To cite this article: A. Hatzidaki, M.W. Jones, M. Santesteban & H.P. Branigan , Language and Cognitive Processes (2013): It's not what you see: it's the language you say it in, Language and Cognitive Processes, DOI: 10.1080/01690965.2013.857782

SHORT ARTICLE

It's not what you see; it's the language you say it in.

A. Hatzidaki[a,b], M.W. Jones[c], M. Santesteban[d], H.P. Branigan[a]

[a]*Department of Psychology, University of Edinburgh, Edinburgh, UK*

[b]*Department of Technology, Universitat Pompeu Fabra, Barcelona, Spain*

[c]*School of Psychology, Bangor University, Bangor, UK*

[d]*Department of Linguistics and Basque Studies,*

University of the Basque Country, Vitoria, Spain

Correspondence should be addressed to Anna Hatzidaki, Center of Brain and Cognition, Department of Technology, Universitat Pompeu Fabra, Barcelona, 08018, Spain. E-mail: ahatzidaki@gmail.com

To cite this article: A. Hatzidaki, M.W. Jones, M. Santesteban & H.P. Branigan , Language and Cognitive Processes (2013): It's not what you see: it's the language you say it in, Language and Cognitive Processes, DOI: 10.1080/01690965.2013.857782

Abstract

In an eye-tracking experiment, we investigated the interplay between visual and linguistic information processing during time-telling, and how this is affected by speaking in a non-native language. We compared time-telling in Greek and English, which differ in time-telling word order (hour vs. minute mentioned first), by contrasting Greek-English bilinguals speaking in their L1-Greek or their L2-English, and English monolingual speakers. All three groups were faster when telling the time for digital than for analogue clocks, and when telling the time for the first than the second half hour. Critically, first fixation and gaze duration analyses for the hour and minute regions showed a different pattern for Greek-English bilinguals when speaking in their L1 versus L2, with the latter resembling that of English monolinguals. Our results suggest that bilingual speakers' eye-movement programming was influenced by the type of time-telling utterance specific to the language of production currently in use.

Keywords: language production; bilingualism; visual cognition; time-telling.

Recent evidence showing a close relationship between the order in which objects are fixated and the order in which they are mentioned (e.g., Gleitman, January, Nappa, & Trueswell, 2007; Griffin & Bock, 2000) has been interpreted as showing a causal relationship between processes of language production and of visual cognition, so that linguistic information needed for production of a particular utterance can determine eye-movement programming (Bock, Irwin, Davidson, & Levelt, 2003).

In this paper we investigate the relationship between visual and linguistic processing from a bilingual perspective: Do bilingual speakers of languages with different linguistic characteristics show eye movements that reflect the properties of the language being produced? We track the eye movements of Greek-English bilinguals producing time-telling expressions in either their first (L1-Greek) or their second (L2-English) language, and English monolinguals producing time-telling expressions in English.

Speakers often have to refer to the visual context. For example, time-telling normally requires speakers to look at a clock or watch. Sometimes, namely mapping from a digital clock display to an *absolute* time expression (*It's nine oh five* to express 9.05), production simply involves retrieving the lexical entry corresponding to each number from memory (Korvorst, Roelofs, & Levelt, 2007; Meeuwissen, Roelofs, & Levelt, 2004). But sometimes, speakers must carry out more complex conceptual and linguistic processing. Producing a *relative* time expression (*It's five past nine* to express 9.05) requires speakers to generate a conceptual representation identifying the point-of-time referent (hour) and the distance-from that referent (minute before/after), and to retrieve relevant lexical and numeric information for structuring in an appropriate syntactic frame.

In languages like English and Dutch, relative time utterances express distance-from information preceding the point-of-time referent, so that minutes are mentioned before the hour. But in languages like Greek, the order of distance-from and point-of-time information is reversed, so that the hour is mentioned before minutes (*Είναι εννέα και πέντε,* 'It's nine plus five' to express 9.05). Assuming incremental processing (e.g., Levelt, 1989; Roelofs, 1998), we might expect these word order differences to be reflected in speakers' uptake of relevant visual information, consistent with research showing a close relationship between order of fixation and order of mention during production of event descriptions (e.g., Griffin & Bock, 2000; see Griffin, 2004, for a review).

Bock et al. (2003) investigated whether English and Dutch native speakers would show a tight temporal relationship between inspection of regions of digital and analogue clocks, and production of elements of time-telling expressions. Both English and Dutch speakers showed patterns of eye movements that mirrored the word order of relative expressions, with both groups fixating first the minute and then the hour region, regardless of display format; both groups also showed longer first fixations in the minute than the hour region. Hence the linguistic information needed for production of a time-telling utterance can determine eye-movement programming, with eyes landing first on the time region that constitutes the starting point of the upcoming utterance.

Bock et al.'s (2003) study provides evidence for a tight coupling between planned utterances and visual attention and makes straightforward predictions about visual attention in L1 speech production. Both English and Dutch speakers fixated clock displays in an order consistent with the minute-hour order mandated by their

L1; by extension, we would expect that speakers of languages that mandate hour-minute order, such as Greek, would show the opposite order of fixation.

But it is less clear what might determine visual attention in bilingual production. In particular, what governs eye-movement planning in bilingual speakers whose two languages mandate different word orders? Eye-movement planning might be determined strictly by the properties of the language being produced (*target language*; cf. Slobin's (1996) notion of *thinking for speaking*). In that case, Greek-English bilinguals' eye movements would depend only on whether they were telling the time in Greek (hour-minute order) or English (minute-hour order). However, recent evidence suggests that the grammatical properties of bilinguals' *non-target language* (language not produced) can interfere with task performance in the target language (Hatzidaki, Branigan, & Pickering, 2011). Furthermore, bilingual speakers may be influenced by prior experience of consistently mapping particular visual elements onto particular linguistic structures in their L1 (e.g., Athanasopoulos, Wiggett, Dering, Kuipers, & Thierry, 2009). Thus it is possible that in bilinguals, the relationship between planned utterances and visual attention might not be so straightforward as proposed for monolinguals; instead, bilinguals' eye movements may be mediated in some way by properties associated with their other language.

The present study therefore explored how visual attention is affected by conflicting syntactic representations in bilingual speakers' languages. We tracked eye movements when a group of Greek-English bilinguals produced relative time expressions in their L1-Greek or their L2-English, and a group of English monolinguals produced relative[1] time expressions in English, while viewing digital and analogue clock displays.

If eye movements are structurally guided by a language-specific linguistic plan tied to the context of language use, then Greek-English bilinguals should tend to fixate the hour region first when producing relative time expressions in their L1-Greek but the minute region first (in the same way as English monolinguals) when producing relative expressions in their L2-English. Alternatively, Greek-English bilinguals might show a consistent tendency to fixate the hour region first, when producing both Greek and English. Such a finding would imply that the relationship between utterance planning and visual attention is not straight forward in bilingual production, and might reflect either an influence from activation of the corresponding time-telling expression in the L1 grammar, or an experience-based preference to fixate first the region that is mentioned first (because Greek-English bilinguals have overwhelmingly more experience in time-telling in L1-Greek than L2-English). Following earlier research, we also predicted longer naming latencies for analogue than digital clocks, and for times that are larger rather than smaller in number magnitude (Bock et al., 2003; Korvost et al., 2007).

Experiment

Method

Participants. Sixty students at the University of Edinburgh were paid to participate (40 Greek-English bilinguals, M= 25 years, SD = 2.41; and 20 English monolinguals, M = 22 years, SD = 2.63). The Greek-English bilinguals were highly fluent in English: Each participant had an English language qualification, such as

[1] The preferred time expression for British English as opposed to American English (see Bock et al., 2003).

IELTS (with a score of ≥ 6.5 in each section); TOEFL iBT 100 (with a score ≥ 23 in all sections); PTE (with a score ≥ 61 in each of the "Communicative Skills" sections); Cambridge CPE Grade B or Cambridge CAE Grade A), and a mean of 7 years' (SD = 1.06) formal instruction in English before moving to the UK. Bilingual participants had also been resident in the UK for 4 years on average (SD = 1.42).

Materials. 256 items were distributed across two blocks of clock-image displays (analogue and digital); each block contained 64 critical items and 64 fillers. Clock times ranged from 2.00 to 9.55 (cf. Korvorst et al., 2007), thus creating eight clock types of five-minute intervals (e.g., 2.05; 2.10 and so on). Images showing an actual-hour, half-hour, and quarter-hour (e.g., 2.00; 2.15; 2.30; and 2.45) acted as fillers. All critical items were seen once and filler items twice within each of the two blocks. The order of clock-time presentation was randomised within each block; blocks of clock-image display were counterbalanced across participants so that half saw analogue-digital order, and half saw the reversed order. Analogue image displays did not feature numbers, but had clearly marked gradations. (Figure 1 shows an example of image displays.)

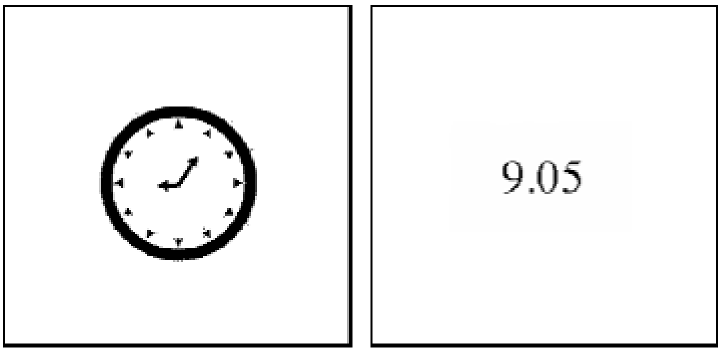

Figure 1. Example of image displays used for analog and digital clocks respectively.

Procedure. Speakers from the Greek-English bilingual group were randomly assigned to a group of *bilingual Greek-speakers* (telling the time in their L1-Greek) or *bilingual English-speakers* (telling the time in their L2-English). Participants were given onscreen instructions in Greek (bilingual Greek-speakers) or English (bilingual English-speakers and English monolinguals). Eye movements were monitored by an SR Research Eyelink 1000 eye-tracker; the experiment was controlled with SR Research Experiment Builder software. Fixation position was sampled at 1000 Hz and saccades prior to critical fixations were detected using a 17-sample saccade detection model with a velocity threshold of 30°/sec, an acceleration threshold of 8000°/sec^2, and a minimum amplitude of 0.5°. Stimuli were presented on a CRT monitor at a viewing distance of 90 cm with a refresh rate of 140 Hz. The displays of analogue and digital clocks subtended a horizontal visual angle of 6.04° and 2.86°, respectively. Viewing was binocular, but only the right eye was tracked.

Adjustment of the infrared cameras attached to the eye-tracker was followed by a calibration procedure (dots viewed in 9 screen locations), repeated whenever measurement accuracy was deemed insufficient (e.g., whenever the participant made a gross head movement). Each trial began with a centrally-presented drift correction (a small circle). The experimenter initiated trial onset following the participant's fixation of the circle. A single trial comprised a centrally displayed fixation point (duration 1000 ms), followed by a clock image display that remained onscreen until participants' verbal response initiation, or time-out (6000 ms for analogue clocks; 4000 ms for digital clocks). A blank screen then appeared for 1000 ms until the fixation point appeared.

Participants were instructed to tell the time displayed on presentation of a clock stimulus, using a relative time expression starting with '*Είναι …*' (bilingual Greek-speakers) or *'It's …'* (bilingual English-speakers and English monolinguals); this ensured a common speech onset. Participants were encouraged to respond quickly, but to ensure they had fully viewed the image before starting to speak. The regions of interest were defined such that each time-region was clearly demarcated for eye-fixation analysis (analogue: hour hand and minute hand regions; digital: regions left or right of the dot). Participants' spoken output was recorded on the PC via an ASIO sound card; recording began automatically at the beginning of the trial and terminated at the end of the trial. A short practice session preceded each block, and a short break was allowed between blocks. Each experimental session lasted approximately 20 minutes.

Scoring and Analysis. Responses were scored as: *Correct* when participants produced the appropriate time using the word order of the target language; *incorrect* when participants produced an inappropriate time; *invalid* when participants produced a speech error, or the voice key was triggered incorrectly; and *omission* when participants produced no response. For naming latencies and accuracy, we based our analysis on *time format* (analogue vs. digital), *distance from referent* (minutes referring to the first half-hour vs. the second half-hour), and *language group* (bilingual Greek-speakers vs. bilingual English-speakers vs. English monolinguals). For eye movement measures (first fixation and first gaze duration before leaving a time region; Bock et al., 2003), we based our analysis on *time format* (analogue vs. digital), *region of interest* (minute vs. hour), and *language group* (bilingual Greek-speakers vs. bilingual English-speakers vs. English monolinguals).

Results

We recorded 1280 responses per time format per group (see Table 1). Because the number of incorrect and invalid responses and omissions was negligible, we focused our analyses on correct responses.

Table 1. *Distribution of correct, incorrect, invalid responses, and omissions for each language group for analogue and digital clocks.*

Language group and Time format	Response Type: Correct	Incorrect	Invalid	Omissions
Bilingual Greek-speakers				
Analogue clock	1126 (88.0%)	36 (2.8%)	24 (1.9%)	94 (7.3%)
Digital clock	1222 (95.5%)	13 (1.0%)	29 (2.3%)	16 (1.2%)
English Monolinguals				
Analogue clock	1105 (86.3%)	39 (3.1%)	28 (2.2%)	108 (8.4%)
Digital clock	1232 (96.2%)	18 (1.4%)	24 (1.9%)	6 (0.5%)
Bilingual English-speakers				
Analogue clock	1083 (84.6%)	54 (4.2%)	35 (2.8%)	108 (8.4%)
Digital clock	1208 (94.4%)	23 (1.8%)	34 (2.6%)	15 (1.2%)

Naming latencies

We conducted separate naming latency analyses for each response language (see Table 2). Bilingual Greek-speakers showed a significant main effect of Time Format ($F(1,19) = 46.74$, $p < .001$), with longer naming latencies for analogue than digital clocks; a significant main effect of Distance from Referent ($F(1,19) = 96.03$, $p < .001$), with longer naming latencies for minutes referring to the second than the first half-hour; and a significant two-way interaction ($F(1,19) = 4.97$, $p < .05$), with a greater difference between second- and first-half hour for digital than analogue clocks.

Bilingual English-speakers and English monolinguals showed a significant main effect of Language Group ($F(1,38) = 4.20$, $p = .05$), with bilingual English-speakers slower to respond than English monolinguals; a significant main effect of Time Format ($F(1,38) = 96.80$, $p < .001$), with longer naming latencies for analogue than digital clocks; and a significant main effect of Distance from Referent ($F(1,38) = 42.39$, $p < .001$),with longer naming latencies for minutes referring to the second than the first half-hour. The only significant interaction was between Time Format and Distance from Referent ($F(1,38) = 9.32$, $p < .01$), with a greater difference between second and first half-hour for digital than analogue clocks.

Table 2. *Latencies (ms) and Standard Deviations in brackets of time-telling for analogue and digital clocks of first half-hour and second half-hour. (Bil. Greek = Bilingual Greek-speakers; Bil. English = Bilingual English-speakers; English Mono. = English Monolinguals)*

Time format and	Language Group		
Distance from referent	Bil. Greek	English Mono.	Bil. English
Analogue clock			
First half	1586 (362)	1656 (398)	1698 (283)
Second half	1723 (420)	1694 (403)	1719 (262)
Difference	*-137*	*-38*	*-21*
Digital clock			
First half	1021 (181)	1018 (109)	1212 (152)
Second half	1238 (238)	1101 (153)	1364 (223)
Difference	*-217*	*-83*	*-152*

Eye movement measures

First fixation analysis (see Figures 2a and 2b) showed a significant main effect of Time Format ($F(1,57) = 39.69$, $p < .001$), with more fixations for digital than for analogue clocks. The main effects of Language Group and Region of Interest were not significant (Fs< 2). Crucially, there was a significant interaction between Language Group and Region of Interest ($F(2,57) = 5.66$, $p < .01$): The two English-speaking groups tended to fixate first the minute region and then the hour region, whereas the Greek-speaking group tended to fixate first the hour region and then the minute region. There was also a significant interaction between Time Format and Region of Interest ($F(1,57) = 57.23$, $p < .001$), with the hour region fixated first more often for analogue than for digital clocks, and the minute region fixated first more often for digital than for analogue clocks. There was also a three-way interaction ($F(2,57) = 3.29$, $p < .05$); separate 3x2 ANOVAs for each clock type yielded a significant interaction between Language Group and Region of Interest in the digital clock analysis($F(2,57) = 5.64$, $p < .01$), showing that the difference in first fixation between minute and hour regions was greater for the two English-speaking groups than for the Greek-speaking group.

Figure 2b indicates that speakers' eye movements tended overall to mirror the structure of linguistic output, but inspection of Figure 2a shows that this effect was mainly driven by the digital clock arrays, with speakers in all groups showing more first fixations in the hour than the minute region (Bock et al., 2003) for analogue clock displays. We suggest that this pattern reflects analogue clocks' richer perceptual space, whereby identifying the hour-referent is more laborious than for digital clock displays. As such, all participants focus initially on identifying the hour hand, overriding any language-based effects. In contrast, identifying regions of interest is less challenging in digital clock displays, allowing effects of target language to manifest clearly.

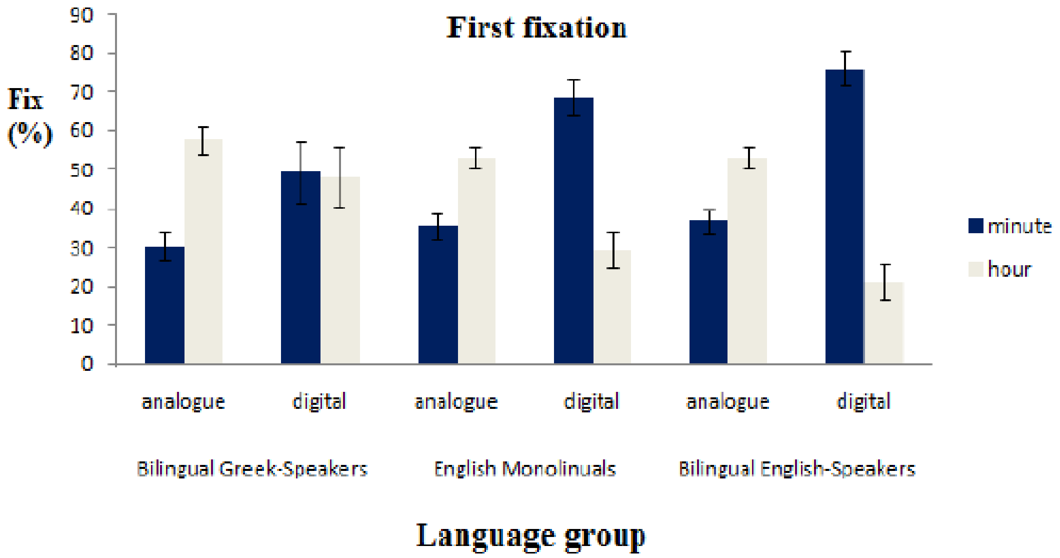

Figure 2a. Percentages of first fixation in min and hour regions of analogue and digital clocks for Bilingual Greek-Speakers, English Monolinguals, and Bilingual English-Speakers. Error bars represent the Standard Error of the Mean.

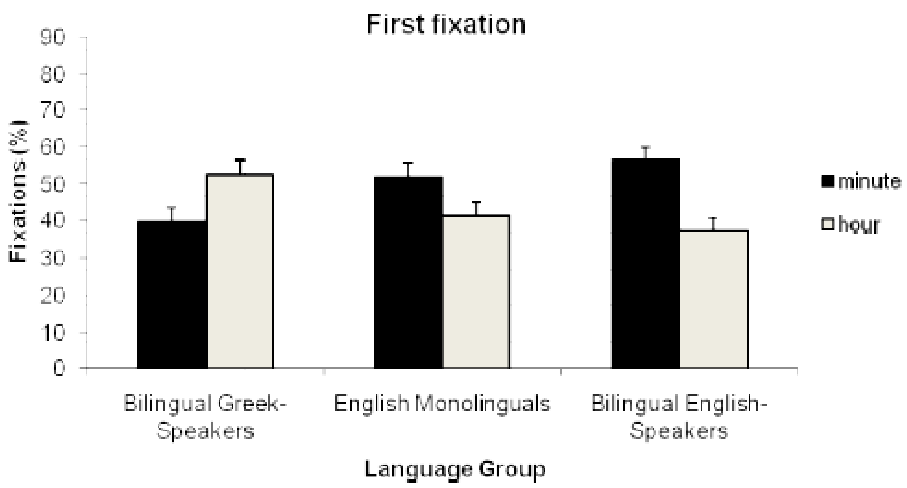

Figure 2b. Percentages of first fixation in min and hour regions collapsed across participants. Error bars represent the Standard Error of the Mean.

The gaze duration analysis (see Figures 3a and 3b) yielded a significant main effect of Time Format ($F(1,57) = 19.65$, $p < .001$), with longer gaze duration for analogue than digital clocks; and a significant main effect of Region of Interest ($F(1,57) = 22.43$, $p < .001$), with longer gaze duration in the minute than the hour region. The effect of Language Group was not significant ($F < 1$). The interaction between Language Group and Time Format was significant ($F(2,57) = 3.77$, $p < .05$); the difference in gaze duration between analogue and digital clocks was greater for bilingual Greek-speakers ($M = 103$) and English monolinguals ($M = 140$) than for bilingual English-speakers ($M = 14$).

There was a significant interaction between Language Group and Region of Interest ($F(2,57) = 4.16$, $p < .05$); the difference in gaze duration between minute and hour regions was greater for the two English-speaking groups than the Greek-speaking group. There was also a significant interaction between Time Format and Region of Interest ($F(1,57) = 39.33$, $p < .001$), with a greater difference in gaze duration between the minute and hour regions for digital than analogue clocks. Finally, there was a significant three-way interaction ($F(2,57) = 4.44$, $p < .05$); separate 3x2 ANOVAs for each clock type yielded a significant interaction between Language Group and Region of Interest in the digital clock analysis only ($F(2,57) = 4.61$, $p < .05$), with the difference in gaze duration between minute and hour regions being greater for the two English-speaking groups than the Greek-speaking group.

The gaze duration analyses showed that English monolinguals consistently spent more time looking at the minute than the hour region. Critically, bilingual English-speakers' overall performance resembled that of English monolinguals rather than bilingual Greek-speakers.

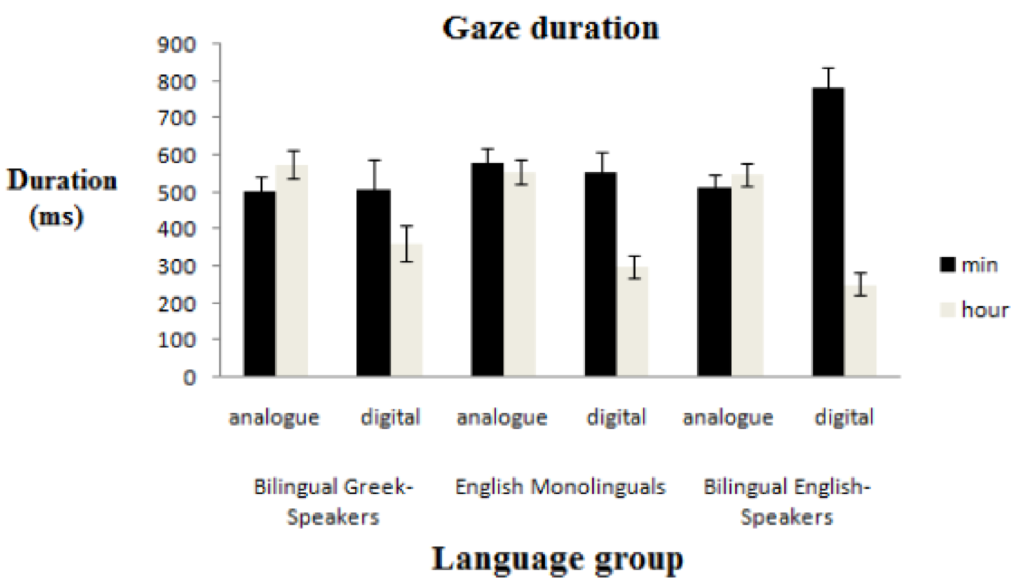

Figure 3a. Duration of gazes in min and hour regions of analogue and digital clocks for Bilingual Greek-Speakers, English Monolinguals, and Bilingual English-Speakers. Error bars represent the Standard Error of the Mean.

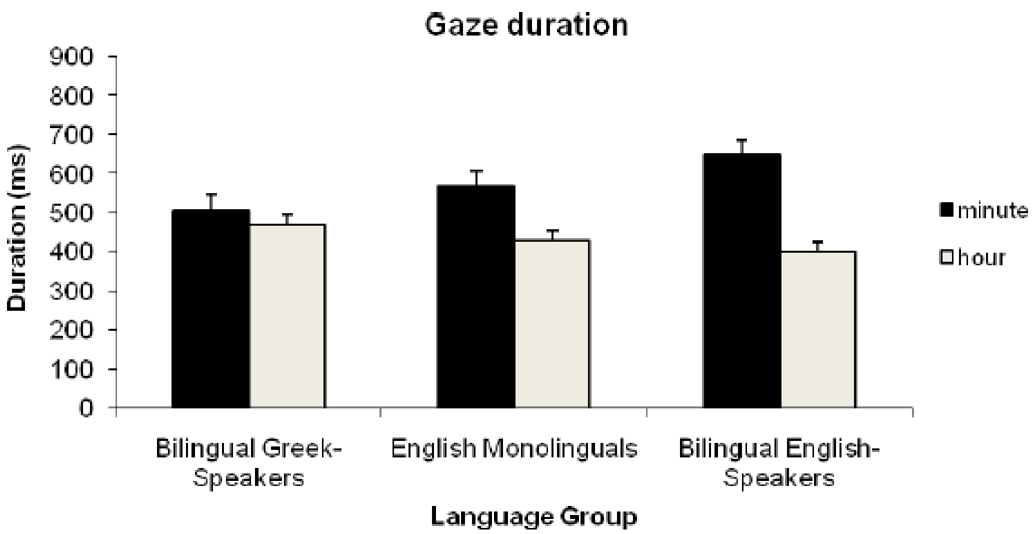

Figure 3b. Duration of gazes in minute and hour regions collapsed across participants. Error bars represent the Standard Error of the Mean.

Discussion

We examined the relationship between language and visual cognition by investigating whether Greek-English bilinguals' eye movements reflected linguistic properties of the target language (language being produced).

Our results replicate and extend previous findings from English and Dutch monolinguals (Bock et al., 2003; Goolkasian & Park, 1980; Korvost et al., 2007). We found increased latencies for analogue than digital clocks, and for minute time-telling relating to the second half-hour compared to the first half-hour, in Greek-English bilinguals as well as in English monolinguals. These results suggest that conceptual operations that are involved in time-telling extend to other languages and other populations, namely bilinguals.

More critically, we found different eye-movement patterns in Greek-English bilinguals telling the time in Greek versus Greek-English bilinguals and English monolinguals telling the time in English, showing that the linguistic characteristics of time-telling in specific languages can affect visual processing. Participants producing Greek tended to fixate the hour region first, whereas participants producing English, whether bilingual English-speakers or English monolinguals, tended to fixate the minute region first, mirroring the order of mention in each language of production. Gaze duration in the minute and hour regions was also similar for the two English-speaking groups, which differed from the Greek-speaking group.

The fact that speakers drawn from the same bilingual population showed different patterns of eye movements depending on the language of production (L1-Greek or L2-English), and no evidence of interference from the contrasting word order of L1 during L2 production, suggests that the influence of language was bound to a particular context of production or to particular utterances, rather than operating as a purely long-term influence on an individual's behaviour (see Papafragou, Hulbert, & Trueswell, 2008). These results are consistent with other evidence showing a close link between the structure of the language being produced (or indeed comprehended) in a particular context, and participants' linguistic and visual behaviour, both in bilingual processing (Kousta, Vinson, & Vigliocco, 2008) and in monolingual processing (Huettig, Chen, Bowerman, & Majid, 2010; Kuchinsky, Bock, & Irwin, 2011).

In sum, we have shown that in bilingual production as in monolingual production, there is a close relationship between visual cognition and linguistic characteristics of the language being produced.

Acknowledgements

The second author is funded by the Coleg Cymraeg Cenedlaethol. The third author is supported by a research grant from the Basque Government (IT665-13).

To cite this article: A. Hatzidaki, M.W. Jones, M. Santesteban & H.P. Branigan , Language and Cognitive Processes (2013): It's not what you see: it's the language you say it in, Language and Cognitive Processes, DOI: 10.1080/01690965.2013.857782

References

Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J-R., & Thierry, G.(2009). The Whorfian mind. *Communicative & Integrative Biology*, *2*, 332–334.

Bock, K., Irwin, D.E., Davidson, D.J., & Levelt, W.J.M. (2003). Minding the clock. *Journal of Memory and Language*, *48*, 653–685.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the *give* and *take* between event apprehension and utterance formulation. *Journal of Memory and Language*, *57*, 544–569.

Goolkasian, P., & Park, D. C. (1980).Processing of visually presented clock times.*Journal of Experimental Psychology: Human Perception and Performance*, *6*, 707–717.

Griffin, Z. M. (2004). Why Look? Reasons for Eye Movements Related to Language Production. In J. Henderson & F. Ferreira, (Eds.), *The Interface of Language, Vision, and Action* (pp. 213–247).New York: Taylor and Francis.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274–279.

Hatzidaki, A., Branigan, H. P., & Pickering, J. M. (2011). Co-activation of syntax in bilingual language production. *Cognitive Psychology*, *62*, 123-150.

Huettig, F., Chen, J., Bowerman, M., & Majid, A. (2010). Do language-specific categories shape conceptual processing? Mandarin classifier distinctions influence eye gaze behavior, but only during linguistic processing. *Journal of Cognition and Culture*, *10*, 39−58.

Korvorst, M., Roelofs, A., & Levelt, W.J.M. (2007). Telling time from analogue and digital clocks: A multiple-route account.*Experimental Psychology*, *54*, 187–191.

Kousta, S. T.,Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 843–858.

Kuchinsky, S. E., Bock, K., & Irwin, D. E. (2011). Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 748–756.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.

Meeuwissen, M.H.W., Roelofs, A.P.A., & Levelt, W.J.M. (2004). Naming analogue clocks conceptually facilitates naming digital clocks. *Brain and Language*, *90*, 434–440.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*, 155−184.

Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb–participle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 904–921.

Slobin, D. I. (1996). From ‘thought and language’ to ‘thinking for speaking’. In J. J. Gumperz and S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 195–217). Cambridge: Cambridge University Press.

Tanenhaus, M. K. (2007). Eye movements and spoken language processing. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 443–469). Oxford: Elsevier.