Cognition 116 (2010) 437–443

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception

Panos Athanasopoulos [a,*], Benjamin Dering [b], Alison Wiggett [b,c], Jan-Rouke Kuipers [d], Guillaume Thierry [b,c,d]

[a] *School of Linguistics and English Language, Bangor University, UK*
[b] *School of Psychology, Bangor University, UK*
[c] *Wales Institute of Cognitive Neuroscience, Bangor University, UK*
[d] *ESRC Centre for Research on Bilingualism in Theory and Practice, UK*

ARTICLE INFO

Article history:
Received 31 October 2009
Revised 21 May 2010
Accepted 31 May 2010

Keywords:
Colour perception
Cultural relativity
Electrophysiology
Visual Mismatch Negativity
Bilingualism

ABSTRACT

The validity of the linguistic relativity principle continues to stimulate vigorous debate and research. The debate has recently shifted from the behavioural investigation arena to a more biologically grounded field, in which tangible physiological evidence for language effects on perception can be obtained. Using brain potentials in a colour oddball detection task with Greek and English speakers, a recent study suggests that language effects may exist at early stages of perceptual integration [Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. (2009). Unconscious effects of language-specific terminology on pre-attentive colour perception. *Proceedings of the National Academy of Sciences, 106*, 4567–4570]. In this paper, we test whether in Greek speakers exposure to a new cultural environment (UK) with contrasting colour terminology from their native language affects early perceptual processing as indexed by an electrophysiological correlate of visual detection of colour luminance. We also report semantic mapping of native colour terms and colour similarity judgements. Results reveal convergence of linguistic descriptions, cognitive processing, and early perception of colour in bilinguals. This result demonstrates for the first time substantial plasticity in early, pre-attentive colour perception and has important implications for the mechanisms that are involved in perceptual changes during the processes of language learning and acculturation.

1. Introduction

Recent studies of the relationship between language and thought have provided substantial behavioural evidence in support of the view that cognitive representation of the world is essentially language-based (for reviews see Casasanto, 2008; Gentner & Goldin-Meadow, 2003; Gumperz & Levinson, 1996; Regier & Kay, 2009; Roberson, 2005; Roberson & Hanley, 2007). More recently, studies have focused on the representation of perceptual categories in bilinguals who speak languages with contrasting lexical and grammatical properties. This line of investigation has been spurred on partly by the fact that multilingualism is becoming increasingly prevalent. Indeed, the majority of the world's population uses more than one language (Cook, 1999, 2003). Moreover, bilingualism presents an ideal opportunity to test a fundamental question of human cognition: is cognitive representation fixed and static, or is it dynamic and flexible, such that exposure to different languages and cultures alters pre-existing representations (Abutalebi & Green, 2007; Green, 1998; Mechelli, Crinion, Noppeney, et al., 2004; Pavlenko, 1999)?

Empirical investigations have shown that additional language learning may indeed alter pre-existing cognitive representations and that bilingual cognition is affected by

---

[*] Corresponding author. Tel.: +44 1248 38 86 68.
E-mail address: panos@bangor.ac.uk (P. Athanasopoulos).

0010-0277/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2010.05.016

the length of cultural immersion in the second language (L2) speaking country (Athanasopoulos, 2009; Cook, Bassetti, Kasai, Sasaki, & Takahashi, 2006), language proficiency (Athanasopoulos, 2007; Bassetti, 2007; Boroditsky, Schmidt, & Phillips, 2003), the language used for task instructions (Boroditsky, Ham, & Ramscar, 2002; Kousta, Vinson, & Vigliocco, 2008), and the age of L2 acquisition (Boroditsky, 2001; Boroditsky et al., 2003).

More recently, under the impulse of neuroscientific methods it has been possible to obtain precise temporal information about the effects of language during online visual processing. Using event related potentials (ERPs), Fontenau and Davidoff (2007) and Holmes, Franklin, Clifford, and Davies (2009) show that colour categorisation occurs between 100–200 ms post-stimulus. In a recent study by Thierry, Athanasopoulos, Wiggett, Dering, and Kuipers (2009), we reported for the first time differences between speakers of different languages in early colour perception. Greek and English speakers were instructed to attend to a target square stimulus (probability 20%) presented within a stream of circle stimuli (probability 80%), some of which varied in luminance (standard circles: probability 70%, deviant circles: probability 10%) while electrical scalp activity was recorded. In two blocks standards and deviants were either light or dark blue, and in two other blocks they were light or dark green. The main analysis focused on an ERP component called Visual Mismatch Negativity (vMMN), a marker considered automatic and unconscious, since it requires no attention directed at the stimulus, and occurs between 170 and 220 ms post-stimulus (Czigler, Balazs, & Winkler, 2002; Winkler et al., 2005). Results showed a vMMN effect of similar magnitude for blue and green standard-deviant contrasts in native speakers of English who do not make a lexical distinction between light and dark blue. Greek participants, on the other hand, displayed greater vMMN modulations for the blue rather than the green standard-deviant contrast, consistent with their language's partition of the blue area of the colour spectrum into a lighter shade (ghalazio) and a darker shade (ble).

These results suggest that language differences may occur very early in the visual processing stream and are likely to be partially perceptual (since the vMMN is considered to be an index of pre-attentive perception originating at 162–232 ms post-stimulus in the Thierry et al., 2009 study) as well as post-perceptual (since Costa, Strijkers, Martin, & Thierry, 2009, have recently shown that lexical access occurs from 200 ms post-stimulus during speech production). However, it is unknown whether all Greek participants in Thierry et al.'s study (who were all students in the UK) pre-attentively distinguished more between blues than greens, or whether there were within-group differences as a function of duration of immersion in the L2-speaking country. Research shows that training participants to learn new artificial categories leads to changes in categorical perception and object recognition at least in the short term (Goldstone & Barsalou, 1998; Notman, Sowden, & Özgen, 2005; Özgen & Davies, 2002). The current paper takes a first step towards investigating the long-term effects of using a second set of categorical divisions (as bilinguals do) on early perceptual processing. Here, we compare the behavioural evidence elicited through a similarity judgement task, the linguistic evidence obtained from a colour naming task, and we re-analyse the electrophysiological data from a colour oddball task originally reported in our previous paper (Thierry et al., 2009) to determine whether cognitive processing, linguistic descriptions, and pre-attentive perception of colour converge.

2. Methods

2.1. Participants

The 20 native Greek speakers that took part in Thierry et al.'s (2009) study were evenly split into two groups based on the amount of time they had lived in the UK, following Athanasopoulos (2009). Both groups were comparable on English proficiency as measured by the Nation (1990) vocabulary test, age of L2 acquisition, years of formal English instruction, and frequency of use of English, measured as hours per week. Table 1 provides a summary of these details[1]. Independent samples *t*-tests confirmed that the groups differed significantly in months of stay in the UK, but not in any of the other measures [$t(18) = -4.63$, $p < .05$ for months in UK, $t(18) = -1.16$, $p > .05$ for English proficiency, $t(18) = 1.70$, $p > .05$ for age of L2 acquisition, $t(18) = -1.71$, $p > .05$ for years of English instruction, and $t(18) = -.29$, $p > .05$ for hours per week using English]. For comparison purposes we also report electrophysiological data from a random selection of 10 of the 20 native English monolingual participants in the Thierry et al. (2009) study. All participants were between 20 and 23 years old, right-handed, had normal or corrected-to-normal vision and were matched for educational level.

2.2. Materials

2.2.1. Oddball task

Four hundred and thirty-two filled circle shapes and 108 filled square shapes subtending approximately 2° of visual angle were presented on a middle gray background on a calibrated CRT monitor. Chromaticity was measured using a Minolta CS-100 Colorimeter. The following Munsell colours were used (CIE 1931 *Y*, *x*, *y* chromaticity coordinates are given in parentheses): dark blue: 5PB/value 4 ($Y = 10.7$, $x = 0.234$, $y = 0.230$), light blue: 5PB/value 7 ($Y = 41.5$, $x = 0.259$, $y = 0.264$), dark green: 5G/value 4 ($Y = 10.7$, $x = 0.259$, $y = 0.397$), light green: 5G/value 7 ($Y = 41.7$, $x = 0.279$, $y = 0.377$). Munsell chroma (saturation) was held constant across stimuli (chroma 6).

2.2.2. Similarity judgements and naming

Seventeen computer-generated Munsell stimuli, falling within the 5B, 10B and 5PB hue ranges, were arranged into all possible pairs of constant hue where distance in luminance (Munsell value) between members in each pair

---

[1] The ranges reported here for age of L2 acquisition and for L2 proficiency are slightly different from the ones reported in the Thierry et al. (2009) paper. This is because of copy-editing oversights in the previous report and the information reported here is accurate.

Table 1
Summary of bilinguals' details.

	Short-stay bilinguals	Long-stay bilinguals
N	10	10
Mean length of stay in the UK (range)	7.2 months (5–12)	42.6 months (18–60)
Sex	4F/6M	6F/4M
Proficiency in English (range)	63/90 (39–84)	71/90 (55–90)
Mean L2 acquisition age (range)	10 years old (6–15)	8 years old (4–12)
Formal English instruction (range)	8 years (6–9)	9 years (5–12)
Frequency of L2 use per week (range)	47 h (20–112)	50 h (10–84)

differed in 2 or 3 steps. Saturation (Munsell chroma) was kept constant at 6. This yielded a total of 19 stimulus pairs. Table 2 shows the 17 Munsell stimuli used for similarity judgements and naming, along with their CIE *Yxy* coordinates.

Table 3 shows the Munsell pairs that were used in the similarity judgement task. Each stimulus subtended a visual angle of approximately 2°.

2.3. Procedure

First, participants performed the oddball shape detection task while their brain activity was measured. Then, after a short break in which the EEG cap was removed, the participants were given the similarity judgement task. Third, participants were tested on a colour naming task to determine their individual linguistic category boundaries. All testing took place in a quiet, darkened room. Below we give a detailed description of what each task entailed.

2.3.1. Oddball task

Participants viewed four blocks of 540 stimuli (two 'blue' and two 'green' blocks). Within each block, one stimulus was frequent (light or dark circle, 70%) and three stimuli were infrequent: luminance deviant (circle with a luminance opposed to that of the frequent stimulus, 10%), standard target (light square, 10%), and deviant target (dark square, 10%). Presentation order was pseudo-randomized, such that two deviants or targets never appeared in immediate succession, and there were at least three standards in a row between two infrequent stimuli. Stimuli were flashed for 200 ms with an interstimulus interval of 800 ms. Participants were instructed to detect squares by pressing the spacebar of a keyboard. Block order was fully counterbalanced between participants. The proportion of hits was high (mean = 95% ± 5). For an example of four experimental blocks please see Fig. 1 in Thierry et al. (2009).

Table 2
Stimuli used in the naming and similarity judgement tasks. The left column shows Munsell designations expressed as Hue/Value (lightness)/Chroma (saturation). The remaining columns show CIE *Yxy* coordinates.

Munsell	*Y*	*X*	*Y*
5PB/2/6	2.89	0.196	0.181
5PB/3/6	5.6	0.216	0.205
5PB/4/6	10.7	0.234	0.23
5PB/5/6	18.2	0.245	0.244
5PB/6/6	29.8	0.253	0.255
5PB/7/6	41.5	0.259	0.264
10B/3/6	5.7	0.196	0.217
10B/4/6	10.9	0.215	0.244
10B/5/6	18.8	0.226	0.253
10B/6/6	29.3	0.238	0.266
10B/7/6	41	0.246	0.272
10B/8/6	56.4	0.253	0.277
5B/4/6	10.2	0.209	0.257
5B/5/6	18.9	0.221	0.27
5B/6/6	29.2	0.232	0.278
5B/7/ 6	41.9	0.24	0.285
5B/8/6	57.4	0.242	0.288

Table 3
All possible 2- and 3-step pairs used in the similarity judgement task.

2-Step pairs	3-Step pairs
5PB/2/6–5PB/4/6	5PB/2/6–5PB/5/6
5PB/3/6–5PB/5/6	5PB/3/6–5PB/6/6
5PB/4/6–5PB/6/6	5PB/4/6–5PB/7/6
5PB/5/6–5PB/7/6	
10B/3/6–10B/5/6	10B/3/6–10B/6/6
10B/4/6–10B/6/6	10B/4/6–10B/7/6
10B/5/6–10B/7/6	10B/5/6–10B/8/6
10B/6/6–10B/8/6	
5B/4/6–5B/6/6	5B/4/6–5B/7/6
5B/5/6–5B/7/6	5B/5/6–5B/8/6
5B/6/6–5B/8/6	

2.3.2. ERP recording and analysis

Electrophysiological data were recorded in reference to Cz at a rate of 1 kHz from 64 Ag/AgCl electrodes placed according to the extended 10–20 convention. Impedances were kept below 7 kΩ. EEG activity was filtered online with a band pass between 0.01 Hz and 200 Hz, and refiltered offline with a 20-Hz low-pass zero phase shift digital filter (slope 48 db/Oct). Eye blinks were mathematically corrected, and epochs with activity exceeding ±75 μV at any cap electrode site were automatically discarded. There was a minimum of 120 valid epochs per condition in every subject. Epochs ranged from −100 to 1000 ms after the onset of the stimulus. Baseline correction was performed in reference to prestimulus activity and individual averages were digitally re-referenced to the global average reference. The vMMN analysis was conducted on individual ERPs elicited by passive standard and deviant circles irrespective of luminance (light and dark circles combined) to discard luminance effects. The vMMN was maximal over the parietooccipital scalp and studied at electrodes IZ, O1, O2, OZ, PO7, PO8, PO9, and PO10.

2.3.3. Similarity judgement task

Participants viewed each of the nineteen pairs of colour chips twice in a fully randomized order, counterbalancing the left/right position of each stimulus in the pair. For each pair of colours participants were instructed to decide by mean of a key press 'how similar or different these colours are' on a scale of 1–9, where 1 represented 'very similar'

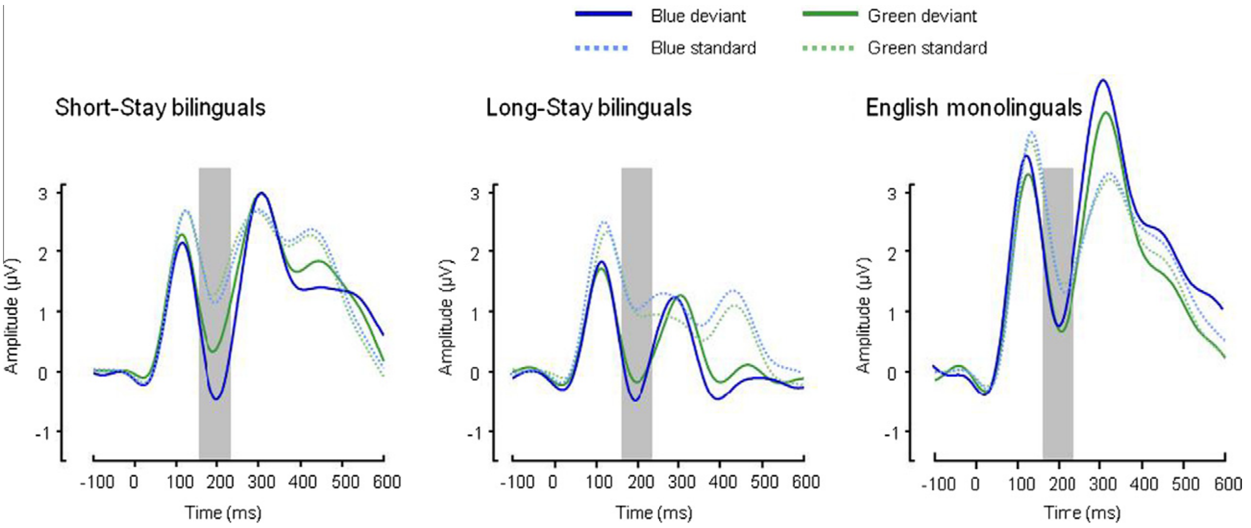

Fig. 1. ERPs elicited by standard circles (standards) and passive deviant circles (deviants) irrespective of luminance over parietooccipital electrodes where the vMMN was maximal (linear derivation of IZ, O1, O2, OZ, PO7, PO8, PO9, and PO10). Mean brain potential amplitude was significantly more negative for deviants than standards between 162 and 232 ms (shaded interval).

and 9 represented 'very different'. The rating scale appeared at the bottom of the screen, immediately below the colour stimuli. Instructions were given to participants in English.

2.3.4. Naming task

Participants were shown the 17 stimuli used in the similarity judgement task in random order and asked to classify each colour either as *ble* or *ghalazio*. Participants were instructed to make all judgments as quickly and accurately as possible. Their verbal response was recorded by the administrator. Instructions were given in Greek.

3. Results

3.1. Oddball task

Fig. 1 shows the interaction between colour (blue vs. green) and Deviancy (deviant vs. standard luminance) in vMMN patterns in the two Greek-English bilingual groups and the English monolingual group. A 3 (Group) × 2 (Colour) × 2 (Deviancy) mixed ANOVA revealed a significant main effect of Deviancy, $F(1, 27) = 49.02$, $p < .05$, $\eta_p^2 = .65$, but no significant main effects of Colour or Group, $F(1, 27) = 2.81$, and $F(2, 27) = .49$, respectively. Crucially, the triple interaction was significant, $F(2, 27) = 3.69$, $p < .05$, $\eta_p^2 = .22$. To probe the interaction further, analyses of variance with the factors Colour and Deviancy were performed for each group. For the English monolinguals, there was a significant main effect of Deviancy, $F(1, 9) = 9.32$, $p < .05$, $\eta_p^2 = .51$, but no significant main effect of Colour, $F(1, 9) < .00$, and no significant interaction, $F(1, 9) = .66$. For the Long-stay bilinguals the main effect of Deviancy was significant, $F(1, 9) = 22.45$, $p < .05$, $\eta_p^2 = .71$, but the main effect of Colour was not significant, $F(1, 9) = .04$. The interaction was also not significant, $F(1, 9) = 1.80$. For the Short-stay bilinguals, there was a significant main effect of Colour, $F(1, 9) = 6.88$, $p < .05$, $\eta_p^2 = .43$, a significant main effect of Deviancy, $F(1, 9) = 17.63$, $p < .05$, $\eta_p^2 = .66$, and a significant Colour × Deviancy interaction, $F(1, 9) = 15.82$, $p < .05$, $\eta_p^2 = .64$, such that the vMMN was significantly greater for blue than green deviants ($p < .05$). This means that Short-stay bilinguals more consistently distinguished blue deviants from blue standards than green deviants from green standards, whereas Long-stay bilinguals treated the difference between blue deviants and standards similarly to that between green deviants and standards, approximating the English monolinguals.

In order to control for possible effects of the other biographical characteristics of the bilinguals on the vMMN patterns reported above, an analysis of covariance (ANCOVA) was performed with the between-subjects factor Group, the within-subject factors Colour and Deviancy, and L2 proficiency, years of formal English instruction and frequency of L2 use as covariates. The triple interaction was statistically significant, $F(2, 24) = 5.06$, $p < .05$, $\eta_p^2 = .30$, even after controlling for these variables. It was not possible to add age of acquisition as a covariate in the above analysis because when the assumption of homogeneity of regression lines was examined for this variable it was found to be untenable, $F(1, 16) = 6.65$, $p < .05$. To overcome this problem, an ANCOVA was run as a Linear Mixed Model using the difference of the vMMN patterns for blue and green standard-deviant contrasts as the dependent variable. This showed that when controlling for age of acquisition, the effect of length of stay remained significant, $F(1, 25) = 3.79$, $p < 0.05$, while there was no significant effect of age of acquisition, $F(1, 25) = .01$.

3.2. Similarity judgement task

A range of Munsell stimuli, arranged in pairs, were used for similarity judgements (see Table 3). However scores were analysed for only a subset of those for each individ-

Table 4
Mean similarity judgements (and standard deviations) of *ble* and *ghalazio* within and cross-category pairs. Figures are rounded to the nearest 2 decimal places.

Groups	Within-category pairs	Cross-category pairs
Short-stay bilinguals	4.43 (1.41)	5.27 (1.37)
Long-stay bilinguals	4.79 (1.08)	4.75 (1.41)

ual, depending on each individual's category boundary between *ble* and *ghalazio* in the naming task, and on the basis that the stimuli cross-cutting the ble–ghalazio boundary were equally distant in lightness from each other and from stimuli within each category. Table 4 shows mean similarity judgement scores for within- and cross-category stimulus pairs in the three groups. A Group × Pairtype mixed ANOVA showed a significant main effect of Pairtype, $F(1, 18) = 4.91$, $p < .05$, $\eta_p^2 = .21$, no significant main effect of Group, $F(1, 18) = .02$, and a significant interaction, $F(1, 18) = 5.80$, $p < .05$, $\eta_p^2 = .24$. Post-hoc *t*-tests revealed that the Short-stay bilinguals distinguished more between cross-category than within-category stimuli in their similarity judgements ($p < .05$) while Long-stay bilinguals did not ($p > .05$)[2].

3.3. Naming task

Fig. 2 shows areas for ble and ghalazio for short-stay and Long-stay bilinguals. Areas were calculated by counting the number of times each stimulus was named *ble* or *ghalazio* and tallying up the number of responses such that a chip is completely shaded ble or ghalazio on the basis of at least 70% of participants within each group naming it so. Otherwise boundaries that cut across the chip indicate that there was some ambiguity within the group indexed by a 60/40% or 50/50% split in naming. As can be seen from the graph, there was less within-group naming agreement at the ble/ghalazio boundary in Long-stay bilinguals than in Short-stay bilinguals.

4. Discussion

The current paper evenly split the Greek participants that took part in Thierry et al.'s (2009) study into two groups based on length of residence in the L2-speaking country in order to investigate to what extent early, pre-attentive components of perception are susceptible to restructuring. We compared behavioural evidence elicited through a similarity judgement task, linguistic evidence obtained through a colour naming task, and electrophysiological evidence from a colour oddball task. Results from each of the tasks showed that cognitive processing, linguistic descriptions, and unconscious perception of colour converge in bilinguals. Specifically, the Short-stay bilinguals showed larger vMMN for blue than for green luminance contrasts, while the Long-stay bilinguals did not, resembling English monolinguals. In addition, Long-stay bilinguals distinguished significantly less between within and cross-category stimuli in their similarity judgements, and displayed significantly less within-group naming agreement at the ble/ghalazio boundary than the Short-stay bilinguals.

These robust differences between bilinguals as a function of length of cultural immersion are in line with psycholinguistic evidence that L2 learners who are also immersed in the L2-speaking country form stronger word-to-referent mappings in their L2 than learners who have only received classroom instruction (Tokowicz, Michael, & Kroll, 2004), and can be attributed to the fact that living in the L2-speaking country provides a physical context in which linguistic and perceptual experience are grounded. The finding that the bilinguals in our study are losing the ble/ghalazio distinction after an average of 3.5 years of immersion in the UK is compatible with psycholinguistic studies on language attrition. For example, Frenck-Mestre (1993) found that in a lexical decision task, answers in the L2 became faster than answers in the L1 after as little as 3 years of immersion. A recent study (Linck, Kroll, & Sunderman, 2009) shows that access to the L1 can be adversely affected after only 3 months immersion in a foreign country. Thus it is possible that even Short-stay bilinguals in our study have been affected by exposure to the L2-speaking environment, however without a control group of Greek monolinguals (or Greeks who have never lived in an English-speaking country before), we cannot at present know whether this is the case or not.

Our results nevertheless suggest that colour categories in adults are quite easily malleable by current experience, and thus more likely to be linguistic constructs rather than hard-wired perceptual constructs. This conclusion is further substantiated by developmental evidence. Because the physical properties of categories like colour or types of objects and the physiology of vision are arguably the same in all language communities, we should not find any cross-linguistic differences in cognition in pre-linguistic infants, and this is indeed what research shows in the domains of colour (Franklin et al., 2008; Goldstein, Davidoff, & Roberson, 2009; Roberson, Davidoff, Davies, & Shapiro, 2004), object types (Imai & Gentner, 1997), space (Hespos & Spelke, 2004), and gender (Martinez & Shatz, 1996). However, all these studies also show that once a child has acquired language, a certain amount of restructuring of previously formed categories occurs, indicating that language is a powerful tool for categorising events and objects in our environment (Levinson, 2001).

Casasanto (2008) proposes a mechanism for language effects on cognition that posits simple associative learning at its core. Language users habitually activate corresponding conceptual categories when using lexical terms or grammatical constructions. The process of language learning then entails strengthening associations between words

---

[2] This analysis follows several recent studies that calculate within and cross-category differences based on individual linguistic boundaries (e.g. Roberson, Hanley, & Pak, 2009; Winawer et al., 2007). Taking the mean boundary for all participants, there was still a significant difference between within and cross-category stimuli for short-stay bilinguals, $t(9) = -3.43$, $p < .01$, that was absent for long-stay bilinguals, $t(9) = .13$, $p = .90$.

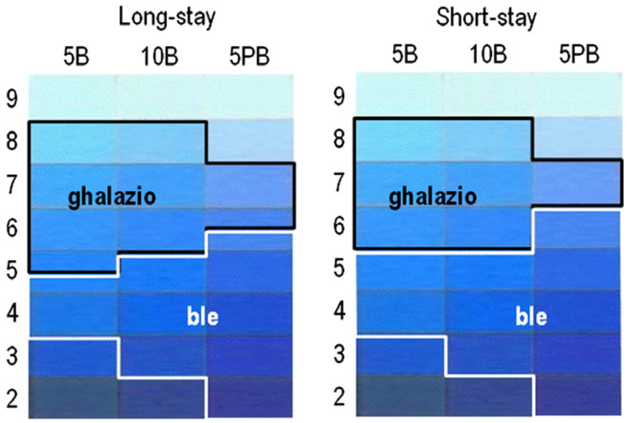

Fig. 2. The three-dimensional Munsell system is shown as a two-dimensional projection of hue (horizontal axis) against lightness (vertical axis). Lines cutting across a colour square indicate within-group naming disagreement (40–60% or 50–50%). Otherwise squares included in the ble or ghalazio boundaries were named consistently at least 70% of the time.

and their corresponding referents in the real world. Since different languages partition the world in different ways, speakers of different languages will necessarily have a different mental map of the world. Crucially, according to Casasanto (2008) the strength of the mappings may be adjusted due to immediate physical and linguistic experience. The current study supports this view by showing that the processes of second language learning and cultural immersion play precisely this role. Through forming and using new associations that contrast with existing ones, previously formed associations are to some extent restructured.

The evidence obtained from brain potentials strongly suggests that neural restructuring of core components of perception (such as those indexed by vMMN modulations) may occur in bilinguals as a function of length of stay in the L2-speaking country. Such an effect has not been found before and may have important implications for the developmental aspect of the links between language, cognition and perception in the human brain. As such, the current study provides a starting point for further studies exploring neural restructuring of early, pre-attentive components of perception. By recruiting larger numbers of bilinguals such studies will have the potential to investigate the precise interaction of a number of variables characterising the bilingual person (e.g. through multiple regression and correlation analyses) and elucidate further the extent to which each may contribute to neural restructuring in the bilingual brain.

Acknowledgments

P.A. is supported by two research development grants from the ESRC Centre for Research on Bilingualism in Theory and Practice. G.T. and J.K. are supported by the Economic and Social Research Council UK (RES-E024556-1); G.T. is supported by the European Research Council (ERC-209704). We thank three anonymous reviewers for their very constructive and insightful criticism of the first draft of our paper, and particularly Reviewer 1 for meticulous feedback regarding data analysis and presentation.

References

Abutalebi, J., & Green, D. W. (2007). The neurocognition of bilingual speech production. *Journal of Neurolinguistics, 20*, 242–275.

Athanasopoulos, P. (2007). Interaction between grammatical categories and cognition in bilinguals: The role of proficiency, cultural immersion, and language of instruction. *Language and Cognitive Processes, 22*, 689–699.

Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition, 12*, 83–95.

Bassetti, B. (2007). Grammatical gender and concepts of objects in bilingual children. *International Journal of Bilingualism, 11*, 251–273.

Boroditsky, L. (2001). Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology, 43*, 1–22.

Boroditsky, L., Ham, W., & Ramscar, M. (2002). What is universal about event perception? Comparing English and Indonesian speakers. In *Proceedings of the 24th annual meeting of the cognitive science society*. Fairfax, VA.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 61–79). Cambridge, MA: MIT Press.

Casasanto, D. (2008). Who's afraid of the Big Bad Whorf? Cross-linguistic differences in temporal language and thought. *Language Learning, 58*, 63–79.

Cook, V. J. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly, 33*, 185–209.

Cook, V. J. (2003). The changing L1 in the L2 user's mind. In V. J. Cook (Ed.), *Effects of the second language on the first* (pp. 1–18). Clevedon: Multilingual Matters.

Cook, V. J., Bassetti, B., Kasai, C., Sasaki, M., & Takahashi, J. A. (2006). Do bilinguals have different concepts? The case of shape and material in Japanese L2 users of English. *International Journal of Bilingualism, 10*, 137–152.

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. In *Proceedings of the National Academy of Sciences* (Early Edition).

Czigler, I., Balazs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology, 39*, 869–873.

Fontenau, E., & Davidoff, J. (2007). Neural correlates of colour categories. *NeuroReport, 18*, 1323–1327.

Franklin, A., Drivonikou, G. V., Bevis, L., Davies, I. R. L., Kay, P., & Regier, T. (2008). Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proceedings of the National Academy of Sciences, 105*, 3221–3225.

Frenck-Mestre, C. (1993). Use of orthographic redundancies and word identification speed in bilinguals. *Journal of Psycholinguistic Research, 22*, 397–410.

Gentner, D., & Goldin-Meadow, S. (Eds.). (2003). *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.

Goldstein, J., Davidoff, J., & Roberson, D. (2009). Knowing color terms enhances recognition: Further evidence from English and Himba. *Journal of Experimental Child Psychology, 102*, 219–238.

Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting conception and perception. *Cognition, 65*, 231–262.

Gumperz, J. J., & Levinson, S. C. (Eds.). (1996). *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.

Green, D. W. (1998). Bilingualism and thought. *Psychologica Belgica, 38*, 253–278.

Hespos, S., & Spelke, E. (2004). Conceptual pre-cursors to language. *Nature, 430*, 453–456.

Holmes, A., Franklin, A., Clifford, A., & Davies, I. (2009). Neurophysiological evidence for categorical perception of color. *Brain and Cognition, 69*, 426–434.

Imai, M., & Gentner, D. (1997). A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62*, 169–200.

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.

Levinson, S. C. (2001). Covariation between spatial language and cognition, and its implications for language learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 566–588). Cambridge: Cambridge University Press.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language. *Psychological Science, 20*, 1507–1515.

Martinez, I., & Shatz, M. (1996). Linguistic influences on categorization in preschool children: A crosslinguistic study. *Journal of Child Language, 23*, 529–545.

Mechelli, A., Crinion, J. T., Noppeney, U., et al. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature, 431*, 757.

Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House/Harper Row.

Notman, L., Sowden, P. T., & Özgen, E. (2005). The nature of learned categorical perception effects: A psychophysical approach. *Cognition, 95*, B1–B14.

Özgen, E., & Davies, I. R. L. (2002). Acquisition of categorical color perception: A perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General, 131*, 477–493.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition, 2*, 209–230.

Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences, 13*, 439–446.

Roberson, D. (2005). Color categories are culturally diverse in cognition as well as in language. *Cross-Cultural Research, 39*, 56–71.

Roberson, D., Davidoff, J., Davies, I., & Shapiro, L. (2004). The development of color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General, 133*, 554–571.

Roberson, D., & Hanley, J. R. (2007). Color categories vary with language after all. *Current Biology, 17*, 605–606.

Roberson, D., Hanley, J. R., & Pak, H. (2009). Thresholds for colour discrimination in English and Korean speakers. *Cognition, 112*, 482–487.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. (2009). Unconscious effects of language-specific terminology on pre-attentive colour perception. *Proceedings of the National Academy of Sciences, 106*, 4567–4570.

Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition, 7*, 255–272.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on colour discrimination. *Proceedings of the National Academy of Sciences, 104*, 7780–7785.

Winkler, I. et al. (2005). Preattentive binding of auditory and visual stimulus features. *Journal of Cognitive Neuroscience, 17*, 320–339.