Eye-tracking Deaf and hearing viewing of sign language interpreted news broadcasts

Jennifer Wehrmeyer
North West University

In this study, the viewing habits of deaf and hearing adults are investigated using eye tracking while they watched interpreted news broadcasts. The study shows that deaf viewers primarily focus on the interpreter and secondarily access picture material, but make very little use of subtitles or lip-reading. In contrast, hearing viewers prioritise pictorial content but also spend significant proportions of time examining subtitles, lip-reading and even watching the interpreter. Viewing patterns are dependent on pictorial information density rather than comprehension. The study confirms the precedence of the interpreter as primary source for deaf viewers, but also questions the efficiency of subtitling as an alternative information source for deaf viewers if an interpreter is present.

Keywords: Eye-tracking, sign language interpreting, subtitling, cognitive psychology, divided attention

The Deaf community

A sociocultural view of deafness. Although most hearing people view deafness merely as a pathological disorder, those (usually but not necessarily profoundly) deaf persons for whom sign language is the main language of communication typically view themselves as belonging to a cultural minority. Those who hold this sociocultural view of deafness refer to themselves as Deaf.

The concept of a Deaf community. Deaf people associate in closed, close-knit local communities linked to schools or churches and may participate in Deaf clubs and organisations that promote Deaf interests (cf. Bidoli, 2009: 132; Humphrey & Alcorn, 1996: 79). They are united by their common use of sign language, which functions as a symbol of identity, medium of interaction and basis of cultural knowledge (Baker-Shenk & Cokely, 1981: 55; Lane et al., 1996: 124-130, 148-9; Lombard, 2006: 14-15, 27-29; Lawson, 2002: 32). The processes of industrialization and globalization have led to interaction between Deaf communities and the formation of national and international organisations that advocate and protect Deaf interests, e.g. DeafSA and the World Federation of the Deaf (WFD), a recognised body within the United States.
Nations. From the 1980s, the rights of the Deaf community have increasingly obtained international recognition and concomitantly, sign language interpreting recognised as a profession.

**The South African Deaf community.** The South African Deaf community is well-established. The primary advocacy organisation, DeafSA (formerly *South African National Council for the Deaf*), was established in 1929 and has nine provincial chapters throughout South Africa (cf. DeafSA, 2008, 2012; SADA, 2012). Deaf people’s interests are also protected by the National Council for Persons with Physical Disabilities in South Africa (NCPPDSA) (Berke, 2009; NCPPDSA, 2011; Newhoudt-Druchen, 2006: 9). However, accurate statistics on the size of the Deaf population still requires research. Although DeafSA’s estimates of 600 000 profoundly deaf signers (SADA, 2012: 9; cf. Berke, 2009; Signgenius, 2009; Olivier, 2007) and one million hearing-impaired persons are generally accepted as the most realistic figures (DeafSA, 2009: 5; PMG, 2007), the recent general 2011 census in South Africa calculated a profoundly deaf population of 0.1% of the South African population (StatsSA, 2012), i.e. only about 52 000 people. Other sources reporting population estimates of hearing-impaired persons vary from 400 000 (SA Yearbook, 2009: 213; StatsSA, 2009) to 4 million (Berke, 2009), with an estimated 12 000 (Lewis, 2009) to 2 million profoundly deaf SASL signers (Olivier, 2007). Since a large population size enhances Deaf lobbying that South African Sign Language (SASL) be recognised as an official South African language, it is not surprising that the census results are not accepted.

As noted by Neves (2008), the Deaf community is not an homogeneous whole, but is stratified according to level of education and literacy, age of acquisition of first language, degree of deafness, use of hearing aids, knowledge of spoken languages, level of general knowledge, etc. (cf. Montero & Soneira, 2010). Such stratification is also true of the South African Deaf community. Based on extrapolation from a statistically representative survey (Wehrmeyer, 2013, forthcoming), the average Deaf South African is born to hearing parents either profoundly deaf at birth or becomes deaf in early childhood and learns sign language on entering a school for the deaf. If white, s/he as an adult has at least Grade 10, is literate and employed in a white-collar job, uses a hearing aid, can speak and lip-read when communicating with hearing persons and mouths when signing with other Deaf. If black, s/he has on average Grade 8, has learnt a trade, can express him/herself in writing using short phrases, does not speak or lip-read well and does not mouth when signing, but relies extensively on facial expression to convey meaning and grammar.

**The status of South African Sign Language**

Another factor that influences the context of TV interpreting in South Africa is the status of SASL.

**Roots.** SASL primarily derives from Irish Sign Language (introduced by Dominican nuns who founded a school for deaf children in Cape Town in 1863), but also has roots in indigenous sign systems as well as influence from the artificial Paget-Gorman sign-supported speech system (Aarons & Akach, 2002: 131-133; Leeson & Saeed, 2012: 44; Lotriet, 2011; Penn, 1992). The South African signing context is further complicated by the strict segregation on the basis of race and clan affiliation in schools (both deaf and hearing) during the apartheid era (Aarons & Akach, 2002: 132), which meant that Deaf groups developed in isolation to each other.

**The Milan resolution.** As was the situation in many other countries, the 1880 Milan resolution to ban sign language as medium of instruction in schools for the deaf (Akach & Lubbe, 2003: 123; Bidoli, 2009: 133) had devastating consequences in South Africa as well. Although implemented in 1920, it was only strictly adhered to in schools for White and Indian children from the 1960s by the apartheid government (Van Herreweghe & Vermeerbergen, 2010).

**Formal recognition of SASL.** It was only in 1994 as a result of both international trends and the new Mandela democracy that SASL was formally recognised as the primary language of the South African Deaf population (Heap & Morgans, 2006: 143; Magongwa, 2012; Reagan, 2008; SA Schools Act, 1996; Constitution, 1996a,b; Ganiso, 2012). In 2001, a National Language Unit was instituted for the development of SASL, and the South African Qualifications Authority established its Standards Generating Body for SASL and SASL interpreting (Heap & Morgans, 2006; Reagan, 2008). Deaf organisations are also lobbying for SASL to be included in the new language policy bill currently being compiled, but is not yet recognised as an official language (PMG, 2009; Reagan, 2012).
**One sign language?** Much research is still needed on the exact nature of SASL. The line adopted by DeafSA and some prominent academics is that there is only one sign language (Aarons & Akach, 2002: 143; Reagan, 2008). Although dialectal variations are acknowledged (DeafSA, 2009: 6; Morgan & Arons, 1999: 356; PMG, 2009: 3, 2007: 4) (viewed as resulting from the apartheid-era isolation of Deaf groups), it is held that they are mutually intelligible and therefore constitute variants of a single language. Proponents of the single language claim that although lexis may differ, the same grammatical constructions are used by all South African Deaf signers, “irrespective of age, ethnicity or geographical region” (Morgan, 2001; cf. Heap & Morgans, 2006: 143; Penn & Reagan, 1994). This adamant approach is fuelled by the need to present a standardised form to the government to be proposed as an official language, rather than being grounded on scientific inquiry.

In contrast, most Afrikaans (by clan) Deaf people regard their sign system as being a separate language (Aarons & Akach, 2002: 135; Wehrmeyer, 2013). The differences in grammatical structure between Afrikaans Sign and that of other SASL users was observed by Vermeerbergen et al. (2007). Leeson and Saeed (2012: 45) also observe that the signing of the Wittebome Deaf community (constituted primarily of Afrikaans-speaking Coloured people) contains handshapes and lexical items characteristic of ISL that do not appear in other SASL forms. Similarly, Newhoudt-Druchen (2006: 8-9) reported that at least five sign language interpreters were needed at national South African Deaf meetings (cf. Morgan, 2008: 107). Moreover, the fact that the Pan South African Language Board (PanSALB) was given a constitutional mandate to develop a standardised form of SASL (Newhoudt-Druchen, 2006: 8-9) also implies that this did not previously exist.

On the other hand, there are some that assert that a number of sign systems/languages exist in South Africa. The HSRC-funded Dictionary of Southern African Signs compiled by Penn et al. (1992) records at least 11 sign systems at that time. Lewis (2009) also reports nine sign language systems in South Africa, which he claims are derived from BSL, Auslan and ASL.

**The relevance of an interpreter**

**Understanding interpreters in general.** The relevance of an interpreter has been questioned in the literature as well as by Deaf viewers themselves. Research reported in the literature indicates that Deaf audiences struggle to comprehend hearing interpreters in general (Jacobs, 1977: 10-14; Jackson et al., 1997: 172-184; Marschark et al., 2004: 345-368; Marschark et al., 2005: 1-34). Initially, incomprehension was ascribed to the deaf person’s weak metacognitive and/or metalinguistic processing skills, lack of education, experience and general knowledge and even to weak signing skills. However, Bidoli (2009: 134) ascribed this lack of comprehension to the problem of divided attention. Moreover, studies also suggests that no significant improvement in comprehension was evidenced for interpretation using natural sign language above that using signed English (i.e. a pidgin form of signing which follows the spoken language structure) (Marschark et al., 2004: 345-368; Marschark et al., 2005: 1-34).

**Understanding TV interpreters.** Comprehension of hearing sign language interpreters on television similarly appears to be problematic. This is consistently indicated in studies undertaken in both the UK (Steiner, 1998; Kyle, 2007; Stratiy, 2005; Stone, 2009) and China (Xiao & Yu, 2009: 155; Xiao & Yu, 2012: 43; Xiao & Li, 2013: 100). In particular, Kyle’s (2007) and Stone’s (2009) studies revealed that Deaf viewers preferred subtitles instead. In these studies, lack of comprehension was ascribed to: poor interpreting skills, poor interpreter signing skills, interpreters focusing on transmitting information and not on producing a coherent message (Stratiy, 2005), the use of dialect, the too small size of the interpreter picture, too fast signing (Kyle, 2007; Xiao & Yu, 2009: 155; Xiao & Li, 2013).

**Understanding South African TV news interpreters.** Similar lack of comprehension of hearing interpreters was confirmed for South African Deaf TV viewers (Wehrmeyer, 2013, forthcoming). (In South Africa, only hearing interpreters are used on TV.) The small picture size and interpreter use of dialect were identified as the two main factors. Other factors included: inconsistent or incorrect interpreter use of SASL features such as facial expression and mouthing; careless, phonologically incorrect or too fast signing; inadequate clothing contrast; incorrect spelling; limited vocabulary; inadequate syntactic constructions; ignorance of Deaf discourse norms; inadequate interpreting strategies such as over-condensation or over-simplification; omission of vital information or syntactic components; (in the case of two
interpreters) weak SASL skills. However it also identified that the Deaf viewer population used numerous dialects/sign systems and that at least 14% had weak signing skills.

It has also been suggested that the South African TV news interpreters used signed English rather than natural SASL (Reagan, 2012). This may well be true of some, but given the present dearth of research into the linguistic features of natural SASL, it is difficult to confirm this. It is rather suggested that the stress of simultaneous interpreting leads to interpreters focusing on information content at the expense of syntactic structure. In fact, the majority of the TV sign language interpreters are children of deaf adults (CODAs) and therefore have native SASL skills.

Interpreting versus lip-reading

A lip-reading culture. Since White and Indian children were taught speech and lip-reading during the apartheid era, it was expected that for these segments of the viewing audience, lip-reading presenters and interviewees would constitute a viable alternative information source. Moreover, research did not prove better comprehension of signed interpretation over lip-reading (Jackson et al., 1997: 172-184). On the other hand, it was hypothesised that Black viewers would not consider lip-reading to be an alternative information source should they fail to understand the interpreter, since speech-training and lip-reading were neglected in Black schools in the apartheid era.

Knowing the language. To be able to lip-read effectively, the Deaf person would have to know the language in which the news was presented or which the interviewees were using. Ironically, deaf children in apartheid South Africa were taught either English or Afrikaans, but no African languages, regardless of the racial categories into which they were forced. Black deaf children were taught in English. Hence Black Deaf South Africans seldom share a common language with their families and clans, unless their families or friends had made a special effort to teach them.

The programme languages for news channels in which signed interpretation is offered are Siswati, Ndebele, Zulu, Xhosa, South Sotho and English. Despite the fact that many Deaf South Africans are literate in Afrikaans, Afrikaans is not offered as programme language on any interpreted news bulletin, although there is a daily uninterpreted Afrikaans news bulletin.

Interpreting versus reading

Subtitles as a source of information. Studies also showed no appreciable difference between viewing an interpreted lecture and reading a printed version of the lecture (Napier et al., 2010). Indeed, research indicates that Deaf audiences advocate subtitles as their preferred means of communication on TV news broadcasts (Kyle, 2007; Stone, 2009; Wehmeyer, 2013). This suggests that reading subtitles or other on-screen text may present alternative sources of information to the Deaf viewer, especially should s/he not understand the interpreter. Moreover, at least for hearing viewers, d’Ydewalle and De Bruycker (2007) showed using eye-tracking that reading subtitles is largely automatic, so that subtitles are read even when they are not essential for comprehension.

TV subtitling. TV subtitling is not necessarily done specifically for Deaf or hard-of-hearing audiences, but may be also offered to hearing audiences who for some reason benefit from a written script (e.g. immigrants). Initially done using stenographers, it is usually achieved by re-speaking using voice recognition software packages to enable live subtitling for news or sports broadcasts (Higgs, 2006; Lambourne, 2006; Marsh, 2006; Pederson, 2010; Remael & Van der Veer, 2006). Although not without problems, this type of service has improved with greater software accuracy and training of respeakers (Ribas & Romero Fresco, 2010). This form of subtitling is usually presented as scrolling subtitles, with a reading rate of approximately 180 words per minute (wpm) (Higgs, 2006; Romero Fresco, 2011) and is adopted in the UK, Denmark, Spain and the Netherlands, among others (Baaring, 2006; de Korte, 2006; Higgs, 2006; Marsh, 2006; Orero, 2006).

Subtitles for hearing-impaired. On the other hand, research into subtitling specifically for Deaf and hard-of-hearing (SDH) audiences has revealed that while Deaf audiences typically demand verbatim subtitling, better comprehension is obtained by edited or summarised versions (Neves, 2008; Pederson, 2010; Romero Fresco, 2009). The BBC policy is to edit where timing constraints require, but to allow verbatim otherwise (British Broadcasting Company, 2009). They are also sensitive to needs of lip-readers when editing out material, attempt to optimally synchronise speech and subtitles and acknowledge
the need to extend subtitle timings if the main picture contains a lot of activity. SDH also pay attention to font, font size, use of capitals versus small letters, contrast between text and background (e.g. the BBC uses white text on black background), colour-coding, on-screen location (Bartoll & Tejerina, 2010; British Broadcasting Company, 2009; Lorenzo, 2010; Pereira, 2010; Utray et al., 2010). They also attempt to convey emotions, whispering, emphasis, shouting, foreign accents, drunken or slurred speech, non-verbal sounds and music in an effort to create a holistic experience for the Deaf viewer (British Broadcasting Company, 2009; Lorenzo, 2010; Pereira, 2010).

**Literacy limitations.** One of the biggest drawbacks of subtitles for deaf audiences is their typically low levels of literacy (Kyle & Harris, 2006; Neves, 2008; Torres & Santana, 2005). This, coupled with concomitant slow reading speed, means that deaf audiences are not always able to derive maximum benefit from subtitles. Similar low levels of literacy are reported for South African deaf adults (Berke, 2009; Ganisio, 2012; Morgan, 2008). Notwithstanding, studies have also shown that video comprehension with subtitles is better than without (Lewis & Jackson, 2001; Neves, 2008).

Eye-tracking evidence has shown that deaf viewers tend to spend more time looking at subtitles than do hearing viewers, possibly due to the greater effort that they need to expend (Kreitz et al., 2013; Szarkovska et al., 2011), but that scrolling subtitles are of little benefit to deaf audiences, with block subtitles allowing easier reading and greater comprehension (Romero Fresco, 2011). Moreover, studies have also shown that Deaf viewers struggle to read comprehensively (whether scrolling or block subtitles) at speeds above 170 wpm, with the recommended reading speed set at 144 wpm (Jensema, 1998; cf. Romero-Fresco, 2009). Moreover, Jensema et al. (2000) found that the higher the subtitle speed, the more deaf viewers focus on them.

**Subtitles on South African news channels.** On South African TV news broadcasts, subtitles are not primarily designed for the hearing-impaired viewer. In South Africa, pre-prepared scrolling subtitles appear on the private ETV channel at the bottom of the main picture to provide a summary of news for those (e.g. business executives) who do not wish to watch the whole broadcast. These subtitles do not necessarily coincide with the immediate story being discussed, so the viewer (whether deaf or hearing) is forced to choose to ignore them and concentrate on the news presenter, or focus on them and ignore the news presenter (cf. Lång et al., 2013). This 5-7 minute summary is repeated throughout the news broadcast. This script is in capitals in black lettering at a reading rate of approximately 180 wpm, corresponding to international practice. Simplified black-lettering block subtitles is also used on both ETV and SABC (the government channel) to provide gist summaries. News captions introducing persona giving the name and position or title of the person being interviewed or discussed appear frequently. These are in capitals using red lettering on a white background on ETV and black lettering on the normal background on SABC channels. Both captions and block summaries tend to stay visible for approximately ten seconds at a time before fading gradually. While captions and block subtitles are offered in the program language, the ETV scrolled summary is only offered in English.

As with lip-reading, text can only be a useful source of information if the Deaf viewer understands the language. Given the lack of knowledge of African languages, it was expected that the Deaf participants would only benefit from onscreen English text.

**The problem of divided attention**

As noted above, according to Bidoli (2009), the need to divide attention between multiple simultaneous information sources detracts from the efficiency of accessing an interpreter on news broadcasts. Attention was initially perceived as unitary (i.e. tasks are processed serially), but is now considered multidimensional (Nebel et al., 2005: 760; Pashler, 1989: 471; Spelke et al., 1976: 219).

**Focused versus divided attention.** The two main aspects of attention are focused and divided attention. Focused attention describes the ability to attend only to relevant stimuli and to ignore distracting ones (Nebel et al., 2005: 760), thereby preventing our sensory and cognitive systems from being overwhelmed by the wealth of visual information in the environment (Tedstone & Coyle, 2004: 277). Although distracting signals are identified in parallel by the perceptual system, only the main signal is transmitted to decision and response initiation processes (Duncan, 1980a; Pashler, 1989: 498). Overt attention covers a very small area of approximately one degree of visual angle around the target (De Valois & De Valois, 1990: 53). According to Ericsson and Ericsson.
distractions outside this area do not slow reaction time for processing and executing tasks, whereas distractions within this area do.

On the other hand, divided attention denotes the skill to distribute limited mental resources to different sources of information or stimuli that arrive simultaneously or nearly simultaneously and process the information into distinct channels (Nebel et al., 2005; 760; Spelke et al., 1976: 219). Divided attention tasks require more attention, are more difficult and demand more working memory load than focused attention tasks (cf. Shlesinger, 2000). Nebel et al. (2005: 770) show that reaction times to complete tasks are consequently slower for divided attention than for focused attention.

Interference effects. Detecting a signal on one channel causes the attention to be brought to that channel. Other channels are still processed (Miller, 1982: 272; Pashler, 1989: 480), but not efficiently, so that observers are likely to miss a subsequent signal on the other channel, especially when time constraints exist (Duncan, 1980b; Miller, 1982: 252). These so-called interference effects are exacerbated if the visual stimuli are large or complex, or if multiple stimuli must be processed in a single coherent task (Paschler, 1989: 480) and impair both tasks, forcing the observer to focus on one at the expense of the other. Spelke et al. (1976), however, demonstrate that it is possible to decrease interference effects with training and practice and thereby train the human brain to divide its attention consciously between two simultaneous complex tasks. This is the primary function of interpreter training programs (cf. Gile, 1995; Shlesinger, 2000). Notwithstanding, for an untrained person such as the average TV viewer, divided attention significantly reduces speed and accuracy in performing visual processing (Pashler, 1989: 478).

Attention models. Nebel et al. (2005: 770) used NMR spectroscopy to demonstrate that visual tasks requiring focused or divided attention demonstrate a common neuronal basis. Under normal conditions, i.e. while performing relatively easy tasks, the ventral frontal regions of the brain are utilised, whereas under high mental demands during divided or focused task processing, dorsal and bilateral areas of the brain are recruited. These findings seem to confirm capacity models which propose that divided attention effects stem from sharing a common pool of resources. If there are not sufficient resources for secondary tasks or component stages of tasks, the efficiency with which each task operates is reduced (Pashler, 1989: 478). Gile’s (1995) effort model, used to account for difficulties experienced while interpreting, is an example of a capacity model. In contrast, postponement models (based on the observation that task response time increases with increasing number of stimuli) propose that a single processing mechanism simply queues tasks in order of stimulus appearance (Pashler, 1989: 469). In other words, when a person tries to perform two tasks at once (such as comprehending both a signed and written message simultaneously), performance is slowed and the secondary task delayed until the primary task has been completed (Duncan, 1980b; Pashler, 1989: 471).

Based on these models, a viewer is able to perceive the interpreter, subtitles and pictorial content simultaneously, but may not be able to successfully divide his/her attention between them and thus may be forced to focus on one information channel at a time. On the other hand, when multiple stimuli redundantly indicate the same response (e.g. coinciding of subtitles with signed message or pictorial content), they activate the response simultaneously rather than queuing (Paschler, 1989: 481). This would explain the semi-automatic response to subtitles in films observed by d’Ydewalle and De Bruycker (2007).

Auditory versus visual primary task. According to Pashler (1989: 498,504-8), when the first task is a complex visual task, accuracy in performing a second visual task is reduced by approximately 20%. In contrast, when the first task is auditory, second task accuracy is unaffected by the time between stimuli even when the second task requires complex perceptual judgment (Pashler, 1989: 504). This means that hearing viewers, whose primary task involves processing a verbal message, are better able to process a secondary visual task (e.g. reading subtitles that do not coincide with pictorial information) than deaf viewers whose primary task involves processing a complex visual message (i.e. the signed interpretation).

Method

Eye-tracking technology using the Tobii T60 eye-tracker was employed to determine Deaf and hearing persons’ viewing patterns.
Procedure

After signing a written consent form and filling in a form with their personal information, participants were seated in front of the eye-tracker fixed-screen monitor. After automatic 9-point calibration, participants were requested to watch a video consisting of short excerpts from three news broadcasts. All the participants watched the videos with sound presented through the inbuilt speakers. The researcher sat in the adjoining observer room (a sound-proof room separated from the test area by a thick glass window). After viewing the entire video, participants were asked which interpreters (if deaf) or program languages (if hearing) they understood and to give their opinions on each broadcast. For the Deaf participants, this was done in either English or Afrikaans as well as in SASL.

The eye-tracker used for the experiment was the Tobii T60, owned by the School of Computing at the University of South Africa in Pretoria. The Tobii T60 display hardware has a resolution of 1280x1024 pixels, a sampling rate of 60 Hz and a response time of 5-16 ms. On the Dell T5400 XP desktop computer connected to the eye-tracker, for a single video stimulus, the offset mean time between stimulus and Tobii display is 20 ± 4 ms and between display and computer graphics card is 33± 7 ms.

Content of video material.

The video material comprised a ten-minute recording of excerpts from three interpreted news broadcasts aired on 21 November 2010, namely the 17.30 news on SABC1 (T1) in SiSwati with interpreter A, the 18:00 news on ETV (E6) in Ndebele with interpreter B and the 22:00 news on ETV (E10) in English with interpreter C. Each excerpt was approximately three minutes long. Care was taken to choose English interviewees on the three channels to maximise the possibility of participants’ lip-reading. T1 only had news captions in Siswati, whereas E6 had news captions in Zulu but scrolling subtitles in English. E10 had all on-screen text in English. The composite video was compiled using Windows Movie Maker and then loaded as an .avi file onto the Tobii project.

All (hearing) interpreters claimed to use standardised, natural SASL. However, at this stage of the research, it was already evident from extensive national questionnaire research (cf. Wehrmeyer, 2013: 179, forthcoming) that interpreter C was understood by most respondents, interpreter B by most Black (but not White) respondents and interpreter A not understood by most respondents.

Participants

Sample groups. Two groups were constituted, namely a hearing control group and a Deaf group. All Deaf participants were profoundly deaf and did not wear any hearing aid for the duration of the test, i.e. they had no access to the auditory message. Although twenty participants were invited for each group, only thirteen Deaf participants were able to attend, either due to transport difficulties or because they were not given permission to take leave from their workplace. This reflects the lower socio-economic conditions under which many deaf South Africans live. Both groups were representative in terms of age, race and gender.

Reading proficiency. In order to ensure sufficient levels of literacy to read and comprehend any written text/subtitles in the video, all participants had at least Grade 10. Because of the low standard of education at deaf schools, the Deaf participants’ literacy levels were also pre-tested by asking them to read the consent form and fill in a form. Their comprehension was then ascertained in a signed dialogue. Although the Deaf participants were selected on the basis that they had at least conversational knowledge of English, only five regarded English as their main spoken language, whereas the remaining six were more proficient in Afrikaans.

Signing proficiency. Only two of the group were born deaf to a deaf parent and thereby qualify as native signers (Johnston, 2010: 107). However, the fact that international statistics indicate that 90% of the Deaf community are born to hearing parents and are therefore not native signers (Akach & Lubbe, 2003: 107; DeafSA, 2012; Marschark, 1997: 47-49; Stone, 2010: 2) means that the television interpreting services should be accessible to this majority (Wehrmeyer, 2013). Hence it was regarded as a sufficient selection criterion that the Deaf participants use sign language confidently as their first language choice. Hence all Deaf participants are active members of their local Deaf communities and were also well known to the Deaf research assistant who recruited them. The six Afrikaans-speaking Deaf participants were all members of the Pretoria Deaf community and described their signing dialect as Afrikaans Sign Language. The five English-speaking Deaf participants were members of the Johannesburg Deaf community and four were also active
DeafSA members. They regarded themselves as using standardised forms of SASL. (Afrikaans Deaf communities are not active in DeafSA since they feel that their sign language is threatened by DeafSA’s attempts to produce a standardised SASL.)

**Comprehension profiles.** Both groups showed similar comprehension profiles for their primary information sources in that only one hearing participant understood the SiSwati language on T1 and only one Deaf participant understood interpreter A (whom he said used a Venda dialect that he had learnt as a boy), whereas five hearing participants understood the Ndebele language and four Deaf participants understood interpreter B on E6. All participants understood either the English language (if hearing) or interpreter C (if Deaf) on the E10 broadcast.

**Eye-tracking analysis**

**Experimental constraints in eye-tracking Deaf people.** Greater prior preparation was required for the Deaf participants, since all instructions had to be conveyed beforehand using SASL and lip-reading. Deaf participants were also highly susceptible to visual interruption. Due to a non-individualistic group philosophy and the silent nature of sign language, greater care also had to be taken to ensure that Deaf participants who had undergone the experiment did not reveal information to those still waiting. For this purpose, the Deaf research assistant monitored conversation in the reception area.

**Data selection.** For the analysis, 18 suitable scans were obtained from the hearing group and 11 from the deaf group, primarily based on sufficiency of eye-tracking data. It is difficult to detect gazes of participants who demonstrate eye deficiencies (such as excessive blinking, squinting or far-sightedness), have nervous tics, are heavy-lidded or wear certain lenses (especially bifocals). Scans of three people who moved frequently and one from a hearing person who admitted to contrived viewing patterns had to be discarded. Generally, Deaf participants lost more data due to body movement than hearing participants, possibly an indication of the greater stress placed on their visual concentration.

**Preparing data.** Fixations at the beginning and end of the video material were discarded to eliminate irregularities. The data was also cleaned for instances where titles of the news broadcast had been inserted, since they did not appear in the original TV news bulletins, and for instances where the eye-track was lost, e.g. when participants leant too far forward. The data were complemented by heat maps.

**Areas of interest.** Four areas of interest (AOIs) were identified, namely: interpreter, picture, text and mouths. The category text includes all legible text superimposed on the screen by the broadcasting corporation, i.e. scrolling subtitles, block subtitles, news captions and legible sports scores. It did not include, however, text that was integrated into the main picture, such as excerpts from documents, financial indicators, weather maps, etc. All areas outside the designated locations, i.e. fixations on the logo, clock, background image and off-screen glances were categorised as other.

**Results**

**Differences between deaf and hearing viewing patterns**

The overall results of the eye-tracking scans were analysed in order to determine differences between Deaf and hearing patterns in terms of the areas of interest.

The heat maps in Figures 1 and 2 below depicting fixation count during a presenter scene for E10 illustrate the typically different foci for deaf and hearing participants:

![Figure 1: Heat map for Deaf group (presenter scene)](image)

(The fixations at the bottom of the screen mark the band where the scrolling subtitles appear.)

Figure 1 demonstrates that the Deaf participants typically fixated on the interpreter (and interestingly, more on
the face than on the hands, although this requires further research). Mean Deaf fixation durations on the interpreter were greater than those for the hearing group (Deaf $M_{df}(\text{interpreter}) = 0.84$ s, SD = 0.908s; Hear $M = 0.417$ s, SD = 0.318s). The high standard deviation reflects the non-homogeneity of the group (cf. Neves 2008). In contrast, Deaf participants paid less attention to picture or text AOIs (Deaf $M_{df}(\text{picture}) = 0.253s$, SD = 0.131; Deaf $M_{df}(\text{text}) = 0.182$, SD = 0.114s).

Table 1: Deaf and hearing total fixation durations as a percentage of video time.

<table>
<thead>
<tr>
<th>Area of interest</th>
<th>Group</th>
<th>Interpreter</th>
<th>Picture</th>
<th>Text</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hearing</td>
<td>12% (7%)</td>
<td>70% (9%)</td>
<td>14% (5%)</td>
<td>4%   (2%)</td>
<td></td>
</tr>
<tr>
<td>Deaf</td>
<td>75% (11%)</td>
<td>18% (7%)</td>
<td>5% (3%)</td>
<td>2%   (1%)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard deviations are given in parenthesis.

**Differences between the two groups.** The data shows that Deaf and hearing participants view the interpreted news broadcasts differently. Hearing participants focused 70% of the time on the main picture, but also gave attention to on-screen text (14%) and the interpreter (12%). In contrast, Deaf participants concentrated mainly on the interpreter (75%), with secondary attention given to the main picture (18%) and little attention given to on-screen text (5%). All three AOI categories displayed significant differences for the two groups (Interpreter: df = 27, $t = 16.5$, $P = 4.96 \times 10^{-11}$; Picture: df = 26, $t = 13.7729$, $P = 1.87 \times 10^{-13}$; Text: df = 26, $t = 3.3548$, $P = 0.0025$)

**Text availability.** Since text was only available 79% of the total viewing time, its use was also calculated in terms of availability. These calculations showed that Deaf participants looked at 7% of the available text, whereas hearing participants looked at 18%, i.e. hearing participants read more than twice the amount of text that the Deaf participants read. Qualitative analysis of the video material showed that the text viewed by the Deaf participants were mainly news captions introducing interview persona and that the scrolling subtitles were mostly ignored except for occasional brief glances (which may indicate that these were perceived to be more of a distraction than a secondary information source).

**Mouth fixations.** When mouth fixations were analysed, it was found that Deaf participants focused on mouths 3% of the time, whereas hearing participants focused 15% of the time. Mouth foci were only available 58% of the time; hence, in terms of availability, Deaf participants focussed on mouths 12% and hearing participants 53% of the available time. This indicates that hearing participants made more use than Deaf viewers of lip-reading as an alternative information source. These differences between Deaf and hearing groups are, however, not significant (df = 20, $t = 0.9176$, $P = 0.3698$).

Figure 2: Heat map for hearing group (presenter scene)

Figure 2 illustrates that hearing participants looked mainly at the main picture but also looked at the text and interpreter AOIs. This is also reflected in their mean fixation durations (Hear $M_{df}(\text{interpreter}) = 0.417s$, SD = 0.381s; Hear $M_{df}(\text{picture}) = 0.403$ SD = 0.216s; Hear $M_{df}(\text{text}) = 0.277s$, SD = 0.136s). If a close-up shot of a human was present (as in Figure 2), hearing participants focused on the face. They were also sensitive to minor changes in the logo or time displays. Their fixations tended to be short and their gaze plots characterised by movement between all areas of interest, demonstrating that the hearing participants monitored all information sources.

Participants’ average total visit duration for each AOI for all three clips are depicted in Figure 2 below for each group as a percentage of the total cleaned viewing time (580s), with the standard deviations in each category given in Table 1 below:

Table 1: Deaf and hearing total fixation durations as a percentage of video time.
Comprehension challenges

The data for each bulletin was then analysed separately in order to determine what effect comprehension had on viewing patterns. Overall, variations in total visit durations for each channel compared to the overall AOI patterns discussed above were not significant for Deaf participants. This indicates that Deaf participants’ fixation patterns are independent of their comprehension of a particular bulletin. However, ANOVA analysis of total visit durations for AOI picture categories for hearing participants showed T1 > E6 > E10 (df = 53, F_{calc}=4.389 > F_{crit}=3.179), indicating that the less hearing participants comprehended the spoken language, the more attention they invested in the picture. When AOI picture data was analysed according to various content scenes (discussed below), it was found that this pattern was mainly evident in presenter scenes (scenes containing only the newsreader) (df = 50, F_{calc}=7.473 > F_{crit}=3.179), indicating that the less hearing participants comprehended a speaker, the more attention they invested on him/her.

On the other hand, the start of scrolling subtitles on E6 increased textual fixations for both groups of participants for this bulletin at the expense of fixations on other information sources. Nevertheless, text remained a tertiary source of information on all bulletins for both deaf and hearing participants.

The data was also analysed in order to determine whether content influenced viewing patterns.

Content variation

Types of scenes. The pictorial content was divided into presenter scenes in which only the newreader is present, interview scenes which depicted interviews or speeches from prominent figures or members of the public, action scenes such as sport or stories that contained a lot of movement in the main picture and textual scenes in which the picture contained legible text, such as excerpts from documents, financial indicators and weather maps.

The resulting total visit duration analyses for the Deaf participants are depicted in Figure 3 below:

Figure 3: Averaged scene analysis for Deaf participants

The fixation analyses for hearing participants are depicted in Figure 4 below:

Figure 4: Averaged scene analysis for Deaf participants

For both groups, fixations on the picture increased as pictorial information density increased, with consequent decrease of fixations on other areas of interest. The results also revealed that Deaf participants look at interviewees more than at presenters. Secondly, Deaf participants seldom fixated on mouths except for interview scenes, whereas hearing participants do. In fact, hearing participant mouth fixations on non-speaking human figures in action scenes are similar to those of Deaf participants for interview scenes. Thirdly, for both groups, textual fixations are the highest for human subjects, i.e. the interviewee or presenter, and decrease as pictorial content increases. Although hearing participants...
consistently accessed text more than Deaf participants across all scene types, both groups relegated text to a tertiary information source.

The differences between the Deaf and hearing groups were consistently found to be significant at 95% probability levels for the AOI categories of interpreter and picture for all scenes, whereas differences were only significant for the AOI mouth category in presenter and interview scenes and were not significant for the AOI text category for any scene.

**Discussion**

The study showed that the interpreter was the primary information source for the Deaf participants, who fixated on the interpreter for long periods of time and seldom monitored alternative sources of information. This reliance was found to be dependent upon the information content in the picture and not the level of comprehension of the interpreter. Apart from the interpreter, only pictorial information is accessed significantly, whereas textual or lip-reading information is seldom accessed.

In contrast, the hearing participants spent most of their viewing time looking at the picture, but also constantly monitored all information sources, seldom fixating long on a single object. Similarly to the Deaf participants, hearing pictorial fixations increased with increasing pictorial information density at the expense of other information sources. Moreover, pictorial fixations also increased with lack of comprehension of a programme language, with hearing participants showing particular increased visual attention to newsreaders when the language is not comprehended. On the other hand, reading text and lip-reading are only utilised as tertiary information sources. Notwithstanding, hearing participants utilised them more than did the Deaf participants, although these differences were not always significant.

That hearing viewers are better able to manage multiple sources of information confirms Pashler’s (1989: 504) findings that divided attention is more efficient if the primary task is auditory rather than visual. Thus hearing viewers use their visual cognitive capacity to maintain the picture as preferred secondary information source and primary visual source and still monitor other information sources (text, interpreter and human mouths). On the other hand, the Deaf viewer’s primary task is visual and complex; hence the extra load on the visual cognitive capacity meant that only one secondary information source (the picture) could be effectively monitored and other sources such as text or human mouths are rarely accessed.

The secondary task selected by both groups was examination of pictorial information. The results therefore indicate that viewers adhere to a hierarchy of information sources. The interpreter, if present, is always retained as primary source by Deaf viewers, regardless of comprehension levels. Picture, text and mouths are secondary sources of information which compete with each other for participant attention. It is suggested that since reading and lip-reading require more processing effort than looking at pictures, they are relegated to third priority and only accessed if pictorial information is insufficient. Moreover, since they are not always available, they may be regarded as less reliable sources requiring even further effort to monitor their availability.

Notwithstanding Deaf groups’ demands for subtitles, the results show that if an interpreter is used, very little visual cognitive capacity remains for processing on-screen text. In particular, the study finds that the present methods of subtitling employed on South African news programs is not beneficial to the Deaf audience. It may well be that synchronising subtitles or introducing SDH would increase the attractiveness of text as an alternative information source for these viewers. Within the present paradigm, however, it is recommended that subtitles be restricted to scenes where pictorial information density is minimal, i.e. presenter and interviewee scenes, for both audience groups.

Although differences between Deaf and hearing participants in this sample were often found to be significant, the small sample sizes of both groups preclude extrapolation to the larger population, necessitating further research. Secondly, participants were selected from a 100 km radius around Pretoria, which also influences the validity of results in terms of national population. The small sample size also meant that although the standard deviations did suggest differences in the abilities of Deaf participants, further study is required in order to ascertain how these differences affect the results. Non-homogeneity of the Deaf audience is made even more complex in South Africa due to the multilingual and multicultural nature of its population, both regarding
spoken languages used as well as variations in the signed languages/dialects.

Conclusion

The study confirmed the main hypothesis that the primary source of information for South African Deaf viewers of interpreted TV news broadcasts is the interpreter, almost to the exclusion of any other visual information source except the picture and regardless of comprehension of the interpreter. It also revealed differences in viewing behaviour between Deaf and hearing participants.

The results contribute to our knowledge of Deaf and hearing viewing patterns on the one hand, and to our understanding of cognitive patterns in divided attention on the other. They also establish a platform for further research in exploring optimal use of multiple information sources on TV news broadcasts.

Acknowledgements

The author wishes to thank the University of South Africa’s School of Computing for the use of their Human-Computer Interaction laboratory.

References


