

Bilingual Speakers in the Lab: Psychophysiological Measures of Emotional Reactivity

Catherine L. Harris

Psychology Department, Boston University, Boston, MA, USA

Bilingual speakers report experiencing stronger emotions when speaking and hearing their first language compared to their second. Does this occur even when a second language is learned early and becomes the dominant language? Spanish–English bilinguals who had grown up in the USA (early learners) or those who were first exposed to English during middle childhood while residing in a Latin American country (late learners) listened to words and phrases while skin conductance was monitored. Stimuli included taboo words, sexual terms, childhood reprimands ('Go to your room!') and single words which functioned as a neutral baseline. Consistent with the hypothesis that a second language is less emotional for the late learners, emotional expressions (i.e. reprimands) presented in the first language elicited larger skin conductance responses than comparable expressions in the second language. For the early learners, no such difference was obtained, indicating that age of acquisition of the second language and proficiency modulate speakers' physiological reaction to emotional language.

Keywords: language, skin conductance, bilingualism, emotion, Spanish–English bilinguals, auditory processing

Introduction

Second language (L2) users commonly report that the subjective experience of language use differs systematically for their first (L1) and second languages (Bond & Lai, 1986; Dewaele & Pavlenko, 2002; Ervin, 1964; Gonzalez-Reigosa, 1976; Pavlenko, 1998, 1999, 2002). A well known example is the case of taboo words, which reportedly generate more anxiety when spoken in one's L1 (Dewaele, this issue; Ferenczi, 1916; Greenson, 1950; Javier, 1989). Using self-report questionnaires with closed and open-ended questions, Dewaele (this issue) asked bilingual and multilingual speakers to rate the emotionality of swear words in their various languages. Swear words in the native language were rated as most forceful, with perceived forcefulness declining monotonically with age of acquisition and the languages' rank-order of acquisition (i.e. swear words in a multilingual's L1 were perceived as more forceful, on average, than swear words in the L2, L3 or later-learned languages). Naturalistic learning contexts also lead to more perceived emotional force than formal instruction. Similar age and learning context trends were found for preferences in expressing anger (Dewaele & Pavlenko, 2003). Interview data and written comments confirmed the following hypothesis about why multilinguals sometimes prefer using their L1 for cursing, and at other times prefer a later-acquired language: if emotionality is desired, then the native

language is preferred, but if the emotionality of swear words is uncomfortable, multilinguals choose a later-learned language.

Differences in emotionality have been documented by soliciting oral or written narratives, speech samples and autobiographical memories (Clachar, 1999; Dewaele & Pavlenko, 2002; Ervin, 1964; Javier *et al.*, 1993; Pavlenko, 1998). Bilingual speakers feel freer to discuss embarrassing topics in their L2 (Bond & Lai, 1986). Studies of codeswitching describe how a change to the L2 often serves a distancing function (Gumperz & Hernandez, 1971; Javier & Marcos, 1989). The bilingual therapist Gonzalez-Reigosa (1976) described Spanish–English bilingual patients who employed English when discussing anxiety-arousing topics, and used English for portraying a persona of self-confidence, calm and emotional reserve (see also Altarriba & Santiago-Rivera, 1994). Immigrants' childhood memories were more emotionally charged when described in their native language (Schrauf, 2000). However, research has not invariably found the L1 to be the language which facilitates emotional expression. Koven (2001, 2003) found greater affect and more complex story telling (e.g. use of quoted speech) when the children of immigrants used the language of the host country, which was also the language of peer socialisation.

One of the few laboratory studies of emotional responsiveness to words in the L1 versus L2 is that of Anooshian and Hertel (1994), who used word recall as their dependent measure. Recall of words is known to be influenced by emotionality (Rubin & Friendly, 1986). Anooshian and Hertel (1994) hypothesised that emotion words in the L2 would lack the emotional connotations which render them easy to recall. They did in fact find an effect of emotionality on recall only in the L1. However, a recent attempt to replicate this failed. Turkish sojourners in the USA who were late learners of English showed similar recall of emotion words in L2 English rather than in L1 Turkish (Ayçiçeği & Harris, *in press*). Because recall of emotion words has led to contradictory findings, it would be beneficial to use a measure of emotion processing which more closely reflects physiological aspects of emotional arousal.

Why Study Bilingualism Psychophysiologicaly?

There are several reasons for using psychophysiological techniques to verify the intuition that the subjective, emotional experience of using a L2 differs from that of using a L1. The theoretical perspective adopted is that subjective experience is mediated by physical states of the nervous system, and that experiences which are commonly reported by many people will have measurable physiological correlates. This will be referred to as the 'brain-based perspective'. Being able to quantify emotional aspects of language with psychophysiological data will thus contribute to a science of the subjective experience of cognitive processing (following Damasio, 1999; O'Regan & Noe, 2001; Schumann, 1997). If age of acquisition of a language influences autonomic arousal, then researchers will have a new source of data with which to evaluate claims about the sensitive period for language. One might be able to show that words learned very early in life are associated with

greater autonomic responses than are later-learned words which are similar in other respects. Adopting a brain-based perspective on the emotional experience of language also allows one to make more sophisticated hypotheses than simply that the L1 will be experienced as more emotional. In particular, aspects of the language-learning history which have consistent effects across many individuals are likely to do so via common brain mechanisms. They are thus candidates for having an effect on the emotionality and the autonomic nervous system. For example, the dominance of the L1 can be supplanted by the L2 through constant exposure and use, as when children acquire the L2 following immigration (Köpke, 2003). In these cases, the dominance of the L2 has a physical basis, which can be assessed via neuroimaging (e.g. Kim *et al.*, 1997) or behavioural tasks (e.g. reaction time tasks). As proficiency and dominance of the L2 have physiological correlates, the 'brain-based perspective' makes the predictions that proficiency and dominance will affect emotional reactivity and can be measured with autonomic nervous system arousal. Indeed, this was the rationale in the current study for seeking individuals whose L2 had become their dominant language.

The brain-based perspective has heuristic value. It specifies that subjective experiences have their origin in measurable brain states. What conclusions should be drawn if individuals' subjective reports are at odds with psychophysiological measures? Such a finding would mandate new theoretical development. This would falsify the brain-based perspective or would force one to develop an explanation for why subjective and physiological reports differed. For example, one could propose that informants were relying on stereotypes or had fashioned an incorrect or incomplete story about their emotional states.

Language and Psychophysiology

The autonomic nervous system responds to danger or threat cues by readying systems of the body to take action, including fighting or fleeing (Hugdahl, 1995). Physiological responses include sweating of palms and fingertips, a signal that can be quantified by measuring the transient increase in electrical conductivity of the skin. The increase in conductivity to a specific stimulus, called a skin conductance response (SCR), typically occurs within 1–1.5 seconds following appearance of the stimulus, and may last for 2–6 seconds.

The phasic amplitude of the SCR is most sensitive to threatening stimuli, but may also index relevance of a stimulus. Thus, even a photograph of the face of an acquaintance, when embedded in a stream of unfamiliar faces, will elicit heightened responsiveness (Channouf & Rouibah, 1997; Tranel *et al.*, 1985). Experiments dating from the mid-20th century have found that reading or hearing taboo-words elicits a larger SCR than reading or hearing neutral words (Bingham, 1943; Gray *et al.*, 1982; Manning & Melchiori, 1974; Mathews & MacLeod, 1985; Mathews *et al.*, 1989; McGinnies, 1949). Recent studies have found that taboo words activate the amygdala and other brain structures which mediate the arousal which accompanies detection of threat (LaBar & Phelps, 1998). Among monolinguals, even emotion-laden words,

such as 'cancer' and 'kill', elicit higher responses than neutral words (Dinn & Harris, 2000).

Language stimuli have been most frequently used to study fear sensitisation, conditioning and personality differences in orienting and general autonomic reactivity (Barry, 1980; Grings & Zeiner, 1965; Mathews *et al.*, 1989; Stelmack *et al.*, 1983a, b). For example, researchers have found that the phrases 'strong shock', 'medium shock', 'weak shock' and 'no shock' elicit skin conductance amplitudes which differ according to the intensity suggested by the words (Grings & Zeiner, 1965). Patients with anxiety disorders show larger SCRs to anxiety-related words, particularly those related to their fears (e.g. spider phobics react to the word 'spider', whether printed or spoken). When hearing ambiguous words such as 'bug', anxiety-prone individuals appear to operate with a bias to interpret words as having a threatening meaning (Mathews *et al.*, 1989).

Lacking in the electrodermal literature are systematic manipulations of the variables of interest to psycholinguists. Researchers haven't designed experiments to test whether SCRs are greater to single words than words in context, to low frequency versus high frequency words, or to the first occurrence of a word or a phrase compared to a latter occurrence. Indeed, current 'state-of-the-art' reviews, such as the chapter by Dawson *et al.* (2000) and the book by Boucsein (1992), do not have sections or index items on language.

In contrast with the dearth of work on autonomic reactivity to language, electrical activity across the scalp, as measured via electroencephalogram, is a common tool for language researchers. Event-related potentials (ERPs) elicited by language stimuli constitute a major subfield of ERP research and the most frequently used method of studying language within cognitive neuroscience. Several studies have used ERPs to examine bilingual language processing but only one known to this researcher has focused on emotional valence of the words. Kim (1993) employed ERPs to index emotional responsiveness to words heard by Korean-English bilinguals speakers in their L1 and L2. The P300 amplitude was selected as the dependent measure, following the literature that this ERP component is sensitive to the incentive or emotional value of a stimulus. Kim recorded ERPs to neutral, positive and negative English words from monolingual English speakers and Korean speakers who had varying degrees of English-language competency. However, no differences were found in P300 amplitude as a function of words' emotional valence or participants' English proficiency. Kim concluded that her emotion words (words with negative connotations such as 'steal' and positive connotations such as 'truth') had probably been insufficiently arousing, since prior studies documenting the sensitivity of P300 amplitude to emotional stimuli used highly evocative stimuli such as pictures and slang expressions (Vanderploeg *et al.*, 1987).

Kim's (1993) study describes an arena in which electrodermal monitoring may be a superior technology to ERPs. ERPs have excellent temporal resolution (milliseconds versus seconds required to measure a phasic increase using skin conductance) and are known to be exquisitely sensitive to myriad lexical and grammatical factors (Kluender & Kutas, 1993; Kutas & King, 1996). However, Kim (1993) found that ERPs were not sensitive to emotional valences

of words. SCRs are well known to be sensitive to differences in the same types of words Kim used in her ERP study (e.g. Dinn & Harris, 2000), suggesting that words' emotionality is a case where electrodermal recording may be preferred to ERPs, despite the latter's superior temporal resolution.

Bilingualism and Electrodermal Activity

Differential responsiveness to emotional terms in L1 and L2 was recently studied in a sample of late learners of English (Harris *et al.*, 2003). Native speakers of Turkish who had moved to the USA for schooling or work after age 18, read or heard a variety of word types in Turkish and English. Participants responded to items by rating them for pleasantness while skin conductance activity was monitored via fingertip electrodes. Items included taboo words (curse words, body parts and sexual terms), reprimands ('Don't do that!') and neutral words ('column', 'table'). The reprimands were selected to be the type that parents might use in an admonishing tone to children, and included two that have strong childhood associations: 'Shame on you!' and 'Go to your room!'.

Figure 1 shows mean phasic SCR for the taboo words, reprimands and neutral words. SCRs were stronger to L1 Turkish emotional expressions than to L2 English expressions, especially in the auditory modality. This corroborates the intuitions of L2 speakers that emotional expressions are more intense when they occur in the native language.

Two additional results are noteworthy. The largest difference between L1 and L2 response patterns occurred not for taboo words, but for reprimands. During debriefing, several participants said that they could hear, in their mind, a family member saying a Turkish reprimand to them. This is consistent with the work of Schrauf and Rubin (1998, 2000, 2003) who prompted adult Spanish–English speakers to retrieve autobiographical memories. For memories that had come to them in words rather than images, speakers were aware of whether Spanish or English was used in the memory.

Second, modality effects differed across the languages, with Turkish, but not English, showing larger amplitudes for auditory stimuli. English did not show an advantage of the auditory modality. This sample of late learners of

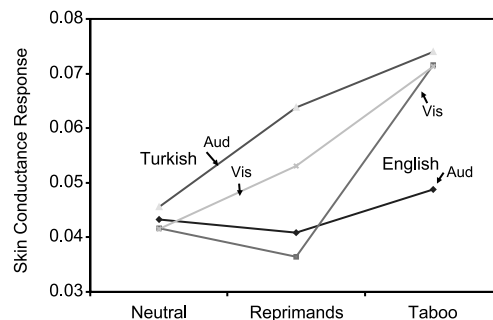


Figure 1 Turkish–English bilinguals' SCRs are graphed according to modality of presentation.

English mostly acquired English in a formal classroom setting (after age 12). This means that the early context of learning English may have emphasised the written form. Taboo words elicited larger SCRs when they appeared printed on the screen than when they were played through the computer's loudspeakers. Words and short phrases presented in isolation can be difficult to discern. Indeed, three taboo terms, 'pee', 'whore' and 'raped', were sometimes not understood by several respondents, who rated these words with a 0 to indicate lack of comprehension (items rated as 0 were excluded from the electrodermal trials summarised in Figures 1 and 2). The relative unfamiliarity of taboo words heard in isolation could have hindered lexical identification, whereas the visual modality allowed easy access to the meaning of these stigmatised words and consequently larger SCRs.

To rule out the possibility that the English words, phrases and taboo items were less emotionally arousing because of their inherent meaning or the manner in which they were spoken, the electrodermal activity of 28 monolingual English speakers (college students) was collected using the same protocol, after removing the Turkish items (Harris & Ayçiçeği, unpublished data). As shown in Figure 2, the monolingual speakers showed a strong taboo word effect that was similar to that of Turkish speakers in their native language. An unexpected finding was that the US college students did not have heightened responsiveness to reprimands. This raises the question of whether culturally specific patterns of childhood discipline influenced Turks' elevated reactivity to reprimands.

This first study indicates the importance of modality and cultural factors in modulating the heightened reactivity to emotional expressions in L1 compared to L2. However, perhaps the most important question is determining if the L1 (acquired in the home from parents) invariably elicits stronger electrodermal responses than later-learned languages. It would be helpful to recruit

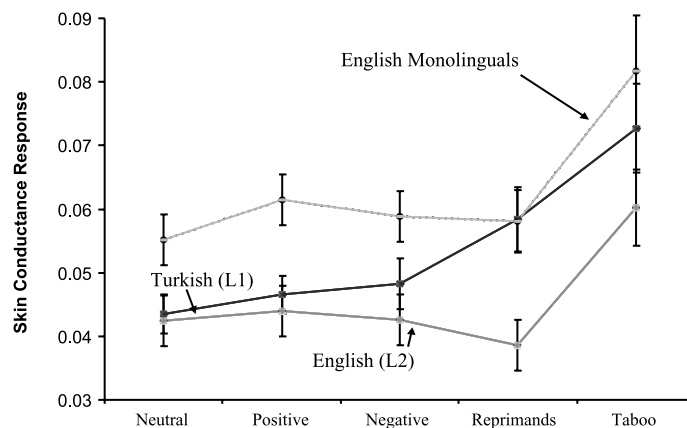


Figure 2 The two lower lines show the Turkish–English bilinguals' SCRs to words and phrases in their L1 and L2. The responses of English monolingual speakers to the English stimuli are also included. Overall SCRs were higher for English monolinguals because they only encountered half as many stimuli as the Turkish participants, and thus their SCRs were not as reduced by habituation to the stimuli and testing situation.

participants who are equally or more proficient in their L2. Young adults who were born in the USA to immigrant parents are an ideal population. In the present study, Spanish–English bilinguals were chosen as the research population. Prior interviews with Spanish-speaking students at Boston University indicated that many viewed English as their most proficient language, but Spanish as their L1, as it was the language acquired from parents during early childhood. If proficiency is the most important variable for heightened autonomic activity to emotion words, then English stimuli should elicit the largest SCRs. If age of acquisition is the most important variable, then Spanish stimuli will be most evocative.

Method

Participants

Participants were undergraduate or graduate students at Boston University who volunteered their time or received course credit. Recruitment proceeded via ads and emails in the psychology department requesting Spanish–English bilinguals who had learned Spanish as a first language. Students who responded to these ads were only run through the experimental protocol if they learned Spanish in the home from Spanish-speaking parents, and if English was the second language acquired. This information was ascertained during a preliminary phone call or when participants appeared at the laboratory.

Language history interview

The experimental protocol included an extensive interview about participants' language-learning history, which was administered during a break in the electrodermal recording session. The interview was relaxed and informal, with the goal of obtaining a detailed and nuanced understanding of factors that could influence the participant's bilingual proficiency and emotional reactivity. Participants were asked their age, country of birth, residence in different countries, age of arrival in the USA or English-speaking countries, different languages encountered and spoken, language spoken in the home by parents and siblings, and use of English and Spanish (or other languages if relevant) at school. Experimenters also used their own intuitions to solve 'mysteries' about the participants' language-learning history. For example, the experimenter might inquire, 'Your family didn't immigrate to the US until you were 16, yet you have nearly accentless English.' The answer could be that the family had vacationed in the USA for many summers and the student had attended an English-language school in the home country and interacted with friends or relatives who were native English speakers. Participants self-rated themselves for English and Spanish proficiency on a seven-point scale (7 = native speaking abilities, 6 = near native, 5 = very good, 4 = good, 3 = fair, 2 = poor, 1 = minimal). Separate ratings were made for speaking, understanding, reading and writing, and then the average across four judgements was taken as the proficiency score.

A group of early learners of English were identified based on the language history interviews ($n = 15$). Participants were born in the USA ($n = 6$) or immigrated between ages three and seven ($n = 9$). All of these participants considered Spanish to be their *first* language, as it was learned from parents and spoken in the home during early childhood. However, all identified English as their *best* language, as it was the language of past and present schooling and peer culture. Two participants who immigrated at age seven rated themselves as having native-speaker abilities in both languages, although English was judged their best language because of their greater facility in written English. Three participants who immigrated before age seven judged themselves to have less than native-speaker abilities in both languages. None of the participants who were born in the USA rated themselves as having native-speaker abilities in Spanish (although they did rate themselves as having near-native abilities). Mean values and ranges for language-history variables appear in Table 1.

A group of late learners ($n = 21$) was also highly proficient in English, but grew up in South or Central America. These respondents generally began to learn English in middle childhood via bilingual classes, and then moved to the USA (or an English-speaking country) in their mid-teens when the family immigrated, or at age 18 or in their early 20s to attend college or graduate school in the USA. The late age-of-arrival for some participants does not indicate poor English abilities. The ability to gain acceptance to a competitive college requires English and generally good scholastic abilities. Many of these respondents had parents who frequently spoke English and had worked or studied in English-speaking countries. Two were born in the USA while parents were working or studying in a US city, and then returned with the family to South America by age two. Many individuals in this group had visited the USA for summer vacations and had used English while travelling abroad. The main difference between the two groups is thus being reared in the USA (early learners) versus acquiring English in school and from family members while growing up in South or Central America (late learners). While

Table 1 Learning history for Spanish–English bilingual participants

	<i>Age Acquired English</i>		
	<i>Early Childhood</i>	<i>Middle-late Childhood</i>	<i>Statistical significance</i>
Age of arrival in the US*	3.1 (0–7)	17.9 (13–25)	$p < 0.001$
Age of exposure to English	3.7 (0–7)	7.9 (1–16)	$p < 0.001$
Length of stay in the US	16.2 (11–19)	2.4 (0.17–20)	$p < 0.001$
Spanish proficiency [†]	5.9 (4.25–7)	6.8 (6–7)	$p < 0.001$
English proficiency [†]	7.0 (6.25–7)	5.6 (4–7)	$p < 0.001$
Age at testing	20.3 (19–32)	20.7 (18–36)	Age n.s

Notes: Values are averages, with minimum and maximum in parentheses. *or other English speaking country. [†] 7 point scale, 7 = native speaker abilities.

all respondents in the latter group rated their Spanish as superior to their English, four individuals who had attended bilingual or English-language schools while growing up in a Spanish-speaking country rated either their reading or their writing in Spanish as being less than native ability, rendering their Spanish proficiency scores 6.25–6.75, rather than 7 (7 indicated native-speaker ability). Ratios of Spanish-to-English proficiency were 0.86 for the early learners and 1.2 for the late learners.

As shown in Table 1, the two groups differed significantly on all demographic and proficiency variables except for age at the time of testing. Correlational statistics were run across the two groups as a whole. Not surprisingly, the learning history variables were intercorrelated. Having an early age of arrival in the USA meant an early age of first significant exposure to English, $r = 0.48$, $p < 0.01$, and of course longer total length of residence in the USA, $r = -0.96$. The self-ratings of proficiency obtained in the current study are consistent with findings of Birdsong and Molis (2001). Early age of arrival was a better predictor of increased English proficiency and decreased Spanish proficiency than either age of significant exposure or length of stay. Age of arrival was strongly correlated with self-rated proficiency in English, $r = -0.71$, and with decreased proficiency in Spanish, $r = 0.62$. Length of stay was strongly correlated with English proficiency, $r = 0.73$, and moderately correlated with decreased Spanish proficiency, $r = 0.57$. First age of significant exposure to English was weakly correlated with English proficiency, $r = -0.38$, $p < 0.03$, and with decreased Spanish proficiency, $r = 0.31$, $p < 0.08$.

Because learning-history variables were intercorrelated, the current study makes no claims about whether the causal factor for electrodermal reactivity is age of arrival in the USA, significant exposure to English or number of years residing in the USA. Other researchers have studied these factors and have concluded that age of arrival is the most important for ultimate attainment of native-like speaking abilities (Birdsong & Molis, 2001). Disentangling the impact of these variables on autonomic reactivity could be pursued in future research.

Materials

Four categories of emotional expressions were used: taboo words, reprimands, endearments and insults, with eight Spanish items and eight English items in each category. Candidate English items in each category were drawn from the previous study (Harris *et al.*, 2003) or were suggested by the undergraduate laboratory assistants. To avoid offending undergraduates,¹ the strongest taboo words in either Spanish or English were not used. Items which are not strictly taboo were also included, if they had strong sexual connotations, such as 'raped' and 'breast' (a complete list of stimuli appear in the Appendix). The category label 'taboo words' is thus a shorthand label for 'taboo words and sexual terms'. It could be useful in future work to separately compare taboo words, including the strongest, most socially stigmatised words, and sexual terms.

A native Spanish-speaking research assistant identified possible Spanish equivalents. For example, the English phrase 'You are everything to me!' was

matched with the Spanish phrase '*i Eres mi vida!*' ('You are my life'). Because repetition of meaning could lead to habituation, not all pairs were direct translation equivalents. Three Spanish-speaking students who were naive to the goals of the study rated a larger pool of English and Spanish emotional expressions for emotional intensity and frequency of use. The final set of items was selected so that Spanish and English items were approximately similar on these dimensions. As shown in the Appendix, when phrases appeared visually, they were presented with initial capital letter and final exclamation point, period or question mark, to convey emotional vividness. This differs from the previous study, in which phrases (which were only reprimands) appeared without punctuation.

Participants encountered a total of 64 emotional expressions, half in English, half in Spanish, with half of the items in each language being presented visually, and the other half auditorily. These 64 emotional expressions were intermixed with 36 less emotional items. These were 36 single words selected from the prior study (Harris *et al.*, 2003), and translated into Spanish by a native Spanish speaker and verified by an additional Spanish speaker. These 36 single words included 12 items each in the categories of positive words, negative words and neutral words. These words had been previously selected using the *Handbook of Semantic Word Norms* (Toglia & Battig, 1978), which contains pleasantness norms for a wide variety of English words. The reason to include positive and negative words in a category which essentially serves as a neutral baseline is to encourage participants to think about the emotional meaning of all items while they are rating stimuli for unpleasantness. Also, the prior study revealed that positive words ('joy', 'bride') and negative words ('cancer', 'kill') differed only minimally from neutral words in electrodermal responsiveness.

Single words may be experienced as less emotional than taboo words and emotional expressions because of lack of context. While single words such as 'cancer' or 'joy' may be recognised as referring to emotion-laden events, they may not readily elicit an emotional reaction unless one conjures up a specific emotional event (e.g. 'I learned I have cancer'). One might wonder if short sentences (and the emotional phrases were actually all sentences) would elicit larger electrodermal responsiveness than single words simply because of their length. At present there is no study comparing electrodermal responsiveness to single words and the same words in a neutral or emotional context. However, in Harris *et al.* (2003) taboo words appeared as single words, and elicited the strongest skin conductance amplitudes. One might also wonder if the single words used were lower frequency than the words in the emotional phrases, and whether frequency differences would explain differences in electrodermal responsiveness. However, taboo words are generally the lowest frequency items and also the ones that elicit the greatest skin conductance amplitudes. Indeed, word frequency was not related to responsiveness in the prior study (Harris *et al.*, 2003). The question of why the single-word stimuli elicit lower SCRs than emotional expressions isn't, however, central to the current study. Instead, the goal is to measure responsiveness to emotional stimuli across L1 and L2. Doing this requires a relatively nonemotional type of stimulus to serve as a neutral baseline.

For each participant, stimuli were presented in both languages and in both the auditory and visual modality. For the single-word items, participants encountered either the Spanish or English item (not both). For the English–Spanish pairs of emotional expressions, participants encountered both the English member and the Spanish member of the pair (e.g. participant encountered both ‘You are everything to me!’ and ‘*iEres mi vida!*’). Since participants were told that their bilingualism was being studied, they may have noticed the translation equivalents. They may have translated the item to compare it to the earlier item, or they may have responded more strongly or more weakly because of their memory for the earlier item. Unfortunately, it is not helpful to compare size of skin conductance amplitudes to the first and second occurrence of a pair, since electrodermal amplitudes decline substantially over the course of a 30-minute experimental session. The second occurrence will generally be much weaker than the first because items later in the session will have dampened amplitudes. The most that can be done is to arrange the order of items so that half the participants received the Spanish item in the first half of their list, while the other half encountered the English item first. Four materials sets were constructed so that items in the single-word category could be assigned to appear either in English or Spanish, and so that all items could be assigned to appear in either the auditory or visual modality.

Equipment and data preprocessing

All items were recorded in English and Spanish by a female native Spanish speaker whose family moved to the USA when she was age seven. She had a native-speaker accent in both languages; her Spanish accent was readily identifiable as Central American.² The emotional expressions were uttered with an emotional tone appropriate to that item (speaking with flat or neutral intonation renders the emotional expression anomalous and unrepresentative of prior experience). Stimuli were recorded using SoundEdit Version 16 on a Power Macintosh G3 computer.

Stimuli were presented using PsyScope experimental software, developed by Cohen *et al.* (1993). Phasic electrodermal activity was recorded using the Davicon C2A Custom Skin Conductance Monitor (NeuroDyne Medical Corporation). Davicon Psychophysiological Assessment Software subtracts the basepoint from the maximum score during each 10-second recording interval, yielding a numeric value in micromhos, which is the amplitude of the phasic SCR.

Enormous variability exists within and across individuals in both tonic and phasic electrodermal levels (Lykken & Venables, 1971). For example, a participant with baseline values which hover around 6.0–7.0 may have phasic amplitudes of 0.40 for a large response and 0.15 for a small response. For a participant whose baseline values are 3.5–4.0, 0.15 may be a large response. To correct for these differences in range and to increase homogeneity of variance in the data, participants’ scores were converted to *z*-scores, following the recommendations of Ben-Shakhar (1985). This means that each individual’s set of scores had a mean of 0 and a standard deviation of 1. Values more than 2.5 standard deviations from an individual participant’s mean score were

truncated to 2.5. Resulting scores were then multiplied by 100 to avoid scrutinising mean scores in which significant digits are several places to the right of the decimal point. While individual scores thus ranged from -250 to 250, averaging together a large set of scores with a mean of 0 results in a value that converges towards 0. When viewing graphs of the *z*-scored data, scores of 0 do not indicate a null response, but instead indicate a response of average amplitude. Scores above 0 indicate a response of greater than average amplitude.

When collecting electrodermal data, the ideal trial yields an increase in electrodermal activity which occurs 1–1.5 seconds after stimulus onset and appears as a spike or wave (Hugdahl, 1995). However, many trials are flat or reveal a decline in skin conductance level. There is considerable debate about how to interpret these non-ideal trials (Dawson *et al.*, 2000). Declines generally occur because skin conductance levels were elevated on a previous trial. The return to baseline may not be completed by the end of the 10-second recording window, but continues into one or two subsequent trials. One method is to record a negative amplitude for these trials. That is, if skin conductance levels fall from 3.56 to 3.20, the trial could be assigned the value of -0.36. This method effectively misattributes a negative value to a trial that was unlucky enough to follow a high-response trial. That same stimulus might have produced an ordinary SCR had it appeared in a different context. An alternative is to delete these trials, although this has the disadvantage that no information is gained from the trial. An intermediate position is to record the value of 0 for these trials. This method was adopted because it corresponds to the fact that a particular participant did not respond to the evoking stimulus.

Procedure

Each participant was tested individually. Ethical guidelines of the American Psychological Association were followed. Informed consent forms describing the experiment were given prior to testing; participants had the option to end their participation at any time or to request rest breaks. No participant found the experiment unduly long or taxing. Electrodes were placed on the index and middle fingers of the dominant hand to record changes in electrical conductivity. Participants were instructed to rate each word or phrase for unpleasantness on a 1–7 scale (1 being most pleasant, 4 neutral and 7 being most unpleasant) by typing the appropriate key on a standard computer keyboard. The instructions stated that 0 was to be typed if they did not know the meaning of the word or phrase. Items that were rated as 0 amounted to less than 1% of trials and were discarded from analysis.

The unpleasantness scale was displayed on the computer screen for the entire 10-second recording interval. Participants were told they had 10 seconds to make their rating, and to respond whenever they felt ready during that time. They were additionally urged to think about the meaning of the word or phrase for the full 10 seconds, given that the physiological response can take several seconds to manifest itself. For visual presentation, the word or phrase was displayed on the screen until a response was made. For auditory

presentation, a question mark remained on the screen until a response was made. Participants were instructed to minimise movement or talking while their skin conductance was being monitored. The experimenter (who was either the author or a trained research assistant) sat next to participants in front of a second computer displaying the electrodermal activity, and made notes on possible artefacts. Artefacts were defined as either (1) an electrodermal response which could not have been elicited by the stimulus, because it occurred immediately with stimulus onset rather than the 1 second or more which is required to measure a change in skin conductivity, or (2) a large or unusual-looking SCR which coincided with movement made by participant (e.g. cough, scratch or stretch).

A well known but sometimes frustrating aspect of measuring skin conductance is that amplitudes decline and become flat as the experimental setting and nature of the task become familiar to participants (Dawson *et al.*, 2000). This is similar to the reduction in arousal that accompanies a sedentary activity such as listening to a lecture. That is, arousal is high following the physical activity that brings one to the lecture hall, and then usually declines as one habituates to the lecture hall and topic matter. Psychophysiologicalists must thus grapple with what can amount to a disappearing dependent variable. Indeed, some researchers use as their dependent variable number of trials until participants' SCRs asymptote at a level value (Dawson *et al.*, 2000). A method used in the current study was to provide participants with a break midway through the 98 trials. This break was used to conduct the language history interview (described earlier). Most participants did dishabituate after the break.

Results

Both common-sense intuitions and findings from the study of Turkish–English bilinguals (Harris *et al.*, 2003) suggest that higher electrodermal responsiveness to auditory stimuli in a language (compared to visual stimuli) indicates comfort and proficiency in that language. Would the early learners of English show an auditory advantage for Spanish, English or both languages? Modality effects were assessed by conducting 2×2 repeated measures analysis of variance (ANOVA) on the two levels of modality (auditory versus visual) and the two languages. As shown in Figure 3, stimuli in the auditory modality elicited higher SCRs than did visual stimuli, but only for the early learners, $F(1,14) = 5.7, p < 0.03$. The late learners had comparable reactivity to visual and auditory stimuli, $F < 1$. No other modality effects or interactions were significant.

To determine how reactivity to emotional expressions differed from the single word condition, 2×5 repeated measures ANOVAs were conducted separately for the two participant groups (two levels of language, five stimulus types). ANOVA on the early learners revealed a main effect of stimulus type, $F(4,56) = 4.2, p < 0.005$. As shown in Figure 4, this effect was specific to taboo words. English taboo words differed from the single word condition, $F(1,14) = 7.7, p < 0.02$, as did Spanish taboo words, $F(1,14) = 4.6, p < 0.05$.

For the late learners, the main effect of stimulus type was significant at $F(4,80) = 4.2, p < 0.004$. Taboo words elicited elevated SCRs (compared to the single word condition) in both English, $F(1,20) = 10.4, p < 0.005$, and Spanish, $F(1,20) = 4.3, p < 0.05$. The reprimands also elicited SCRs that were greater than the single word condition, but only in Spanish $F(1,20) = 4.4, p < 0.05$ (see Figure 5).

Speakers also rated items for unpleasantness. As shown in Figure 6, participants' ratings of unpleasantness were broadly similar across languages, modalities and age-of-acquisition groups. As expected, endearments were rated as highly pleasant and insults, reprimands and taboo words were rated as unpleasant. The single word condition included equal numbers of positive,

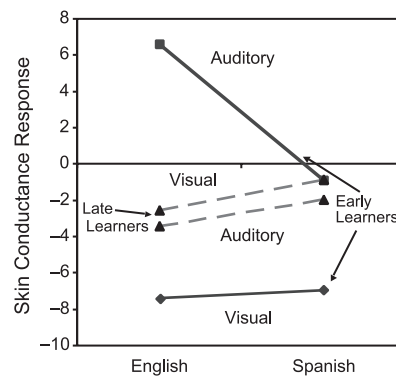


Figure 3 Items in all stimulus categories were averaged and plotted according to auditory or visual presentation. Early learners had larger SCRs to auditory stimuli than to visual stimuli, consistent with naturalistic, conversation learning mode for both languages.

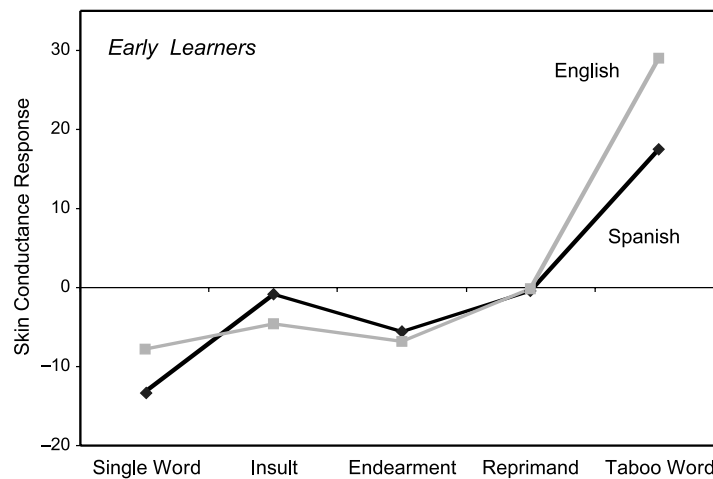


Figure 4 SCRs plotted for the different stimulus categories for the early learners. Taboo items were statistically greater than the single-word category.

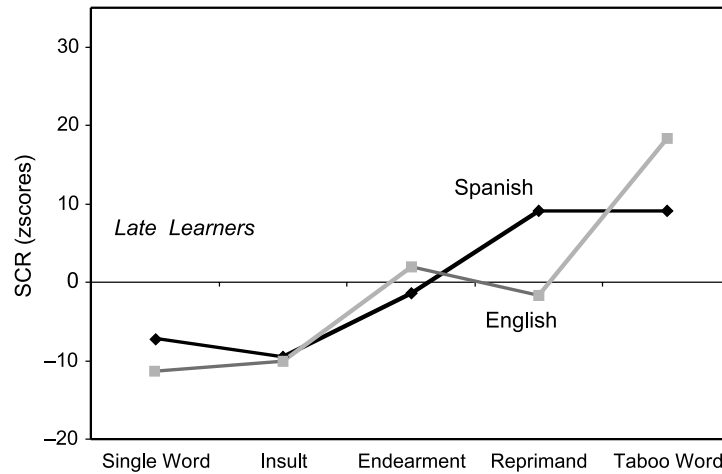


Figure 5 SCRs plotted for the different stimulus categories for late learners. Taboo items were statistically greater than the single-word category, as were reprimands presented in Spanish.

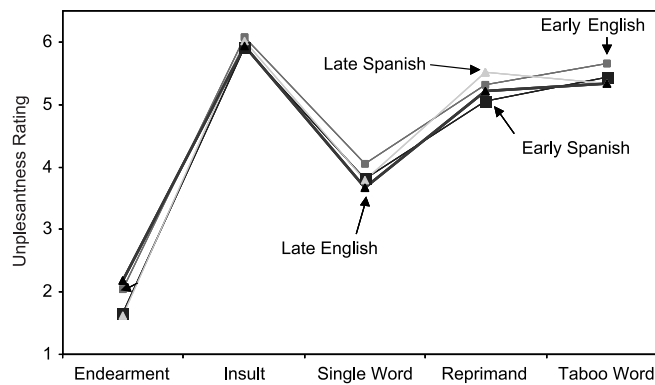


Figure 6 Unpleasantness ratings provided by the two groups of participants while they heard and read words and phrases.

negative and neutral words and thus ended up with a mean score at the midpoint of the seven-point unpleasantness scale.

ANOVA conducted on the ratings with independent variables, such as modality, learning group and stimulus type, revealed a significant language \times stimulus type interaction. Pairwise comparisons localised this difference to the case of endearments. Averaging over learning group, Spanish endearments were rated as more pleasant (i.e. less unpleasant) than English endearments, $F(1,35) = 7.6, p < 0.005$.

Pairwise comparisons were conducted between the English and Spanish ratings separately for the early and late learners. The only significant difference occurred for the late learners' ratings of reprimands. Reprimands were rated as more unpleasant in Spanish than in English $F(1,19) = 4.7,$

$p < 0.05$. Note that this is consistent with the findings for the electrodermal data: the late learners, but not the early learners of English, had heightened reactivity to reprimands.

Did individual participants' ratings of unpleasantness correlate with their skin conductance amplitudes within the categories of emotion words? Weak but statistically significant correlations were obtained for taboo items, across all participants and combining over languages, $r = -0.18$, $t(497) = -4.0$, $p < 0.001$. There was also a weak but significant correlation obtained for endearments, $r = 0.10$, $t(531) = 2.2$, $p < 0.03$. Correlations for the other categories were smaller than these and not significant.

Discussion

Taboo words in both languages elicited the largest skin conductance amplitudes. Surprisingly, insults elicited SCRs that were no higher than the single-word category (which functioned as the relatively nonemotional baseline). Endearments were also somewhat ineffective in eliciting a SCR. Reprimands were the stimulus type which elicited a different pattern between L1 and L2. It was with the reprimands also that a difference was found between the early and later learners of English. Late learners showed heightened reactivity to reprimands in Spanish, but not in English. The heightened SCRs to reprimands in L1 will thus be termed the 'reprimand effect'. The bilinguals who acquired English early (who were born in the USA or immigrated by age seven) did not show a reprimand effect. They responded similarly to reprimands in both languages, and responses were not significantly elevated about the single-word condition.

The current study confirms and extends prior findings about bilinguals' responses to one category of emotional language, reprimands. The Spanish-English bilinguals who learned English later in life, resemble the Turkish-English bilinguals, who acquired English after age 18 (Harris *et al.*, 2003). Like the late learners in the current study, the Turkish-English bilinguals had elevated skin conductance amplitudes for reprimands in L1 but not in L2. Together, the results of the current and the previous study indicate that one category of emotional expressions (reprimands) reliably elicits less autonomic arousal when items are presented in L2 which was not acquired until middle childhood or the teen years.

The linguistic specificity of the effect (i.e. only reprimands in L1 elicited heightened SCRs) is consistent with findings that adult bilinguals readily classify autobiographical memories as having been encoded in L1 or L2 (Schrauf & Rubin, 1998, 2000, 2003). However, reprimands do not invariably evoke heightened skin conductance amplitudes in L1. As noted previously, an English monolingual group did not respond to English reprimands, although they did respond to taboo words. The early bilinguals of the current study also did not respond to the reprimands, thus resembling their monolingual peers at Boston University. Considering the English monolingual data together with the data collected from the early learners of English raises the question of whether factors specific to growing up in the USA render reprimands innocuous, even when encountered in the native language. For example,

reprimands may not be as salient in North American child-rearing as they are in some other speech communities.

Both the early and late learners responded more strongly to taboo words presented in English (see Figures 4 and 5). The early learners may have responded more strongly to taboo words in English than in Spanish because English is their dominant language. If taboo words generally elicit stronger reactions in L1 than in L2, we would expect the late learners to have higher SCRs to taboo words in Spanish, mirroring the effects found with Turkish–English bilinguals in the earlier study (Figure 1). One interpretation of the late learners' data is that cultural factors may have influenced their responses to taboo words. During debriefing, participants frequently mentioned that taboo words are commonly used in colloquial Spanish, and consequently these items elicited little visceral charge. Several participants reported that taboo words 'sound harsher' in English. This could be the reason why both groups showed higher emotional responses to taboo words in English. Furthermore, English taboo words may be more relevant to the day-to-day life of all of the participants, since all of them are college students living in an urban North American city. The late learners (and possibly both groups) might demonstrate a greater response to Spanish taboo words after months of immersion in a Spanish-speaking peer group.

Aspects of the data, such as the modality effects, were novel and thus interpretation is difficult. However, given the heuristic value of the current work, some discussion would be useful. The previous study demonstrated that the auditory modality had an advantage (i.e. elicited larger SCRs) in the L1. In the current study, for early learners of English, auditory stimuli did elicit larger SCRs than visual stimuli, in both Spanish and English, with a trend for a greater advantage of auditory stimuli in English. This is consistent with the proposal that early learning of both Spanish and English created an auditory advantage in both languages. The slightly stronger auditory effect for English is consistent with English dominance, which is expected for the children in North America (Köpke, 2003).

This pattern of the early learners of English is thus readily interpretable. Not so the pattern for the late learners. One might expect the late learners would show an auditory advantage in their L1 Spanish, but not their L2 English, as this was the finding for the Turkish late learners. Instead, the late learners had no auditory advantage in either language. I propose the following interpretation, recognising that it is highly speculative. Identifying isolated words and phrases auditorily is difficult when items appear in one's L2 (Mayo *et al.*, 1997). When auditory discrimination is difficult, the visual modality will be relied on more, leading to heightened attention to visual stimuli. This may have created a general visual-reliance strategy for the late learners which influenced responses both to Spanish and English, thus negating what should have been an auditory advantage for Spanish. Recent work on memory for words presented either auditorily or visually indicates that modality can have strategic effects, and that auditory presentation enhances the distinctiveness of words (Smith & Hunt, 1998, 2000).

One can also ask why no language effects or age effects were found for the endearments and the insults. The most that can be said at present is simply

that these stimuli were ineffective in eliciting large SCRs. Endearments may not have elicited strong SCRs because electrodermal recording is mainly sensitive to threat and punishment (Fowles, 1980). Comments during debriefing suggested some reasons for the ineffectiveness of the insults. Participants commented that the insults sounded silly or could be readily laughed off. For example, the insults included 'You are so fat' and 'You are so stupid'. Although the goal had been to find statements that would be effective in insulting college students, some participants reported that these items were frequently used in light-hearted contexts and thus could have little noxious impact.

Note however that the insults were rated as highly unpleasant (Figure 6). If participants reported during debriefing that the insults felt innocuous, why did the insults receive the highest unpleasantness ratings of any stimuli? The participants presumably know that the insults literally refer to an unpleasant situation (e.g. being told that one is fat, being told that one is stupid). The insults are rated for unpleasantness by consulting knowledge of the meaning of the phrases. However, the electrodermal responsiveness is sensitive to whether the insults actually caused a visceral reaction. Here the postdebriefing comments are relevant. Possibly because they have become habituated to such phrases, college students find statements such as 'You are so stupid' unthreatening, and thus these stimuli did not elicit heightened SCRs.

This apparent dissociation between cognitive knowledge and autonomic response is interesting. Such dissociations have been studied in work on decision-making using a betting card game (Bechara *et al.*, 1997). Electrodermal activity was sensitive to the riskiness of choosing cards from a deck stacked with bad outcomes ('lose \$50') before players could consciously articulate which was the risky deck. In the current study, electrodermal activity was not sensitive to cognitive knowledge, but (apparently) only to stimuli such as taboo words and reprimands which carried deep-felt negative or socially stigmatised associations.

The proposal that cognitive knowledge can dissociate from autonomic response suggests a methodological improvement. Bilinguals could be asked to imagine a phrase spoken to them in an appropriate interpersonal situation. For example: 'Imagine the phrase "You are so fat" spoken in a context where it would be wounding. Then rate the emotional intensity of your feeling.' This rating task could be used to identify stimuli (in different languages) which would activate the autonomic nervous system. Substituting this task for the current task (rating phrases for unpleasantness) would likely also increase autonomic arousal.

It is well known that the same words and phrases carry different emotional connotations for different language communities (Harré, 1986; Kovecses, 1994; Wierzbicka, 1999). Some of our taboo words or reprimands could have appeared relatively neutral to some participants. This variability will likely create variability in electrodermal responsiveness, thus decreasing statistical power. Restricting participants to a specific culture, and choosing stimuli whose emotional resonances are known for that culture, would reduce variability and thus increase statistical power, but would limit the ability to generalise beyond the selected culture.

Conclusions

Early L2 learners of English show heightened SCRs to both Spanish and English taboo words, consistent with the proposal that when two languages are learned early, emotion-laden terms activate the autonomic nervous system equally. The data are inconsistent with a strong variant of the 'L1 is more emotional' thesis. Because the early learners reported greater proficiency in their L2, the data are consistent with the proposal that the L1 is more emotional when it is the more proficient language. Supporting this are the data from late learners, who demonstrated heightened SCRs to Spanish reprimands, but not English reprimands. This extends the prior finding of a reprimand effect in Turkish to a different language and culture. The linguistic specificity of the reprimand effect (that it is found only with the native language, not a second language) contributes to prior documentation of 'internal languages of retrieval' (Schrauf & Rubin, 1998, 2000, 2003). Measuring skin conductance appears a useful way to explore the extent to which linguistic phrases are mentally represented with the emotional associations that accompanied them during childhood learning.

Acknowledgements

Ayse Ayçiçeği collected the data on English monolinguals during our collaboration on the study of Turkish–English bilingual speakers, and supervised initial data collection for the current study. I thank her for this and for years of collaboration and support on myriad research projects. Dialma Miranda assisted with translation, selection and recording of both Spanish and English stimuli. Elena Isaacs, Karen Meersohn and Tabitha Pancharatnam ran many of the participants through the protocol. Bruce Mehler of Neurodyne Medical Corporation offered technical advice on interpreting the skin conductance data. Wayne Dinn assisted with electrodermal equipment and provided comments on an earlier version of this paper.

Correspondence

Any correspondence should be directed to Dr Catherine L. Harris, Psychology Department, Boston University, 64 Cummington St., Boston, MA 02215, USA (charris@bu.edu).

Notes

1. During debriefing participants maintained that they would not have been offended had the strongest taboo words been used. These have been recently used in psycholinguistic research by MacKay *et al.* (2002).
2. No literature exists on whether emotional responses are stronger or weaker to auditory stimuli which match or mismatch the accent of the listener.

References

- Altarriba, J. and Santiago-Rivera, A.L. (1994) Current perspectives on using linguistic and cultural factors in counseling the Hispanic client. *Professional Psychology: Research and Practice* 25, 388–397.

- Anooshian, J.L. and Hertel, T.P. (1994) Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion* 8, 503–514.
- Ayçiçeği, A. and Harris, C.L. (in press) Bilinguals' recall and recognition of emotion words. *Cognition and Emotion*.
- Barry, R.J. (1980) Electrodermal responses to emotive and non-emotive words as a function of personality differences in affect level. *Biological Psychology* 11, 161–168.
- Bechara, A., Damasio, H., Tranel, D. and Damasio, A.R. (1997) Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295.
- Ben-Shakhar, G. (1985) Standardization within individuals: A simple method of neutralizing individual differences in skin conductance. *Psychophysiology* 22, 292–299.
- Bingham, W.E. (1943) A study of the relations which the galvanic skin response and sensory reference bear to judgments of the meaningfulness and importance of 72 words. *Journal of Psychology* 16, 21–34.
- Birdsong, D. and Molis, M. (2001) On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language* 44, 235–249.
- Bond, M.H. and Lai, T. (1986) Embarrassment and code-switching into a second language. *Journal of Social Psychology* 126, 179–186.
- Boucsein, W. (1992) *Electrodermal Activity*. New York: Plenum.
- Channouf, A. and Rouibah, A. (1997) Suboptimal familiar faces exposure and electrodermal reactions. *Anuario de Psicología* 74, 85–97.
- Clachar, A. (1999) It's not just cognition: The effect of emotion on multiple-level discourse processing in second-language writing. *Language Sciences* 21, 31–60.
- Cohen, J.D., MacWhinney, B., Flatt, M.R. and Provost, J. (1993) PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers* 25, 257–271.
- Damasio, A. (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.
- Dawson, M.E., Schell, A.M. and Fillion, D.L. (2000) The electrodermal system. In J.T. Cacioppo and L.G. Tassinari (eds) *Handbook of Psychophysiology* (2nd edn) (pp. 200–223). Cambridge: Cambridge University Press.
- Dewaele, J.-M. and Pavlenko, A. (2002) Emotional vocabulary in interlanguage. *Language Learning* 52, 263–322.
- Dewaele, J.-M. and Pavlenko, A. (2003) 'I wasn't taught how to express anger': Emotion repertoires of multilingual speakers. Paper presented at the International Pragmatics Association (IPrA) conference, July 15, Toronto, Canada.
- Dinn, W. and Harris, C. (2000) Neurocognitive function in antisocial personality disorder. *Psychiatry Research* 97, 173–190.
- Ervin, S.M. (1964) Language and TAT content in bilinguals. *Journal of Abnormal and Social Psychology* 5, 500–507.
- Ferenczi, S. (1916) *Contributions to Psychoanalysis*. Boston: Badger.
- Fowles, D.C. (1980) The three arousal model: Implications of Gray's two-factor learning theory for heart rate, electrodermal activity, and psychopathy. *Psychophysiology* 17, 87–104.
- Gonzalez-Reigosa, F. (1976) The anxiety arousing effect of taboo words in bilinguals. In C.D. Spielberger and R. Diaz-Guerrero (eds) *Cross-cultural Anxiety* (pp. 89–105). Washington, DC: Hemisphere.
- Gray, S.G., Hughes, H.H. and Schneider, L.J. (1982) Physiological responsivity to a socially stressful situation: The effect of level of moral development. *Psychological Record* 32, 29–34.
- Greenson, R. (1950) The mother tongue and the mother. *International Journal of Psychiatry* 31, 18–23.
- Grings, W.W. and Zeiner, A. (1965) Autonomic responses to words modified by sensitizing and conditioning experiences. *Journal of Psychosomatic Research* 8, 373–378.

- Gumperz, J.J. and Hernandez, E. (1971) *Bilingualism, Bidialectalism, and Classroom Interaction in Language and Social Groups*. Palo Alto: Stanford University Press.
- Harré, R. (1986) *The Social Construction of Emotion*. Oxford: Basil Blackwell.
- Harris, C.L., Ayçiçeği, A. and Gleason, B.J. (2003) Taboo words and reprimands elicit greater autonomic reactivity in a first than in a second language. *Applied Psycholinguistics* 4, 561–578.
- Hugdahl, K. (1995) *Psychophysiology: The Mind–body Perspective*. Cambridge, MA: Harvard University Press.
- Javier, R.A. (1989) Linguistic considerations in the treatment of bilinguals. *Journal of Psychoanalytic Psychology* 6, 87–96.
- Javier, R.A., Barroso, F. and Muñoz, M.A. (1993) Autobiographical memory in bilinguals. *Journal of Psycholinguistic Research* 22, 319–338.
- Javier, R.A. and Marcos, L.R. (1989) The role of stress on the language-independence and code-switching phenomena. *Journal of Psycholinguistic Research* 18, 449–472.
- Kim, M. (1993) The effects of emotional connotations of English words on event-related potentials (ERPs) in Korean–English bilinguals and English-speaking monolinguals. Unpublished doctoral dissertation, University of Georgia, USA (available from University Microfilm).
- Kim, K.H.S., Relkin, N.R., Lee, K-M. and Hirsh, J. (1997) Distinct cortical areas associated with native and second languages. *Nature* 388, 171–174.
- Kluender, R. and Kutas, M. (1993) Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5, 196–214.
- Köpke, B. (2003) Neurolinguistic aspects of attrition. *Journal of Neurolinguistics* 17, 3–30.
- Kovecses, Z. (1994) *Emotion Concepts*. New York: Springer-Verlag.
- Koven, M. (2001) Comparing bilinguals' quoted performances of self and others in tellings of the same experience in two languages. *Language in Society* 30, 513–558.
- Koven, M. (2003) *French–Portuguese bilinguals' affective performances in two languages*. Paper presented at the International Pragmatics Association (IPrA) conference, July 15, Toronto, Canada.
- Kutas, M. and King, J.W. (1996) The potentials for basic sentence processing: Differentiating integrative processes. In I. Ikeda and J.L. McClelland (eds) *Attention and Performance* (Vol.16) (pp. 501–546). Cambridge, MA: MIT Press.
- LaBar, K.S. and Phelps, E.A. (1998) Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science* 9, 490–493.
- Lykken, D.T. and Venables, P. (1971) Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology* 8, 656–672.
- MacKay, D.G., Shafto, M., Taylor, J.K., Marian, D.E., Abrams, L. and Dyer, J. (2002) *The taboo Stroop and related effects of emotion on memory and attention*. Presented at the Annual Meeting of the Psychonomics Society, Kansas City, MO.
- Manning, S.K. and Melchiori, M.P. (1974) Words that upset urban college students: Measures with GSRs and rating scales. *Journal of Social Psychology* 94, 305–306.
- Mathews, A. and MacLeod, C. (1985) Discrimination of threat cues without awareness in anxiety states. *Journal of Abnormal Psychology* 95, 131–138.
- Mathews, A., Richards, A. and Eysenck, M. (1989) Interpretation of homophones related to threat in anxiety states. *Journal of Abnormal Psychology* 98, 31–34.
- Mayo, L.H., Florentine, M. and Buss, S. (1997) Age of second-language acquisition and perception of speech in noise. *Journal of Speech and Hearing Research* 40, 686–693.
- McGinnies, E. (1949) Emotionality and perceptual defense. *Psychological Review* 56, 244–251.
- O'Regan, K.J. and Noe, A. (2001) A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24, 939–973.
- Pavlenko A. (1998) Second language learning by adults: Testimonies of bilingual writers. *Issues in Applied Linguistics* 9, 3–19.
- Pavlenko, A. (1999) New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition* 2, 209–230.
- Pavlenko, A. (2002) Bilingualism and emotions. *Multilingua* 21, 45–78.

- Rubin, D.C. and Friendly, M. (1986) Predicting which words get recalled: Measures of free recall, availability, goodness, emotionality, and pronunciability for 925 nouns. *Memory and Cognition* 14, 79–94.
- Schrauf, R.W. (2000) Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture and Psychology* 6, 387–417.
- Schrauf, R.W. and Rubin, D.C. (1998) Bilingual autobiographical memory in older adult immigrants: A test of cognitive explanations of the reminiscence bump and the linguistic encoding of memories. *Journal of Memory and Language* 39, 437–457.
- Schrauf, R.W. and Rubin, D.C. (2000) Internal languages of retrieval: The bilingual encoding of memories for the personal past. *Memory & Cognition* 28, 616–623.
- Schrauf, R.W. and Rubin, D.C. (2003) On the bilingual's two sets of memories. In R. Fivush and C.A. Haden (eds) *Autobiographical Memory and the Construction of the Narrative Self: Developmental and Cultural Perspectives* (pp. 121–145). Mahwah, NJ: Erlbaum.
- Schumann, J. (1997) *The Neurobiology of Affect in Language Learning*. Boston: Blackwell.
- Smith, R.E. and Hunt, R.R. (1998) Presentation modality affects false memory. *Psychonomic Bulletin & Review* 4, 710–715.
- Smith, R.E. and Hunt, R.R. (2000) The effects of distinctive processing on retrieval induced forgetting. *Memory & Cognition* 28, 503–508.
- Stelmack, R.M., Plouffe, L. and Falkenberg, W. (1983a) Extraversion, sensation seeking and electrodermal response: Probing a paradox. *Personality and Individual Differences* 4, 607–614.
- Stelmack, R.M., Plouffe, L. and Winogron, W. (1983b) Recognition memory and the orienting response: An analysis of the encoding of pictures and words. *Biological Psychology* 16, 49–63.
- Toglia, M.P. and Battig, W.F. (1978) *Handbook of Semantic Word Norms*. Hillsdale, NJ: Erlbaum.
- Tranel, D., Fowles, D.C. and Damasio, A.R. (1985) Electrodermal discrimination of familiar and unfamiliar faces: A methodology. *Psychophysiology* 22, 403–408.
- Vanderploeg, R.D., Brown, W.S. and Marsh, J.T. (1987) Judgments of emotion in words and faces: ERP correlates. *International Journal of Psychophysiology* 5, 193–205.
- Wierzbicka, A. (1999) *Emotions Across Languages and Cultures: Diversity and Universals*. Cambridge: Cambridge University Press.

Appendix: English and Spanish Stimuli

Spanish and English items were direct translations except where noted.

<i>Reprimands</i>	
Shut up!	<i>¡Cállate!</i>
Now you're in trouble!	<i>¡Sabes que me lo vas a pagar!</i> (You're going to pay for that)
Shame on you!	<i>¡Avergüénzate!</i>
Stop that!	<i>¡Deja eso!</i> (Leave that)
Don't do that!	<i>¡No hagas eso!</i>
No hitting!	<i>¡Sabes que estas caliente conmigo!</i> (You're hot with me, more intense in Spanish)
Go to your room!	<i>¡Vete a tu cuarto!</i>

That's not nice!	<i>¡Estas castigado!</i> (You are punished)
Insults	
Are you crazy?	<i>¿Eres loca?</i>
I never want to see you again!	<i>¡Nunca más te hablaré!</i> (I will never talk to you again)
You are such a loser!	<i>¡Que bruta eres!</i> (You're so stupid)
You suck!	<i>¡Idiota!</i> (You idiot)
What a moron!	<i>¡Canto de estúpida!</i> (Piece of stupid person)
You are so fat!	<i>¡Qué gorda eres!</i>
You disgust me!	<i>¡No quiero verte la cara!</i> (I don't want to see your face)
I hate you!	<i>¡Te odio!</i>
Endearments	
I love you more than anything!	<i>¡Te amo!</i>
You are everything to me!	<i>¡Eres mi vida!</i> (You are my life)
I've missed you so much!	<i>¡Me haces falta!</i>
When will I see you again?	<i>¡Siempre te recordaré!</i> (I will always remember you!)
I would die for you!	<i>¡Sin ti no puedo vivir!</i> (I can't live without you)
I can't wait to see you!	<i>¡Me alegra verte!</i> (It makes me happy to see you)
Hey, sweetie!	<i>¿Mi corazón, cómo estás?</i> (My heart, how are you)
You are so beautiful!	<i>¡Eres tan bella!</i>
Taboo and sexual terms	
asshole	<i>cabrón</i>
pussy	<i>críca</i>

dick	<i>maricón</i> (slang term: gay)
pee	<i>mear</i>
shit	<i>mierda</i>
bitch	<i>puta</i>
breast	<i>seno</i>
raped	<i>violada</i>
<i>Single word category, sorted by neutral, aversive and positive words</i>	
<i>Neutral (n =12)</i>	
box	<i>caja</i>
street	<i>calle</i>
column	<i>columna</i>
finger	<i>dedo</i>
table	<i>mesa</i>
name	<i>nombre</i>
number	<i>numero</i>
part	<i>parte</i>
door	<i>puerta</i>
branch	<i>rama</i>
chair	<i>silla</i>
envelope	<i>sobre</i>
<i>Aversive (n =12)</i>	
murder	<i>asesinato</i>
anger	<i>cólera</i>
crime	<i>crimen</i>
cruel	<i>crueidad</i>
pain	<i>dolor</i>
disease	<i>enfermedad</i>

war	<i>guerra</i>
kill	<i>matar</i>
death	<i>muerte</i>
fight	<i>pelea</i>
danger	<i>peligro</i>
grave	<i>tumba</i>
<i>Positive (n = 15)</i>	
joy	<i>alegría</i>
friend	<i>amigo</i>
love	<i>amor</i>
kiss	<i>beso</i>
home	<i>casa</i>
happy	<i>feliz</i>
freedom	<i>libertad</i>
mother	<i>mamá</i>
honey	<i>miel</i>
father	<i>papá</i>
laugh	<i>risa</i>
smile	<i>sonrisa</i>