Languages flex cultural thinking*

CERI ELLIS
Division of Neuroscience & Experimental Psychology, University of Manchester
GUILLAUME THIERRY
School of Psychology, Bangor University
AWEL VAUGHAN-EVANS
School of Psychology, Bangor University
MANON WYN JONES
School of Psychology, Bangor University

Recent studies have revealed remarkable interactions between language and emotion. Here, we show that such interactions influence judgments made regarding cultural information. Balanced Welsh–English bilinguals categorized statements about their native Welsh culture as true or false. Whilst participants categorized positive statements as true when they were true, they were biased towards categorizing them as true also when they were false, irrespective of the language in which they read them. Surprisingly, participants were unbiased when categorizing negative statements presented in their native language Welsh, but showed a reverse bias - categorizing sentences as false, even when they were true - for negative statements when they read them in English. The locus of this behavior originated from online semantic evaluation of the statements, shown in corresponding modulations of the N400 peak of event-related brain potentials. These findings suggest that bilinguals perceive and react to cultural information in a language-dependent fashion.

Keywords: bilingualism, culture, emotion, semantics, ERP

Introduction

Recent evidence has shown that language affects basic aspects of human cognition (Boroditsky, Schmidt & Phillips, 2003; Boutonnet, Athanasopoulos & Thierry, 2012; Thierry, Athanasopoulos, Wigget, Dering & Kuipers, 2009). Recent data moreover show effects of language on cultural identity, modulating the processing of objectively verifiable information (Ellis, Kuipers, Thierry, Lovett, Turnbull & Jones, 2015) as well as subjective beliefs and cultural stereotypes (Briley, Morris & Simonson, 2005; Danziger & Ward, 2010; Ogunnaike, Dunham & Banaji, 2010). Language that refers to cultural membership is often emotionally laden. For example, the word "foreigner" in English is derived from the Latin "person outside", and by speaking the word, one aligns oneself, however temporarily, with a specific in-group (Ogunnaike et al., 2010).

How can the bilingual mind then accommodate different perspectives, which originate from the different languages spoken? Previous research suggests that emotions processing plays a key role in shaping conceptual knowledge, judgment and behavior via its interaction with language. For instance, unconscious access to the native language (L1) when bilinguals read in their second language (L2; Thierry & Wu, 2007) is repressed when the words are negative (Wu & Thierry, 2012). Furthermore, risk-taking behavior – associated with high gains or losses – in L1 is characterized by greater impulsivity and intuitive bias than in L2, which is in turn characterized by greater rationality and risk-aversion (Costa, Foucart, Amon, Aparici & Apesteguia, 2014; Gao, Zika, Rogers & Thierry, 2015; Keysar, Hayakawa & An, 2012). Thus, emotionally marked scenarios differentially interact with each language, resulting in quantifiably different behaviors.

However, current findings are unable to elucidate whether language-emotions interactions affect bilinguals' real-world semantic knowledge: that is, their perception and verification of true information. Here, we examined this question by manipulating the truth status of information pertaining to bilinguals' native culture, which provided an emotionally charged instance of semantic knowledge. Welsh–English bilinguals read objectively true and false statements that presented Wales, and Welsh culture, in either a positive or a negative light, written either in Welsh or English, and made truth-value judgments. Recent studies on social identity have shown

Address for correspondence:
Dr Ceri Ellis, Division of Neuroscience & Experimental Psychology, Faculty of Biology, Medicine & Health, Room 3.316, Jean McFarlane Building, University of Manchester, Oxford Rd, Manchester, M13 9PL
ceri.ellis@manchester.ac.uk

Table 1. *Experimental design and example of a statement 'set'*

Set a	Premise	Valence
Wales has the richest, most affluent community of ***farmers.***	False	Positive
A deeply Welsh and noble way of life is represented by our ***farmers.***	True	Positive
In Wales, supermarkets get the cheapest milk directly from ***farmers.***	False	Negative
Young Welsh people are discouraged from becoming ***farmers.***	True	Negative
Set b	Premise	Valence
Young Welsh men become very rich in their careers as ***farmers.***	False	Positive
The highest quality lamb meat in Britain is produced by our ***farmers.***	True	Positive
A shameful way of life is represented by our ***farmers.***	False	Negative
Wales has a problem with poverty in some communities of ***farmers.***	True	Negative

evidence for in-group favoritism, manifest in a greater tendency towards cooperative behavior with other group members (cf. Balliet, Wu & De Dreu, 2014), as well as stronger implicit in-group bias (Danziger & Ward, 2010). We therefore expected that our Welsh native participants would be generally biased – in both languages – towards assessing positive statements about Welsh culture as true, regardless of truth-value and we expected them to show the reverse bias for negative statements. Furthermore, we expected that these biases would be more pronounced in their L1 Welsh than L2 English, given evidence suggesting a stronger link between L1 and emotions processing (Altarriba, 2008; Dewaele, 2004; Pavlenko, 2008). We used event-related brain potentials (ERPs) to validate the locus of the effect at a semantic level based on modulations of the classical N400 peak (Kutas & Hillyard, 1980), and to assess potential links between behavioral observations and semantic integration.

Methods

Participants

Sixteen highly proficient Welsh–English bilinguals (14 females; $M_{age} = 22.56$, $SD = 7.17$) were included in the final analyses. Five participants were excluded due to poor data quality. All participants were right-handed, had normal or corrected vision, and reported no language impairments. All participants had been exposed to the Welsh language from birth, and had acquired English at an early age ($M = 5.13$, $SD = 3.16$). Participants' self-ratings of language proficiency (on a scale of $1 =$ not literate, to $10 =$ very literate) for reading, writing, speaking and comprehension were high for both Welsh (*Grand M* $= 9.16$, $SD = 1.48$), and English (*Grand M* $= 8.64$, $SD = 1.25$). Participants self-reported more daily use of Welsh ($M = 74.69\%$, $SD = 18.02$) compared with English ($M = 24.69\%$, $SD = 18.39$) in our Language History Questionnaire, which probes

participants' general language experience across several domains (e.g., reading, writing, socializing, watching television, etc.). The Multigroup Ethnic Identity Measure (MEIM; Phinney, 1992, Roberts, Phinney, Masse, Chen, Roberts & Romero, 1999) revealed a strong sense of Welsh cultural belonging ($M = 3.40$, $SD = 0.50$, $\alpha = .88$: $1 =$ indifferent response to $4 =$ strong cultural response). Participants provided informed consent and took part in the experiment in return for payment or course credit. Ethical approval was granted by the School of Psychology ethics committee at Bangor University.

Stimuli

Three hundred and twenty statements in English and their Welsh translations were constructed. Within each language, the statements were divided into 40 sets of 8, which ended in the same final word. Participants were presented with four statements from the English sets, and four statements from the Welsh sets that were not the translation of the English selection (Table 1). Thus, for any given participant, experimental sentences were never repeated, not even by way of translation. Therefore, the experimental design involved three factors: Language (English, Welsh), Emotional valence (positive, negative), and Truth-value (true, false). Valence and Truth-value were counterbalanced across languages.

Norming of stimuli

Twenty balanced Welsh–English bilinguals ($M_{age} = 27.15$, $SD = 12.87$; 100% reported L1 Welsh) participated in a separate pre-test to validate statements for valence (rated on a scale from $1 =$ positive, to $7 =$ negative) and plausibility ($1 =$ plausible, $0 =$ not plausible). Prior to the norming study, three native speakers of Welsh independently verified statements as true or false, such that only statements on which raters agreed were included in the study. In the norming study, cloze probability was obtained by asking participants to provide three
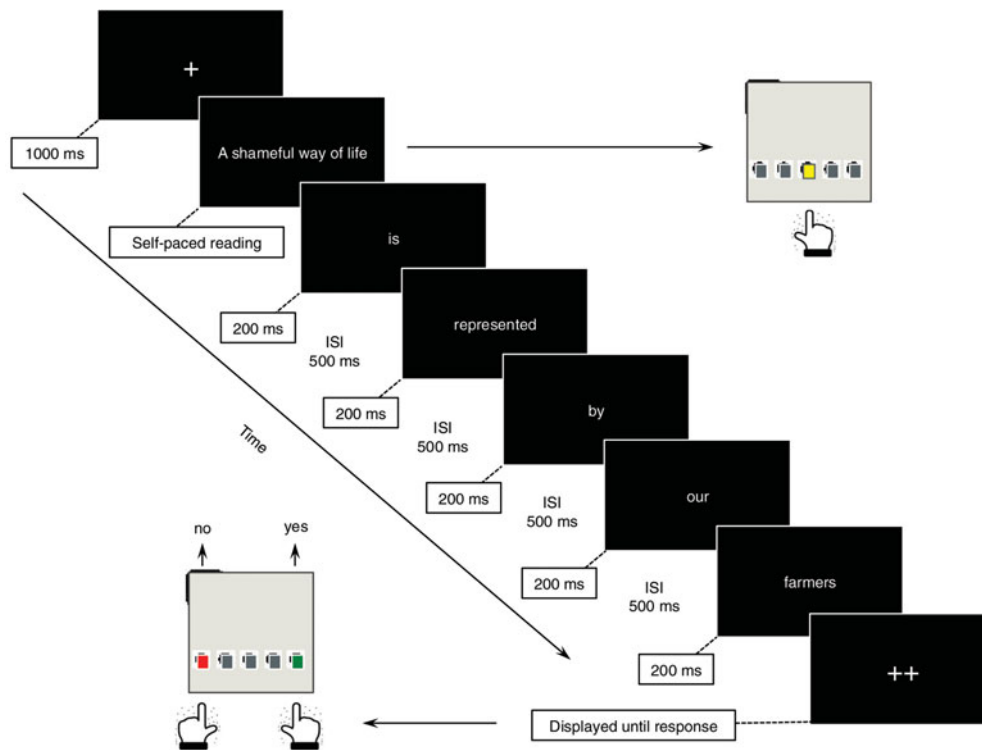
Figure 1. (Colour online) An example of how statements were presented

possible completions for each statement. If one of these matched our actual target word, a score of 1 was given. All other responses were scored 0. Mean values across all statements (52%) succeeded our threshold of 40% (Coulson, Urbach & Kutas, 2006), and did not differ between conditions (all *ps* > 0.05). Moreover, target words were controlled for frequency and word length in both Welsh and English (Welsh: Cronfa Electroneg o Gymraeg, Ellis, O'Dochartaigh, Hicks, Morgan & Laporte, 2001; English: CELEX lexical database, Baayen, Piepenbrock & van Rijn, 1993). Statements constructed a priori as 'positive' and 'negative' were validated (Positive Welsh: $M = 1.91$, $SD = 0.43$; Positive English: $M = 1.99$, $SD = 0.52$; Negative Welsh: $M = 6.28$, $SD = 0.17$; Negative English: $M = 6.19$, $SD = 0.41$), yielding a significant effect of valence ($F = 1684$, $p < 0.0001$, $\eta_p^2 = .99$) but no differences between languages ($p = 0.702$) or Truth-value ($p = 0.510$). The median plausibility of the statements was high (95%).

Procedure

Stimuli were presented at center screen position, in white, courier new, 18-point font on a black background of a 19-inch cathode ray tube (CRT) monitor with a refresh rate of 75 Hz, using E-prime 1.0 software. Reading of the first clause of each statement was self-paced, followed by single-word presentation of the final clause at a rate

of 200 ms per word and an inter-stimulus interval of 500 ms (Fig. 1). Following the presentation of the final word, participants were prompted ('++') to make a true/false judgment by providing a button press response. This method of prompt was used in order to minimize eye-movements, and to provide consistency across blocks when the testing language switched. Responses were self-paced such that the prompt remained on screen until participants responded. Participants were briefed verbally at the outset of the experiment to make the true/false judgment as quickly as possible following the prompt. Three practice trials preceded the experimental trials. The experiment was divided into two parts; four blocks of statements presented in Welsh, and four blocks presented in English, with a break from the experiment to complete the LHQ and MEIM questionnaires between segments. Blocks were randomized within-language, and language order was counterbalanced across participants. The presentation order of statements was pseudorandomized, such that participants would not encounter the same final word within a single block.

Behavioral Data Analysis

Accuracy and reaction times (RT) were modeled as a function of three within-participant factors: Language (Welsh, English), Valence (positive, negative), and Truth-value (true, false). For accuracy data, a binomial
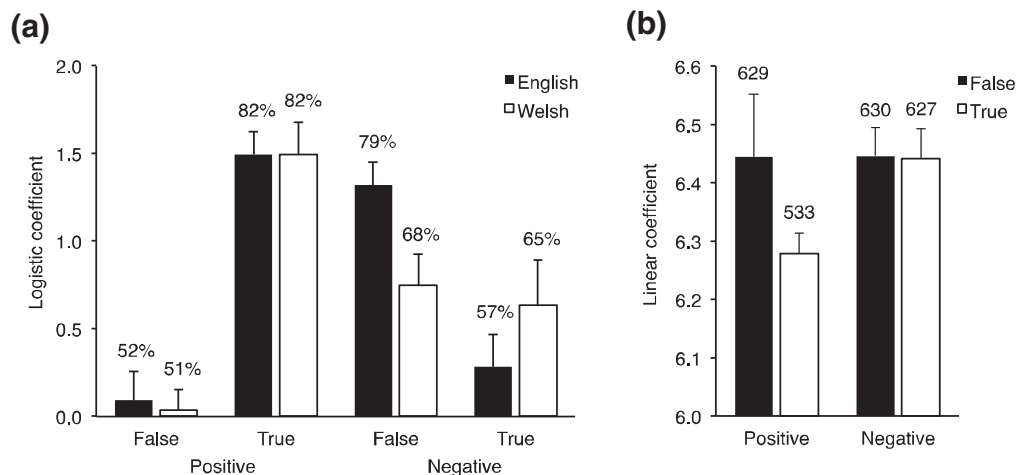
Figure 2. Behavioral results. **(a)** Accuracy scores (average % provided above each bar) for truth judgments as a function of Language and Truth-value, split by Valence. *Note*: Errors bars represent SEs. **(b)** Reaction Time (average provided above each bar) for truth judgments as a function of Truth-value and Valence. *Note*: Errors bars represent SEs.

logistic regression was implemented. Reaction time data were log transformed, and examined with linear mixed effects analyses. Analyses were run in the R software environment (R Development Core Team, 2008) using the lme4 library (Bates, Maechler & Dai, 2008; Baayen, 2008). β-values are reported, and tested at $p < 0.05$.

ERP Analysis

Electroencephalogram activity was continuously recorded from 64 Ag/AgCl electrodes according to the extended 10/20 convention, referenced to the Cz electrode at a rate of 1 kHz. Impedances for all electrodes were kept below 5 kΩ. The EEG was filtered online, with a band-pass filter between 0.1 and 200 Hz and re-filtered offline using a low-pass zero phase shift digital filter with a cut-off frequency of 20 Hz. Both EEG and behavioral data were collected simultaneously. Eye blink artefacts were corrected mathematically (based on an algorithm developed by Gratton, Coles & Donchin, 1983), and remaining artefacts were removed manually upon visual inspection of the data using Scan 4.4 software. Epochs ranged from −100 to 1000 ms after final word onset. Epochs with activity exceeding ± 75 μV at any electrode site over the scalp were discarded. Baseline correction was performed in reference to pre-stimulus activity and individual averages were digitally re-referenced to the common average reference. ERPs time-locked to the final word of each statement were visually inspected, and mean amplitudes were measured in temporal windows determined based on variations of the mean global field power measured across the scalp (Picton, Bentin, Berg, Donchin, Hillyard, Johnson, Miller, Ritter, Ruchkin, Rugg & Taylor, 2000). The N400 was maximal over central electrodes (C1, Cz, C2, CP1, CP2, CPz), in which it is

classically observed (Kutas & Hillyard, 1980; Hagoort, Hald, Bastiaansen & Petersson, 2004). Peak latency detection was locked to electrode Cz (site of maximal amplitude).

Results

Behavioral Results

For accuracy data (Fig. 2a), the full (Language*Truth*Valence) interaction model was found to provide the best fit for the data, compared with lower-order interaction models, $X^2 = 274.63$, df = 10, p < .0001 (Barr, Levy, Scheepers & Tily, 2013). Including a by-subject random slope for each of the Language, Truth and Valence factors led to non-convergence in the model, so we simplified the final model to include random intercepts for subjects and items. Collinearity was not an issue in this model: Fixed-effects correlations (|*r*|) were less than .7 for all predictors. The intercept represents the average likelihood that participants were accurate in the English/Positive/False condition. Each coefficient compares the average for a different combination of fixed factor levels against this intercept.

As expected, participants displayed a bias for positive statements, such that true statements were accurately categorized, whereas false statements were miscategorized as true ($b = 1.40$, $z = 10.59$, $p < .0001$). Accuracy was moreover identical in English and Welsh, for both false statements ($b = -0.05$, $z = -0.46$, $p = .641$) and true statements ($b = 0.05$, $z = 0.29$, $p = .770$). Also as expected, participants displayed a reverse bias in response to English negative statements, such that they were more likely to accurately categorize false statements

($b = 1.32$, $z = 9.55$, $p < .001$), whereas true statements were miscategorized as false ($b = -2.44$, $z = -13.15$, $p < .001$). Contrary to our hypotheses, however, negative statements read in Welsh did not elicit a similar bias: The significant three-way interaction showed that participants tended to be less accurate in rejecting false statements ($b = -0.51$, $z = -2.94$, $p = .003$) and more likely to accept true statements ($b = 0.86$, $z = 3.37$, $p < .001$) for negative statements read in Welsh, compared with other combinations of factors.

Participants' reaction time data (Fig. 2b) showed that a lower order interaction model (Language+Truth*Valence) contributed unique variance beyond the additive model ($X^2 = 12.71$, $df = 6$, $p = .022$). The model included by-subjects intercepts and slopes (1+Language+Truth*Valence|Participant), and the by-item intercept (1|Item). Fixed-effects correlations ($|r|$) were less than .7 for all predictors. The intercept represents the average estimated RT in the English/Positive/False condition.

For positive statements, participants were faster to respond to true compared with false information ($b = -0.16$, $t = -3.43$, $p = .001$), and response time was identical in English and Welsh ($b = 0.08$, $t = 0.89$, $p = .372$). Participants' RT to false information did not differ between negative and positive statements ($b = 0.00$, $t = 0.06$, $p = .948$), whereas responses to true information were significantly slower compared to false statements ($b = 0.16$, $t = 3.11$, $p = .002$). [1]

Electrophysiological Results

Repeated measures ANOVAs were implemented with Language (English vs Welsh), Valence (Positive vs Negative) and Truth (True vs False) as independent factors. In the first analysis, N400 mean amplitudes were analysed for both correct and incorrect trials (average of 38 trials per condition, SD = 2). Given the asymmetry in accuracy data for positive and negative statements (along with an interaction with language), participants' 'incorrect' responses plausibly comprised strategic responses in addition to simple errors. Our initial analysis therefore gave all responses the same consideration. We found a main effect of Truth ($F_{(1,15)} = 5.67$, $p = .030$), such that false statements elicited greater negativity relative to true statements (Fig. 3). No other effects emerged (Language, $F_{(1,15)} = 1.93$, $p = .185$; Valence, $F_{(1,15)} = 0.59$, $p = .454$; Language*Truth,

$F_{(1,15)} = 0.27$, $p = .610$; Language*Valence, $F_{(1,15)} = 2.29$, $p = .151$; Truth*Valence, $F_{(1,15)} = 0.02$, $p = .896$; Language*Truth*Valence, $F_{(1,15)} = 0.00$, $p = .996$).

We then analyzed N400 mean amplitudes for trials in which participants correctly discerned true and false statements ($M = 25$ trials per condition; $SD = 4.58$; Fig. 4). We found no main effect of Language ($F_{(1,15)} = 0.89$, $p = .360$), Valence ($F_{(1,15)} = 0.46$, $p = .506$), or Truth ($F_{(1,15)} = 1.25$, $p = .282$). However a Language*Truth interaction emerged ($F_{(1,15)} = 5.05$, $p = .040$). Post hoc analysis split by Language revealed a significant difference between true and false statements presented in English ($F_{(1,15)} = 7.87$, $p = .013$), but no differences emerged for statements presented in Welsh. No other significant effects were found (English: Valence ($F_{(1,15)} = 0.03$, $p = .865$), Valence*Truth interaction ($F_{(1,15)} = 1.04$, $p = .323$); Welsh: Truth ($F_{(1,15)} = 0.33$, $p = .575$), Valence ($F_{(1,15)} = 1.20$, $p = .291$), Valence*Truth ($F_{(1,15)} = 1.55$, $p = .232$)).

In one further post hoc analysis, we examined whether such effects were accompanied by the kind of late ERP modulations often found for emotional words (see e.g., Citron, 2012). Late Positive Potential (LPP) mean amplitudes (time range: 520–660 ms; electrodes: Cz, C1, C2, Pz, P1, P2) also yielded a significant Language*Truth interaction ($F_{(1,15)} = 9.75$, $p < .01$) in keeping with the modulations found in the N400 range (see Fig. 4). Crucially, we found no significant modulations by Valence.

Indices of "Pride" and "Defense"

In order to further understand the effects found, and relate behavioral and ERP data more directly, we made the a priori decision to define two descriptive indices: (a) a 'pride index' measuring the bias towards accepting positive information regardless of truth-value; and (b), a 'defense index', measuring the bias towards rejecting negative information (Fig. 5). We compared difference in N400 mean amplitude to the behavioral indices by calculating difference waves between false and true conditions for trials that elicited a correct response only.

The Pearson correlation between the difference in the magnitude of the defense index (negative sentences only) across languages and corresponding mean N400 amplitudes was significant ($r = .74$, $p = .001$, two-tailed). Unsurprisingly, no such correlation emerged for the pride index ($r = -.32$, $p = .222$, two-tailed), given that differences between conditions were very small.

Discussion

In this study, we examined whether perception of culturally relevant statements is modulated by language in early Welsh–English bilinguals. We found that positive

---

[1] We conducted a control analysis, in which we centered the IVs using dummy coding. In this model, which significantly reduced collinearity, our results were maintained: (Intercept: $b = .75$, Language: $b = -.03$, Valence: $b = -.016$, Premise: $b = .21$, Language*Valence: $b = -.02$, Language*Premise: $b = .12$, Valence*Premise: $b = -.49$, Language*Valence*Premise: $b = .11$).
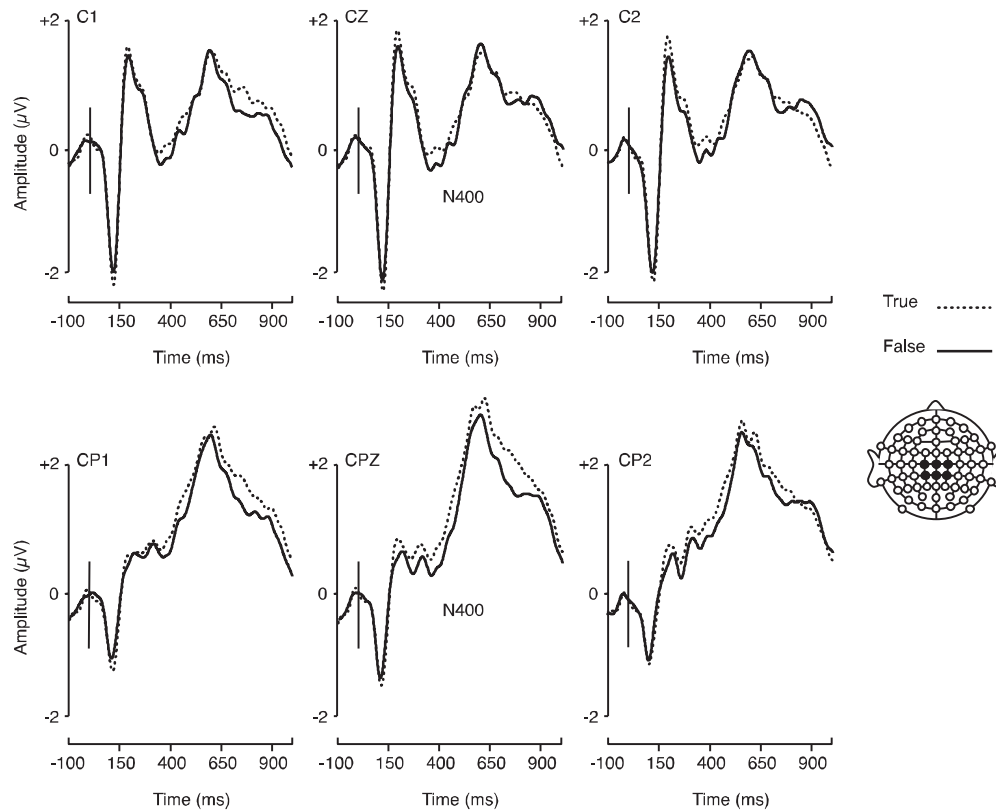
Figure 3. ERP responses to true vs. false statements, collapsed across Language and Valence. Waveforms depict averaged brain potentials at the six electrodes included in the analysis (C1, Cz, C2, CP1, CP2, CPz). The grey bar indicates the analysis time window (300–500 ms post-stimulus).
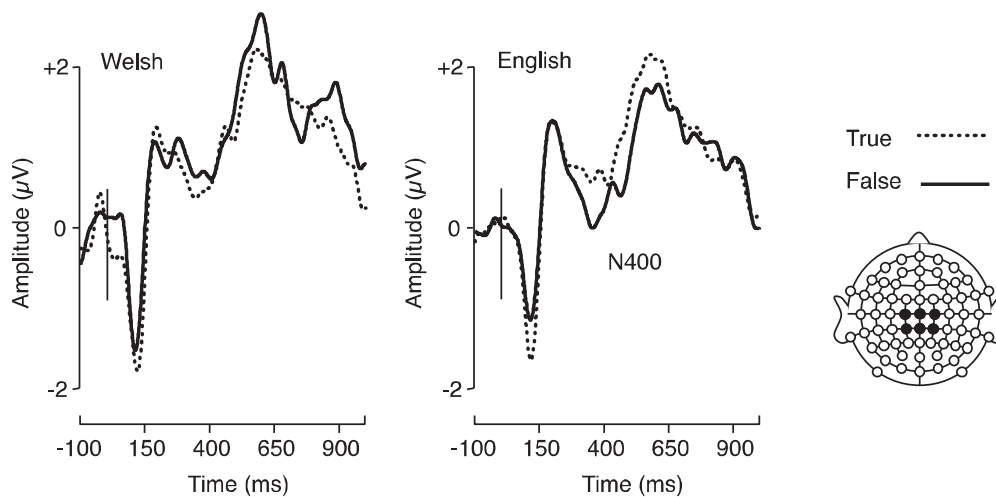

Figure 4. ERPs elicited by correct responses to true and false statements presented in the native (Welsh) and second (English) language. Waveforms depict averaged brain potential variations over the 6 electrodes where N400 amplitude was maximal (C1, Cz, C2, CP1, CP2, CPz).

statements were accurately categorized when they were true, and at chance level when they were false, an expected bias indicative of 'Welsh pride', which was not affected by the language in which the statements were presented. Conversely, participants displayed the expected reverse bias when dealing with negative statements (i.e., showing an increased tendency to categorize true statements as false), perhaps to minimize the impact of negative facts, but this bias was only observed in the second language English. Thus, whereas the second language appears
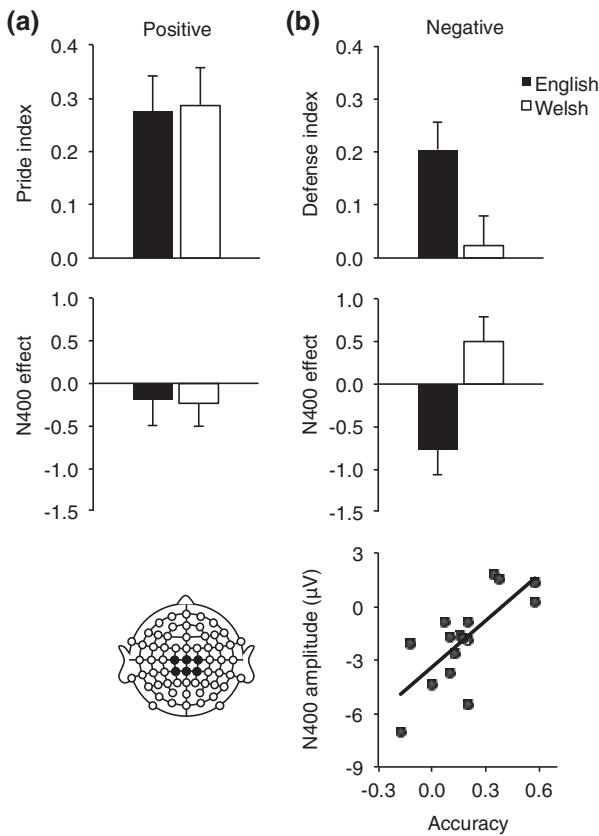
Figure 5. Relationship between "Pride" and "Defense" index on the one hand and N400 mean amplitude modulations by truth-value on the other. **(a)** Positive statements. Top, Difference in accuracy between true and false conditions, i.e., the "Pride index". Bottom, N400 mean amplitude difference between false and true conditions. **(b)** Negative statements. Top, Difference in accuracy between false and true conditions, i.e., the "Defense Index". Centre, N400 mean amplitude difference between false and true conditions. Bottom, Correlation between cross-language difference in defense index and mean N400 amplitudes.

to shield the bilingual from detrimental information regarding her culture, the native language does not.

Despite these behavioral differences, participants were overall able to distinguish true from false statements from a semantic integration point of view, as evidenced by the main effect of Truth-value on N400 mean amplitude (Hagoort et al., 2004; Ellis et al., 2015). The N400 is known to reflect the extent to which a target word fits within its preceding semantic context, such that greater mean amplitudes index a greater semantic integration effort (Kutas & Hillyard, 1980, 1984). Importantly, the weakness of the N400 modulation observed here is unsurprising given than no strong expectations could be formed by the reader as regards the sentence-final words (Martin, Thierry, Kuipers, Boutonnet, Foucart & Costa, 2013; Kutas & Federmeier, 2011). Indeed, cloze

probability was necessarily low ($M = 60\%$) in this experiment because the focus was on truth-value rather than semantic expectancy.

However, in the case of trials that yielded a correct response – that is, those trials in which participants perceived the contrast between true and false statements more clearly – the N400 was modulated by Truth-value only in English. This suggests that a second language context favors rational processing as compared to the native language, in which participants display more semantic uncertainty.

Previous studies have suggested that emotions are more strongly linked with L1 than L2 in bilinguals (Altarriba, 2008; Dewaele, 2004; Pavlenko, 2008). Emotion words are arguably comparatively better visualized and contextualized than neutral words in L1 (Altarriba & Bauer, 2004), and are also better recalled than in L2 (Aycicegi & Harris, 2004). Recent findings moreover show that such asymmetric language-emotion links affect cognition more generally in bilinguals. For instance, using event-related brain potentials, Wu and Thierry (2012) showed that Chinese–English bilinguals unconsciously access the native Chinese translations of positive and neutral words presented in English, but not that of negative words. The common modulations by affective valence often observed in the P600 range (for a review, see Citron, 2012), however, were absent in our data since amplitudes in the P600 range were only affected by the same language x truth value interaction affecting the N400 range. Therefore, the P600 differences observed can be construed as a carry-over effect of the differential amplitudes elicited in the N400 range. The absence of a modulation of P600 mean amplitudes by affective valence is not very surprising given that the critical words used in our study were not inherently emotional and that the affective manipulation concerned the statements as a whole rather than their final word.

Using similar paradigms in which the emotional manipulation concerned statements in their entirety rather than specific words in isolation, Keysar et al. (2012) and Costa et al. (2014) showed a reduction of the 'framing effect' in L2: Bilingual participants faced with making a decision (e.g., a forced-choice between two medical treatments) are more sensitive to the positive ('*you can save the lives of 200,000 people*') or negative ('*400,000 people will die*') framing of the situation when presented with the information in their L1. A more normative behavior in L2 suggests more rational evaluation of situations, owing to weaker links between L2 and emotion (see also Jończyk, Boutonnet, Musiał, Hoemann & Thierry, 2016).

Our data shows that such language-emotions asymmetry affects even the perception and verification of real-world semantic knowledge. Bilinguals processing negative cultural information in their second language

retrieve meaning more objectively, and thus are more likely to deny undermining comments regarding their culture. In contrast, when participants are faced with such information in their native language, negative statements confuse the semantic system to a greater extent, blurring the contrast between true and false information, and thus causing them to drop their guard.

The mechanism underlying such language-culture dissociation effects must involve interactions between brain structures involved in language-selection networks (e.g., Abutalebi & Green, 2007; Luk, Green, Abutalebi & Grady, 2012), basic emotion generation (e.g., limbic areas, Damasio, Grabowski, Bechara, Damasio, Ponto, Parvizi & Hichwa, 2000; Dalgleish, 2004) and regions of the brain implementing higher-order semantic processing (e.g., temporal poles, Lambon-Ralph, Pobric & Jefferies, 2009; Bonner & Price, 2013). Further research using functional neuroimaging is required to characterize the neural organization of such networks.

To conclude, we set out to examine how bilinguals might perceive verifiable information differently in the native and the second language, but unexpectedly found that semantic evaluation of negative content is selectively disturbed in the native language. Thus, bilinguals are more susceptible to emotional interference in their native language, but better able to withstand cultural criticism in their second. These results extend language contextual effects beyond the realm of decision-making to the domain of objective information assessment.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(*3*), 242–275.

Altarriba, J. (2008). Expressions of emotion as mediated by context. *Bilingualism: Language and Cognition, 11*, 165–167.

Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *Am J Psychol,* 389–410.

Aycicegi, A., & Harris, C. (2004). Bilinguals' recall and recognition of emotion words. *Cognition and Emotion, 18*(*7*), 977–987.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.

Baayen, R.H., Piepenbrock, R., & van Rijn, H. (1993). CELEX Lexical Database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin, 140*(*6*), 1556.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(*3*), 255–278.

Bates, D., Maechler, M., & Dai, B. (2008). lme4: Linear mixed-effects models using s4 classes [Computer software manual]. Retrieved from http://lme4.r-forge.r-project.org (R package version 0.999375-28)

Bonner, M. F., & Price, A. R. (2013). Where is the anterior temporal lobe and what does it do?. *J Neurosci, 33*(*10*), 4213–4215.

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax and semantics. *Language in Mind: Advances in the Study of Language and Cognition,* 61–79.

Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Res, 1479*, 72–79.

Briley, D. A., Morris, M. W., & Simonson, I. (2005). Cultural chameleons: Biculturals, conformity motives, and decision making. *Journal of Consumer Psychology, 15*, 351–362.

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language, 122*(*3*), 211–226.

Costa, A., Foucart, A., Arnon, I., Aparici, M., & Apesteguia, J. (2014). "Piensa" twice: On the foreign language effect in decision making. *Cognition, 130*, 236–254.

Coulson, S., Urbach, T. P., & Kutas, M. (2006). Looking back: Joke comprehension and the space structuring model. *Humor–International Journal of Humor Research, 19*(*3*), 229–250.

Dalgleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience, 5*(*7*), 583–589.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience, 3*(*10*), 1049–1056.

Danziger, S., & Ward, R. (2010). Language changes implicit associations between ethnic groups and evaluation in bilinguals. *Psychological Science, 21*, 799–800.

Dewaele, J. M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development, 25*, 204–222.

Ellis, C., Kuipers, J, Thierry, G., Lovett, V., Turnbull, O., & Jones, M. W. (2015). Language and culture modulate online semantic processing. *Social, Cognitive and Affective Neuroscience*. DOI:10.1093/scan/nsv028

Ellis, N. C., O'Dochartaigh, C., Hicks, W., Morgan, M., & Laporte, N. (2001). Cronfa Electroneg o Gymraeg (CEG): a 1 million word lexical database and frequency count for Welsh. Available: http://www.bangor.ac.uk/canolfanbedwyr/ceg.php.en (accessed January 2015).

Gao, S., Zika, O., Rogers, R. D., & Thierry, G. (2015). Second language feedback abolishes the "hot hand" effect during even-probability gambling. *J Neurosci, 35*(*15*), 5983–5989.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol, 55*(*4*), 468–484.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*(*5669*), 438–441.

Jończyk, R., Boutonnet, B., Musiał, K., Hoemann, K., & Thierry, G. (2016). The bilingual brain turns a blind eye to negative statements in the second language. *Cognitive, Affective, & Behavioral Neuroscience, 16*(*3*), 527–540.

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect thinking in a foreign tongue reduces decision biases. *Psychological Science, 23*, 661–668.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annu Rev Psychol, 62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161–163.

Lambon-Ralph, M. A., Pobric, G., & Jefferies, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: convergent evidence from rTMS. *Cereb Cortex, 19*(*4*), 832–838.

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes, 27*(*10*), 1479–1488.

Martin, C. D., Thierry, G., Kuipers, J. R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language, 69*(*4*), 574–588.

Ogunnaike, O., Dunham, Y., & Banaji, M. R. (2010). The language of implicit preferences. *Journal of Experimental Social Psychology, 46*, 999–1003.

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition, 11*, 147–164.

Phinney, J. (1992). The Multigroup Ethnic Identity Measure: A new scale for use with adolescents and young adults from diverse groups. *Journal of Adolescent Research, 7*, 156–176.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology, 37*(*2*), 127–152.

R Development Core Team. (2008). An Introduction to R. *Network Theory Limited, Bristol*.

Roberts, R. E., Phinney, J. S., Masse, L. C., Chen, Y. R., Roberts, C. R., & Romero, A. (1999). The structure of ethnic identity of young adolescents from diverse ethnocultural groups. *The Journal of Early Adolescence, 19*(*3*), 301–322.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences, 104*, 12530–12535.

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*, 4567–70.

Wu, Y. J., & Thierry, G. (2012). How reading in a second language protects your heart. *J Neurosci, 32*, 6485–6489.