Evaluation of techniques to improve the legibility of bilingual Variable Message Signs

Samantha L. Jamson*

Institute for Transport Studies, University of Leeds, UK, LS2 9JT

* Corresponding author:
Tel: +44 113 3436606
Fax: +44 113 3435334
e-mail: sjamson@its.leeds.ac.uk

Abstract

This study evaluated a number of techniques that could be employed to reduce the amount of time drivers spend searching and reading bilingual signs. Using a tachistoscope, monolingual and Welsh bilingual participants were presented with various configurations of bilingual signing. The amount of information was varied (i.e. the number of lines) and a number of language-differentiation techniques were implemented. These techniques attempted to aid the perception and recognition of the relevant language and relied either on manipulating the position of the two languages, or by using demarcation (colour, font etc.). With regards to the amount of information presented, it was found that the reading response time for a single line of relevant text within a two-line bilingual sign was not significantly different to the reading response time for a one-line monolingual sign. Thus, participants were able to extract the relevant language from the bilingual sign with no decrement in performance. However, reading response time for a message of two lines of relevant text in a four-line bilingual was significantly longer than the reading response time for a two-line monolingual sign. Thus the amount of information (even if irrelevant) impacted on their performance. With regards to the positioning techniques, grouping the lines by language resulted in a decreased reading response time compared to when the text was grouped by content. In addition, positioning the user's dominant language at the top of the sign improved reading times for both one and two-line messages on bilingual signs. All the demarcation techniques were successful in reducing reading times on four-line bilingual signs, and it was found that having established a particular pattern of presentation, an unexpected change significantly increased reading time.

Keywords: Bilingual; Variable Message Signs; Reading time; Traffic safety

1. Introduction

Across Europe and beyond, there are a number of countries that can call themselves bilingual, i.e. where two languages are given equal status (e.g. Belgium, Canada, New Zealand). The Welsh language is the oldest living language of Great Britain and one of the oldest in Europe. At the start of the 20th century, the Welsh language was spoken by almost half the population of Wales. However, since then the number of Welsh speakers has decreased steadily until fairly recently. By 1991, although the number of people able to speak the language was still more than half a million, this represented just 18.7% of the population. The 1991 Census results suggested that this decline has stabilised, particularly among young people. It is likely that the next Census will show that there has been a further increase, because nearly one-third of all primary school children in Wales are now receiving their education in Welsh-medium or bilingual schools. A survey published by the Welsh Office in 1995, shows that 21% of the population of Wales speak Welsh of which 55% of them are first-language speakers, (i.e. they spoke more Welsh than English as a child). 13% of the population of Wales claimed to be fluent in Welsh whilst 66% claim no knowledge of Welsh at all.

In July 1997, the UK Government published a White Paper, A Voice for Wales, which outlined its proposal for the devolution of Wales. Subsequently, Parliament passed the Government of Wales Act 1998 which established the National Assembly for Wales (NAW). This enabled the transfer of the devolved powers and responsibilities from the Secretary of State for Wales to the NAW. As a result, both Welsh and English are official languages of the NAW. The NAW is required to treat both languages equally, so far as is practicable, in conducting all its business and all legislation must be in both languages. Traffic signs are part of this legislation and this creates the need for all signs, including Variable Message Signs (VMS), to be in a bilingual format. This means that drivers encounter signs with double the amount of text, thus presenting a number of interesting research questions.

There exists a fair amount of literature that attempts to measure the time taken for drivers to read traffic signs in order to produce design guidelines (Moore and Christie, 1963; Johansson and Backlund 1970; Forbes, Snyder and Pain, 1965). Hall, McDonald and Rutley (1991) reported that reading times increased in a linear manner with the number of names on a sign. With regards to bilingual signs, whilst drivers need only read half of the information presented in order to glean the appropriate message, research has confirmed that reading times increased compared to their monolingual counterparts (Rutley, 1972; Lesage, 1981; Harjula, Luoma and Rämä, 1998). Rutley (1972) reported that it took drivers 10-15% longer to read bilingual signs. Additional research (Rutley, 1972; Rutley, 1974) found that overall reading time was minimised when the preferred language was positioned uppermost. Current practice in Wales is to use the dominant language on the first line of the sign and the secondary language on the second line. Thus, Welsh text typically appears at the top of signs in North Wales where the Welsh language is more widely spoken, while English appears at the top in South Wales (where English is dominant).

Although bilingual traffic signs have been used in Wales for more than a decade, the move to bilingual VMS is a relatively recent one. VMS can convey a wide range of messages, which may be varied temporally in response to the prevailing traffic and environmental conditions making them an extremely powerful traffic management tool. Research at a European level has attempted to produce guidelines for VMS best practice (e.g. TROPIC). However, most of the research on VMS has concentrated on legibility criteria in terms of character luminance (Kerr, Snelgar, Jordan, Emmerson and Linfield, 1987; Padmos, van den Brink, Alferdinck and Folles, 1987; Mace, 1988; Colomb and Hubert, 1991; Upchurch, Armstrong, Baaj and Thomas, 1992; Ullman and Dudek, 2001). There has been little work carried out which attempts to discover the optimum method of displaying the bilingual text. Where research has taken place, the scope is fairly limited and tends to concentrate on only one implementation of a sign, usually due to financial constraints (e.g. Anttila, Luoma and Rama, 2000).

With regards to traffic safety, an increase in sign reading time could be accompanied by a decrease in the amount of visual attention paid to the road. Whether this leads to an increase in accident risk depends on the context of the driving situation (e.g. vehicles ahead, relative speed, proximity of vehicles in the adjacent lane etc.). When considering the visual distraction of in-car systems, some guidelines exist as to how much distraction is acceptable (ISO, 1999; SAE, 1999). However there is no baseline against which we can evaluate the level of visual distraction posed by road signs. As a result, researchers and policy makers have simply sought to minimise the overall reading time.

In order to quantify the success of a particular sign configuration, it is necessary to consider how easy it is to read and understand. In a dynamic environment, such as the driving task, drivers are required to simultaneously process a large amount of information pertaining to the vehicle's controls, the road conditions and situation. It is therefore logical that the amount of time drivers devote to reading roadside messages should be kept to a minimum, whilst at the same time providing sufficient information for the driver to proceed safely and efficiently.

It is therefore of interest to evaluate the differing amounts of time that drivers spend searching for the relevant information, depending on the sign configuration. This can be measured by recording the reading time of the sign in question, i.e. the time it takes a driver to locate and read the appropriate information. Typically, individual performance on this task varies by parts of seconds. It is also expected that the differences between the different sign layouts will be small and possibly smaller than the differences between subjects. Mean reading times over a range of bilingual sign configurations (Rutley, 1972) were between 1 and 2 seconds. However the small sample size resulted in standard deviations of up to 0.67sec. In such cases it is not possible to draw any statistically robust conclusions.

Although these differences are small they could have an impact on driver performance and workload when considered over the range of driving conditions and the rate of exposure. At any moment in time there will be many thousands of drivers on the road reading information from signs. For most there are no problems. It is the rare lapses in performance or recognition that impact on the safety and efficiency of the traffic system. Whether one sign configuration is better or worse depends on very small changes often at the extremes of the performance distribution.

2. Issues with bilingual signing

2.1 Message comprehension

Literature in the field of visual perception suggests that the comprehension of a message can be thought of as being partitioned into three stages (Anderson, 1990). The first stage comprises the perceptual processes by which the written message is encoded. The second stage is termed the parsing stage whereby the words in the message are transformed into a mental representation of the combined meaning of the words. The third stage is the utilisation stage, in which the mental representation of the message meaning is used. If the message is an assertion, the reader may simply store the meaning in memory; if it is an instruction, they may act according to the information provided.

While the latter two stages relate to understanding and the use of the message, the perceptual processes describe how users obtain and form the words of a message. It is this process that is the focus of the study reported here. Given the identical nature of the character sets for Welsh and English, drivers have to scan numerous lines of text in order to extract the relevant information. This search could be aided by the clear differentiation between the two character sets. This is particularly important given the fact that VMS can accommodate up to four lines of text (two in each language).

2.2 Language differentiation techniques

The specification of traffic sign formats is intended to aid perception and recognition so as to develop a more automated process that requires less attention. To minimise the overall reading time for bilingual signs, two approaches have traditionally been adopted. Firstly, the position or sequence of the languages has been optimised; secondly, attempts have been made to use demarcation so that drivers can identify the appropriate portion of the message more efficiently.

2.2.1 Language positioning

Research on static bilingual road signs in English and Welsh (Rutley, 1972; Rutley, 1974) found that overall reading time was minimised when English was placed uppermost on the sign. There was one exception, a laboratory experiment that did not involve driving. In this case, bilingual subjects who chose to read Welsh exhibited shorter reading times when viewing signs with the Welsh language uppermost. Rutley concluded that "in order to achieve the shortest reading time

for the vast majority of people in Wales, the English names should be placed above their Welsh equivalents".

Further evidence for this was found in a larger Canadian study involving four population types (Lesage, 1981). They were either monolingual (French or English) or bilingual with expressed preferences for French or English. Again, it was found that the placement of the reader's dominant language at the top of a sign resulted in significantly improved legibility. This effect was most pronounced for monolingual readers, since bilingual readers appeared to respond to whichever language was in the uppermost location.

Finland uses this system in a formalised manner, due to the spread of Swedish and Finnish speakers across the whole country. Currently, over 90% of the population use Finnish as their native language, while the remainder use Swedish. However, the Swedish language is fairly well known among the general population due to the educational system. Finnish road signs are either mono- or bilingual depending on the language status of the district in which they are situated. A district's language status is evaluated every ten years and depends on the current population of the district. If 8% or more (or at least 3000) of the population has Swedish as mother tongue, the district is considered bilingual. The dominant language of the district is then presented first on the traffic signs.

2.2.2 Language demarcation

There are a number of potential ways of improving drivers' ability to distinguish between the two languages on a bilingual sign. While at first, as with any new concept, drivers may not make an immediate connection between distinguishing features and the appropriate language, over time such connections should become automatic and search patterns more focussed. The possible demarcation possibilities are as follows:

a) Colour***:*** For traffic signs that have legal significance, the colour of the text, background and borders are specified. However VMS signs typically provide warnings and information and thus do not have statutory significance. It could therefore be possible to use colour as a distinguishing mechanism. The colours selected would have to be of neutral or equal meaning to drivers. For example, the use of red should be avoided as this implies dominance or urgency. Colours that have statutory associations with other types of traffic signs would also have to be avoided since these may serve as a distraction to the alternative language. The chosen colours would also need to have equal luminance characteristics under varying light and weather conditions.
b) Font type: Variations in font can be achieved by changes in the "stroke" width used to form characters as well as the overall dimensions of the characters. Changes to the stroke width on VMS characters are achieved by illuminating more or less pixels. This affects the light output and consequently the target value of the character (Staplin et al., 1997). The use of different font sizes is not acceptable as this impacts on the reading distance.
c) Case: Although differentiation of language by case is used on bilingual signs in New Zealand (English in capital letters, Maori in lower case) no literature has been found to support this decision. The main drawback with this approach is that it has been widely shown that sentence font (i.e. upper and lower case) facilitates reading. Another disadvantage is that the language written in capitals could be construed as having priority.
d) Separation: Providing a clear spatial separation between languages could assist drivers' search process. Lesage (1981) found that separation between two languages improved glance legibility, particularly for longer messages. It was also found that monolingual readers benefited most from separation and that the improvement was consistently greater when the driver's preferred language was not dominant.

Studies into the impact of manipulating character variables as a means of improving language differentiation are scarce. Those cited have not found significant differences in legibility, the exception being separation of the languages. The experimental work reported here had two aims. The main aim of the study was to examine if bilingual signs were more difficult to read than their monolingual counterparts and to quantify the most appropriate method of demarcation. Additionally, the study evaluated the effects of learning and expectancy on reading times. This second aim arose due to the fact that based on the research outlined above, presenting the locally dominant language first on a bilingual sign minimises reading times. But what happens when a driver travels through an area with a different presentation pattern? Research into the impact of

changing patterns of presentation is conspicuously absent from the literature. Thus, the second aim of the study was to evaluate the impact of inconsistency in presentation.

3. Method

These trials were designed to examine the most appropriate configuration of bilingual signs in a controlled manner to determine their relative readability. Readability was measured using reading time, i.e. the amount of time taken to locate and read the appropriate information.

3.1 Tachistoscope

A tachistoscope was used to present the material to the participants. A tachistoscope is a method of presenting successive stimuli to participants and recording their response time to those stimuli. The stimuli can comprise text, pictures or moving images. Traditionally, tachistoscopes are operated manually by a researcher, allowing a limited number of trials to be undertaken in any one session. However, it is more common today for such experiments to be PC-based, thus allowing greater flexibility in design, portability of the experiment and less incidence of operator error/fatigue. The tachistoscope program used in this experiment was developed in-house and was used to manipulate a wide number of variables including text configuration/colour, display duration and pause length. This allowed the experimenter full control over the variables.

The program was developed using Microsoft Visual Studio, in order to ensure portability across Windows platforms and screen resolution. The code was designed in order to allow easy changes to be made to an individual line's colour, case and text. This allowed different signs to be quickly constructed from the lines. A high resolution timer (8^{-7} s) was used to measure reaction (reading) time.

3.2 Participants

Thirty participants were recruited for this study. The sample included 16 males between the ages of 19 and 54 [M= 36 years] and 14 females between the ages of 20 and 49 [M= 34 years]. Eighteen participants reported understanding no Welsh at all and are referred to as Monolingual English. The remaining twelve reported good comprehension of Welsh, for speaking reading and writing, and were happy to use either language. These participants are referred to as Bilingual Welsh. All were familiar with using a keyboard.

Monolingual English participants were given the following instructions:
"You will be presented with a number of stimuli on the computer screen. These stimuli will be in the form of road signs. Some of these signs will be monolingual (English or Welsh), and others will be bilingual (English and Welsh). For each of the signs that are presented to you, please read aloud the English component of the signs and then press Yes. The next sign will then appear. If there is no English component, then just press No".

Bilingual Welsh participants were given the following instructions:
"You will be presented with a number of stimuli on the computer screen. These stimuli will be in the form of road signs. Some of these signs will be monolingual (English or Welsh), and others will be bilingual (English and Welsh). For each of the signs that are presented to you, please read aloud the message in whatever language you feel most comfortable with. Then press Yes. The next sign will then appear."

Participants were allowed a period of familiarisation with the response buttons. A total of 196 signs were presented over the course of approximately ten minutes.

3.3 Manipulated variables

There were a number of variables of interest that were manipulated in the tachistoscope trials. These are described in more detail in the following sections along with examples of the stimuli presented.

3.3.1 Number of lines

Studies have reported that reading time increases with the amount of information presented (e.g. Hall, McDonald and Rutley, 1991). This seems perfectly logical and by including this variable in the trials a reliability check was made on the data. These reading times were also used as a baseline measure against which the different sign configurations were compared. A typical set of signs used in this comparison are [A] and [B].

[A] *[B]*

The number of lines ranged from one to four. Each of the lines had a maximum of 12 characters (regardless of language type).

3.3.2 Number of languages

It has also been reported that readers take longer to extract information from a bilingual sign than from a monolingual sign (when the amount of information to be extracted remains constant). The tachistoscope trials aimed to establish if there was any difference in reading times for signs such as [C] and [D].

[C] *[D]*

Even though the amount of information displayed in the English language was the same in each of the signs, drivers had to engage in additional visual search in the case of [D] in order to locate the relevant information and confirm that the second line was irrelevant.

3.3.3 Language position

According to current policy, the position of each of the languages depends on the language dominance in the surrounding population. The tachistoscope trial therefore manipulated this variable such that the signs appeared in each of the two possible configurations as in [E] and [F].

[E] *[F]*

3.3.4 Language demarcation

A number of techniques were used to differentiate between the languages shown on the bilingual signs. It was hypothesised that such demarcation may aid drivers in their visual search technique when trying to locate the appropriate language.

i. The first of these methods was the use of colour. It was deemed important to use two colours that would not induce prioritisation. The colours yellow and green were used, although it is recognised there may be inherent difficulties in the display of these colours on VMS. These colours were chosen merely as a demonstration of what this type of demarcation is capable of achieving. The use of the two colours was balanced, i.e. on some occasions the English text was in green and sometimes in yellow, and the same was true for the Welsh text, see [G]. The reading time for this sign was compared to one without colour demarcation [H].

[G] *[H]*

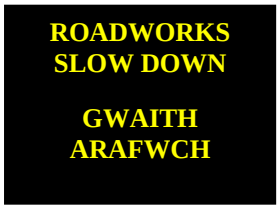
ii. The second method of demarcation involved the use of different cases to distinguish between the languages. It was thought the most obvious demarcation would be between languages displayed in upper case and those displayed in sentence case.

[I]

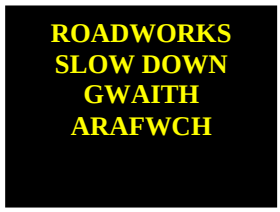
[J]

Therefore, on some occasions the English text was in upper case and sometimes in lower case, and the same was true for the Welsh text. The reading time for [I] was compared to one without case demarcation [J].

iii. The final method used was a separation line between the two languages (in the case of four line signs only). The reading times of, for example, [K] and [L] were compared.



[K]

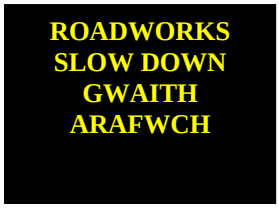


[L]

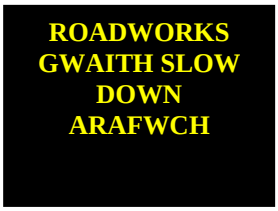
3.3.5 Message sequencing

There are two possible ways of displaying a bilingual four-line sign (two lines in English, two lines in Welsh) to drivers.

Firstly, the four lines of text could be placed on one sign. This then leads to the issue of whether the text should be arranged by message language or by message content. For example should the format of Sign [M] (arranged by language) or Sign [N] (arranged by content) be used?



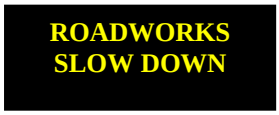
[M]



[N]

The advantage of [M] is that drivers may have less visual searching to undertake in order to perceive the whole message. On the other hand, with message [N] drivers have to scan all four lines in order to retrieve the appropriate information. The tachistoscope trials attempted to establish if reading times differed between these two options.

Secondly, the four lines of text could be split over two sequenced signs (each with two lines of text). As for the four-line signs there is a further decision to make with regards to the type of sign used here. Either the signs are monolingual (and thus all the message content is contained on one sign, for that particular subset of the population) or are bilingual (such that all drivers receive half of the message content). See signs [O1 then O2] and [P1 then P2] for the respective examples.



[O1]

[O2]

[P1]

[P2]

Each of these options has its drawbacks. The [O1 then O2] option means that monolingual drivers will not comprehend the first message if it is not in their language. If the message is in the form of an instruction or contains a critical warning then those drivers who have understood the sign will

begin to react to the sign in a manner that may be incomprehensible to the other drivers. The message will of course become clear when they reach the sign in their language but this confusion may lead to safety problems (e.g. unpredictable lane changes).

On the other hand, if the option [P1 then P2] is employed, all drivers only receive half the message at first. This may not be a problem if each element of the message is self-contained, as in the example above (i.e. each of the messages "Roadworks" and "Slow Down" could exist on their own without confusion). However, if the message when split is incomprehensible, as in signs [Q1 then Q2], then confusion may occur as drivers attempt to search for the rest of the message on the following line. Drivers may attempt to make sense of the second line of the message, possibly causing unnecessary driver distraction.

[Q1] *[Q2]*

The tachistoscope trials established reading times for these different characterisations of four line signs.

3.3.6 Learning and expectancy

Research shows that the speed of cognitive processes can be increased if they are of an automatic nature as opposed to a controlled nature (Schneider and Shiffrin, 1977; Shiffrin and Dumais, 1981). This automaticity arises as a result of learning and in the case of road signs, should be as a result of intuitive and consistent design and presentation. Thus, any sign configurations (such as language dominance or demarcation) should be incorporated into policy in order for drivers to learn the appropriate arrangement. Any change in or variation from the policy may lead to confusion and/or missed messages.

The tachistoscope trials investigated the likely impact of inconsistent information on reading times. This was achieved by firstly priming participants with a consistent pattern of sign configuration. Immediately following this sequence there appeared a sign that was inconsistent with the previous sequence. It was hypothesised that reading times would decrease as participants became familiar with the pattern and then increase (to exceed the initial reading time) when presented with the inconsistent sign configuration.

[R1] *[R2]*

[R3] *[R4]*

It was hypothesised that if reading time for [R3] was shorter than for [R1], then this was an indication that learning had taken place. If it was subsequently found that the reading time for [R4] was longer than that of [R1], then it could be that the presence of an inconsistent sign is detrimental to comprehension. Several options were tested whereby the pattern relied on differences in cases, colour and language position.

4. Results

Sign configuration was used as a repeated-measure factor, with varying levels depending on the comparisons made. Language was included as a between-subject factor having two levels (Monolingual English and Bilingual). The data were checked for normality and homogeneity of

variance using the Kolmogorov-Smirnov and Levene tests respectively and subjected to parametric testing (ANOVA or t-tests) to identify changes in response times across the configurations.

4.1 Number of lines

The reading times for one, two, three and four-line (monolingual) signs were compared. There was found to be a significant effect of sign length on reading times ($F_{[3,87]}$=307.97, p<.001). Post hoc testing revealed there to be a linear trend, such that each time another line was added to a sign, reading time increased significantly($F_{[1,29]}$=433.8, p<.001). Figure 1 presents these results.

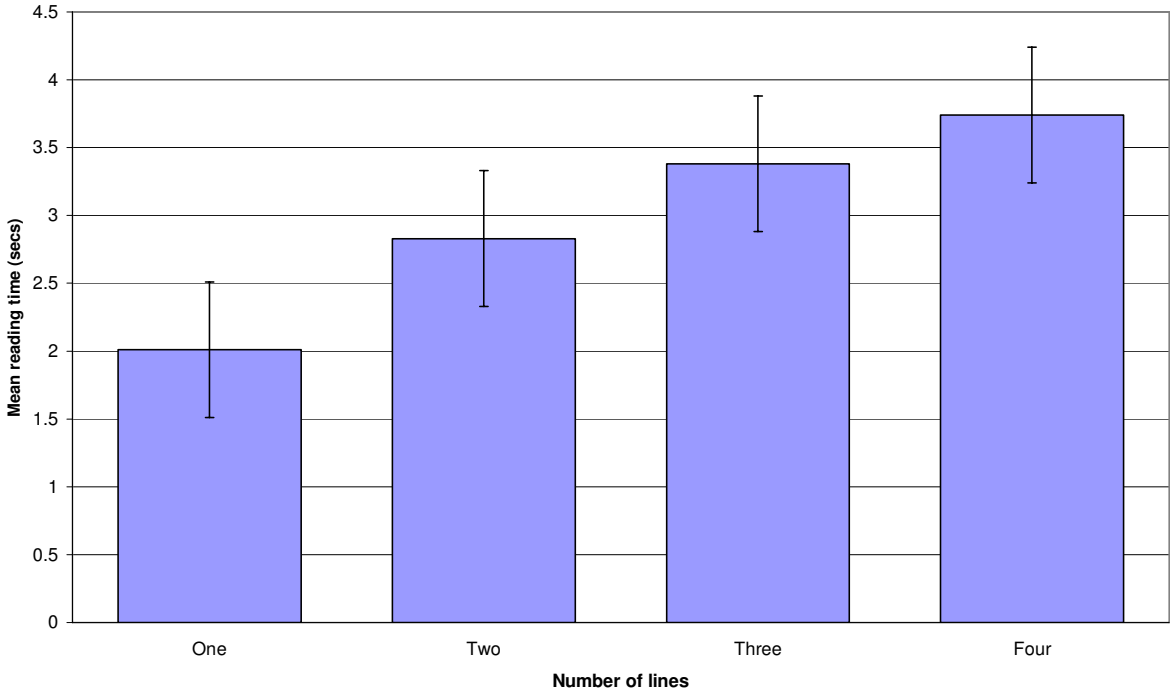


Figure 1 Effect of sign length on mean reading times (monolingual signs)

It can be seen that reading times almost doubled when moving from one-line to four-line signs. These results suggests validity of the experimental technique. There were no effects of Language.

4.2 Number of languages

The reading times for monolingual signs (with either one or two lines of information) were compared to the reading times for bilingual signs (with either one or two lines of information, in each language). It was hypothesised that even though the amount of information displayed in the English language remained constant, with a bilingual sign drivers have to engage in additional visual search in order to confirm that the second line is irrelevant. The mean reading times are shown in Figure 2.

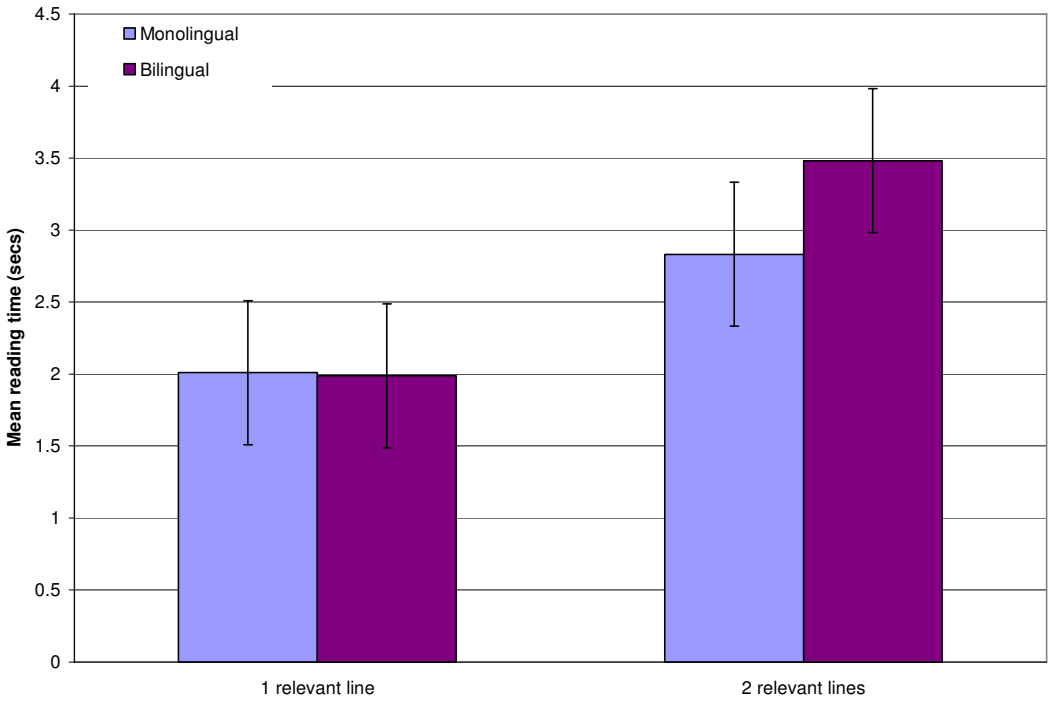


Figure 2 Effect of additional language on reading times

An analysis was carried out using both Number of Languages and Number of Lines as repeated measures factors. There was found to be a significant interaction between the two factors ($F_{[1,29]}$=132.51, p<.001), but no effect of Language. Figure 2 shows that with one line of relevant line of text, reading times did not increase when subsequently embedded in a bilingual sign. However, there was a significant increase in reading time when two lines of relevant text needed to be extracted from a bilingual sign, compared to the reading time for extracting two lines of text from a monolingual sign.

With regards to comparing the reading times for a four-line bilingual sign to that of a four-line monolingual sign, from Figure 1 it can be seen that the average reading time for the latter is 3.74 seconds. There was no statistically significant difference between these two types of sign. It therefore appears that when participants were required to read a four-line bilingual sign, with only two relevant lines of text, their reading times increased to the extent that they were comparable to those found when they actually read four lines of monolingual text. This represents a rather poor performance with four-line bilingual signs.

4.3 Language position

It was hypothesised that participants would be able to locate the information in their dominant language more rapidly, if that information appeared in a dominant position (i.e. on the top of the sign). Reading times were compared between instances when the preferred language appeared on the top and when it appeared on the bottom. This was carried out for signs with one relevant line of information (per language) and for those with two lines of information (per language). The results are shown in Figure 3.

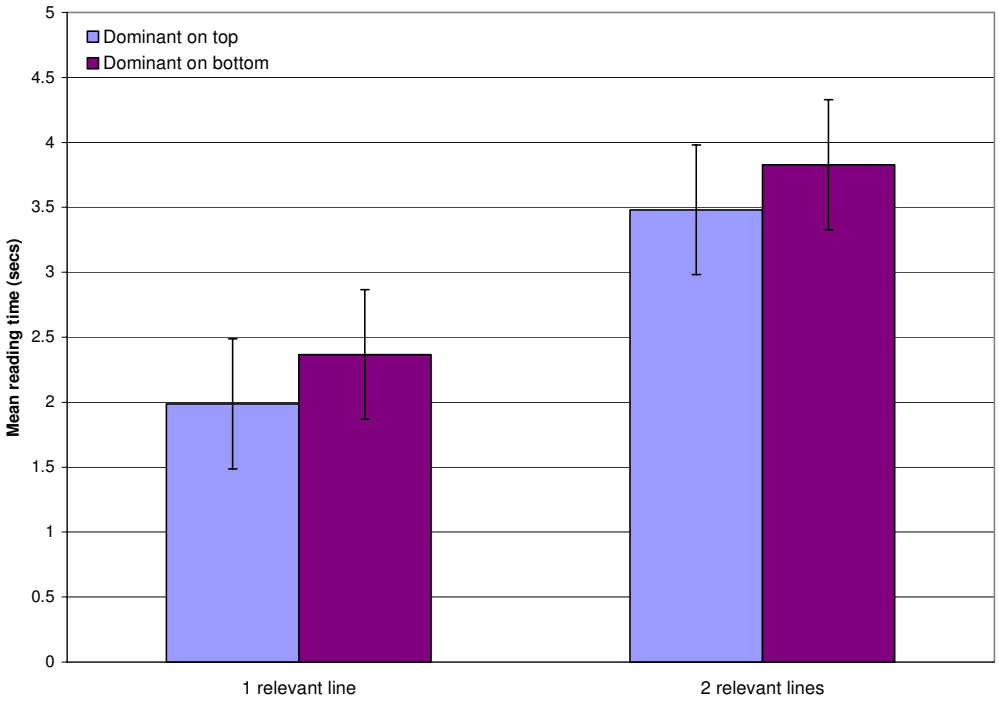


Figure 3 Effects of the position of the preferred/dominant language

There was found to be a main effect of Language Position such that mean reading times were shorter if the preferred language was uppermost. This held true both for signs with one line of relevant text ($t_{[29]}$=13.29, p<.001) and two lines of relevant text ($t_{[29]}$=4.05, p<.001).

4.4 Language demarcation

The findings above suggest that four-line bilingual signs with two relevant lines of text take as long to read as a four-line monolingual sign. A number of techniques were tested to try to alleviate this problem.

The use of colour, case and a separation line were used to distinguish between the two languages on bilingual signs. None of the separation techniques were found to reduce reading time for bilingual signs with one relevant line of text. However, when participants were required to extract two lines of information from a four-line bilingual sign, the use of colour ($t_{[29]}$=12.67, p< .01), case ($t_{[29]}$=8.26, p< .01) and a separation line ($t_{[29]}$=13.03, p< .01) all significantly reduced reading times, see Figure 4. There were no effects of Language.

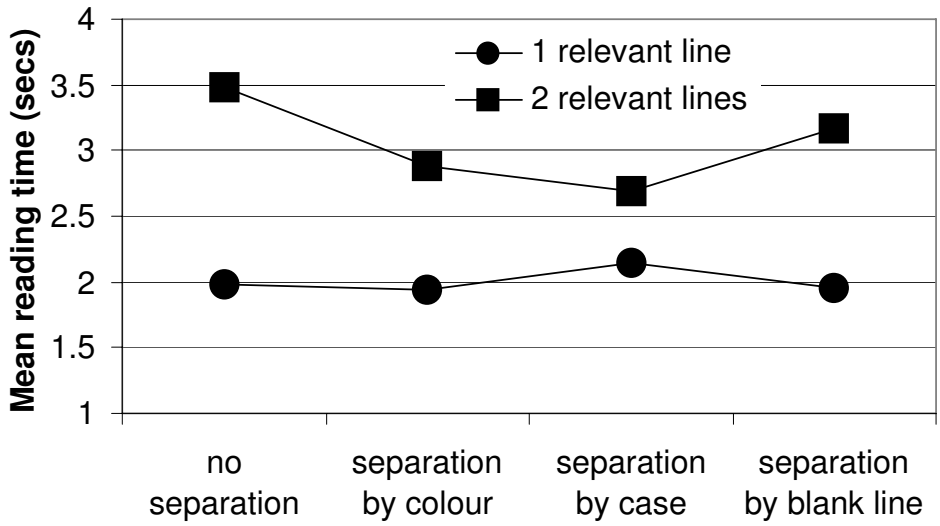


Figure 4 Effects of different separation techniques on reading times

4.5 Message sequencing

The two possible ways of displaying a bilingual four-line sign (two lines in English, two lines in Welsh) were compared.

4.5.1 Four-line bilingual signs

First of all, the experiment studied how a single four-line bilingual sign should be configured, with regards to the shortest reading times. Comparisons were made to discover if the text should be arranged by message language or by message content. The results are shown in Figure 5. It can be seen that when participants were presented with the information arranged by language (i.e. 2 rows of one language followed by two rows of another), the reading times were significantly shorter than if presented with the same information arranged by message content ($F_{[1,29]}$=13.67, p<.001).

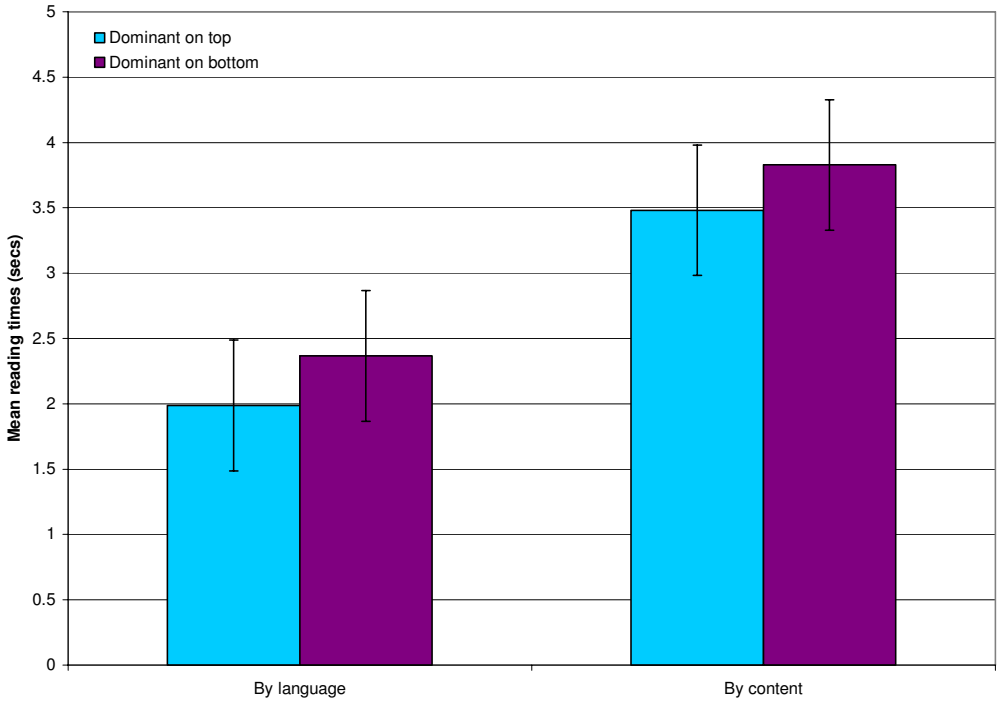


Figure 5 Effect of message configuration, in terms of content and language, on reading times

Again a significant effect was also found for the position of the dominant language. When the preferred language was placed above the non-preferred language (regardless of whether the message was arranged by content or language), reading times were shorter ($F_{[1,29]}$=38.88, p<.001).

This result suggests that participants engaged in more visual search if the languages were split over the four lines of the sign. In practical terms this could mean that visual distraction could be higher in this situation.

4.5.2 Four line message split over two signs

The four lines of text could be split over two sequenced signs (each with two lines of text). However, if the message when split is incomprehensible, as in signs [Q1 then Q2], then confusion may occur as drivers attempt to search for the rest of the message on the following line. It was hypothesised that if the information in the bilingual signs was self-contained, then reading times would be the same as other bilingual signs. However, if the message was split in such a way that drivers tried to complete the rest of the message by searching in the non-preferred language text, then reading times would be increased.

The reading times were compared for ambiguous [Q1 and Q2] and unambiguous signs [P1 and P2].

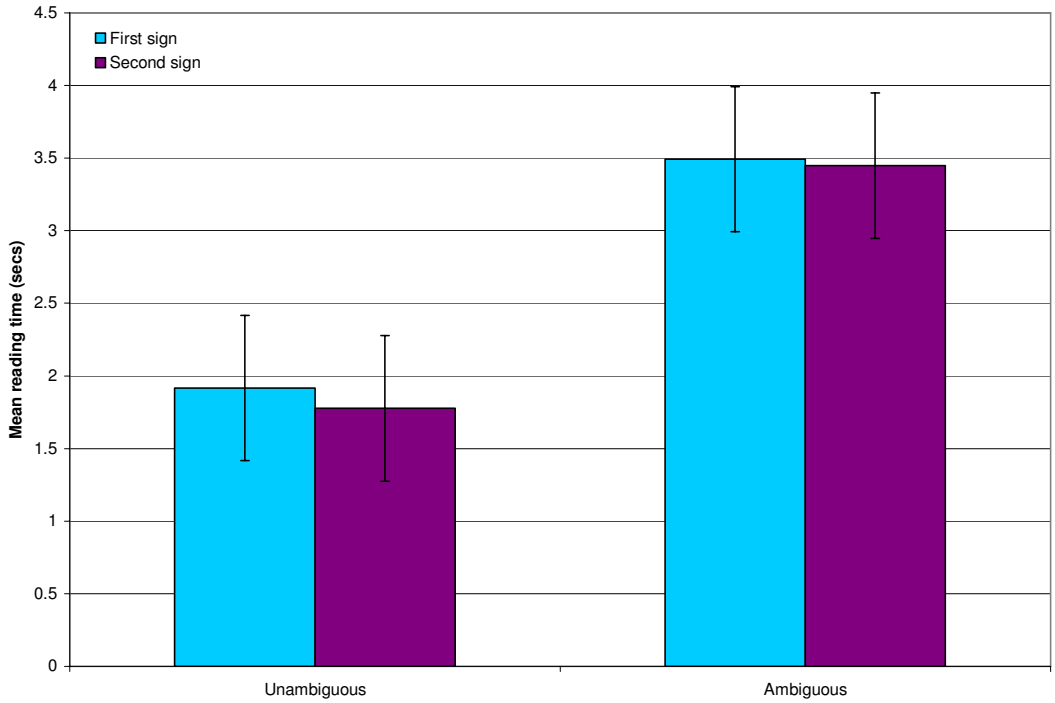


Figure 6 Effect of message ambiguity on reading times

There were significant increases in reading times when comparing the ambiguous and unambiguous signs for both the first ($t_{[29]}$=25.07, p< .001) and second ($t_{[29]}$=13.85, p< .001) signs. Therefore, participants were spending longer searching for the missing information in the first sign, and there seems to be additional confusion when presented with the remaining information. Language was not a significant factor.

4.6 Learning and expectancy

A number of signs appeared in sequence; this sequence was designed to test firstly the ability of a participant to learn a pattern and secondly to evaluate the effects of a breakdown in the sequence on reading times. Three types of sequences were studied.

a) In the case type sequence, the first three signs showed the English component in uppercase and the Welsh component in sentence case (colour was constant). The fourth sign showed the opposite.
b) In the colour sequence, the first three signs showed the English component in yellow font and the Welsh component in green font (case type was constant). The fourth sign showed the opposite.
c) In the language position sequence, the first three signs showed the English component on the top and the Welsh component underneath (case type and colour were constant). The fourth sign showed the opposite.

Mean reading times were calculated for the three sequences of case type, colour and language position and are shown in Figure 7.

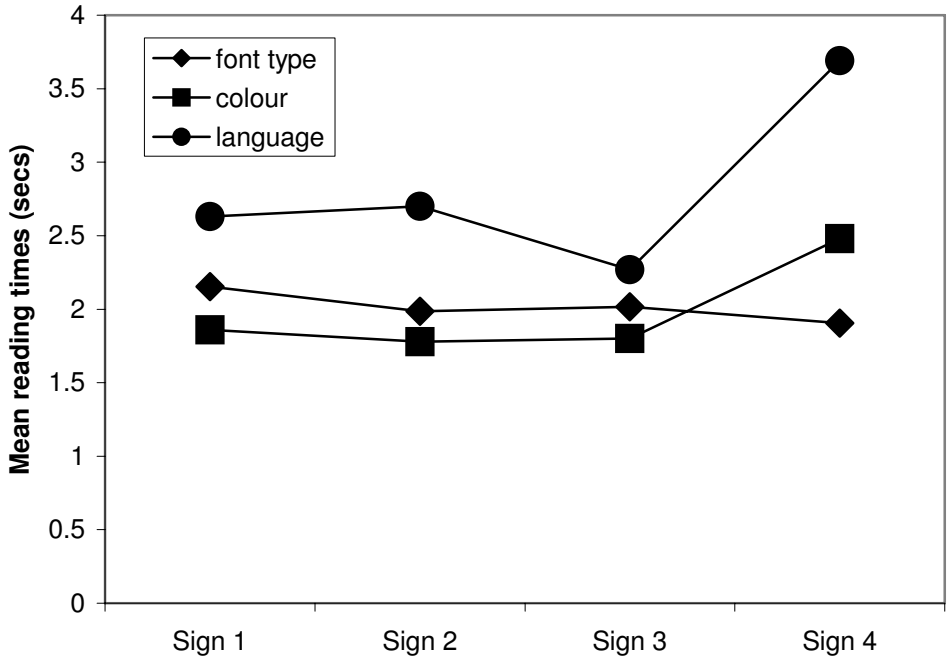


Figure 7 Effect of learning and expectancy on reading times

With regards to the ability of participants to learn a pattern of sign configurations, it appears that demarcation by case type was not readily recognised. All the reading times were approximately 2 seconds. There was no improvement in reading times as the sequence progressed, but on the other hand there was no subsequent deterioration on presentation of the non-standard sign configuration.

With regards to colour, again there was no learning effect as evidenced by the lack of improvement in the reading times from Sign 1 to Sign 3 (although at an average of 1.8 seconds they were shorter than those found in the previous sequence). However, when presented with the non-standard sign configuration, the reading time significantly increased to over 2 seconds ($t_{[29]}$=7.89, p< .001).

In the case of language position, the results are much clearer. Firstly there was a significant decrease in reading times (Sign 1 v Sign 3) after the learning period ($t_{[29]}$=8.25, p< .001). In addition, on the presentation of the non-standard sign, reading times dramatically increased to over 3.5 seconds ($t_{[29]}$=13.56, p< .001). This represents an increase in reading time of approximately 40% compared to the initial reading time in Sign 1. So not only is the learning effect completely eliminated, but additional performance decrements are observed. This suggests that the participants learnt the pattern so effectively that when presented with a non-standard sign, it invoked confusion. These results were not dependant on Language type.

5. Discussion

This study compared the legibility of various VMS configurations in a controlled laboratory setting using a tachistoscope to record reading response times. The signs were viewed by monolingual (English) and bilingual (Welsh and English) participants under ideal viewing conditions. The results are summarised in Figure 8. This figure also shows how far drivers would have travelled (in metres) at a motorway speed of 70mph, given the reading times obtained for each of the sings.

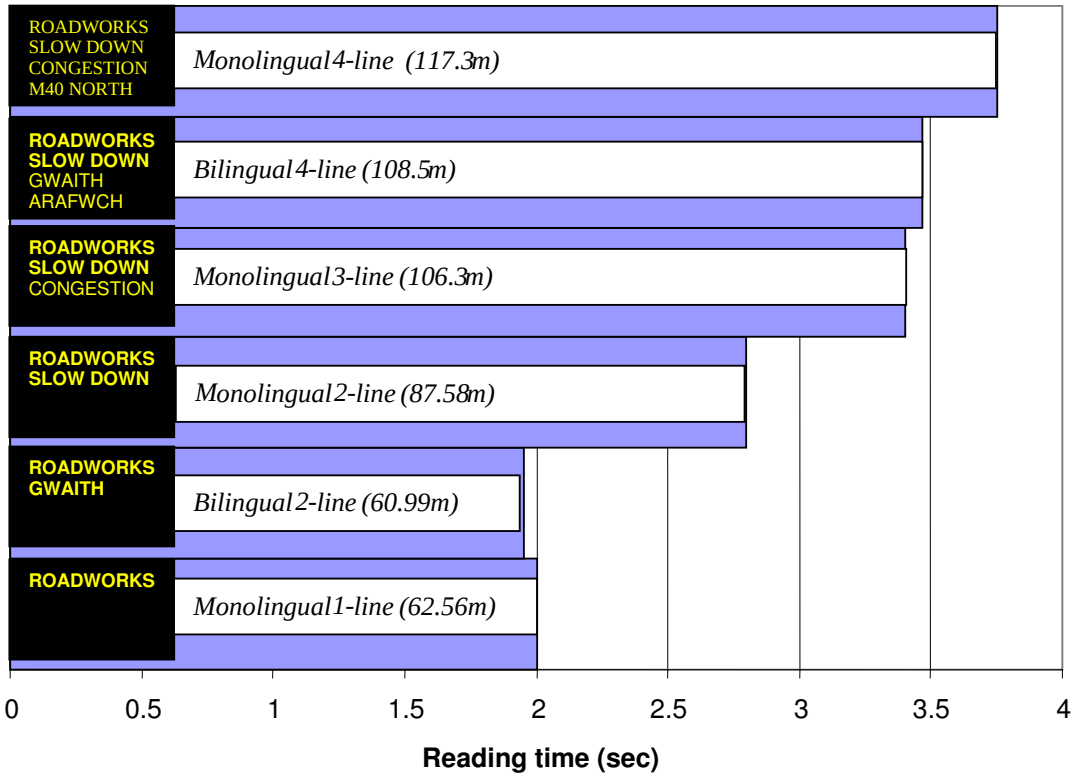


Figure 8 Summary of reading times

As expected, reading response times for monolingual signs increased with the number of lines of text each sign contained. With one line of text, participants responded within 2 seconds; this increased linearly to 3.75 seconds for four lines of text. With bilingual signs, it was found that participants were able to respond just as quickly if they were only required to locate one line of relevant text. However, if this increased to two-lines of relevant text, embedded in a bilingual sign, they responded significantly slower (compared to a monolingual two-line sign). Thus four-line bilingual signs could pose a problem in terms of visual distraction.

It is important to put these results into the context of traffic safety. Could the increase in reading time be associated with an increased accident risk? This is not a straightforward relationship as it would depend on additional factors such as traffic-related variables (e.g. density), driver-related variables (e.g. reaction times) and vehicle-related factors (e.g. brake performance). There are a number of thresholds of visual demand to be found in the literature which define the maximum amount of time a driver should have their eyes off the road. These are commonly referred to as the "2-second rules". However, neither ISO 15007-1 (1997) nor SAE J2364 (1999) state the background for claiming their thresholds.

Whilst extended glances towards roads signs reduce the amount of information gleaned from the forward view, it is not always true that this will cause safety problems. A driver could attend to a road sign for 4 seconds when driving at a low speed in low traffic density; this may not lead to safety problems. On the other hand, a shorter glance away at a higher speed in heavy traffic may increase accident risk. Using single glance duration as a threshold for driving safety is inappropriate unless task difficulty is taken into account. Therefore whilst the results from this laboratory experiment suggest that both bilingual and monolingual four line signs are associated with glances exceeding 2 seconds, it should be noted, that the study was carried out in ideal laboratory conditions – where reading the sign was the primary task. This is not representative of the driving task, where drivers would be required to engage in visual search of the signs whilst maintaining their performance on the primary task of driving. Now that baseline data has been collected, a more realistic testing regime should be employed, e.g. driving simulator trials. In

addition, the way in which participants were required to respond (i.e. verbal response plus reaction time) is not ecologically valid. More commonly, VMS provide information or an instruction (e.g. move into the left hand lane). Taking these two limitations together, it seems logical to suggest that the next appropriate stage of testing would involve the use of either driving simulator or on-road trials. The former is more appropriate as it would allow tightly-controlled testing in a safe environment. Whilst it might be argued that the additional task of driving will only serve to increase reading times, drivers may effectively "chunk" information into more manageable pieces.

A way of aiding and shortening this process of chunking would be to emphasise which information on a bilingual sign is relevant to the driver. In the study reported here, a number of alternative display methods were tested in an attempt to reduce the reading response times for four-line bilingual signs. Firstly it was investigated whether the lines of text should be grouped by language or by message content. Not surprisingly, grouping the lines by language resulted in significantly faster reading times than when the text was grouped by content. Furthermore, locating the user's dominant language at the top of the sign improved reading times for both one and two line messages on bilingual signs.

The study also investigated a range of demarcation techniques, which tried to assist the participants in locating the text applicable to them. Language demarcation using colour, case, and a separation space were all tested. For one line of relevant text on a two-line bilingual sign, none of the demarcation techniques resulted in a significant improvement in reading times. This suggests that participants were already responding as quickly as they could. However for two lines of relevant text on a four-line bilingual sign, each technique significantly improved reading times. The greatest improvement was found using case (upper/lower), then colour then a separation line. Although colour and case were the most powerful demarcation tools, there could be problems to do with implied priority and light output which may make these two tools inappropriate for use in the field.

The improvement in reading time that results from learning a pattern or sequence was also investigated. Significant improvements in reading time were found using demarcation by colour and language position. More importantly, it was found that having established a particular pattern of language, an unexpected change significantly increased reading time. This has two implications; first that there will be a benefit in ensuring that the language order on the VMS should be the same as that on the surrounding static traffic signs. Secondly that travelling from North Wales where Welsh is the first language on traffic signs to South Wales where English is the first language on traffic signs could cause problems.

As an alternative to demarcation, the study investigated whether four-line messages could be split into two sequenced messages. Such an arrangement would allow the existing two-line VMS signs to be retained and used to display longer messages. As would be expected the total reading time for a two-line message split over two signs was longer than a two-line message contained on a single sign. More importantly if the message was ambiguous, the reading time for each component increased by more than 50%. This therefore does not seem a viable option. Other options could include using alternating VMS signs, which display the two languages in turn (Anttila et al, 2000).

A limitation of the study is in the use of reading times as the dependent measure. Whilst this is adequate in such a laboratory setting, where the purpose of the experiment was to compare sign configurations in a tightly controlled manner, a real world investigation would also have to consider whether the sign had been understood, as well as perceived. For example Dewar (1976) instructed drivers to classify signs (e.g. into warning or regulatory signs) whilst Lajunen (1996) measured behavioural response (changes in speed). These would seem a sensible next step, having used this basic laboratory study to select the "best contenders".

6. Conclusions

Results from this study suggest that embedding text in a bilingual road sign increases reading time, compared to a situation where the text was presented separately in a monolingual sign. This increased reading time represents an increase in the search time required to locate the relevant information and could have implications for driver distraction. A number of demarcation techniques were found to aid drivers in their visual search, with the proviso that, if used, these techniques should be implemented at a national level so as to minimise driver confusion. Further work should concentrate on systematically incorporating the task of sign-reading into the driving task in order to evaluate various sign configurations.

Acknowledgements
The Research was supported by the National Assembly for Wales. Thanks to Andrew Bailey for programming the tachistoscope.

References

Anderson, J.R. (1990). Cognitive psychology and its implications. W.H. Freeman and Company: New York.

Anttila, V., Luoma, J., Rämä, P. (2000). Visual demand of bilingual message signs displaying alternating test messages. Transportation Research Part F 3, 65-74.

Colomb, M., Hubert, R. (1991). Legibility criteria of Variable Message Signs. Proceedings of the third International Conference on Vision in Vehicles (pp 309-316). Elsevier Science: North Holland.

Dewar, R., Ells, J., Mundy, G. (1976). Reaction time as an index of traffic sign perception. Human Factors 18(4), 381-391.

Forbes, T.W., Snyder, T.E., Pain, R.F. (1965). Traffic sign requirements. A review of factors involved, previous studies and research needed. Highway Research Record, 70, 48-56.

Hall, R.D., McDonald, M., Rutley, K.S. (1991). An experiment to assess the reading time of direction signs. Proceedings of the third International Conference on Vision in Vehicles, (pp 333-350). Elsevier Science: North Holland.

Harjula, V., Luoma, J., Rämä, P. (1998). Effects of bilingual variable message signs on information overload. TROPIC project. EC Contract RO-96-SC.303/2.

International Standards Organisation, (1999). Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems. Part 2: Equipment and procedures. IOS/CD 15007-2, ISO/TC 22/SC 13/WG8 N216 Rev2.

Johansson, G., Backlund, F. (1970). Drivers and road signs. Ergonomics 13(6), 749-759.

Kerr, J. Snelgar, R. Jordan, T. Emmerson, P., Linfield, P. (1987). Optimum display factors for light emitting variable message road signs. Proceedings of the third International Conference on Vision in Vehicles (pp 277-288).

Lajunen, T., H, P., and Summala, H.(1996). The ergonomics of road signs: explicit and embedded speed limits. Ergonomics. 39(8) 1069-83.

Lesage, P.B. (1981). Design and comprehension of bilingual traffic signs. (Report Number: CR 8101). Ottawa, Canada: Transport Canada Research and Development Centre, Road and Motor Vehicle Traffic Safety Branch.

Mace, D.J. (1988). Sign legibility and conspicuity. Transportation Research Board Special Report (218) 270-293.

Moore, R.L., Christie, A.W. (1963). Research on traffic signs. Proceedings of the Engineering for Traffic Conference (pp113-122). Road Research Laboratory.

Padmos, P., van den Brink, T., Alferdinck, J., Folles, E. (1987). Matrix signs for motorways: system design and optimum light intensity. In A.G. Gale, M.H. Freeman, C.M. Haslegrave, P. Smith & S.P. Taylor (Eds.), Vision in vehicles-II (pp. 269-276). Amsterdam: North-Holland.

Rutley, K. (1972). An investigation into bilingual (Welsh/English) traffic signs. TRRL Report LR 475. Transport Research Laboratory, Crowthorne, UK.

Rutley, K. (1974). A second investigation into bilingual (Welsh/English) Traffic signs. TRRL Supplementary Report 34 UC. Transport Research Laboratory, Crowthorne, UK.

Schneider, W., Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. Psychological Review, 84, 1-66.

Shiffrin, R.M., Dumais, S.T. (1981). The development of automatism. In J.R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Society of Automotive Engineers, (1999). Definition and experimental measures related to the specification of driver visual behavior using video-based techniques. Recommended Practice – SAE J-2396 (Draft).

Staplin L., Gish K.W., Decina L.E., Lococo K.H., Harkey D.L., Tarawneh M.S., Lyles R., Mace D., Garvey P. (1997). Synthesis of human factors research on older drivers and highway safety, Volume II: Human Factors and Highway Safety Research Synthesis FHWA-RD-97-095.

Ullman, G.L., Dudek, C.L. (2001). Effect of roadway geometrics and large trucks on variable message sign readability. National Research Council. 80th Transportation Research Board Meeting, Washington, D.C.

Upchurch, J., Armstrong, J.D., Baaj, M.H., Thomas, G.B. (1992). Evaluation of variable message signs: target value, legibility, and viewing comfort. Transportation Research Record (1376) 35-44.