Chapter 1

Introduction to Eye-Tracking

1.1 What Is Eye-Tracking?

Have you ever watched someone's eyes as they read or look at a scene? What do you notice about the eyes' behaviour from simple observation? There are descriptions of eye-movements, as well as what they might tell us, that go back as far as Aristotle (Wade, 2010). However, this kind of observation only provides a limited understanding of eye-movements, and historically led to the erroneous conclusion that eyes 'glide over scenes and to alight on objects of interest, which they would fix with unmoved accuracy' (Wade and Tatler, 2011, p. 37). It was not until the late nineteenth and early twentieth centuries, when early eye-trackers were developed, that we gained insight into the rapid discontinuous nature of eye-movements.

> We fancy that we can move our eyes uniformly, that by a continuous motion like that of a telescope we can move our eyes along the sky-line in the landscape or the cornice of a room, but we are wrong in this. However determinedly we try to do so, what actually happens is, that our eyes move like the seconds hand of a watch, a jerk and a little pause, another jerk and so on; only our eyes are not so regular, the jerks are sometimes of greater, sometimes of less, angular amount, and the pauses vary in duration, although, unless we make an effort, they are always short. During the jerks we practically do not see at all, so that we have before us not a moving panorama, but a series of fixed pictures of the same fixed things, which succeed one another rapidly. (Brown, 1895, pp. 4–5)

The 'twitches' and 'jerks' of the eye – in other words their movements – are what we commonly refer to as *saccades*. The interval between the eyes' movements, when the eyes 'stop', are called *fixations*. Both are a type of automatic, physiological response, which means that they are not under our conscious control (Rayner et al., 2012). In reading, saccades do not always move the eye forward in a text. About 10–15 per cent of the time, readers move their eyes back (regress) to previously encountered sections of text. These backward movements are referred to as *regressions*. Regressions can be short or long. Short regressions are usually due to overshooting a target. If we look at the first example in Figure 1.1, the intended eye-movement was to the word 'normally'. However, the eye 'overshot' and landed on 'safe'. The short regression moves the eye back to the intended target. Longer regressions are usually attributed to the difficulty of the text, which can be due to a range of factors. Looking at the second example in Figure 1.1, a reader might not associate 'infect' with 'violence' and go back and check that the word was indeed 'violence' and not perhaps 'virus' instead.

In very basic terms, when we read or look at a scene or images, our eyes stop to process the information at that location and then move to another point where other information is available. During fixations, the cognitive system perceives and processes the visual input, as well as planning when and how far to move the eyes next. Under most normal circumstances, during a saccade the eyes move so quickly that we do not obtain new visual information (Rayner, 2009). While new visual information is not encoded during saccades,

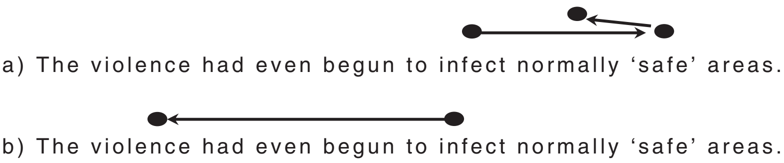


Figure 1.1 Examples of regressions due to different factors: (a) overshooting the target; and (b) difficulty. The oval depicts a fixation, while the arrows illustrate saccades, with the right-pointing arrow indicating a forward saccade (generally simply referred to as a saccade) and the left-pointing arrows indicating a backward saccade (referred to as a regression).

the processing of already perceived information can continue (Irwin, 1998; Irwin and Carlson-Radvansky, 1996). Fixations, saccades and regressions generally occur ‘automatically’, without our conscious awareness. Thus, tracking eye-movements provides a window into a largely unconscious behaviour. Crucially, *eye-tracking* technology tells us where people’s eyes land, how many times they land in that position or region (*fixation count*), and how long each fixation lasts (*fixation duration*), as well as measuring saccade duration and length. To sum this up simply, eye-tracking is a technology that measures fixations, saccades and regressions in response to visual input, while an eye-tracker is the device that does this. The data that is produced by an eye-tracker will depend on what we are using it for, and we go into much more detail about the different eye-tracking measures throughout the book, and qualitative analyses (heat maps, cluster maps, scan paths) that can be undertaken in Chapter 7.

Why would being able to track and measure people’s eye-movements be of any interest to applied linguists? It appears that in complex processing tasks, like reading and scene perception, eye location provides an index of attention (Rayner, 2009). This means that our eyes indicate what we are paying attention to and how much cognitive effort is being expended to process the input at the fixation point. Thus, the difficulty and complexity of what the eyes are looking at influences fixations and saccades (Castelhano and Rayner, 2008). When the input is more difficult, fixation durations and regressions increase, while saccade size decreases. This means that in reading, more difficult texts elicit more and longer fixations and regressions, while saccades get shorter. When looking at scenes or images that are more crowded, cluttered or dense, fixations also get longer and saccades get shorter. It is important to note that the difficulty we are talking about here is in ‘global’ terms – so it is a property of a text, image or scene as a whole. We can also talk about ‘local’ properties, and in reading we can look at effects of individual words or short sections of text. For example, we can look at eye-movements to a specific word, like ‘infect’ in Figure 1.1. There are a set of standard measures that are used in eye-tracking to look at local effects; these measures will be presented in detail in Section 3.2 and will be the main focus of the discussion in the remainder of the book when talking about reading tasks. It is important to keep in mind that factors other than difficulty can influence both global and local eye-tracking measures – for example, the reader’s/viewer’s goals (Rayner and Pollatsek, 1989). Reading a text for understanding produces a different pattern of eye-movement behaviour from skimming a text, as does looking at images on a page to understand what is being depicted compared to memorising images and their location.

Global eye-tracking measures give us an indication of typical looking behaviour. An overview of fixation time and saccade length and duration for silent reading, oral reading and scene perception is presented in Table 1.1. It is important to keep in mind that

Table 1.1 ‘Global’ characteristics of fixations and saccades in reading and scene perception for skilled English readers (based on Castelhano and Rayner, 2008; Rayner, 1978, 1998), with line indicating unspecified information.

Task	Mean fixation duration (ms)	Mean saccade size	Mean saccade duration (ms)
Silent reading	225–250	2° / 7–9 letter spaces	30
Oral reading	275–325	1.5° / 6–7 letter spaces	—
Scene perception	260–330	4°–5°	40–50

this table simply depicts average behaviour. For example, saccades can span anywhere from two to more than eighteen characters and fixation duration can vary from 50 to 600+ ms for a single reader in one passage (Rayner, 1998, 2009). The table highlights that there is a considerable range in terms of fixations and saccades depending on the visual input and the task. For skilled, adult readers, a scientific paper will elicit longer fixations and shorter saccades than a Harry Potter novel. Crucially, these global measures will underscore group differences. If we consider a standard newspaper article intended for a general audience, the text should not prove challenging for educated native-speaker adult readers with no history of vision or language processing difficulties. This same text will be more challenging for ten-year-olds, less proficient non-native speakers, participants with dyslexia, etc. For these groups, we would expect more and longer fixations, more regressions, and shorter saccades. Although it has not been rigorously investigated (yet), it is likely that global eye-tracking measures can provide a metric for identifying the skill of readers. In other words, because global measures provide a metric of ‘difficulty’, we would expect a highly proficient group of non-native speakers to have fewer and shorter fixations, fewer regressions and longer saccades than less proficient non-native speakers on a particular text.

As Table 1.1 indicates, eye-movement behaviour varies depending on whether participants are reading silently or orally or viewing a scene. As Rayner (2009) points out, ‘It is actually somewhat hazardous to generalise across these tasks in terms of eye-movement behaviour’ (p. 1459). He goes on to hypothesise that the differences in eye-movements are due to the cognitive mechanisms involved in the different tasks, as well as because of the interaction between the cognitive and oculomotor systems differing as a function of the task. Again, we can ask ourselves why this might be important for applied linguists. In general, applied linguistics work with ‘real-world’ and ‘authentic’ materials. This might involve presenting children or second language learners with a reading test, a range of readers with a story or novel, storybooks accompanied by images, films with subtitles, etc. Chapters 4–6 will discuss some of these types of authentic materials and the methodological considerations when using them. Crucially, if we use materials that have image and language components, we would expect different patterns of fixations and saccades for the image and text portions of the stimuli simply because the eyes behave differently when they encounter these sources of information. This makes any direct comparison of images and text somewhat complicated.

Further, Table 1.1 points to some specific differences between tasks. First, fixations tend to be longer in oral reading than silent reading. Skilled readers can read words more quickly than they can say them aloud, which means that our eyes outpace our word production. In oral reading, our eyes fixate longer and make shorter saccades, likely so that they do not get too far ahead of what we are saying. Second, fixations and saccades in scene perception

210 172 386 172 133 187 180
The most recent statistics showed that crime rates were

253 400 335 197 260 285 225
increasing. The epidemic was out of control. The violence had

179 333 172 179 375 267
even begun to infect normally `safe´ areas. The police were

180 130 192 375
desperately looking for a cure to the problem.

Figure 1.2 Hypothetical fixations for a skilled reader during silent reading. Fixations (in ms) are indicated above the word where they occurred.

tend to be longer than in either type of reading because the eyes take in information from a wider area in an image than in reading. Again these differences highlight potential issues with comparing eye-movement behaviour across task types and different visual sources.

It is important to note that the averages presented in Table 1.1 and the discussion thus far have been focused on skilled readers in English. Crucially, eye-movement behaviour is comparable in skilled readers of other similar alphabetic languages. However, in a language like Hebrew, where information is more densely packed than English, largely because vowels are not generally spelled out, there tend to be shorter saccades (about 5.5 letter spaces), while fixation durations are similar (Pollatsek et al., 1981). In Chinese, which has a very different writing system, again average fixation durations are quite similar to those presented in Table 1.1 and regression rates do not differ markedly (Rayner, 2009). However, average saccade length is much shorter and is typically only two to three characters.

The global effects that have been discussed thus far with regard to reading are informative because they tell us something about overall performance and they can be used to differentiate difficulty in terms of texts and for various groups of readers. However, they are somewhat limited because they do not tell us where a difficulty may arise. As mentioned above, they do not reveal any local effects. More precisely, in a text, some words will be read more than once, while others will not have a fixation associated with them at all – in other words they are skipped. There is good reason to believe that skipped words are processed on the fixation prior to the skip (Rayner, 2009). English readers can acquire useful information from an asymmetrical region around the fixation point that extends about three to four letter spaces to the left of the fixation and fourteen to fifteen character spaces to the right. Information that can be used for word identification is obtained from a smaller region extending to about seven to eight character spaces to the right of a fixation. If we look at the text in Figure 1.2,[1] this means that when the eyes land on the word ‘even’, they will also see and identify the word ‘begun’. This will make a fixation on the word ‘begun’ unnecessary. Unfortunately, global measures (e.g. number and length of fixations on a text as a whole) do not provide this level of information. It is the local measures that provide more precise information about reading and viewing behaviour.

[1] This passage is loosely based on the materials from Allbritton, McKoon and Gerrig (1995). They investigated metaphorical language processing, but did not make use of eye-tracking. We will use this example passage, or parts of it, throughout the book.

Figure 1.2 can demonstrate some important differences between a global and local examination of eye-movement data. We can see that there are twenty-four individual fixations to the text and can calculate the mean fixation duration, which is 241 ms. Based on the information in Table 1.1, this is within the expected range of fixation duration for a skilled reader during silent reading. The global data would be useful for comparing skilled and unskilled readers. Thus, the expectation would be that unskilled readers would have more fixations and that their mean fixation duration would be longer. However, what a global assessment will not tell us is that for our (hypothetical) reader, the words 'statistics', 'epidemic' and 'infect' all have two fixations, and that the fixations at the end of sentences, where sentence integration occurs, are all above the expected average. Local measures – the fixation pattern to individual words or small regions – would be used to explore processing effort within the text itself.

Finally, before turning to the next section, we will briefly discuss scene perception. As we will see in Chapter 5, we do not talk about eye-movement measures in terms of being global and local when examining looking patterns for images. However, as with reading, the average values depicted in Table 1.1 vary depending on the task and the exact nature of the scene. Viewers do not fixate every part of a scene (Rayner, 2009). It appears that viewers can very quickly obtain the gist of a scene in a single glance in as short as 40 ms (Castelhano and Henderson, 2008; De Graef, 2005). This initial fixation is used to point to the appropriate or interesting regions for subsequent fixations (Rayner, 2009). Overall, viewers primarily fixate on informative areas of a scene (Antes, 1974; Mackworth and Morandi, 1967), as well as on salient aspects, which are generally defined in terms of features like contrast, colour, intensity, brightness, spatial frequency, etc. (Mannan, Ruddock and Wooding, 1995, 1996; Parkhurst and Niebur, 2003). Viewing is also influenced by the task and by real-world knowledge. For example, in a visual search task of a scene containing the sky and a road, when participants are asked to look for a 'jeep', fixations are largely constrained to the road, which is where a jeep is likely to be found (Neider and Zelinksy, 2006). Finally, viewers tend to fixate near the centre of an object (Henderson, 1993), and when looking at scenes they look at people (or characters like Mickey Mouse) and concentrate their fixations on the face (Henderson and Hollingworth, 1999).

1.2 Why Use Eye-Tracking?

> It is quite apparent that research utilizing eye-movements to examine cognitive processing tasks is burgeoning as more and more researchers have started to use eye tracking techniques in the last few years. (Rayner, 2009, p. 1458)

The quote by Rayner describes the rapid increase in the use of eye-tracking in cognitive psychology at the turn of the twenty-first century. This sentiment is echoed by Liversedge, Gilchrist and Everling (2011) who point out that in the early 1980s the number of eye-tracking laboratories in psychology departments in UK universities could be counted on the fingers of one or two hands, while three decades later almost all such departments had an eye-tracking device. What led to the rise of eye-tracking research in psychology, and why might this be relevant to applied linguists? Part of the explanation for the spread of eye-tracking is a practical one. Eye-tracking systems have become more readily available, cheaper (although they can still be quite expensive), and more user-friendly (even though they might not seem so upon a first encounter). While affordability and usability are not reasons to use eye-tracking technology, they mean that eye-tracking systems are increasingly accessible to applied linguists.

The main driver behind the rise in eye-tracking research in psychology and psycholinguistics is the belief that there is a tight relationship between eye-movements and cognitive processing. Put another way, there is a belief that the eyes provide a 'window' into the mind, which is sometimes referred to as the *eye–mind hypothesis* or *eye–mind assumption* (Just and Carpenter, 1980). This means that eye-movements tell us about cognitive processing. Importantly, as was discussed in the previous section, eye-movements reflect cognitive processes that operate automatically, giving eye-tracking some important advantages over other behavioural measures (discussed in more detail in Scherr, Agaus and Ashby, 2016).

Advantages of eye-tracking:

1 Eye-tracking provides a 'direct' measure of processing effort during a task, rather than at the output of a decision, recall or production task, which are often subject to strategic effects.
2 Eye-movements have little variance due to individual differences in memory, deliberate decision-making processes and recognition strategies, which are generally implicated in explicit judgement tasks.
3 The temporal precision of eye-tracking provides a record of behaviour from the first moment a visual stimulus (text, image, scene, etc.) is perceived until the stimulus is removed or a participant stops looking at it. For example, eye-tracking allows us to measure the cognitive effort that is expended the first time readers encounter a word, as well as when and if they go back to re-read the word at some point.
4 Although it generally occurs in a laboratory setting, eye-tracking allows readers and viewers to engage with visual stimuli as they normally would (when they are presented on a computer screen). This means that participants can read and re-read at their own pace and look where they want, *without* the need to impose an additional task.

Because of its advantages, eye-tracking research has become the 'gold standard' for studying reading in psycholinguistics and cognitive psychology (Rayner, 2009, p. 1474), and an ever-increasing number of researchers in applied linguistics are beginning to use it. However, it is important to note that the eye–mind assumption relies on two underlying beliefs (Pickering et al., 2004). First, there is the supposition that what is being fixated is what is being considered. This means that when the eyes fixate the word 'infect' in Figure 1.3, the processing system is working to decode and understand this word and not the word 'even' that occurred three words before. In other words, readers try to interpret words as they are encountered. However, this assumption is somewhat of an

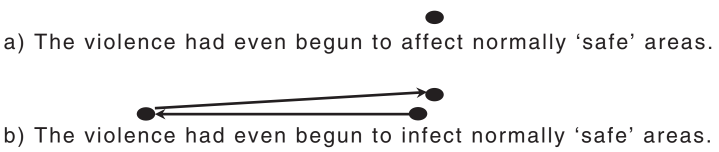


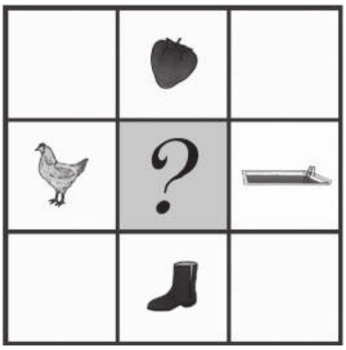
Figure 1.3 Fixations (depicted by circles) used to assess cognitive effort to two matched words: 'affect' and 'infect'. The word 'affect' has one fixation, while 'infect' has two, indicating that greater cognitive effort was expended in processing it.

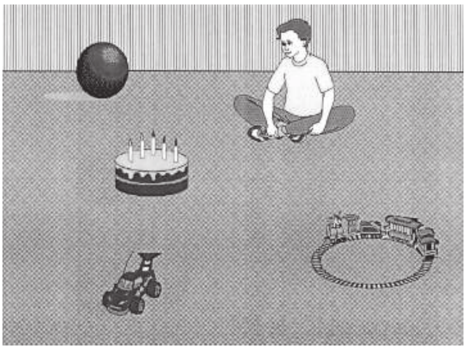
oversimplification (Ehrlich and Rayner, 1983). For example, in a sentence like 'Jill sold her horse to Jack because she decided to quit riding,' when the eyes land on the word 'she', in order to interpret it, the processing system will need to consider previously encountered entities that could be potential referents of the pronoun. Thus, when the eyes land on 'she', the processing system is indeed working on this word; however, it is also considering previous elements of the sentence that could be potential referents (e.g. 'Jill').

The second part of the eye–mind assumption is that the amount of time spent fixating an item or region reflects the cognitive effort required to process it. This means that longer and more fixations indicate greater processing effort, and shorter fixations and/or more skipping indicate less processing effort. Importantly, these terms are relative: longer/shorter duration and more/less processing effort need to be in comparison to something. In general, cognitive effort is linked to a particular region of interest (ROI, sometimes referred to as area of interest (AOI)) using local measures. Thus, if we want to explore how using metaphorical language influences processing, readers could be presented with sentences like those in Figure 1.3. In this example the ROIs are the words 'affect' and 'infect'. The number (fixation count) and length (fixation duration) of fixations to the ROIs would be measured to determine processing effort. In the example, we see that 'infect' has two fixations, while 'affect' only has one; this could be taken as an indication that 'infect' requires more processing effort.

So far much of the discussion and the examples have been focused on reading. While investigations of reading can tell us a lot about language representation and processing, they have clear limitations. For example, we cannot use reading studies to tell us about the linguistic skills of pre-literate children. Further, reading is only one element of language competence; thus limiting ourselves to reading studies will never give us a complete picture of linguistic ability. Importantly, in addition to presenting written text, eye-tracking technology can present static images and visual scenes (often referred to as the *visual-world paradigm*), as well as visual media like films and television programmes. Such stimuli are often used to explore auditory processing – or listening. In this type of research, eye-movements and fixations to visual input are time-locked to a particular linguistic variable (a word, pronoun, syntactic structure, etc.) that is presented auditorily. The different areas/images on the screen are defined as ROIs. Data are usually reported in terms of total fixation duration, as well as the proportion of saccades directed to or the proportion of time spent looking at a target ROI compared to competitor ROIs (see Section 5.1 for a discussion). As with reading, the eye–mind assumption holds. Thus, what is being looked at is what is being processed and the number of fixations and fixation time indicate processing effort.

Figure 1.4 provides some examples of how the visual-world paradigm works (from Altmann and Kamide, 1999 and Chambers and Cooke, 2009). In Example A, static images are presented on a visible grid, while in B the scene is more 'natural'. In both cases eye-movements and fixations to the depicted objects are measured and assessed relative to each other. As can be seen in A, Chambers and Cooke (2009) presented English–French bilinguals with a screen containing static images of a swimming pool, a chicken ('la poule' is a near homophone of pool) and two distractor objects (a strawberry and a boot). Participants listened to sentences like 'Marie va décrire la poule' ('Marie will describe the chicken') and their eye-movements were monitored time-locked to the noun 'poule'. So what happens when English–French bilinguals hear the word 'poule'? Do they only think about the French word 'poule', or do they also consider the homophone 'pool' from their other language? To assess this, Chambers and Cooke (2009) calculated both the





Example A
'Marie va décrire la poule'
'Marie will describe the chicken'

Example B
'The boy will eat the cake'

Figure 1.4 Examples of the visual-world paradigm presenting images and auditory stimulus. In Example A, the images are presented in a visible grid, while in B they are presented as a more 'natural' scene. The ROIs would be the individual objects and entities.

mean number of saccades and the probability of fixating the target (pool) and near homophone (chicken). In Example B, Altmann and Kamide (1999) wanted to know what people think about when they hear the word 'eat'. If they primarily consider edible things, there should be more looks to the cake relative to the other objects on the screen. Both of these studies utilise eye-tracking in the visual-world paradigm as a way to investigate the behaviour of listeners in real time. Finally, it is important to note that ROIs can be defined for other types of visual material. Thus, for films the ROIs could be some aspect of the image that appears at a given point in the video and the subtitle regions during the same time period.

In general, what the eyes are looking at *is* what is being processed. However, we have highlighted some caveats to this assertion; and it is important to note that this belief has been challenged by some. Central to the eye–mind hypothesis is the view that the amount of time and number of times that the eyes look at something provide an indication of processing effort. Because of this, a significant effort has been made to try to relate local fixation measures to specific cognitive processes. To do this we would need very specific evidence linking particular eye-movement measures with underlying cognitive events. Importantly, at this point we do *not* have such evidence and we do *not* know how to uniquely map cognitive events to the different eye-movement measures (Pickering et al., 2004). In other words, we cannot use the eye-movement record to pinpoint, for example, exactly when a word's meaning is accessed, or when it was integrated into our unfolding understanding of a sentence or discourse. What exactly we can conclude from various eye-tracking data and measures will be touched upon in the next section and will be the focus of Chapters 4, 5 and 6.

1.3 Basic Considerations When Doing Eye-Tracking

It is probably already clear that eye-tracking is potentially a useful technology. Importantly, it is increasingly being used by researchers in applied linguistics. The growth of eye-tracking means that researchers who have not previously encountered the technology are coming into contact with it in presentations and papers, as well as possibly wanting to try it out for themselves. In order to establish it on a sound footing in the field, it is important to come to an understanding about (1) what eye-tracking is; (2) what it can and cannot tell us; and (3) how it might be meaningfully used in applied linguistics research. Over the course of this book we will address all of these.

By this point, we should have a basic idea of eye-tracking. Eye-tracking systems, as well as their associated software, differ in certain ways. An important consideration when we start out doing eye-tracking research is choosing the 'right' system or understanding the capabilities of one that we may have access to. Chapter 2 will discuss some of the basic properties of eye-tracking systems and software, as well as things to consider when choosing or using one. The remainder of the book, as well as this chapter, will provide insight into what eye-tracking technology can tell us, how to use it and how we might integrate it into our research in a meaningful way. As we have seen, while the technology can provide a window into the mind, we need to be cautious about overstating what it tells us. Importantly, eye-tracking's capacity to tell us anything about cognitive processing is largely dependent on designing good studies. Thus, experimental design, as well as the methodological considerations associated with different paradigms, will be a predominant focus of the book.

As was touched on in Section 1.2, eye-tracking can be used for the examination of participants' processing of different types of verbal and non-verbal stimuli, and a combination of different types of stimuli. Importantly, as we have seen, one of the attractions of eye-tracking is its ability to tell us what the mind is thinking about – which is whatever we are looking at – and how much cognitive effort is being expended on this. In general, more and longer fixations indicate that more processing effort is needed, while fewer and shorter fixations and/or skipping indicate that less processing effort is required. As mentioned previously, short and long are relative terms. If we think about this in relation to a 'real-world' comparison, this becomes clearer. If we want to know whether a three-year-old boy is of tall/short/average height for his age, which of the following statements would help us assess this: 'the three-year-old boy is short compared to giraffes' or 'the three-year-old boy is short compared to other three-year-old boys'? Clearly to say anything meaningful about the child's height we want to assess this with a reasonable comparison set – other three-year-old boys and not giraffes (which are apparently already six feet tall at birth).

Thus, if we want to know whether more/less/equal processing effort is required for words that we encounter frequently versus words that we rarely encounter, we need to first make sure that our comparison set is right. In Example 1.1 sentences (a), (b) and (c) provide potential stimuli to address our question; they will allow us to compare the reading pattern for the high-frequency word 'house' versus low-frequency words 'protractor' or 'louse'. While in (a) the word 'house' is more frequent than 'protractor', 'protractor' is also a longer word. Thus, greater processing time for 'protractor' could be due to its frequency, length or both. In other words, 'protractor' does not provide convincing evidence that we spend longer reading low-frequency words than high-frequency ones. In the second example, the words 'house' and 'louse' are well matched for length, as well as for phonological properties, and simply differ in frequency. However, in (b) the words are in very different sentence frames. Any longer reading times for 'louse' in (b) could be due to its frequency, the general parsing difficulty for the sentence or both. Finally, in (c) we see a 'fair' comparison, where the only thing that differs is the frequency of the words 'house' and 'louse'; thus shorter fixations for 'house' can convincingly be attributed to less processing effort being required for higher-frequency words.

Example 1.1 Hypothetical stimuli exploring the effect of word frequency on reading times when words *are not* well matched on other properties like word length (a) or sentence structure (b) and when they *are* well matched on these properties (c).

(a) `She found a house yesterday.` vs `She found a protractor yesterday.`
(b) `She found a house yesterday.` vs `I was told that what was found yesterday by the barber was a louse and not dandruff.`
(c) `She found a house yesterday.` vs `She found a louse yesterday.`

Potentially more challengingly, we may want to determine whether test-takers spend longer, in other words expend more processing effort, reading gist questions or questions asking about details. If we conduct a study and find that on average test-takers spend 500 ms longer reading detail questions, we might conclude that detail questions require greater cognitive processing. However, if the detail questions in the study are systematically four words longer than the gist questions, the difference in fixations could be entirely due to the length of the questions, as sentences or questions that have more words take longer to read. Thus, in this case the comparison is not reasonable because differences in question length and type are confounded. In order to ensure that the conclusions that are drawn from the study are warranted, we would either need to match our question types on factors like length and syntactic complexity, which are known to influence reading time, or use an analysis technique that could take into account these differences (e.g. mixed-effects modelling or multiple regression analyses), which will be discussed in Section 7.2.

Typically, in eye-tracking studies involving reading we are interested in particular phenomena, for example the influence of word frequency on processing, as in the 'house'/'louse' example. We show readers complete sentences or texts and not simply the individual words 'house' and 'louse'. This means that we are interested in eye-movements made to a very specific ROI in a sentence or longer discourse. For instance, if we wanted to look at how using metaphorical language influences reading and comprehension we could show readers passages like those in Example 1.2. (Usually the two versions would be presented in different experimental lists, so that a single participant would only see one version, but each version would be seen by an equal number of participants – see Section 3.1.2 for more on counterbalancing in eye-tracking studies.) In each version keywords would be set as ROIs, which are indicated by the underlining in the example.[2] The ROIs correspond to words that make the metaphor of 'crime as a disease' apparent, which are matched to ensure that comparisons are fair/appropriate. We would count and measure the number of fixations to our ROIs. This would show us whether there is a difference in reading time, in other words processing effort, for the metaphorical and non-metaphorical language. We could also explore whether readers performed better on a comprehension task following the different versions and examine whether this correlated with reading times. We could also look at whether metaphorical versus non-metaphorical language impacted reading time in the rest of the passage. Because processing effort is relative, an analysis will determine whether more or less processing effort is required for the metaphorical compared to the non-metaphorical

[2] In all of the examples in this chapter, the ROIs are underlined to make them easier to identify, but they would not be indicated in any way to participants in a study.