing text: Talker-specific auditory imagery in reading. *Journal of Experimental Psychology: Human Perception and Performance*. **34**(2), pp.446–459.

Aloni, M., Beaver, D. and Clark, B. 1999. Focus and topic sensitive operators. In: Dekker, P. and Kerdiles, G. eds. *Proceedings of the 12th Amsterdam Colloquium, Amsterdam*. Amsterdam, pp.55–60.

Arnold, J. E., Kaiser, E., Kahn, J. M. and Kim, L. K. 2013. Information structure: Linguistic, cognitive, and processing approaches. *Wiley Interdisciplinary Reviews: Cognitive Science*. **4**(4), pp.403–413.

Ashby, P. 2017. A preliminary study of written intonation. *Journal of the English Phonetic Society of Japan*. **21**, pp.39–62.

Ashcraft, M. 2006. *Cognition* 4th ed. London: Pearson.

Auer, P. 1999. From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism*. **3**(4), pp.309–332.

Auguié, B. 2017a. *gridExtra: Miscellaneous functions for "grid" graphics*. [Online] [Accessed 30 January 2018] Available from: https://CRAN.R-project.org/package=gridExtra

Auguié, B. 2017a. *egg: Extensions for 'ggplot2', to align plots, and set panel sizes*. [Online] [Accessed 30 January 2018] Available from: https://CRAN.R-project.org/package=egg

Baines, P. and Haslam, A. 2002. *Type and typography*. London: Lawrence King.

Balota, D. A., Yap, M. J. and Cortese, M. 2006. Visual word recognition: The journey from features to meaning (a travel update). In: Traxler, M. and Gernsbacher, M. eds. *Handbook of psycholinguistics*. London: Academic Press, pp.285–376.

Banks, I. 1992. *Crow Road*. London: Abacus.

Barton, J. J., Hanif, H. M., Björnström, L. E. and Hills, C. 2014. The word-length effect in reading: A review. *Cognitive Neuropsychology*. **31**(5–6), pp.378–412.

Bates, D., Mächler, M., Bolker, B. and Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. **67**(1), pp.1–48.

Battistella, E. L. 1990. *Markedness: The evaluative superstructure of language*. New York: State University of New York Press.

Benatar, A. and Clifton, C., Jr. 2014. Newness, givenness and discourse updating: Evidence from eye movements. *Journal of Memory and Language*. **17**, pp.1–16.

Bernsen, N. O. 2002. Multimodality in language and speech systems—from theory to design support tool. In: Granström, B., House, D. and Karlsson, I. eds. *Multimodality in language and speech systems*. Springer, pp.93–148.

Bever, T. G. 1970. The cognitive basis for linguistic structures. In: Hayes, J. ed. *Cognition and the development of language*. New York: Wiley, pp.279–362.

Bhatia, V. K. 1997. Genre analysis today. *Revue belge de philologie et d'histoire*. **75**(3), pp.629–652.

Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Bigelow, C. and Holmes, K. 1993. The design of a Unicode font. *Electronic Publishing*. **6**(3), pp.289–305.

Binder, K. S. and Rayner, K. 1998. Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review*. **5**(2), pp.271–276.

Birch, S. and Clifton, C., Jr. 1995. Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech*. **38**(4), pp.365–391.

Birch, S. and Garnsey, S. 1995. The effect of focus on memory for words in sentences. *Journal of Memory and Language*. **34**, pp.232–267.

Birch, S. and Rayner, K. 2010. Effects of syntactic prominence on eye movements during reading. *Memory & Cognition*. **38**(6), pp.740–752.

Birch, S. and Rayner, K. 1997. Linguistic focus affects eye movements during reading. *Memory & Cognition*. **25**(5), pp.653–660.

Birner, B. J. 1994. Information status and word order: An analysis of English inversion. *Language*. **70**, pp.233–259.

Bloomfield, L. 1987. Literate and illiterate speech. In: Lockett, C. F. ed. *A Leonard Bloomfield anthology*. London: University of Chicago Press, pp.84–93.

BNC 2007. National corpus, version 3 (BNC XML Edition). [Online] [Accessed 30 January 2018] Available from: https://corpus.byu.edu/bnc/

Bock, J. K. and Mazzella, J. R. 1983. Intonational marking of given and new in-

formation: Some consequences for comprehension. *Memory & Cognition*. **11**(1), pp.64–76.

Boland, J. E. 2004. Linking eye movements to sentence comprehension in reading and listening. In: Carreiras, M. and Clifton, C., Jr. eds. *The on-line study of sentence comprehension: Eyetracking, ERPS and beyond*. New York: Psychology Press, pp.51–76.

Bolinger, D. L. 1958. A theory of pitch accent in English. *Word*. **14**(2–3), pp.109–149.

Bolinger, D. L. 1961. Contrastive accent and contrastive stress. *Language*. **37**(1), pp.83–96.

Bösch, P. 1983. *Agreement and anaphora: A study of the role of pronouns in syntax and discourse*. New York: Academic Press.

Box, G. and Cox, D. 1964. An analysis of transformations. *Journal of the Royal Statistical Society*. **26**, pp.211–252.

Braun, B., Dainora, A. and Ernestus, M. 2011. An unfamiliar intonation contour slows down online speech comprehension. *Language and Cognitive Processes*. **26**(3), pp.350–375.

Brennan, J. 2016. Naturalistic sentence comprehension in the brain. *Language and Linguistics Compass*. **10**(7), pp.299–313.

Bringhurst, R. 2002. *The elements of typographic style: Version 2*. Point Roberts: Hartley & Marks.

Brown, G. 1983. Prosodic structure and the given/new distinction. In: Cutler, A. and Ladd, D. eds. *Prosody: Models and measurements*. Berlin: Springer, pp.67–77.

Brysbaert, M., Drieghe, D. and Vitu, F. 2005. Word skipping: Implications for theories of eye movement control in reading. In: Underwood, G. ed. *Cognitive processes in eye guidance*. Oxford: Oxford University Press, pp.53–77.

Büring, D. 2007. Semantics, intonation and information structure. In: Ramchand, G. and Reiss, C. eds. *The Oxford handbook of linguistic interfaces*. Oxford: Oxford University Press, pp.445–474.

Calhoun, S. 2009. What makes a word contrastive? Prosodic, semantic and pragmatic perspectives. In: Barth-Weingarten, D., Déhe, N. and Wichmann, A. eds. *Where prosody meets pragmatics*. Bingley: Emerald, pp.53–77.

Catach, N. 1986. The grapheme: Its position and its degree of autonomy with respect to the system of language. In: Augst, G. ed. *New trends in graphemics and orthography*. Berlin: De Gruyter, pp.1–10.

Chafe, W. 1987. Cognitive constraints on information flow. In: Tomlin, R. ed.

Coherence and grounding in discourse: Outcome of a symposium, Eugene, Oregon, June 1984. Amsterdam: John Benjamins, pp.21–52.

Chafe, W. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: Li, C. ed. *Subject and topic*. New York: Academic Press, pp.27–55.

Chafe, W. 1988. Punctuation and the prosody of written language. *Written Communication.* **5**(4), pp.396–426.

Chomsky, N. 1971. Deep structure, surface structure, and semantic interpretation. In: Steinberg, D. D. and Jakobovits, L. A. eds. *Semantics*. Cambridge: Cambridge University Press, pp.183–216.

Clark, H. H. and Brennan, S. E. 1991. Grounding in communication. In: Resnick, L. B., Levine, J. M. and Teasley, S. D. eds. *Perspectives on socially shared cognition*. Washington: American Psychological Association, pp.129–130.

Clifton, C., Jr., Staub, A. and Rayner, K. 2007. Eye movements in reading words and sentences. In: van Gompel, R., Fisher, M., Murray, W. and Hill, R. eds. *Eye movement research: A window on mind and brain*. Oxford: Elsevier, pp.341–372.

Collins, P. C. 1991. *Cleft and pseudo-cleft constructions in English*. Oxon: Routledge.

Connors, R. J. 1993. *Actio*: A Rhetoric of Written Delivery (Iteration Two). In: Reynolds, J. F. ed. *Rhetorical memory and delivery: Classical concepts for contemporary composition and communication*. London: Routledge, pp.65–78.

Cook, V., Vaid, J. and Bassetti, B. 2009. Writing systems research: A new journal for a developing field. *Writing Systems Research.* **1**(1).

Coulmas, F. 1986. Reported speech: Some general issues. In: Coulmas, F. ed. *Direct and indirect speech*. Berlin: De Gruyter, pp.1–29.

Coulmas, F. 2003. *Writing systems: An introduction to their linguistic analysis*. Cambridge: Cambridge University Press.

Coulson, S., King, J. W. and Kutas, M. 1998. Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes.* **13**(1), pp.21–58.

Cowles, H. W. 2003. *Processing information structure: Evidence from comprehension and production.* Ph.D. thesis, University of California, San Diego.

Cowles, H. W. 2012. The psychology of information structure. In: Krifka, M. and Musan, R. eds. *The expression of information structure*. Berlin: De Gruyter, pp.287–318.

Cruttenden, A. 1997. *Intonation* 2nd ed. Cambridge: Cambridge University Press.

Crystal, D. 1998. Toward a typographical linguistics: Keynote address for the ATypI 1997 Conference, University of Reading. *Type.* **2**(1), pp.7–23.

Cutler, A. 1976. Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*. **20**(1), pp.55–60.

Cutler, A. and Foss, D. J. 1977. On the role of sentence stress in sentence processing. *Language and Speech*. **20**(1), pp.1–10.

Cutler, A., Oahan, D. and van Donselaar, W. 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech*. **40**(2), pp.141–201.

Dahl, R. 1979. *Tales of the unexpected*. Harmondsworth: Penguin.

Daniels, P. T. 2013. The history of writing as a history of linguistics. In: Allan, K. ed. *The Oxford handbook of the history of linguistics*. Oxford: Oxford University Press, pp.53–69.

Derrida, J. 1967. *Of grammatology*. Baltimore: Johns Hopkins University Press.

Dickens, C. 1839. *Nicholas Nickleby*. London: Penguin.

Dik, S., Hoffman, M. E., Jong, J. R. de, Djiang, S. I., Stroomer, H. and Vries, L. de 1981. On the typology of focus phenomena. In: Hoekstra, T., van der Hulst, H. and Moortgat, M. eds. *Perspectives on functional grammar*. Dordrecht, Netherlands: Foris, pp.41–74.

Ditman, T., Holcomb, P. J. and Kuperberg, G. R. 2007. An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*. **44**(6), pp.927–935.

Donchin, E. and Coles, M. G. 1988. Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*. **11**(3), pp.357–374.

Douglas, P. 2009. Encoding intonation: The use of italics and the challenges for translation. In: Mahlberg, M., González-Díaz, V. and Smith, C. eds. *Proceedings of the Corpus Linguistics Conference, 20–23 July 2009, Liverpool*. Liverpool.

Drieghe, D., Rayner, K. and Pollatsek, A. 2005. Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*. **31**(5), pp.954–969.

Duchowski, A. T. 2017. *Eye tracking methodology: Theory and practice*. London: Springer.

É. Kiss, K. 2017. Deriving the properties of structural focus. In: Lee, C., Kiefer, F. and Krifka, M. eds. *Contrastiveness in information structure, alternatives and scalar implicatures*. Springer, pp.47–61.

É. Kiss, K. 1998. Identificational focus versus information focus. *Language*. **74**(2), pp.245–273.

Engbert, R., Longtin, A. and Kliegl, R. 2002. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Re-*

search. **42**(5), pp.621–636.

Engbert, R., Nuthmann, A., Richter, E. M. and Kliegl, R. 2005. SWIFT: A dynamical model of saccade generation during reading. *Psychological Review.* **112**(4), pp.777–813.

Engle, R. A. 1998. Not channels but composite signals: Speech, gesture, diagrams and object demonstrations are integrated in multimodal explanations. In: Gernsbacher, M. A. and Derry, S. J. eds. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society, 1–4 August 1998, Wisconsin,* pp.321–326.

Erickson, T. D. and Mattson, M. E. 1981. From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior.* **20**(5), pp.540–551.

Erteschik-Shir, N. 2007. *Information structure: The syntax–discourse interface.* Oxford: Oxford University Press.

Feng, S., D'Mello, S. and Graesser, A. C. 2013. Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review.* **20**(3), pp.586–592.

Ferreira, F. and Clifton, C., Jr. 1986. The independence of syntactic processing. *Journal of Memory and Language.* **25**(3), pp.348–368.

Ferreira, F. and Lowder, M. W. 2016. Prediction, information structure, and good-enough language processing. *Psychology of Learning and Motivation.* **65**, pp.217–247.

Ferreira, F. and Patson, N. D. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass.* **1**(1–2), pp.71–83.

Ferreira, F., Bailey, K. G. and Ferraro, V. 2002. Good-enough representations in language comprehension. *Current Directions in Psychological Science.* **11**(1), pp.11–15.

Féry, C. and Samek-Lodovici, V. 2006. Focus projection and prosodic prominence in nested foci. *Language.* **82**, pp.131–150.

Filik, R. and Barber, E. 2011. Inner speech during silent reading reflects the reader's regional accent. *PloS ONE.* **6**(10), p.e25782.

Fodor, J. A. 1983. *The modularity of mind: An essay on faculty psychology.* Cambridge: MIT Press.

Foss, D. J. 1969. Decision processes during sentence comprehension: Effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior.* **8**(4), pp.457–462.

Fraundorf, S. H., Benjamin, A. S. and Watson, D. G. 2013. What happened (and what didn't): Discourse constraints on encoding of plausible alternatives. *Journal of Memory and Language.* **69**, pp.196–277.

Fraundorf, S. H., Watson, D. G. and Benjamin, A. S. 2010. Recognition memory re-

veals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language.* **63**(3), pp.367–386.

Frazier, L. and Fodor, J. D. 1978. The sausage machine: A new two-stage parsing model. *Cognition.* **6**(4), pp.291–325.

Frazier, L. and Rayner, K. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology.* **14**, pp.178–210.

Garrod, S. 2006. Psycholinguistic research methods. In: Brown, K. ed. *The encyclopedia of language and linguistics.* New York: Elsevier, pp.251–257.

Gelderman, M. 1999. Fonts: A short introduction to font terminology. *TUGboat.* **20**(2), pp.81–88.

Gibson, E. and Pearlmutter, N. J. 2000. Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research.* **29**(2), pp.231–240.

Glanzberg, M. 2005. Focus: A case study on the semantics–pragmatics boundary. In: Szabó, Z. G. ed. *Semantics versus pragmatics.* Oxford: Clarendon Press, pp.72–110.

González, M. de los Á. G. 2001. *The theme–topic interface: Evidence from English.* John Benjamins.

Graesser, A. C., McNamara, D. S. and Louwerse, M. M. 2003. What do readers need to learn in order to process coherence relations in narrative and expository text? In: Sweet, A. P. and Snow, C. E. eds. *Rethinking reading comprehension.* New York: Guilford, pp.82–98.

Grice, H. 1975. Logic and conversation. In: Cole, P. and Morgan, J. eds. *Syntax and semantics 3: Speech acts.* New York: Academic Press, pp.41–58.

Grice, H. P. 1989. *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Grosjean, F. and Collins, M. 1979. Breathing, pausing and reading. *Phonetica.* **36**(2), pp.98–114.

Gunter, T. C., Stowe, L. A. and Mulder, G. 1997. When syntax meets semantics. *Psychophysiology.* **34**, pp.660–676.

Gussenhoven, C. 1999. On the limits of focus projection in English. In: Bosch, P. and van der Sandt, R. eds. *Focus: Linguistic, cognitive, and computational perspectives.* Cambridge: Cambridge University Press, pp.43–55.

Gussenhoven, C. 1983. Testing the reality of focus domains. *Language and Speech.* **26**(1), pp.61–80.

Gussenhoven, C. 2006. Types of focus in English. In: In Lee, C., Gordon, M. and Büring, D. eds. *Topic and focus: Cross-linguistic perspectives on meaning and into-*

nation. Berlin: Springer, pp.83–100.

Guterson, D. 1995. *Snow falling on cedars.* London: Bloomsbury.

Hadding-Koch, K. and Studdert-Kennedy, M. 1964. An experimental study of some intonation contours. *Phonetica.* **11**(3–4), pp.175–185.

Hajičová, E. 1991. Topic-focus articulation and coreference in models of discourse production. *Journal of Pragmatics.* **16**(2), pp.157–166.

Hajičová, E. and Sgall, P. 2004. Degrees of contrast and the topic-focus articulation. In: Steube, A. ed. *Information structure: Theoretical and empirical aspects.* Berlin: De Gruyter, pp.1–15.

Hale, J. 2003. The information conveyed by words in sentences. *Journal of Psycholinguistic Research.* **32**(2).

Halliday, M. 1967. Notes on transitivity and theme in English: Part 2. *Journal of Linguistics.* **3**(2), pp.199–244.

Hamblin, C. 1973. Questions in Montague English. *Foundations of Language.* **10**, pp.41–53.

Haspelmath, M. 2006. Against markedness (and what to replace it with). *Journal of Linguistics.* **42**, pp.25–70.

Hedberg, N. and Sosa, J. M. 2008. The prosody of topic and focus in spontaneous English dialogue. In: Lee, C., Gordon, M. and Büring, D. eds. *Topic and focus: Cross-linguistic perspectives on meaning and intonation.* Dordrecht: Springer, pp.101–120.

Heim, I. 1983. File change semantic and the familiarity theory of definiteness. In: Bauerle, R., Schwarze, C. and von Stechow, A. eds. *Meaning, use and interpretation of language.* Berlin: De Gruyter, pp.164–189.

Henderson, J. M., Singer, M. and Ferreira, F. 1995. *Reading and language processing.* Hove: Taylor & Francis.

Henderson, L. 1985. On the use of the term 'grapheme'. *Language and Cognitive Processes.* **1**(2), pp.135–148.

Henry, L. and Wickham, H. 2017. *purrr: Functional Programming Tools* [Online]. [Accessed 30 January 2018] Available from: https://CRAN.R-project.org/package=purrr

Hirschberg, J. 1993. Pitch accent in context: Predicting intonational prominence from text. *Artificial Intelligence.* **63**(1–2), pp.305–340.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. and van de Weijer, J. 2011. *Eye tracking: A comprehensive guide to methods and measures.* Oxford: Oxford University Press.

Hoop, H. de 2004. On the interpretation of stressed pronouns. In: Blutner, R. and

Zeevat, H. eds. *Optimality theory and pragmatics*. Basingstoke: Palgrave Macmillan, pp.25–41.

Horn, L. 1981. Exhaustiveness and the semantics of clefts. In: Burke, V. and Pustejowski, J. eds. *Proceedings of the North Eastern Linguistic Society (NELS), Amherst*. Amherst: University of Massachusetts, pp.125–142.

Huey, E. 1908. *The psychology and pedagogy of reading*. New York: Macmillan.

Hulme, C., Snowling, M., Caravolas, M. and Carroll, J. 2005. Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading*. **9**(4), pp.351–365.

Hunt, R. R. 1995. The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review*. **2**(1), pp.105–112.

Hyönä, J., Lorch, R. F., Jr. and Kaakinen, J. K. 2002. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology*. **94**(1), pp.44–55.

Inhoff, A. W. and Radach, R. 1998. Definition and computation of oculomotor measures in the study of cognitive processes. In: Underwood, G. ed. *Eye guidance in reading and scene perception*. Amsterdam: Elsevier, pp.29–54.

Inhoff, A. W. and Topolski, R. 1992. Lack of semantic activation from unattended text during passage reading. *Bulletin of the Psychonomic Society*. **30**(5), pp.365–366.

Ippolito, M. 2007. On the meaning of *only*. *Journal of Semantics*. **25**, pp.45–91.

Jackendoff, R. S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Janda, L. A. 1996. Unpacking markedness. *Cognitive Linguistics in the Redwoods: The Expansion of a New Paradigm in Linguistics*. **6**, pp.207.

Jones, D. 1956. *An outline of English phonetics*. New York: Heffer.

Jury, D. 2007. History of typography. In: Bazerman, C. ed. *Handbook of research on writing: History, society, school, individual, text*. New York: Lawrence Erlbaum.

Just, M. A. and Carpenter, P. A. 1980. A theory of reading: From eye fixations to comprehension. *Psychological Review*. **87**(4), pp.329.

Just, M. A., Carpenter, P. A. and Woolley, J. D. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*. **111**(2), pp.228.

Katz, J. and Selkirk, E. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*. **87**(4), pp.771–816.

Kenesei, I. 2006. Focus as identification. In: Molnár, V. and Winkler, S. eds. *The architecture of focus*. Berlin: De Gruyter, pp.137–168.

Khordoc, C. 2007. Visual sound effects in *Asterix*. In: Varnum, R. and Gibbons, C. T. eds. Jackson: University Press of Mississippi, pp.156–173.

Kliegl, R. and Engbert, R. 2003. SWIFT explorations. In: Radach, R., Hyönä, J. and Deubel, H. eds. *The mind's eye: Cognitive and applied aspects of eye movement research.* Amsterdam: Elsevier, pp.391–412.

Kliegl, R., Nuthmann, A. and Engbert, R. 2006. Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General.* **135**(1), pp.12.

Kohler, K. J. 2006. What is emphasis and how is it coded? In: Hoffman, R. and Mixdorff, H. eds. *Proceedings of the Third International Conference on Speech Prosody, 2–5 May 2006, Dresden.* Dresden: International Speech Communication Association, pp.748–751.

Kretzschmar, F., Pleimling, D., Hosemann, J., Füssel, S., Bornkessel-Schlesewsky, I. and Schlesewsky, M. 2013. Subjective impressions do not mirror online reading effort: Concurrent EEG-eyetracking evidence from the reading of books and digital media. *PloS ONE.* **8**(2).

Krifka, M. 2007. Basic notions of information structure. In: Féry, C., Fanslow, G. and Krifka, M. eds. *Working papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6.* Potsdam: Universitätsverlag Potsdam, pp.13–55.

Krifka, M. and Musan, R. 2012. Information structure: Overview and linguistic issues. In: Krifka, M. and Musan, R. eds. *The expression of information structure.* Berlin: De Gruyter, pp.1–44.

Kuipers, A. 1975. On symbols, distinctions and markedness. *Lingua.* **36**, pp.31–46.

Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, **82**(13), pp.1–26.

Ladd, D. 1980. *The structure of intonational meaning: Evidence from English.* Bloomington: Indiana University Presss.

Ladd, D. R. and Morton, R. 1997. The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics.* **25**(3), pp.313–342.

Lambrecht, K. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents.* Cambridge: Cambridge University Press.

Lawrence, S., Nolan, F. and McDougall, K. 2008. Acoustic and perceptual effects of telephone transmission on vowel quality. *International Journal of Speech, Language & the Law.* **15**(2).

Leinenger, M. 2014. Phonological coding during reading. *Psychological Bulletin.*

140(6), pp.1534–1555.

Levinson, S. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lewis, R. L. 2000. Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research*. **29**(2), pp.241–248.

Lewis, R. L. 1998. Reanalysis and limited repair parsing: Leaping off the garden path. In: Fodor, J. D. and Ferreira, F. eds. *Reanalysis in sentence processing*. Dordrecht: Springer, pp.247–284.

Lupton, E. 2010. *Thinking with type: A critical guide for designers, writers, editors, & students* 2nd ed. New York: Princeton Architectural Press.

Mangen, A., Walgermo, B. R. and Brønnick, K. 2013. Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*. **58**, pp.61–68.

Margolin, S. J., Driscoll, C., Toland, M. J. and Kegler, J. L. 2013. E-readers, computer screens, or paper: Does reading comprehension change across media platforms? *Applied Cognitive Psychology*. **27**(4), pp.512–519.

Marian, V., Bartolotti, J. and Chabal, S. 2012. CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE*. **7**(8), p.e43230.

Matić, D. 2015. Information structure in linguistics. In: Wright, J. D. ed. *The international encyclopedia of social and behavioral sciences*. Amsterdam: Elsevier, pp.95–99.

Matić, D. and Wedgwood, D. 2012. The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics*., pp.1–37.

Mattingly, I. 1992. Linguistic awareness and orthographic form. In: Frost, R. and Katz, L. eds. *Orthography, phonology, morphology, and meaning*. Amsterdam: Elsevier, pp.11–26.

McAteer, E. 1989. emphTypeface effects in written language. Ph.D. thesis, University of Glasgow.

McAteer, E. 1992. Typeface emphasis and information focus in written language. *Applied Cognitive Psychology*. **6**(4), pp.345–359.

McConkie, G. and Rayner, K. 1975. The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*. **17**(6), pp.578–586.

McConkie, G. W. 1983. Control of eye movements. In: Rayner, K. ed. *Eye movements in reading: Perceptual and language processes*. New York: Academic Press, pp.67–80.

Meseguer, E., Carreiras, M. and Clifton, C., Jr. 2002. Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition.* **30**(4), pp.551–561.

Meyer, D. E. and Schvaneveldt, R. W. 1971. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology.* **90**, pp.227—234.

Miles, W. 1930. Ocular dominance in human adults. *Journal of General Psychology.* **3**, pp.412–430.

Mitchell, D. C. 2004. On-line methods in language processing: Introduction and historical review. In: Carreiras, M. and Clifton, C., Jr. eds. *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond.* New York: Psychology Press, pp.15–32.

Molnár, V. 2002. Contrast—from a contrastive perspective. *Language and Computers.* **39**(1), pp.147–161.

Morris, R. K. and Folk, J. R. 1998. Focus as a contextual priming mechanism in reading. *Memory & Cognition.* **26**(6), pp.1313–1322.

Müller, K. and Wickham, H. 2017. *tibble: Simple Data Frames* [Online]. [Accessed 30 January 2018] Available from:
https://CRAN.R-project.org/package = tibble

Nagy, W. E. and Anderson, R. C. 1995. *Metalinguistic awareness and literacy acquisition in different languages.* University of Illinois at Urbana-Champaign, Center for the Study of Reading.

Norton, C. W. 2012. *Processing italics.* MA thesis, University of Leeds.

Nunberg, G. 1990. *The linguistics of punctuation.* Menlo Park: Centre for the Study of Language and Information.

Osterhout, L. and Holcomb, P. J. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language.* **31**, pp.785–806.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R. and Inoue, K. 2004. Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In: Carreiras, M. and Clifton, C., Jr. eds. *The on-line study of sentence comprehension: Eyetracking, ERPS and beyond.* New York: Psychology Press.

Otterbein, S., Abel, C., Heinemann, L. V., Kaiser, J. and Schmidt-Kassow, M. 2012. P3b reflects periodicity in linguistic sequences. *PLoS One.* **7**(12), p.e51419.

Partee, B. 1991. Topic, focus and quantification. *Semantics and Linguistic Theory.* **1**, pp.159–188.

Perfetti, C. A. 2003. The universal grammar of reading. *Scientific Studies of Read-*

ing. **7**(1), pp.3–24.

Perrone-Bertolotti, M., Kujala, J., Vidal, J. R., Hamame, C. M., Ossandon, T., Bertrand, O., Minotti, L., Kahane, P., Jerbi, K. and Lachaux, J. -P. 2012. How silent is silent reading? Intracerebral evidence for top-down activation of temporal voice areas during reading. *The Journal of Neuroscience.* **32**(49), pp.17554–17562.

Perrone-Bertolotti, M., Rapin, L., Lachaux, J. -P., Baciu, M. and Lœvenbruck, H. 2014. What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research.* **261**, pp.220–239.

Pierce, J. R. 1979. *An introduction to information theory: Symbols, signals and noise* 2nd ed. New York: Dover.

Pierrehumbert, J. B. 1980. *The phonology and phonetics of English intonation.* Ph.D. thesis, Massachusetts Institute of Technology.

Plöchl, M., Ossandón, J. P. and König, P. 2012. Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience.* **6**.

Pollatsek, A., Raney, G. E., Lagasse, L. and Rayner, K. 1993. The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale.* **47**(2), pp.179–200.

Pollatsek, A., Reichle, E. D. and Rayner, K. 2006. Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology.* **52**, pp.1–56.

Price, C. J. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage.* **62**(2), pp.816–847.

Prince, E. 1981a. Topicalization, focus-movement, and yiddish-movement: A pragmatic differentiation. In: Alford, D. K., Hunold, K. A., Macaulay, M. A. and Walter, J. eds. *Proceedings of the Seventh Annual Meeting of the Berkeley Linguistics Society, 14–16 February 1981, Berkeley.* Berkeley, pp.249–264.

Prince, E. 1981b. Toward a taxonomy of given-new information. In: Cole, P. ed. *Radical pragmatics.* New York: Academic Press, pp.223–255.

R Core Team 2017. *R: A language and environment for statistical computing* [Online]. [Accessed 30 January 2018] Available from: https://www.R-project.org/

Radach, R. and Kennedy, A. 2004. Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology.* **16**(1/2), pp.3–26.

Rankin, I. 2001. *The falls*. London: Orion.

Rapp.B., Benzing, L. and Caramazza, A. 1997. The autonomy of lexical orthography. *Cognitive Neuropsychology*. **14**(1), pp.71–104.

Rastle, H., Kathleen and Coltheart, M. 2002. 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology: Section A*. **55**(4), pp.1339–1362.

Rastle, K. 2007. Visual word recognition. In: Gaskell, G. ed. *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press, pp.71–87.

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*. **124**(3), pp.372–422.

Rayner, K. and Clifton, C., Jr. 2009. Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology*. **80**(1), pp.4–9.

Rayner, K. and Reichle, E. 2010. Models of the reading process. *Wiley Interdisciplinary Reviews: Cognitive Science*. **1**(6), pp.787–799.

Rayner, K., Ashby, J., Pollatsek, A. and Reichle, E. D. 2004. The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*. **30**, pp.720–732.

Rayner, K., Carlson, M. and Frazier, L. 1983. The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*. **22**(3), pp.358–374.

Rayner, K., Chace, K. H., Slattery, T. J. and Ashby, J. 2006. Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*. **10**(3), pp.241–255.

Rayner, K., Li, X. and Pollatsek, A. 2007. Extending the E-Z Reader model of eye movement control to Chinese readers. *Cognitive Science*. **31**, pp.1021–1033.

Rayner, K., Pollatsek, A., Ashby, J. and Clifton, C., Jr. 2012. *Psychology of reading* 2nd ed. London: Taylor & Francis.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R. and Clifton, C., Jr. 1989. Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*. **4**(3–4), pp.21–49.

Rayner, K., Slattery, T. J. and Bélanger, N. N. 2010. Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review*. **17**(6), pp.834–839.

Reichle, E. D., Pollatsek, A., Fisher, D. L. and Rayner, K. 1998. Toward a model

of eye movement control in reading. *Psychological Review.* **105**, pp.125–157.

Reichle, E. D., Rayner, K. and Pollatsek, A. 1999. Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research.* **39**, pp.4403–4411.

Reichle, E. D., Warren, T. and McConnell, K. 2009. Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review.* **16**(1), pp.1–21.

Repp.S. 2010. Defining 'contrast' as an information-structural notion in grammar. *Lingua.* **120**(6), pp.1333–1345.

Ritter, R. M. 2002. *The Oxford guide to style.* Oxford: Oxford University Press.

Roberts, C. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics — Ohio State University Department of Linguistics.*, pp.91–136.

Robinson, D. 2017. *broom: Convert Statistical Analysis Objects into Tidy Data Frames* [Online]. [Accessed 30 January 2018] Available from: https://CRAN.R-project.org/package=broom

Rogers, H. 2005. *Writing systems: A linguistic approach.* Oxford: Blackwell.

Rooth, M. 1992. A theory of focus interpretation. *Natural language semantics.* **1**(1), pp.75–116.

Rooth, M. 2016. Alternative semantics. In: Féry, C. and Ishihara, S. eds. *The Oxford handbook of information structure.* Oxford: Oxford University Press, pp.19–40.

Rooth, M. 1985. *Association with focus.* Ph.D. thesis, University of Massachusetts.

Rosenthal, R. 1979. The 'file drawer problem' and tolerance for null results. *Psychological Bulletin.* **86**(3), pp.638–641.

Saldanha, G. 2011. Emphatic italics in English translations: Stylistic failure or motivated stylistic resources? *Meta: Journal des traducteurs / Meta: Translators' Journal.* **56**(2), pp.424–442.

Sampson, G. 1985. *Writing systems: A linguistic introduction.* London: Hutchinson Education.

Sanders, T. and Canestrelli, A. R. 2012. The processing of pragmatic information in discourse. *Cognitive Pragmatics.* **4**, pp.201.

Sanford, A. J., Sanford, A. J., Molle, J. and Emmott, C. 2006. Shallow processing and attention capture in written and spoken discourse. *Discourse Processes.* **42**(2), pp.109–130.

Sanz, M. J. P. 2010. Weblogs as a subgenre of political communication/Marketing. *Revista de Lingüística y Lenguas Aplicadas.* **2**(1), pp.57–68.

Saussure, F. de 1959. *Course in general linguistics.* New York: Philosophical Library.

Schooler, J. W., Reichle, E. D. and Halpern, D. V. 2004. Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In: Levin, D. T. ed. *Thinking and seeing: Visual metacognition in adults and children.* Cambridge, MA: MIT Press, pp.203–226.

Schotter, E. R., Angele, B. and Rayner, K. 2012. Parafoveal processing in reading. *Attention, Perception, & Psychophysics.* **74**(1), pp.5–35.

Selkirk, E. 2002. Contrastive *FOCUS* vs. presentational *focus*: Prosodic evidence from right node raising in English. In: Bel, B. and Marlien, I. eds. *Speech Prosody 2002: Proceedings of the First International Speech Prosody Conference, 11–13 April 2002, Aux-en-Provence.* Aux-en-Provence: Laboratoire Parole et Langage, pp.643–646.

Selkirk, E. 2007. Contrastive focus, givenness and the unmarked status of 'discourse-new'. In: Féry, C., Fanselow, G. and Krifka, M. eds. *Working papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6.* Potsdam: Universitätsverlag Potsdam, pp.125–145.

Selkirk, E. O. 1995. Sentence prosody: Intonation, stress and phrasing. In: Goldsmith, J. ed. *Handbook of phonological theory*. Cambridge: Blackwell, pp.550–569.

Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal.* **27**(3), pp.379–423.

Shields, J. L., McHugh, A. and Martin, J. G. 1974. Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology*. **102**(2), pp.250–255.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. and Hirschberg, J. 1992. ToBI: A standard for labelling English prosody. In: *Proceedings of 2nd International Conference on Spoken Language Processing, 12–16 October 1992, Banff.* Banff, pp.286–290.

Simola, J., Kuisma, J., Öörni, A., Uusitalo, L. and Hyönä, J. 2011. The impact of salient advertisements on reading and attention on web pages. *Journal of Experimental Psychology: Applied.* **17**(2), pp.174–190.

Sloman, S. A., Hayman, C., Ohta, N., Law, J. and Tulving, E. 1988. Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* **14**(2), pp.223–239.

Slowiaczek, M. L. and Clifton, C., Jr. 1980. Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior.* **19**, pp.573–582.

Smith, F. 1994. *Understanding reading: A psycholinguistic analysis of reading and learning to read* 5th ed. Hillsdale, NJ: Lawrence Erlbaum.

Sproat, R. 2002. The consistency of the orthographically relevant level in Dutch. In: Neef, M., Neijt, A. and Sproat, R. eds. *The relation of writing to spoken language*. Berlin: De Gruyter, pp.35–48.

Stalnaker, R. 2002. Common ground. *Linguistics and Philosophy*. **25**(5), pp.701–721.

Stalnaker, R. 1977. Pragmatic presuppositions. In: Rogers, A., Wall, B. and Murphy, J. P. eds. *Proceedings of the Texas conference on performatives, presuppositions, and implicatures, Arlington*. Arlington: Center for Applied Linguistics, pp.135–148.

Staub, A. and Rayner, K. 2007. Eye movements and on-line comprehension processes. In: Gaskell, G. ed. *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press, pp.327–342.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C. and Rayner, K. 2010. Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology. Human Perception and Performance*. **36**(5), pp.1280–1293.

Steedman, M. 2000. Information structure and the syntax–phonology interface. *Linguistic Inquiry*. **31**(4), pp.649–689.

Sternberg, R. J. and Sternberg, K. 2017. *Cognitive psychology* 7th ed. Boston: Cengage Learning.

Stöckl, H. 2005. Typography: Body and dress of a text — a signing mode between language and image. *Visual Communication*. **4**(2), pp.204–214.

Stolterfoht, B., Friederici, A. D., Alter, K. and Steube, A. 2007. Processing focus structure and implicit prosody during reading: Differential ERP effects. *Cognition*. **104**(3), pp.565–590.

Storey, D. 1960. *Flight into Camden*. Harmondsworth: Penguin.

Swales, J. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Taft, M. and van Graan, F. 1998. Lack of phonological mediation in a semantic categorization task. *Journal of Memory and Language*. **38**, pp.203–224.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. and Sedivy, J. C. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science*. **268**(5217), pp.1632–1634.

Terken, J. and Nooteboom, S. G. 1987. Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive processes*. **2**(3/4), pp.145–163.

Theroux, P. 1967. *Waldo*. London: Penguin.

Trask, R. 1992. *A dictionary of grammatical terms in linguistics*. London: Taylor &

Francis.

Traxler, M. J. 2014. Trends in syntactic parsing: Anticipation, Bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences*. **18**(11), pp.605–611.

Truss, L. 2003. *Eats, shoots and leaves*. London: Profile.

Umbach, C. 2004. On the notion of contrast in information structure and discourse structure. *Journal of Semantics*. **21**(2), pp.155–175.

Vallduví, E. 1993. *The informational component*. New York: Garland.

Vallduví, E. and Engdahl, E. 1996. The linguistic realization of information packaging. *Linguistics*. **34**(3), pp.459–520.

Van Gompel, R. P. and Majid, A. 2004. Antecedent frequency effects during the processing of pronouns. *Cognition*. **90**(3), pp.255–264.

Van Herten, M., Kolk, H. H. and Chwilla, D. J. 2005. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*. **22**(2), pp.241–255.

Van Heuven, W. J., Mandera, P., Keuleers, E. and Brysbaert, M. 2014. SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*. **67**(6), pp.1176–1190.

Van Leusen, N. 2004. Incompatibility in context: A diagnosis of correction. *Journal of Semantics*. **21**(4), pp.415–441.

Velleman, D., Beaver, D., Destruel, E., Bumford, D., Onea, E. and Coppock, L. 2012. It-clefts are IT (inquiry terminating) constructions. In: Chereches, A. ed. *Proceedings of the 22nd Semantics and Linguistic Theory Conference, 18–20 May 2012, Chicago.*, pp.441–460.

Vitu, F. 2005. Visual extraction processes and regressive saccades in reading. In: Underwood, G. ed. *Cognitive processes in eye guidance*. Oxford: Oxford University Press, pp.1–32.

Von Restorff, H. 1933. Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychologische Forschung*. **18**, pp.299–342.

Von der Malsburg, T. and Vasishth, S. 2011. What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language*. **65**(2), pp.109–127.

Von der Malsburg, T., Kliegl, R. and Vasishth, S. 2015. Determinants of scanpath regularity in reading. *Cognitive science*. **39**(7), pp.1675–1703.

Wagner, M. and Watson, D. G. 2010. Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*. **25**(7–9), pp.905–945.

Walker, S. 2001. *Typography and language in everyday life: Prescriptions and practices*. Harlow: Longman.

Waller, R. 1996. Typography and discourse. In: Barr, R., Kamil, M. L., Mosen-

thal, P. B. and Perason, P. D. eds. *Handbook of reading research.* Mahwah, NJ: Lawrence Erlbaum, pp.341–380.

Waller, R. H. 1980. Graphic aspects of complex texts: Typography as macropunctuation. In: Kolers, P. A. ed. *Processing of visible language.* Springer, pp.241–253.

Walsh, G. 2016. Screen and paper reading research — a literature review. *Australian Academic & Research Libraries.* **47**(3), pp.160–173.

Watson, D. G., Tanenhaus, M. K. and Gunlogson, C. A. 2008. Interpreting pitch accents in online comprehension: H* vs. L + H. *Cognitive Science.* **32**(7), pp.1232–1244.

Weber, A., Braun, B. and Crocker, M. 2006. Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech.* **49**(Pt 3), pp.367–92.

Wickham, H. 2007. Reshaping data with the reshape package. *Journal of Statistical Software.* **21**(12).

Wickham, H. 2009. *ggplot2: Elegant graphics for data analysis.* New York: Springer-Verlag.

Wickham, H. and Henry, L. 2017. *tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions* [Online]. [Accessed 30 January 2018] Available from: https://CRAN.R-project.org/package = tidyr

Wickham, H., Francois, R., Henry, L. and Müller, K. 2017. *dplyr: A Grammar of Data Manipulation* [Online]. [Accessed 30 January 2018] Available from: https://CRAN.R-project.org/package = dplyr

Wolf, M. and Katzir-Cohen, T. 2001. Reading fluency and its intervention. *Scientific Studies of Reading.* **5**(3), pp.211–239.

Yap, M. J. and Balota, D. A. 2009. Visual word recognition of multisyllabic words. *Journal of Memory and Language.* **60**(4), pp.502–529.

Zimmermann, M. 2007. Contrastive focus. In: Féry, C., Fanselow, G. and Krifka, M. eds. *Working papers of the SFB632, Interdisciplinary Studies on Information Structure (ISIS) 6.* Potsdam: Universitätsverlag Potsdam, pp.147–159.

Zimmermann, M. and Onea, R. 2011. Focus marking and focus interpration. *Lingua.* **121**, pp.1651–1670.

Appendix A: Stimuli for Experiment 1

Test items

1. A: Amanda's getting married to this guy she met last year. She said his name was John. B: She said his name was *Stephen*, and I'm sure of that because I actually know him. True or false: Speaker B knows Amanda's fiancé (true).
2. A: There's one food Sarah can't eat for fear of an allergic reaction. She can't eat peanuts safely. B: She can't eat *almonds* safely, everything else is fine. True or false: Sarah does not have any allergies (false).
3. A: There's a particular tool that every gardener needs. They should buy a rake before anything else. B: They should buy *shears* before anything else, those are the most useful tool by far. True or false: They are talking about tools (true).
4. A: Someone asked me what the most common farm animal is. You'll always find horses at a farm. B: You'll always find *chickens* at a farm, as a rule. True or false: They are talking about jobs (false).
5. A: If you want to design jewellery, go with the most versatile precious metal. You can do anything with platinum in the workshop. B: You can do anything with *silver* in the workshop, I reckon. True or false: Speaker A thinks platinum is versatile (true).
6. A: I'm going to start playing the ultimate team sport. Anyone can see that basketball is the best. B: Anyone can see that *cricket* is the best, and always will be. True or false: Speaker B dislikes cricket (false).
7. A: I need to tell the visitors about the best public transport option for getting here. From the airport, they're best off taking the bus. B: From the airport they're best off taking the *train*, which is cheaper and faster. True or false: Speaker B thinks the train is the best option (true).
8. A: We need to paint this room in the most uplifting colour there is. We

all agreed on yellow being very cheery. B: We all agreed on *orange* being very cheery last time we talked, actually. True or false: They are discussing depressing colours (false).

9. A: Conducting a survey is pointless, because the nation's favourite Italian dish is obvious. Everybody loves spaghetti here. B: Everybody loves *lasagne* here, I'm surprised you're claiming otherwise.

10. A: I was wondering what the toughest baking challenge was. People say it's hard to make cakes properly. B: People say it's hard to make *cookies* properly, other stuff is rarely as challenging.

11. A: I have strong opinions on which household task to do first. You must always do hoovering before anything else. B: You must always do *dusting* before anything else, it's crucial.

12. A: Someone was talking about America's favourite music genre on the radio today. Americans listen to rock all the time. B: Americans listen to *hip-hop* all the time, it's way bigger than other genres.

13. A: When looking at a used car, one visual feature is crucial. Check the headlights before anything else. B: Check the *bodywork* before anything else, it's the best indicator of wear and tear.

14. A: We need to get rid of one of the pieces of office equipment. We don't need the photocopier any more. B: We don't need the *printer* any more, everything else should stay.

15. A: A single drug causes more misery than all the others put together. They should never have legalised tobacco. B: They should never have legalised *alcohol* in this country.

16. A: The refurbishment only made a real difference to a single aspect of the room. Everyone is going to notice the ceiling. B: Everyone is going to notice the *flooring* in there, they won't be looking up.

17. A: He got a promotion and has a higher military rank. They've given him the position of commander. B: They've given him the position of *general*, you can tell by the uniform.

18. A: The space probe mission can only afford to visit one planet. They decided to go to Venus. B: They decided to go to *Mercury* on this mission, which is a very wise choice.

19. A: My son asked for the best possible phone for his birthday. The best phones are made by Apple. B: The best phones are made by *Samsung* these days, get him one of them.

20. A: Of course, the bride wants all the bridesmaids in the same type of cloth-

ing. She wants them all to wear suits. B: She wants them all to wear *dresses* as per tradition.

21. A: I need the most beautiful species of tree in my garden. Nothing looks better than an oak at the end of a lawn. B: Nothing looks better than a *willow* at the end of a lawn, anything else is second best.

22. A: Experts recommend a specific type of footwear for sightseeing. You can't go wrong with boots on a tour. B: You can't go wrong with *sandals* on a tour, in my experience.

23. A: I've been reading in my guidebook about the best type of eating place in Rome. We should go to restaurants for every meal. B: We should go to *cafes* for every meal, we need to save cash!

24. A: I know the best time of day for getting real work done. We are most productive in the evening. B: We are most productive in the *morning* before we've had lunch.

25. A: The museum has an exhibit on the deadliest type of weapon. More people have been killed by guns than anything else. B: More people have been killed by *swords* than anything else, historically.

26. A: We can afford one more piece of furniture for this room. It would really be improved by a table next to the bed. B: It would really be improved by a *wardrobe* next to the bed.

27. A: When starting a business one should consider the best vehicle to buy. Starting off with a van is the sensible option. B: Starting off with a *bicycle* is the sensible option, it's cheapest.

28. A: The guy who owns the pub comes from a different country. He was born in Russia originally. B: He was born in *Finland* originally, although his accent is a bit unusual for a Finn.

29. A: The instrument children should learn first is clear. Playing the guitar is easy and fun. B: Playing the *trumpet* is easy and fun, and children pick it up far quicker.

30. A: He had to have one of his internal organs removed. They took his kidney out during surgery. B: They took his *spleen* out during surgery, because that was what was causing the problem.

31. A: I'm going to go to the bar and get the strongest spirit they have. They sell vodka that could knock out a horse. B: They sell *whiskey* that could knock out a horse, their other drinks are watered down.

32. A: One educational institution has more impact than the others. People are shaped by university to a huge extent. B: People are shaped by *school* to a

huge extent, right through their lives.

33. A: I'm going to visit the best European city for culture. The literature points to Berlin as being the most cultured. B: The literature points to *Geneva* as being the most cultured city on the continent.

34. A: Computer recognition systems must be able to understand one shape before the others. The simplest thing to recognise is the triangle, surely. B: The simplest thing to recognise is the *circle*, I think you'll find.

35. A: We need to pick the hot drink that will sell best. The vast majority of people drink tea. B: The vast majority of people drink *coffee* in the modern world, that's the safe choice.

36. A: Someone asked me about the most beautiful thing in the solar system. Nothing out there is more stunning than the sun. B: Nothing out there is more stunning than the *Earth*, it's like a jewel in space.

37. A: I'm telling my son to study the best language for an international career. It's invaluable to learn Portuguese right now. B: It's invaluable to learn *Chinese* right now, it gives you a competitive edge.

38. A: I'm going to go shopping for the nicest watch I can find. I expect that will be a Rolex. B: I expect that will be a *Breitling* if you're going to the new mall for it.

39. A: Tattoo artists must be getting tired of doing that single same image again and again. Everyone is getting roses all the time. B: Everyone is getting *skulls* all the time, it's unbelievable.

40. A: The tourist board asked me what the most varied part of the UK was. There's no debate that Scotland is the most diverse. B: There's no debate that *England* is the most diverse, it has more of everything.

Filler items

1. A: There was a big fuss about the comedian who appeared on that chat show. People didn't think his jokes were acceptable. B: People didn't think his jokes were acceptable, but people are so easily offended these days. True or false: A comedian appeared on a chat show (true).

2. A: He talks a lot about moving to the world's nicest island. He must be thinking of Australia. B: He must be thinking of Australia, yes, it's a lovely place to be. True or false: Speaker B thinks Australia is unpleasant (false).

3. A: In families, there's one kind of relative who tends to get neglected. Not many people have much time for uncles. B: Not many people have much time for uncles these days, no. True or false: They are discussing families

(true).

4. A: One of the pets must have eaten the plant. Suspicion falls on the dog. B: Suspicion falls on the dog, yes, especially given that it's already tried to eat the other plants. True or false: They are wondering which pet ate a cake (false).

5. A: It sounds like Ian has been overdoing one of his exercises. He's done too many crunches, judging from his complaints about pain. B: He's done too many crunches, yes, you can tell. True or false: Ian has been complaining (true).

6. A: Do you know what the oldest currency in the world is? The pound is the oldest currency still in use. B: The pound is the oldest currency, I knew that. True or false: Speaker B claims the dollar is the oldest currency (false).

7. A: Wasn't that guy really into one of those American sports? He was always going on about baseball, if I remember right. B: He was always going on about baseball, I think so. True or false: Speaker A suggests that the man liked baseball (true).

8. A: We should get dad a cookery book by his favourite celebrity chef. He really enjoys Rick Stein's stuff, so let's get him something by him. B: He really enjoys Rick Stein's stuff, yes, good idea. True or false: Speaker B suggests a book by Delia Smith (false).

9. A: What did you make of the coverage of the royal wedding? It feels like that's all they have been talking about lately. B: To be honest I'm **really** sick of it all. True or false: Speaker B is fed up with the royal wedding coverage (true).

10. A: Mum used to tell me off for not drying my hair when I was young. I used to lie on the sofa and make it damp. B: That is **extremely** irritating behaviour, you were a little brat! True or false: Speaker B thinks Speaker A's behaviour was fine (false).

11. A: I wanted to buy a candle, so I popped into that new shop. There's no way I'm able to afford anything they have in there. B: I heard that everything they had was **insanely** expensive—so that's true? True or false: Speaker A found the shop expensive (true).

12. A: Why is everyone so nostalgic about that terrible song? It was awful at the time, and it's still awful now. B: I'm afraid I **totally** disagree, it's an incredible song! True or false: They are talking about radio stations (false).

13. A: The results of the assessment were better than we expected. All departments have improved, as far as the inspectors are concerned. B: We have

done **exceptionally** well, considering the challenges we've been facing. True or false: They are talking about an assessment (true).

14. A: The population of sea birds has increased. This is really good news in terms of our conservation goals, right? B: The population has **moderately** increased, but that doesn't mean we're out of the woods. True or false: Speaker B thinks everything is fine (false).

15. A: We're going to go and visit a local castle this afternoon. I don't suppose you feel like coming with us? B: I appreciate the offer, but I'm afraid I find old buildings **really** boring. True or false: Speaker A is planning a trip (true).

16. A: Did I see you at the fun run the other day? I thought they had fantastic weather for it, so sunny. B: It was good for those of us watching, but it was **terribly** hot if you were actually running. True or false: They are discussing walking (false).

17. A: You can quote Nietzche all you like. It will not change my mind, or my position. B: What if I start quoting Goethe instead, will that make a difference? True or false: Speaker B asks about quoting Goethe (true).

18. A: I'm considering a change of direction. I think I might look into performance poetry, something involving words. B: That's an interesting idea, and I think you'd be good at it. True or false: Speaker A wants to get into poetry (false).

19. A: I heard that, when she moved to London, Linda was surviving by selling her art on the streets. That must have been a difficult time. B: Perhaps, but I imagine it was also quite exciting for her. True or false: Linda sold art in London (true).

20. A: I saw Dale going over to the kids in the park. It looked like he was trying to teach them how to throw their frisbee. B: He can never leave something like that alone, he always has to get involved and show people the 'right' way. True or false: Dale was showing the kids how to trampoline (false).

21. A: For a while in my youth I was an apprentice upholsterer. I didn't see it through, in the end, but I still have an eye for upholstery. B: I wondered why you were giving those chairs such a hard look. True or false: Speaker A knows something about upholstery (true).

22. A: How many times have the window cleaners come around? I feel like they've hardly been here for months. B: They've actually been here about three times in the last two months. True or false: Speaker B says the window cleaners never come (false).

23. A: I'm not sure which website to use to store my photos. I used to use Flickr,

but that seems a bit outdated now. B: I just have everything on Dropbox, but I don't really share my galleries or anything like that. True or false: They are talking about websites (true).

24. A: I think I might order dessert, I'm still hungry. The menu mentions a dessert trolley, but I haven't seen it. B: I saw it earlier, it's definitely worth having a look because it had quite a range of things on it. True or false: Speaker A is not hungry (false).

25. A: We can only afford one type of snack for the function. The majority of people coming will be happy with crisps. B: The majority of people coming, sure, they'll be happy with crisps.

26. A: There's a specific dental product which makes the biggest difference. So many problems can be solved by using floss. B: So many problems can be solved by using floss, yes, I suppose that's true.

27. A: It feels like we're all moving towards the worst political system. Adopting communism never works. B: Adopting communism never ever works, I can't dispute that.

28. A: Rachel has gone to cast her vote in the election. She'll definitely vote for the Greens. B: She'll definitely vote for the Greens, like she always does.

29. A: Let's take the kids to their favourite theme park next year. That's definitely still Disneyland, I think. B: That's definitely still Disneyland, yes, they're still obsessed with that place.

30. A: I've started roleplaying and I've picked a character type. You can't go wrong with a wizard, you know. B: You can't go wrong with a wizard at all, good choice.

31. A: I can only afford to take up one form of exercise. You can get fast results with yoga. B: You can get fast results with yoga for sure, I felt it made a real difference when I took it up.

32. A: Of all things a husband or wife can provide, one stands out. A partner needs to provide support above all else. B: A partner needs to provide support above all else, you are one hundred percent right.

33. A: The lifestyle supplement has an article about the best writing implements. In terms of quality, fountain pens are the best. B: In terms of quality, fountain pens are the best, I agree.

34. A: The forest is a popular destination for a certain outdoor activity. The majority of people go for the hunting. B: The majority of people go for the hunting, and quite a few for fishing too.

35. A: A woman stopped me in the street and asked for the best route to the

beach. The nicest walk is through the park, by the river. B: The nicest walk is through the park, you're right there.

36. A: All the meadows are covered in the same flower. It's boring just to see endless daffodils everywhere. B: It's boring to see endless daffodils everywhere, but it's better than seeing nothing.

37. A: People want to know the essential makeup product for this season. They must wear lipstick above all else. B: Certainly, they must wear lipstick above all else, the redder the better.

38. A: They're showing something on TV about that bird which has really declined. It's very rare to spot a sparrow these days. B: It's very rare to spot a sparrow these days, yes, sadly.

39. A: The experts say a single disastrous weather event will happen soon. They're predicting a tsunami before the end of the year . B: They're predicting a tsunami before the end of the year, yes.

40. A: We should buy some containers for the kitchen. Getting a load of glass jars would be a good start. B: Getting a load of glass jars would be a good start, you're right.

41. A: It's strange to think that the best video website didn't exist when we were kids. It feels like YouTube has always been around. B: It feels like YouTube has always been around, yeah, it's funny that it actually hasn't.

42. A: That dog only eats one kind of meat, it's strange. He won't touch anything if it isn't beef. B: He won't touch anything if it isn't beef, yes, he's very fussy.

43. A: The bush in the church yard actually bears fruit. You can pick blackberries from it in the summer. B: You can pick blackberries from it, yes, did you not know that already?

44. A: The book I'm reading is about one type of mythological creature. The most interesting creatures in mythology are dragons. B: The most interesting creatures in mythology are dragons? Yes, I'll go with that.

45. A: I need to buy the best material to make boxes with. The strongest, cheapest solution is cardboard, right? B: The strongest, cheapest solution is cardboard, right.

46. A: I'm going to take a programming course, I want to learn the how to make websites. That will be worth a little bit of money, won't it? B: That will be worth a little bit of money, yes, but don't expect immediate job offers.

47. A: I have such difficulty with my skin—I wonder what the healthiest type of moisturiser is. Cocoa butter must be pretty healthy, it's so natural. B: Cocoa butter must be pretty healthy, I agree.

48. A: Who do you think the best band of the 60s were? The Rolling Stones were always the greatest. B: The Rolling Stones were always the greatest, I am with you on that.

49. A: I never went to any of the festivals, but I know which is the biggest. Glastonbury has always been the most well-attended, as far as I know. B: Glastonbury has always been the most well-attended, and with good reason.

50. A: I don't know what kind of hat to buy, but I need something casual. A baseball cap is probably the best bet. B: A baseball cap is probably the best bet, if you can find one you like.

51. A: I'm wondering which newspaper would be closest to my political views. In terms of liberal values, it must be The Guardian. B: In terms of liberal values it must be the Guardian, definitely.

52. A: The director of the film festival said they'd be showing the best Kubrick film. The best Kubrick film is 2001, hands down. B: The best Kubrick film is 2001, agreed—that must be what he meant.

53. A: If you're planning a holiday, you need to think about the best season in which to travel. Visiting Thailand in the cool season is the best, it's much nicer then. B: Visiting Thailand in the cool season is ideal, yes, I always go then.

54. A: What was that line from Eliot about the worst month? 'April is the cruellest month', isn't that it? B: 'April is the cruellest month', yes, it's an insightful bit of poetry.

55. A: I was checking out some old computer games magazines, it reminded me of the best games console ever. The original Playstation was just incredible. B: The original Playstation was just incredible, I loved mine.

56. A: Who was your favourite character in that film? The soldier who they met in the desert was fantastic. B: The soldier who they met in the desert was definitely my favourite.

57. A: I'll be bringing the car when we go to the cinema. Are you sure you don't need a lift? B: I'm **absolutely** sure I don't need one, I'm working next door to the cinema all day anyway.

58. A: I'm going back to carring a handbag again. I tried to go without one for a few weeks, decluttering my life, but it hasn't worked. B: Well, I can't live without mine.

59. A: Can you hear the air conditioning in your room? They say that there's actually a health impact, in terms of noise affecting people negatively. B: The air con in our room is **incredibly** loud, it's really annoying.

60. A: I got a new e-reader to replace the one I dropped in the bath. This one has much better contrast than the old one, which is great. B: That's good, the level of contrast is **very** important with e-readers.

61. A: They just unveiled a statue of Mohammed Ali outside the building where I work. Hearing about his life story again, it's inspiring. B: Yes, he **truly** is the greatest.

62. A: I bought a new water bottle at the weekend. It's quite an expensive one, but I think it will be worth it. B: In my experience, it's **very** much worth paying a little more for kit like that.

63. A: When I started teaching, I didn't realise the true nature of the job. I felt at first that I was just there to tell the students facts. B: That's a **really** common misconception, which is a shame.

64. A: Did you hear the guy busking in the station? He was playing the guitar in a way I've never heard before. B: That guy is **insanely** talented, I've never heard anything like it.

65. A: I didn't realise my tablet could accept such huge SD cards. Apparently the maximum limit is something like a terabyte. B: That's **ridiculously** huge, why would you need that on a tablet?

66. A: Remember when the boss promised us a special treat at Christmas? We all expected a bonus, and instead got that trip to the bowling alley. B: I remember all too well, it was **colossally** disappointing.

67. A: I can't stop watching that TV show. I don't know why I'm so hooked, the contestants are all so arrogant! B: That young one with the brown hair, she's **especially** smug.

68. A: The program you wrote for us has saved a lot of time. I can't help but cringe when think of all the time I wasted doing the routine manually. B: Well, it was an **unusually** long-winded process, so I'm glad to have helped.

69. A: I've brought the new mineral samples with me. I think you'll be excited when you see them. B: Yes, I heard that some of them are **remarkably** pure.

70. A: The customer service department are worried about the changes to the regulations. I've had several of them phoning me up with various concerns today. B: Tell them that the changes aren't **radically** different.

71. A: If you want to zone out, I'd recommend buying a fishtank. There's something relaxing about watching fish swim around. B: My problem is that I find fish **amazingly** interesting, and watching them keeps my brain whirring.

72. A: I just saw the film about that mountaineer who had to crawl to safety

with a broken leg. It really made me feel the fear and desperation he must have experienced. B: That man was **astoundingly** resilient; I would have died up there.

73. A: Brian has been complaining that he has no money since the holiday. That's why he's not coming out on Friday. B: He only has no money because he throws it around **excessively**.

74. A: Did you see the photos of dragonflies that my sister took? I think her work is really improving, the latest shots look fantastic. B: They do, and the insects turn out to be **extraordinarily** beautiful in close-up.

75. A: For some reason, my parents are going on holiday to Nevada this summer. They've booked tickets for July. B: They do realise that it's going to be **incredibly** hot there at that time of year, right?

76. A: I think the time has come for me to tell my parents that I'm getting married. As you know, they don't exactly approve of him, so it's going to be hard. B: Please remember that all your friends, including me, support you **fully** in going ahead.

77. A: Do you think that Martin is actually straight? I always assumed he was gay, but I saw him with a girl the other day. B: Well, being **outrageously** camp doesn't always mean someone is gay.

78. A: I've just picked up the keys to the new place. We'll be packing up for the next couple of days, then moving at the weekend. B: Good luck, I find moving house **phenomenally** stressful.

79. A: I couldn't believe what happened in the pub. How often does someone just whip out a gun whilst you're having a pint?! B: I admit, that was **quite** unexpected.

80. A: Maybe we can change Stephanie's mind about this? If she goes ahead with her plans, it's going to be harder for us to achieve our own goals. B: The impression I have is that her mind is **totally** made up, and she won't change it.

81. A: My thinking is that their plan was to steal the goods somewhere in Asia. Stealing them in Europe would have made it more likely our security would catch them. B: I think you are **somewhat** right with that, but not completely.

82. A: I don't think the casting is very realistic. What are the chances that every important figure in a medieval society would be so attractive? B: That king you were talking about, he's **strikingly** good-looking, isn't he.

83. A: If we spend more money on the finer details of the building, we won't

have much left to furnish it. We need good quality desks, chairs, and so on. B: Yes, but attention to detail is **supremely** important in a flagship project like this.

84. A: Louisa said the play was a big disappointment. She was saying that the script was just dull; flat dialogue, boring characters, the whole thing. B: In places it was **almost** innovative and interesting, but in the end, yes, it was disappointing.

85. A: I'm ordering the material for the curtains today. I just want to check with you, do you still think the shade we decided on is the right one? B: I think it might be **slightly** darker than is ideal, but that doesn't really matter.

86. A: Next time you go to the supply cupboard, could you get me some more folders? I'm close to running out of them. B: I'll have a look, but last time I was there the box of folders was **nearly** empty.

87. A: I like getting out into the natural world. Remember when we stayed in that cottage out in Wales? B: At night, it was **totally** black, so dark you couldn't see your hand in front of your face.

88. A: The last candidate we interviewed was the most promising of them all. I was impressed by the confidence she projected. B: She was also **uncommonly** intelligent, in my opinion.

89. A: You can spend too much time improving the workflow. Afer a certain point, it won't make a difference; you just need to start work. B: The tricky thing is to know exactly when that point arrives.

90. A: You wouldn't know it to look at them. They have actually been members of the club for years. B: Yes, that's quite surprising; I thought they weren't members at all.

91. A: You'll look like a princess holding the flowers. Like a Pre-Raphaelite painting. B: And what if I don't want to look like an old painting?

92. A: You've packed too many cartons of juice. You only need seven or eight, at most. B: You say that now, but when everyone is thirsty at breakfast time I think you'll change your mind.

93. A: The lecturer was talking about theories of nutrition. It's a field which is marred by a lot of psuedo-scientific nonsense. B: Yes, it's one of those areas that attracts a lot of new-age types.

94. A: I see endless quotations about 'respect'. I don't think half of the people who post this stuff think about what the word really means. B: Very few of us think deeply about meaning, you're right.

95. A: Have you ever had something really odd for breakfast? A while ago I

thought pasta for breakfast would be a good idea. B: I imagine you realised you were wrong quite quickly.

96. A: I can see some kind of grey smoke up on the horizon. I think that's near that village we visited. B: I suppose they're having some kind of bonfire, perhaps to burn bad crops.

97. A: Why is that machine making so much noise? It's always been relatively quiet, until today. B: It's been running continuously for several years, so it's probably just in need of a service.

98. A: I was accused of theft the other day. I popped into a newsagents carrying a drink I'd bought somewhere else, and they thought that I was trying to steal it. B: It's not a nice feeling to be accused of something you didn't do.

99. A: Have you been following the match? The way our defence is falling apart, I'd say Brazil are going to score again before the end of this half. B: If they do score again, it'll be a foregone conclusion.

100. A: So many aspects of computers have improved in the last twenty years or so. One thing that never seems to change is the basic layout of the keyboard. B: I think that's because it doesn't really require any improvements, it already works.

101. A: What is it that's so good about this album? I don't quite understand the appeal. B: For me the first thing I noticed was how they delay the vocals on most tracks, so it kind of builds anticipation.

102. A: Did you hear about that diplomat who killed someone? It's a very strange idea that you can murder someone, and not be punished at all. B: Yes, I've never really understood the whole concept of diplomatic immunity.

103. A: Do you recognise the style of that public artwork? I think it's the same artist who made the one we saw last week. B: Do you mean the guy who always does the big black cylinders?

104. A: Be careful in the lab when you're handling the glassware. Even the simplest components, like the beakers, are worth quite a lot of money. B: Don't worry, I know that glassware is expensive in this field.

105. A: I felt sorry for that character in the end. I don't think he ever saw himself as the bad guy. B: I think that's what was so clever about the show; they made you feel sympathetic towards him.

106. A: When the police came round, I felt relieved at first. Then, when I saw how young they looked, I couldn't really believe it. B: Absolutely, I've always had a hard time feeling reassured by someone ten years younger than me.

107. A: Tom said he wants to have a major career change. He said something

about conservation, I wasn't quite sure what he meant? B: He's been obsessed with conservation for the last five years, I think he wants to leave academia and be an eco-warrior.

108. A: Did you see the episode when they got stuck in the lift? It was one of those incredibly lazy flashback episodes where they just show old clips. B: I hate that kind of thing, it's such blatant filler.

109. A: She's an archeologist, she works in Peru. She actually gets to go into ancient tombs, open secret passageways, stuff like that. B: She must get to be the first person to see things which have been unseen for centuries.

110. A: Yes, I'm still wearing my pyjamas. After the day I had yesterday, which was just awful in every way, I think I deserve a lazy day. B: Wearing pyjamas for a whole day always makes me feel a bit like I'm in hospital.

111. A: I just can't stand the taste of coconut. That new soft drink they're promoting everywhere, the coconut water thing, it sounds horrible. B: I like the flavour of coconut, I think that's why I like Thai cuisine so much.

112. A: Where the hell have you been? I've logged on about thirty times and looked for you today everywhere, but I haven't seen you once. B: I was trying to have a day where I was off the grid from nine to five.

113. A: I saw that woman who used to present those cookery shows in the 90s. She was just coming out of Harrods. B: Is she the one who used to go on and on about using peanuts as ingredients?

114. A: What kind of cheese do you prefer? I used to be a big fan of cheddar, but these days I prefer French cheeses. B: Well, since I went vegan a few months ago, cheese is just a fond memory.

115. A: This place seems to be bad for mosquitoes. I was bitten so many times last night, despite using lots of citronella. B: Citronella doesn't work as a repellent, you need something with DEET in it.

116. A: The elders say that about two thousand of the tribe have died from an unknown disease. That's an appreciable percentage of the whole tribe. B: It is, but they have survived similar outbreaks before.

117. A: The trouble with the blossom is that it falls on the sundial. I have to go out several times a day to brush it clean. B: You could just leave it, there's a certain charm in letting things stay natural.

118. A: I can't believe this woman is being held up as a symbol of equal rights. Her own politics are utterly at odds with that. B: I suppose people don't really read her original writings, they just go along with majority opinion.

119. A: Do they still make netbook computers? I think a netbook might be the

perfect present for my husband. B: Smartphones and tablets have pretty much wiped out netbooks, but you can still get small laptops.

120. A: What do you think of my beard? I used to have one back when I worked in the factory, but it's been a long time. B: I think it looks good, but I'm surprised to see you following fashion for once.

Appendix B: Stimuli for Experiment 2

Test items

1. Two brothers were renovating their fruit shop, and deciding what kind of containers to put the discounted items in. The younger brother said they should put them/barrels by the tills. The older brother decided that *baskets* would be ideal. True or false: The brothers had a fruit shop (true).

2. Benjamin was out getting ingredients for dinner. He was making his wife's favourite dish, which he remembered as requiring plenty of preparation/sage. When he got back home, his wife said she hoped he'd bought the *cumin* that the recipe required. True or false: Benjamin was going to make lunch (false).

3. The telecommunications company wanted their new campaign to target the 'nerd' demographic. The managing director proposed working references to them/superheroes into the adverts. The marketers did a survey and decided to refer to *gaming* predominantly. True or false: They decided to put gaming references in the adverts (true).

4. A group of people in a retirement home were talking about the food they had been served for lunch last weekend. Most agreed that it had been meat/chicken. Later, someone found a copy of last weekend's menu, which clearly stated *mutton* was the main course. True or false: They were talking about breakfast (false).

5. A new lecturer started at the college, teaching art. From his first lessons, the students thought he had probably worked as a professional artist/painter before. In the final lesson of term he mentioned that he had been a *sculptor* before taking this job. True or false: The lecturer used to be a sculptor (true).

6. Following an inspection, it was decided that most of the road maintenance team required new clothing. The managers prioritised acquiring new gear/trousers, without consulting the workers. It turned out that they had ur-

gently needed their *jackets* to be replaced. True or false: The crew did not need anything new (false).

7. An archaeologist was asked what sorts of artefacts might be found at a new quarry site. She said that they would uncover lots of items/minerals. When digging started, *fossils* were being found regularly. True or false: They found several fossils (true).

8. The producer wanted new enemies in the game that were magic users. The lead designer thought the producer wanted suggestions/wizards, and came up with a few. The producer then sent round a memo saying how much he liked *witches* and how they should try designing those. True or false: The game was based on driving (false).

9. Steph's friend Nigel was allergic to a certain kind of food. Steph was fairly sure that he couldn't eat nuts/peanuts. She spoke to a mutual friend, who said that if Nigel ate *almonds* he'd be very unwell.

10. The personal trainer's client complained about pain when following the trainer's routines. As he expected, she had discomfort in her joints/elbows. Later, a different client told the trainer he'd had problems with his *ankles* after doing it.

11. Alex was buying Robert some jewellery, and was thinking about which precious metal to go for. He was sure that Robert really liked rings/gold. That evening, Robert mentioned how *platinum* was the most attractive metal.

12. Victoria was helping her uncle move boxes of art equipment into his new studio. Victoria had presumed the majority of boxes would be full of things/paints. She ended up spending most of the day lugging boxes full of *brushes* up the stairs.

13. A rehabilitation program was launched that would place ex-offenders in community service roles. Community leaders feared that criminals/murderers would be assigned to their neighbourhoods. They later found out that *burglars* made up the majority of the program.

14. A small city zoo received a large donation with the proviso that it went towards acquiring new animals. The zoo management felt that mammals/bison would provide value for money. Most of the keepers were of the opinion that *camels* would be popular.

15. The new inmate was keen to work out where to buy drugs. His cellmate told him that a lot of deals went down in the prison/yard. A year into his sentence, he knew that the *canteen* was the best place to score.

16. Sarah's dad was a huge F1 fan, and she'd always wondered what in partic-

ular kept him interested. When it was televised, he seemed to show endless fascination with the coverage/cars. One day, when she'd popped round, her dad told her that the *crashes* were what kept him glued to it.

17. The frescoes in the old church included various supernatural beings. The priest invited an art historian to take a look, expecting her to find the artwork/angels especially interesting. The historian spent most of her time looking at the *demons* in the frescoes.

18. Hannah had a real thing for men who worked in stereotypically macho positions. Her best friend thought she'd get married to a musclebound type/policeman. In the end, she was chief bridesmaid when Hannah was married to a *fireman* she met on holiday.

19. The pet shop assistant prided herself on figuring out what sort of product any given customer would want. That morning, she told a small boy approaching her where the pets/
hamsters were before he spoke. He then asked her if the *goldfish* were expensive.

20. Something pale and white was growing on the stones at the cave entrance. Simon guessed it was a type of mushroom/moss and took a sample. When a botanist friend of his saw the sample, she said it looked to her like a *fungus* native to the area.

21. The artificial intelligence algorithm was being trained to recognise cats. The programmers thought that the image/ears of the cat would be the most important identifier. The most successful version they came up with looked for *whiskers* before anything else.

22. A gardener at a stately home was deciding what to use to divide one field into two. The owner said some consideration/fences would be useful. When the time came, he used *hedges* to separate the fields.

23. They would need to bring light sources for the parade. Grace suggested everyone else would have a light/torch of some type. On the night, they saw that nearly everyone had a *lantern* with them.

24. The Spanish farmer wanted to start an orchard. He decided a patch of land close to his house would be best suited to growing citrus/oranges. He told his wife this, who said that *lemons* would be ideal for that position.

25. At the start of the show, the celebrity chef said she'd tell viewers how to make a perfect version of a dessert. Watching, Rachel thought it would be a traditional classic/pavlova. The chef demonstrated how to make a *pancake* at the end of the show.

26. Melissa's parents always knew she would grow up to be a scientist. They were sure she would become a researcher/physicist. Many years later, they were proud when she became a *chemist* of some renown.

27. A journalist was writing a story about lost property at airports. She supposed a lot of things/wallets would be handed in. When she contacted a few airports, they all said that *passports* were the most common item although they were swiftly returned.

28. Laura went to see a talk on the architecture of Greek temples. She was anticipating a lot of information on the way features/lintels were designed. The speaker spent almost the entire talk going into a lot of detail about *pillars* and their importance.

29. A book that examined the role of vehicles in modern culture became very popular. The author anticipated strong reactions to the position he adopted on design/tires. He soon found himself receiving lots of emails about *motors* and his claims about them.

30. Hugh was trying to learn about the items in the big new PC game. The first time he played, he assumed the armour/chainsaw was the key item. In an in-game text chat later, his friend Mike, an expert at the game already, told him the *shotgun* was the best weapon.

31. The TV station had to develop Christmas channel idents, and this year they wanted to use a familiar Christmas character. In the first brainstorming meeting the idea of using a cartoon/reindeer was popular. By the time the idents were animated, a *snowman* was the central character.

32. The school librarian wanted to provide children with rewards for finishing books. She argued that giving the children incentives/sweets would encourage them. A child psychologist, temporarily based at the school for his research, thought *stickers* would be the right choice.

33. Tessa was doing up her bungalow, and wanted to resurface the drive. Friends recommended using contractors/gravel for it. Her cousin, who was in construction, said that *tarmac* was the most suitable surface.

34. The kitchen staff were making a range of cheese sandwiches for the cafe. Going off previous sales, they decided to make a lot of sandwiches with haste/tomato. That lunch, they found they ran out of the sandwiches that had *chutney* in them.

35. A backpacker was wondering where to sleep. He really fancied spending the night in a building/tent for a change. After a chain of complicated events and no small amount of frustration, he found himself sleeping in a *hostel*

that night.

36. The gym sold various accessories that members might have forgotten to bring. Nancy imagined that they shifted a lot of stuff/deodorant. The receptionist told her that *towels* outsold everything.

37. Five friends were planning a walk in the country, and discussing what to wear. The most outspoken one was advising them to wear good raincoats/shoes. It turned out one of the group had done the route a while ago, and they said *wellies* were required.

38. Susan's son was ill and stuck in bed, so she went to the newsagents to get him something to cheer him up. She thought he'd appreciate snacks/magazines, so bought a few. When she was opening the door he shouted down asking if she'd got any *comics* for him.

39. Tina's mother-in-law was visiting, and Tina wanted to stock up on things she liked. As far as Tina remembered, her mother-in-law usually had a drink/sherry before bed. Her husband told her that she liked to drink a *whiskey* last thing at night.

40. A graphic designer had to find an image to symbolise the Netherlands for an advert. He thought he would grab a photo/tulip from an image library, and asked his boss her thoughts. She sent a message saying that a *windmill* would suit the advert well.

Filler items

1. A wedding was being held in a registry office. Only a few family members on both sides were present, and it was clear from the atmosphere in the room that many of the guests were **deeply** suspicious of this union. The bride and groom appeared oblivious to the tension behind them. True or false: The ceremony was happening at a registry office (true).

2. Stephanie's plans for the conference were at best ill-advised, and at worst would put the continued success of the company in real peril. Her subordinates had been trying **unusually** hard to find some way to stop her— but their efforts were doomed. True or false: Everyone was positive about Stephanie's plans (false).

3. The new teacher was nervous about the first assignment of the term. She had been worried that she wouldn't have a standard by which to mark the work, but then a colleague emailed her a set of guidelines. She felt **incredibly** relieved. True or false: The teacher was relieved (true).

4. Some of the magazines in the dentist's waiting room were **incredibly** old.

Whilst he was waiting, twelve year-old Harry picked up a car magazine. The cover date was a couple of months before he was born. True or false: Harry looked at a fashion magazine (false).

5. Diana took her date to a lesser-known part of the park, which had beautiful fountains and tall walls. She was slightly irritated when he pointed out a large rat that was scuttling around the flower beds, but, on the whole, the mood was suitably romantic. True or false: Diana went on a date (true).

6. There was a long running dispute between the two pubs in the village. Each claimed to be the oldest, and each had multiple pieces of evidence that supported the claim. Local historians seemed unable to resolve this. True or false: Everyone knew which pub was the oldest (false).

7. There was a smell in the apartment that nobody had been able to identify. It wasn't unpleasant, but it was unusual, and everyone who visited said that it reminded them of something they couldn't place. True or false: The apartment smelled strange (true).

8. Somebody had left the door to the bar open. It was getting quite cold, and some of the customers were passive-aggressively muttering, but none of them wanted to get up to close it. True or false: The bar was getting hotter (false).

9. A production company started making a new TV show. The show was aimed at young children, and featured an animated bird that interacted with live studio guests. The technology was impressive, but expensive. True or false: The TV show was targeted at children (true).

10. Dave had been having trouble with his emails. Every time he flagged a message on his laptop, it didn't seem to alter the same message when he looked on his phone. It was really irritating him. True or false: Dave never had issues with his email (false).

11. Kerry needed new whiteboard markers. Popping into the supermarket, she was pleased to find some refillable ones. She bought five different colours, and several refill cartridges. True or false: She went to buy new markers (true).

12. When the family were living in New Zealand they were close to a glacier. They would often take a trip to the mouth of the glacier and watch the brightly-coloured specks that were tour groups ascending the lower parts. The sense of scale was awe-inspiring. True or false: The family lived in Papua New Guinea (false).

13. At 6am the astronauts were driven to the launch station. Onlookers might

have expected them to appear nervous, but their training had been so thorough that they were comparatively calm. True or false: The astronauts went to the launch station in the morning (true).

14. One of the most dangerous pieces of equipment in the factory was the high-power laser cutter. After a few near-accidents, management restricted its use to a handful of senior employees. True or false: The laser cutter was not dangerous at all (false).

15. Visiting his parents' house, Billy found a CD wallet in his bedroom. The CDs inside were full of MP3s he had downloaded as a teenager. He loaded them up and was appalled at his youthful music taste. True or false: Billy's music taste had changed since he was younger (true).

16. The court clerk had sat through so many cases that she had developed a good sense for when someone would be found guilty. However, she was still wrong about a third of the time. Of course, she kept her predictions to herself. True or false: The clerk was never wrong about the cases (false).

17. The results of the school inspection were **much** better than predicted. The teachers, who had been preparing for the worst, were cheered to find that overall performance had improved. Addressing them the next day, the Head advised cautious optimism.

18. Brian had been complaining vociferously to anyone who would listen that he was low on cash. He got a little sympathy from the new guy in the office, but most of the old hands knew that Brian was only broke because he spent his money **very** carelessly.

19. The amateur photographer was **extremely** interested in dragonflies, and spent a lot of his time lurking near lakes and ponds to get the perfect shot. He found the insects to be truly beautiful in closeup, and had almost infinite patience in trying to capture images of them.

20. Two girls were chatting on the bus about their plans for summer. One was laughing about her parents choosing to go on holiday in a village an hour from where they lived. Her laughter was **incredibly** loud.

21. The population of sea birds on the estuary had been fluctuating **massively** for a decade. Sanjay, a conservationist working at the nearby university, was of the opinion that climate change was responsible. He started keeping detailed records of temperatures.

22. Martin was one of those men whose flamboyant manner had people labelling him as gay from a young age. He found this **very** irritating, as he considered himself essentially asexual. He wished that people would just

stop speculating.

23. Having run out of space in their first home, the young family were moving into a roomier property a little further out of town. The father was finding the whole process **incredibly** stressful, and had come down with a bad cold that may have been exacerbated by the stress.

24. It was a quiet Wednesday evening in the small country pub. The place was empty except for a couple of regulars and their dogs, which were lazing by the fire. At 9:30, the peace was shattered by the arrival of an **insanely** drunk man brandishing a large plastic sword.

25. Sandra's latest Air BnB guest was a shy middle-aged man of indeterminate nationality. On Friday night, Sandra asked him if he would like to join her and her friends visiting a beach the next day. To her surprise, he accepted the invitation **very** enthusiastically.

26. A small stall had recently opened halfway up the mountain. It was selling **very** cheap bottles of water that had a badly-printed logo which smudged in sweaty hands. Most tourists were dubious about the product, but the stall was doing a good trade.

27. Aaron had moved from Western Canada to Northern England to settle down with his English wife Heather. Shortly after he made the move, the couple were **absolutely** delighted to find out she was pregnant with their first child. Their daughter was born at the end of summer.

28. Headphones were one of those things that invited considerable debate in some circles. The rise of fashionable premium brands had left many audio enthusiasts **hugely** frustrated, as they felt that style was being valued more highly than audio quality. Fans of the brands deemed this snobbery.

29. When Ian first started playing the game, it was just as a way to relax after work. After a while, though, he became **thoroughly** addicted. He came to the point where he was considering installing it on his office computer—and that was when he knew he had to stop.

30. The software the small business used to do their accounting was custom-made, and did the job **very** well. Management were fond of boasting about how much better it was than Excel.

31. A friend of Beckie's consistently posted ridiculous nature photos on Facebook. Beckie had long since given up pointing out that the purported 'natural' pictures were **heavily** photoshopped. She just rolled her eyes whenever one popped up on her feed.

32. Bobby worked at a large warehouse that stocked carpets and soft furnish-

ings. Sometimes he would take his son to the warehouse; his son, who was autistic, **really** loved running his hands over the rolls of carpets, feeling the different textures.

33. The issues in their relationship were, as is so often the case, related to communication. Although they talked to each other constantly, neither of them fully understood the other's perspective. This made for frequent arguments, and both of them feeling isolated and alienated.

34. When the friends arrived in Bali, they discovered their visit coincided with a festival. This was excellent news in terms of seeing some additional local colour, but bad news in terms of most restaurants and bars being busier than they might have been.

35. Martina was a big advocate of using toothpaste to polish things. She had already showed her friends how good it was for removing scratches from watches, but she did have to admit that the minty scent tended to linger.

36. Rainy day commuting had been even worse since his shoes had started to leak. Reluctant to shell out for a new pair so far from payday, he had realised that the damp feeling could be alleviated by going without socks.

37. The new horror film was well-reviewed, and had a decent performance at the box office. Critics and audiences alike seemed to respond to the quality of dialogue, and the reliance on psychological horror over cheap scares.

38. Having spent a long time trying to get his literary fiction published, James decided to try writing a science fiction novel. He tried to write it with a degree of literary sophistication, but also concentrated on including tropes typical of the genre. When he showed it to his friends, the response was mixed.

39. On the first day of the training course, the instructors began with a lecture about the importance of organising one's time. They talked extensively about the idea of 'time thieves', and presented a set of strategies to deal with these people.

40. Every time Michael put on his baseball cap, he went through a moment of doubt. He was never quite sure if it was currently more fashionable to wear it forwards, or backwards. He normally went for backwards.

41. Paul had been a high-flying trader in the City of London for nearly all his adult life. In the last few years, however, he had been experiencing the nagging feeling that the job was at odds with his strong religious beliefs.

42. Sarah was accumulating loyalty cards at an ever-increasing rate. Despite this, whenever she went to the nearest branch of her preferred coffee chain,

she never seemed to have her card with her.

43. Before he got into interior decoration, Barry had been an upholsterer. Clients often saw him giving a long, hard look at their furniture. His designs for rooms often centred on interesting upholstery.

44. When the songwriter first played the band his new song, everyone agreed that it was promising, but seemed to lack something in the verses. The drummer thought a descending bass run might add interest.

45. When he left home one morning, Richard noticed that his neighbours had put a sofa out in the street. Knowing them, he imagined it was going to be there for quite a while. It turned out he was correct.

46. The group had started secular but spiritual. Over time, it began to move towards being more and more religious in tone. Some members started to feel as if they were part of a cult.

47. The radio host was the sort of person who repeatedly claims to hate puns, but repeatedly makes them. After a while the station director noticed an increase in complaints about his pun-heavy chatter.

48. An American family were visiting the UK over Christmas. They were intrigued by the idea of attending a pantomime. They were disappointed to find out that there were no pantomimes showing anywhere near their accommodation.

49. Deborah hadn't spoken to anyone from South Africa before, so when she met Nina she subjected her to a stream of questions. She was particularly interested in South African slang. Nina was pleased to provide some vivid examples.

50. Arthur didn't tell his sister that he was seeing her favourite band, because the show was sold out and his girlfriend had bought him the tickets as a present. This didn't prevent him becoming very guilty about it.

51. Jess didn't want to do the washing up. She knew her mum was going to ask her to do it. She also knew that if she held off long enough, her brother would have to do it.

52. The fortune teller told Jessica that her next relationship would end badly. Other people might have been unsettled. Jessica, however, didn't really believe in psychic predictions.

53. A medium-sized town in the desert was expanding rapidly. People were coming to the area to work in nearby mines. This soon resulted in pressure being put on the limited water supply.

54. Even though Rebecca's friend Charles was quite brilliant in academic terms,

he was terrible at communicating his understanding. When he started applying for lectureships, she knew he would struggle.

55. The spreadsheet program had a variety of options to sort data. One extremely useful option was to format a set of cells as a table, which enabled each column to be sorted with a single click of the mouse.

56. The spiritual guru was lecturing his followers on the concept of 'regret'. His thinking on the matter was rather unorthodox: he advised the group that they should seek to be as regretful as possible.

57. Leaving the club, the group of friends had a debate in the street. They could either wait twenty minutes and get the night bus home, or they could flag down a taxi. They already had bus tickets, but ended up getting the taxi.

58. Having taken redundancy, Ellen was enjoying not getting up early in the morning. Her new routine was to sleep late and then head to a coffee shop. In the shop she'd drink lattes and browse job sites.

59. As a child, Sasha had read a book about Japan. It sparked a long-term fascination with the country's architecture. She grew up to be an architect, and her work frequently drew on the inspiration she took from Japanese buildings.

60. The security guard carried a hipflask with him at all times. Speculation was rife that he drank alcohol throughout the day. In fact, the flask was empty—he just carried it as a good luck charm.

61. When Ray told Claire he believed in the afterlife, she asked what kind of place he expected to live in. She expected him to give a typical account of heaven without thinking. Instead, he fell silent and looked very thoughtful.

62. Liam didn't have a legitimate excuse for the low quality of his work. The rest of his team were fairly sure he would be fired, and most were relieved. He hadn't been a very popular member of staff.

63. On a trip to a safari park, Kelly snapped several pictures of the gorillas. In one, the silverback appeared to be giving a 'thumbs up' sign. Kelly thought the image stood a chance of going viral online.

64. Yasmin's mother owned an old rug. The rug hung on the living room wall for most of Yasmin's childhood. Her mother was unsure of the rug's provenance, but suspected it came from Iran.

65. Germaine's great grandfather was a bareknuckle boxer. Going off the records, it didn't seem like he'd had a lot of success. Despite this, his career lasted for fifteen years.

66. The team's first experiments with adding nitromethane to their fuel were

unpromising. It didn't seem to make any difference to the engine. They gave up and looked into other methods of speeding things up.

67. Julia was listening to a podcast all about different drinking vessels. It was actually quite interesting. The history of the pint glass was particularly exciting.

68. Both halves of the couple had exactly the same anxieties about their future together. Neither of them wanted to get married, but both of them feared the other would ask.

69. Kate read an article about how important it is to reduce the size of your email inbox. She started following the advice, and found that it actually helped. Soon she was far less stressed when opening Outlook.

70. Bradley got heavily into board games when he was at university. He amassed quite a collection of different games, but on completing his studies found it harder to find people who were interested in playing games with him.

71. Surprisingly, the school cafeteria often didn't have a vegetarian choice on the menu. A group of students complained to the headmistress about this, and soon managed to get this fixed.

72. The typical lifestyle in the mountain region had revolved around keeping livestock. Nowadays, however, only a few shepherds remained in the area, and most people worked in the nearest city.

73. At the theme park, Jenny and her mother spent a huge amount of time queuing for the new rollercoaster. Despite the lengthy wait, when they finally got to ride it, their expectations were exceeded. It was genuinely thrilling.

74. For years, the brick sides of the bridge had been plastered with elaborate graffiti. The people living in the area rarely saw the perpetrators, who tended to work under cover of darkness. At least their artwork was quite skillful.

75. Nature documentaries always sent Louis to sleep. Everyone would be talking about the latest incredible footage the next day, and he would have dropped off ten minutes in.

76. The band's guitar tech had to carry a soldering iron with him on tour. The lead guitarist had a knack for damaging the cable jack on his main guitar, and it had to be fixed before most gigs.

77. A very well-respected actor was cast in a cheesy blockbuster. Reviews were universally terrible, but many of them praised the actor's attempts at giving the role some dignity. His reputation had survived.

78. An IT specialist was relocating from Florida to Boston for work. One of the

many things he had to take care of on arrival was buying suitable clothing for the colder climate. Fortunately, he quite liked the cold.

79. At Easter the family had a tradition of devising intricate egg hunts. The hunt typically involved deciphering a set of handwritten clues, which would take the form of rhyming riddles. The younger children were often tempted to dispense with the clues and ransack the house.

80. The small Caribbean nation consisted of one large island and several smaller atolls. The large island was dominated by a huge active volcano. Whilst the volcano had been quiet for nearly thirty years, locals knew that it was simply a matter of time before the next major eruption.

Colophon

This thesis was set in Charis SIL using the XeTeX typesetting engine. It was written in the Pandoc Markdown format, using the Vim text editor. The charts and graphs in Chapters 4 and 5 were created in the R language (R Core Team, 2017), using the packages *ggplot2* (Wickham, 2009), *egg* (Auguié, 2017a), *gridExtra* (Auguié, 2017b) and *extraFont* (Chang, 2014).