Applied Psycholinguistics **24** (2003), 561–579
Printed in the United States of America
DOI: 10.1017.S0142716403000286

Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language

CATHERINE L. HARRIS
Boston University

AYŞE AYÇIÇEĞI
Istanbul University

JEAN BERKO GLEASON
Boston University

ADDRESS FOR CORRESPONDENCE
Catherine L. Harris, Psychology Department, Boston University, 64 Cummington St., Boston, MA 02215. E-mail: charris@bu.edu

ABSTRACT
Second language speakers commonly acknowledge that taboo terms can be uttered with greater ease in their second language (L2) than in their first language (L1). To investigate this phenomenon psychophysiologically, 32 Turkish–English bilinguals rated a variety of stimuli for pleasantness in Turkish (L1) and English (L2) while skin conductance was monitored via fingertip electrodes. Participants demonstrated greater autonomic arousal to taboo words and childhood reprimands ("*Shame on you!*") in their L1 compared to their L2. This finding provides quantifiable support for the subjective experiences of L2 speakers.

Bilingual speakers have reported that it is easier to say some kinds of emotionally charged words or expressions in a language that is not their native language. In this study, we explored the possibility that this difference between a first language (L1) and a second language (L2) has psychophysiological components. In particular, we hypothesized that taboo words in a native language would elicit stronger skin conductance responses than similar taboo words learned later in life in an L2. We also hypothesized that emotional expressions learned early in life in an L1 (such as the kinds of reprimands young children hear) would evoke greater physiological responses, whether the individuals experiencing these responses were fully aware of them or not.

A number of authors have remarked that taboo words appear to generate less anxiety when spoken in a foreign language (Ferenczi, 1916; Greenson, 1950; Javier, 1989). Bilingual speakers also feel freer to discuss embarrassing topics

in their L2 (Bond & Lai, 1986; Sechrest, Flores, & Arellano, 1968). Anooshian and Hertel (1994) noted that age of acquisition of a language appears to be more important than proficiency for at least some cases of how the emotional resonance of a language is experienced. They cite the example of a woman who grew up in a Spanish-speaking home and learned English after 8 years of age. Although as an adult she was most proficient in English, she prayed in Spanish, because praying in English never "felt right." Studies of code-switching report that the L2 is often used to increase emotional distance (Altarriba & Rivera–Santiago, 1994; Bond & Lai, 1986; Gonzalez–Reigosa, 1976; Gumperz & Hernandez, 1971; Javier & Marcos, 1989). In contrast, the native language is more often the language of emotional expressiveness (Dewaele & Pavlenko, 2002). For example, Sechrest et al. (1968) reported that married Filipinos used Tagalog for intimate expression, even though English was habitually spoken at home.

One way to measure the emotional impact of words is through their effect on autonomic reactivity. Electrodermal monitoring (the psychophysiological technique that records skin conductance responses, formerly referred to as galvanic skin response) has long been used to measure differential responsiveness to words. Studies have shown that among monolinguals, emotion and taboo words elicit greater skin conductance responses compared to neutral words (Gray, Hughes, & Schneider, 1982; Dinn & Harris, 2000; McGinnies, 1949). Electrodermal recording has not been previously used to study differential reactivity in the languages spoken by bilingual individuals.

The goal of the current study was to determine if electrodermal recording would demonstrate that words with equivalent semantic content in an L1 and L2 generate different degrees of autonomic reactivity. Relevant research literature includes studies of taboo words and emotion terms in monolinguals and bilinguals. Taboo words are known to be the most emotionally evocative of language stimuli (Gray, Hughes, & Schneider, 1982; Jay, 1992; Matthews & MacLeod, 1985; McGinnies, 1949; Nothman, 1962; Van Lancker & Cummings, 1999). Researchers have used taboo-word presentation to study the working of the brain's emotional systems, because processing of taboo words is believed to activate the amygdala, a structure in the limbic system involved in emotion processing and emotion-mediated learning (Labar & Phelps, 1998). MacKay et al. (2002) used taboo words in a Stroop task. Taboo words were displayed in a salient color, the task being to name the ink color and ignore the word. Participants were slower to name the color of the taboo words compared to the color of the neutral words. This "Stroop interference effect" presumably occurs because the anxiety associated with taboo words (or people's understanding of their status as socially stigmatized words) interferes with attending to the color-naming task. The Stroop interference effect was much stronger than the interference effect found for neutral words, and it persisted for over 100 trials.

Aversive words (such as *cancer, death, stabbed*) resemble taboo words in that they produce orienting and attentional effects (Hill & Kemp–Wheeler, 1989; Wischner & Gladis, 1969), although the magnitude is typically less than that of taboo words or words of strong personal significance (Crane, Dieker, & Brown, 1970). Compared to neutral words, aversive words are easier to identify as

words in a lexical decision task (Hill & Kemp–Wheeler, 1989). Both clinical and nonclinical subjects were better able to form associations to aversive words than to neutral words (Wischner & Gladis, 1969).

Comparing emotional responsiveness to words in an L1 versus an L2 has been examined in relatively few laboratory studies, and only one used taboo words. Gonzalez–Regiosa (1976) asked Spanish–English bilinguals to read lists of 10 Spanish taboo words and 10 English taboo words. Following the reading of each list, they rated themselves on their degree of anxiety using a questionnaire sensitive to "state anxiety" (a transient state). Participants also completed a questionnaire designed to assess their level of "trait anxiety" (i.e., whether they had an anxiety-prone personality). Participants rated themselves as more anxious after reading the taboo words in their L1. This result held for both high- and low-anxiety participants.

Anooshian and Hertel (1994) examined differential emotional connotations in an L1 versus an L2 using a recall test. Recall of words is known to be influenced by emotionality (Rubin & Friendly, 1986). Anooshian and Hertel (1994) hypothesized that emotion words in the L2 would lack the emotional connotations that render words easy to recall. Participants first rated blocks of words for emotional intensity, pronounceability, and how much activity was inherent in the meaning of the word. After the rating task, participants received a surprise recall task. Emotionality effects were strongest for words when they had been rated for emotionality. More emotion words than neutral words were recalled in the L1, and an equivalent number of emotion and neutral words were recalled in the L2. This supports the authors' hypothesis that only words in the L1 would show an emotionality effect. This supports the thesis of the current paper, which is that words in an L1 have greater emotional resonance than words in an L2. However, in an extension and replication of Anooshian and Hertel's (1994) paradigm, two of the current authors found greater recall for emotion words in both the L1 and L2 (Ayçiçeği & Harris, in press). Indeed, a goal of the current paper was to determine if a psychophysiological measure would be a useful methodology for quantifying intuitions about the greater emotionality of words heard or read in the L1.

Although no prior study has employed electrodermal recording to gauge emotional responsiveness, one study used event-related potentials (ERPs). In this psychophysiological technique, electrodes fixed to the scalp measure brain wave activity corresponding to stimulus presentation. It was interesting that the underlying rationale and design for this study was similar to our own: would emotional responsiveness vary for words in an L1 and an L2? Kim (1993) used P300 amplitude as her 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, because prior studies documenting the sensitivity of the P300 ampli-

tude to emotional stimuli used highly evocative stimuli such as pictures and slang expressions (Vanderploeg, Brown, & Marsh, 1987).

Kim's (1993) study motivates the current project by describing an area in which electrodermal monitoring may be a superior psychophysiological technology in comparison to ERPs. ERPs are sensitive to millisecond changes in brain electrical activity and change predictability in response to both lexical and grammatical factors (Weber–Fox & Neville, 1999). However, Kim (1993) found that ERP waveforms did not vary by emotional valence for monolingual or bilingual participants and so concluded that ERPs are not sensitive to the emotional valences of words. Skin conductance responses are known to be sensitive to differences in the same types of words Kim used in her ERP study (e.g., Dinn & Harris, 2000). Words' emotionality may be a case where electrodermal recording is preferred to ERPs, despite the latter's superior temporal resolution.

Recent research has begun to elucidate hypotheses for why emotional resonance may differ between two languages (Bond & Lai, 1986; Dewaele & Pavlenko, 2002; Dewaele & Regan, 2001; Durst, 2001; Pavlenko, 1999, 2002; Schrauf, 2000; Wierzbicka, 1999). An L1 and L2 are frequently learned in different contexts. The L1 is learned in the context of family life and thus generally includes emotional extremes and usually also an individual's earliest encounters with the gamut of human emotions (Schrauf, 2000). An L2 is frequently the language of schooling, work, and professional achievement and thus comes to be associated with emotional control, autonomy, and achievement (Bond & Lai, 1986; Dewaele & Pavlenko, 2002; Ervin, 1964). Bond and Lai (1986) reasoned that, because the L2 is usually mastered in more emotionally neutral settings than the L1, less arousal would be conditioned to L2 words. These authors predicted that bilingual interviewees would be more comfortable discussing embarrassing topics in their L2. Indeed, it was found that the interviewees spoke at greater length about embarrassing topics (but not neutral topics) when they were instructed to respond in their L2.

Differential emotional context of language acquisition appears to be a sufficient explanation for greater emotionality of words and expressions in the L1. However, the implications for taboo words are not obvious. Historically, the anxiety-arousing effect of taboo words has been attributed to punishment from parents for saying these words during childhood (Ferenczi, 1916). This seems an overly simplistic view from a contemporary standpoint. In some households and some cultures, children acquire taboo words outside the home and learn their stigmatized status from peers. The broader concept of societal disapproval may be more important than punishment by parents. Taboo words learned in an L2 may actually be learned in a similar environment to taboo words in an L1 (i.e., both learned from peer or street culture) and thus may have similar emotional associations.

According to this view, elecrodermal recording could reveal high reactivity for taboo words in both an L1 and an L2, given that reasonably proficient L2 learners will be aware of the stigmatized nature of these words. It thus seemed reasonable to identify emotional expressions that resemble taboo words in being associated with personal threat but are unambiguously acquired in early childhood as part of daily family life. Inspired by prior work on language routines

and childhood language socialization (Gleason, 1985), we selected the category of childhood reprimands, including expressions like *Go to your room!* and *Shame on you!*

The current study is the first attempt to examine autonomic reactivity to emotional expressions in the L1 and L2. We thus felt it was acceptable to include an exploratory variable, the variable of presentation modality. Comparing visual versus auditory presentation can be useful for future research for two reasons. Greater skin conductance responses might occur for auditory stimuli because the spoken language likely has more diverse and more numerous emotional associations than written language. Spoken words and phrases may activate specific memories. Spoken language is acquired before visual language (for L1 acquisition). To the extent that linguistic representations that are learned early become connected with emotional regulation systems (Bloom & Beckwith, 1989), auditory language may be more closely tied to emotional arousal than visual language. We also varied the modality of presentation for exploratory methodological purposes. For investigating languages that do not use the Latin alphabet, it could be convenient to restrict data collection to the auditory realm. For languages that use the Latin alphabet, typing in a list of words for visual presentation is less labor intensive than recording and digitizing auditory stimuli, thus reducing time to pilot and test a hypothesis.

METHOD

Bilingual speakers of quite dissimilar languages (English and Turkish) were recruited for this study. The reactivity to comparable words and phrases in their L1 (Turkish) and their L2 (English) was monitored for each participant.

Participants

Turkish–English bilinguals were chosen for this study for several reasons. Although the largest bilingual population residing in Boston undoubtedly consists of Spanish speakers, Spanish–English bilinguals have very varied language-learning histories, including being born in the United States or immigrating at diverse ages. Turks residing in Boston most typically arrived in their 20s to pursue graduate school or professional work. They thus have a good command of English but relatively late acquisition and late immersion in an English-speaking country. This allows us to study a relatively homogenous sample of late learners of English. A second reason was our assumption that Turkish would share few cognates with English. Electrodermal results could vary if words in English activate cognates in Turkish.

The participants were 32 native Turkish speakers (15 males, 17 females) who were students or professionals working in the Boston area. All were late learners of English, having acquired English after 12 years of age. The average age of arrival in the United States was 24 (range = 16–31 years), and the mean length of residence was 4 years. The first in-depth exposure to English for most participants was enrollment in an English-language high school in Turkey at age 12 (13 participants) or enrollment at an English-language university in Turkey at

Table 1. *Participants' demographic variables and fluency*

Mean (Range) of Demographic Variables		Fluency Test and Self-Ratings	Mean and Range	
			Turkish	English
Age	28 (20–47)	Word fluency test (total words produced)	37 (17–56)	29 (17–43)
Years education	18 (12–20)	Spoken (conversation)	5	3.15 (2–4)
Age of immersion	16 (12–31)	Understanding	5	3.35 (2–4)
Age of arrival in United States	24 (16–31)	Reading	5	3.6 (2–4)
		Writing	5	2.55 (1–4)
Years of residence	4 (1–15)			

Note: The age of immersion indicates enrollment in English-language high school, university or arrival in the United States. The word fluency test is the sum of the words produced to letters F, A, and S; 1 min was provided for each letter. Self-ratings on the 1–5 scale indicate poor, fair, good, very good, or native-speaker ability.

age 18 (16 participants). For the remaining three participants, immersion in English began with their arrival in the United States at age 17, 20, or 31, with prior exposure to English consisting of formal classroom instruction in high school or university in Istanbul. The average years of education was 18, which indicates that most participants had completed their undergraduate degree and had 1 or more years of graduate school. Participants rated their own English ability in conversational fluency, reading, understanding, and writing on a 1–5 scale. Numbers on the scale correspond to poor, fair, good, very good, or native ability. Most respondents judged their comprehension of English to be good. Consistent with the preponderance of students and professionals in our sample, average self-rated reading ability was slightly higher than conversational ability. Following Anooshian and Hertel (1994), we also assessed fluency by administering a word fluency task. Participants were given 1 min to produce as many words as possible beginning with a specific letter. Three trials were conducted, using the letters F, A, and S. The score for the word fluency task is the sum of words produced across the three trials (Goodglass & Kaplan, 1972). The number of words generated in English was 29 (range = 17–43) and in Turkish it was 37 (range = 17–56), a statistically significant difference, t (31) = 5.4, $p < .001$. This difference supports bilinguals' reports that they were more fluent in Turkish than English. Means and ranges for demographic variables appear in Table 1.

Design and materials

Participants read on a computer screen or heard via the computer loudspeaker a variety of word types in Turkish (L1) and English (L2) while their skin conductance response was monitored via fingertip electrodes. A $2 \times 2 \times 5$ within-subjects design was used with a Latin-squares crossing: two levels of language, two levels of modality (auditory vs. visual presentation), and five categories of

stimuli: 16 neutral (*door*), 16 positive (*bride, joy*), 16 aversive (*disease, kill*), 9 taboo (*asshole, breast*), and 7 reprimands of the type commonly spoken to children (*Don't do that!* and *Go to your room!*). All factors were counterbalanced across participants. Although participants encountered stimuli in all conditions of the design, they did not see (or hear) the same word in both languages. All stimuli appear in Appendix A.

The positive, aversive, and neutral words were selected and coded from the *Handbook of Semantic Word Norms* (Toglia & Battig, 1978) using the pleasantness scale, which ranges from 1 to 5. Positive words had pleasantness ratings of 3.5 or higher, aversive words had ratings of 2.0 or lower, and neutral words had ratings from 2.5 to 3.4. These three categories of items were selected to have comparable familiarity, as measured by Toglia and Battig's ratings. These items had been used in a previous electrodermal study (Dinn & Harris, 2000, 2003). To select taboo words, we modified the list used by Gonzalez–Regiosa (1976), deleting items that might not be known to the Turkish speakers (*hymen, tampon*). Items were translated into their Turkish equivalents by the second author (A.A.) and verified by a second native speaker of Turkish. Translation equivalents did not exist for all the reprimands or taboo words; thus, we substituted items that had similar meanings and emotional connotations, based on a list of suggestions made by the second author and three faculty members at Istanbul University. For taboo words, substitutions reflected the judgments of these four native Turkish speakers that the taboo word had a sexual connotation or the status of a stigmatized word. Two sexual terms in Turkish are cognates of English words: *seks* (sex) and *masturubasyon* (masturbation). We used the latter in Turkish and did not employ its English cognate. The reprimands were selected to be ones that parents would frequently say to children, although many of these are used in adult–adult contexts as well. When reprimands were displayed visually, they appeared with initial capital letter and final exclamation point (*Shut up!*). This contrasted with other stimuli, which appeared in lower case with no punctuation. Our reason for this visual difference was to use the visual medium to convey the unit of the expression and its emotionality, so that any differences between auditory and visual presentation could be attributed to specifically auditory aspects. This was a concern because, in pilot work, we observed that lower case multiword utterances (e.g., *Shame on you!*) were not always immediately understood to be coherent expressions with emotional force.

After items were selected, we obtained word frequency measures for English stimuli (using the 1982 Francis and Kucera word frequency database) and Turkish stimuli, using Goz' recently compiled listing of Turkish word frequencies (Goz, in press). As listed in Table 2, Turkish words had slightly higher word frequency values than did English items. The positive, negative, and neutral English words had a mean frequency of 127 occurrences per million words and did not differ from each other in word frequency, as measured using the Francis and Kucera (1982) database. The mean frequency for taboo words was substantially less, 4 occurrences per million words. Using Turkish word frequency counts (Goz, in press), a similar pattern was found: taboo words were much lower in frequency than the positive, negative, and neutral words, which did not differ from each other. When comparing frequencies across languages, differ-

Table 2. *Stimulus ratings and response latency*

	Word Frequency		Response Latency		Unpleasant		English Familiarity	
	English	Turkish	English	Turkish	English	Turkish	Current Subjects	Native Speakers
Reprimands	—	—	3.9	3.8	5.4	5.5	4.4*	4.9
Taboo	4	19	3.8	3.9	4.4	4.7	3.4*	4.8
Aversive	89	168	3.4	3.1	5.8	6.2	4.6	4.9
Positive	114*	467	3.0	2.8	1.8	1.8	5.9	6.0
Neutral	175*	448	3.3	3.0	3.2	3.6	5.8	5.7

Note: The word frequency is the number of occurrences in a corpora of 1 million words. The response latency is the number of seconds to rate each item on a 1–7 scale for unpleasantness. Current participants only rated English stimuli for familiarity (1–7 scale). The native speakers were 40 Boston University undergraduates who rated the stimuli as part of a separate investigation.
*Significant difference $p < .05$, between English and Turkish values. (Log frequency was used when conducting the t tests on word frequency.)

ences did emerge (see means in Table 2). Turkish translations of the English words had higher frequency counts. The mean occurrence per million words was 297 for Turkish items and 147 for English items, a significant difference, $t\ (112) = 2.1$, $p < .04$. Because the methodology of database construction likely differs between the Turkish and English databases, it would be premature to draw conclusions at this point. However, the frequency difference could mean that English stimuli were both objectively lower in frequency and probably also subjectively less frequent to our participants, because of their fewer number of years speaking English.

All items were recorded and digitized for computerized presentation. Items were recorded by a female native speaker of each language. For most items, a neutral tone was employed. For reprimands, a lightly admonishing tone was used, appropriate to the meaning of the reprimand.

Apparatus and procedure

Stimuli were presented on a Power Macintosh G3 using PsyScope, which is experimental control software developed by Cohen, MacWhinney, Flatt, and Provost (1993). Participants rated each word for pleasantness on a 1–7 scale (7 = *maximally unpleasant*) by typing the corresponding key on a standard keyboard using their nondominant hand. Electrodermal activity was recorded using the Davicon C2A Custom Skin Conductance Monitor (NeuroDyne Medical Corporation). Electrodes were attached to the tip of the first and second fingers of the subject's dominant hand. A 10-s recording interval began coincident with stimulus onset. Following Hugdahl (1995), we calculated the amplitude of the phasic skin conductance response by subtracting the basepoint from the maximum

score. This amplitude was then divided by the basepoint in order to reduce variation caused by individual differences in baseline skin conductance levels. The resulting value in micromhos was our dependent measure; we will refer to this as the skin conductance response (SCR).

Taboo words, reprimands, and aversive words were distributed throughout the presentation list to be maximally distant from other items in their category. Each participant was presented with 64 intermixed trials. This meant that participants could predict neither the language nor the modality of the next item. The experimenter took notes on trials when artifacts to the electrodermal record may have been introduced (e.g., participants' talking or sneezing). Less than 2% of the data were discarded because of artifacts.

To determine if familiarity with words in a second language influences skin conductance responses, participants rated the English words for familiarity, using the 7-point scale and instructions used by Toglia and Battig (1978). A subset of 22 participants performed the rating task. Eight participants completed the rating task immediately after the electrodermal recording session and 14 completed the rating task on a subsequent day. Forty monolingual native speakers of English also performed the rating task as part of a separate study. We included the native-speaker familiarity ratings in Table 2 as a comparison.

RESULTS

An analysis of variance (ANOVA) was conducted on the three-factor, within-subject design. Significant main effects were obtained for language, $F\ (1, 31) = 10.77$, $p < .002$, and for stimulus type, $F\ (4, 124) = 11.56$, $p < .001$, indicating that SCR amplitudes were higher in Turkish than in English and that amplitudes differed for word types. The omnibus ANOVA also revealed an interaction between stimulus type and language, $F\ (4, 124) = 2.64$, $p < .05$, and between language and modality, $F\ (1, 31) = 4.6$, $p < .05$, but no three-way interaction ($p > .3$). The significant two-way interactions are graphed in Figures 1 and 2.

Planned comparisons were conducted between Turkish and English for each of the stimulus types. The largest difference between languages (and the only statistically significant pairwise comparison) occurred for reprimands, $F\ (1, 31) = 10.78$, $p < .005$, with reprimands in Turkish eliciting stronger SCRs than reprimands in English. This difference held for both auditory and visual presentation. Taboo words elicited larger skin conductance responses in Turkish than in English, but statistical significance was obtained only when the analysis was restricted to the auditory modality, $F\ (1, 31) = 3.88$, $p < .05$. Responses to taboo words displayed in the visual modality did not differ between the two languages.

Table 2 presents mean unpleasantness ratings, latency to rate words for unpleasantness, participants' ratings for familiarity of English words and monolingual English familiarity ratings, and mean word frequency values (from the Kucera and Francis, 1982, corpus of 1 million English words, and the Goz, in press, corpus of 1 million Turkish words). The time to rate the reprimands and taboo words was substantially longer (3.8 s) than rating times for other stimulus categories (3.1 s), as indicated by the main effect of stimulus type in the ANOVA performed on response latencies, $F\ (4, 124) = 18.2$, $p < .001$. This

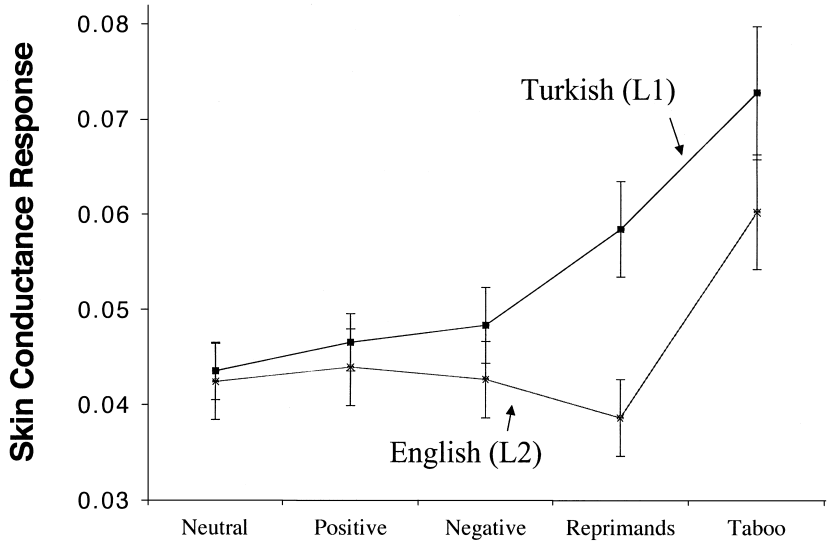

Figure 1. The phasic skin conductance response (μmhos) is shown for the five categories of stimuli. The phasic skin conductance response is obtained by subtracting the minimum from the maximum score and dividing by the base point.

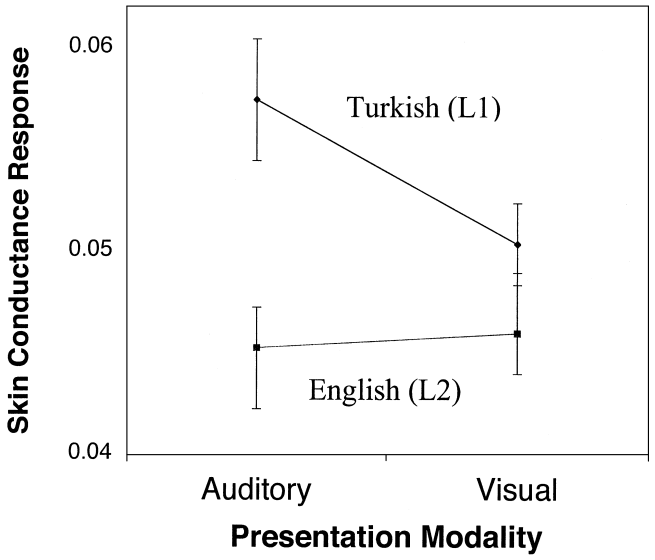

Figure 2. Averaging over stimulus categories, skin conductance responses are plotted separately for auditory and visual presentation.

presumably reflects the fact that the neutral, positive, and aversive words were selected for having specific pleasantness ratings (in English; Toglia & Battig, 1978) and thus could be more easily rated than the taboo words and reprimands. It is important that the time to rate items was similar in both languages (3.5 s in English and 3.4 s in Turkish). Participants rated the aversive, neutral, and positive words in a manner consistent with the Toglia and Battig pleasantness values (i.e., aversive words were rated as more unpleasant than neutral words and neutral words as less pleasant than positive words). As shown in Table 2, our Turkish participants provided very similar ratings for English and Turkish words.

Participants rated English taboo items as the least familiar of all the categories. Taboo items were rated as significantly less familiar than the next most unfamiliar category, the reprimands, $F\ (1, 21) = 7.0$, $p < .02$. The aversive words were rated as less familiar than the neutral words, $F\ (1, 21) = 19.5$, $p < .001$. Turkish participants did rate English taboo words and reprimands as less familiar than did Boston University monolingual undergraduates. It is interesting that the Turkish participants judged positive, negative, and aversive words as moderately to highly familiar, and gave them similar ratings to those of monolingual undergraduates. This indicates that Turkish participants had sufficient length of stay (or sufficient immersion in the English-speaking world) for basic vocabulary items to come to have a similar subjective familiarity to that experienced by the monolingual undergraduates.

In an exploratory analysis, we investigated whether the length of the reprimands influenced skin conductance amplitudes, and whether individual participants' rating was correlated with skin conductance responses. We compared short reprimands (*Stop that!, Shut up!, No!*) to long reprimands (*Shame on you!, Go to your room!, I hate you!, Don't do that!*). The Turkish counterparts to these also had the same length categorization. In Turkish, long auditory reprimands had higher SCRs than short auditory reprimands, $t\ (54) = 2.0$, $p < .05$. Skin conductance amplitudes were similar for long and short reprimands in the visual modality, and no differences were apparent in either modality for English. Long reprimands may have activated stored memories, which increased autonomic responding.

Correlations were conducted on participants' individual pleasantness ratings and their SCRs. It seemed likely that personal ratings of unpleasantness could predict SCRs for the taboo words and reprimands, because it is plausible that individuals differ in whether they find specific items threatening or socially stigmatized. This expectation was partially confirmed. No significant correlations were obtained for the positive, negative, and neutral items ($r = .21$ to $-.07$) or for the taboo items ($r = .19$). However, auditory reprimands manifested a modest correlation with individual unpleasantness ratings ($r = .29$, $p < .05$).

To determine if those words that are more familiar to Turkish speakers elicited stronger SCRs, we correlated participants' familiarity ratings with SCRs elicited by those words. The correlation was near 0 ($r = -.05$), indicating that SCR is not influenced by familiarity. Because taboo words had low familiarity but high SCR, this could mask any subtle familiarity effects when taboo words are pooled with other stimuli. We thus correlated SCR and familiarity separately for the five stimulus categories. None of the correlations were significantly different from zero ($r = -.09$ to $.02$).

A final exploratory analysis concerned whether the following demographic and language-learning variables predicted skin conductance responses: age, age of exposure to English, age of arrival in the United States, length of stay in the United States, self-rated proficiency, English verbal fluency (i.e., fluency scores), and gender. The effects of the age of acquisition on proficiency have frequently been examined with multiple regression when participants have a varied age of acquisition. For example, Birdsong and Molis (2001) found that the age of arrival in the United States was the strongest predictor of English proficiency for immigrants from Spanish-speaking countries. We conducted two separate multiple regressions, both using the above predictor variables; but in one regression the dependent variable was SCRs elicited by English stimuli, and in the other it was Turkish stimuli. Gender was entered as a dummy variable (1 and −1 for females and males). None of these variables were significantly correlated with SCR in the raw correlation matrix, and they did not become significant when entered as copredictors of SCR analysis with multiple regression.

DISCUSSION

The greatest reactivity was to taboo words in both languages, consistent with prior studies with monolingual speakers in which taboo words elicited strong autonomic responses (Gray et al., 1982; McGinnies, 1949; Zajonc, 1962). SCRs to taboo words were slightly stronger in the L1, especially in the auditory modality. This supports earlier reports that bilingual speakers experience more anxiety when encountering taboo words in an L1 compared to an L2 (Ferenczi, 1916; Greenson, 1950; Javier, 1989). Consistent with this, during debriefing, several participants reported "feeling nothing" when uttering or hearing a taboo phrase in their L2. However, Turkish participants nevertheless reacted quite strongly to English taboo phrases. Indeed, the finding that taboo words in the L2 elicited larger responses than reprimands in the L1 indicates that the socioemotive force of deviation from social norms has a large effect. Speakers may report "feeling nothing" when uttering or hearing taboo words in their L2 because they are aware that taboo phrases generate a diminished visceral response. This relatively lower response is dichotomized as a "nothing" response when reflecting on the emotional feeling of taboo expressions.

We did not anticipate that there would be such a marked reaction to childhood reprimands in the native language. Similar expressions in English had very little effect on these native Turkish speakers. The finding of greater SCRs for childhood reprimands in an L1 compared to an L2 does not have a precedent in the literature. We included the childhood reprimands to explore the hypothesis that early language codevelops with emotional regulation systems (Bloom & Beckwith, 1989). The large skin conductance amplitudes for L1 reprimands are consistent with this view or may be due to the family context of learning of the L1 (Bostwick, 1996). Further studies will be needed to substantiate the hypothesis that language learned early in life, especially emotional expressions, elicits a more visceral reaction than language learned later in life.

As shown in Figure 2, in the L1, auditory stimuli elicited larger reactivity than visual stimuli. In the L2, however, mean skin conductance amplitudes were

equivalent for the auditory and visual modality. Why would auditory stimuli elicit greater autonomic arousal than visual stimuli in the L1 but not the L2? This finding could mean that visual and auditory stimuli activate modality-specific lexical representations (Anooshian & Hertel, 1994). Language acquired early in life, prior to the age when reading is the principal source of new vocabulary, is acquired via the auditory modality. This modality-specific vocabulary may be tightly connected to brain systems for emotional arousal, given the proliferation of neural connections in early and middle childhood. Auditory expressions may also have richer associations than representations activated by print, because more diverse and emotionally rich language experiences occur in the auditory modality. A greater number of associations may result in greater autonomic reactivity. In an L2, a number of associations between auditory and visual representations may be more similar to each other. A greater amount of experience in the L2 may take place with print, because of the context of learning in school and use of the language as an adult in university and work settings.

Of course, the Language × Modality interaction graphed in Figure 2 does not require us to posit modality-specific representations. Visual and auditory input could activate modality-independent lexical representations in both Turkish and English. These representations could then lead to a certain amount of autonomic arousal. The auditory qualities of the stimulus, such as tone, cadence, accent, and participants' knowledge that the voice is speaking their L1, may be the critical factors in eliciting high autonomic arousal to auditory stimuli in the L1. How much arousal is conferred by auditory aspects, independent of lexical content, could be explored by filtering speech so that words' identities could not be discerned.

During debriefing, several participants spontaneously mentioned that hearing the reprimands in Turkish led to a memory (or impression of remembering) a family member speaking these phrases. This is consistent with the proposal that modality- and language-specific memories are stored and readily accessible. Our comparison of auditory and visual presentations presents a cautionary note for future research. The difference between an L1 and L2 was reduced for written stimuli. If maximal differences between an L1 and L2 are sought, auditory presentation is probably the more useful experimental approach.

Familiarity and skin conductance

Skin conductance reflects the relevance of a stimulus, such that even a photograph of the face of an acquaintance, when embedded in a stream of unfamiliar faces, will elicit heightened responsivity (Channouf & Rouibah, 1997; Tranel, Fowles, & Damasio, 1985). Could increased electrodermal responses to Turkish words be due, in part, to increased familiarity of Turkish words compared to English words? Our late learners of English would certainly rate most Turkish words as more familiar than most English words. However, there are several reasons to believe that the stronger skin conductance responses to Turkish words reflect greater emotional associations rather than the greater familiarity of Turkish words. Prior work (all with monolingual speakers) has repeatedly found an effect of emotionality on skin conductance responses (Bingham, 1943; Man-

ning & Melchiori, 1974; Mathews, Richards, & Eysenck, 1989), but not of words' familiarity (Ellis, Hadyn, Quayle, & Young, 1999; Jacobs, 1955). English taboo words were rated as less familiar than the English reprimands and positive, aversive, and neutral words, yet these items had the strongest SCRs of the English stimuli. Neutral and positive items were rated as most familiar, yet showed the weakest SCRs. As noted in the results section, across all of the five stimulus conditions, English familiarity ratings were uncorrelated with SCRs elicited by English stimuli.

Electrodermal activity is generally believed to be sensitive to the difference in emotional valence between neutral and aversive words, with higher SCRs usually obtained for aversive words (Bingham, 1943; Manning & Melchiori, 1974; Mathews, Richards, & Eysenck, 1989). Unexpectedly, in neither language did positive and aversive words, such as *bride* and *cancer*, elicit significantly stronger SCRs than neutral words. This is surprising because prior work with English monolinguals, using the same stimuli and apparatus (Dinn & Harris, 2000, 2003), found elevated SCRs for aversive compared to neutral words. Although future work is needed to resolve this, we suggest the following as an explanation for the lack of difference between aversive and neutral words. Electrodermal reactivity to arousing stimuli diminishes as those stimuli become expected (Dawson, Schell, & Filion, 1990). Indeed, Dawson et al. (1990) suggest that one measure of the strength of autonomic response is how many items show a measurable SCR before habituation occurs. We suggest that the arousing effects of negatively valanced words like *cancer* and *war* are reduced when they are encountered in the context of expletives and sexual terms. This hypothesis could be tested in future research.

CONCLUSIONS

Electrodermal recording is a robust and sensitive method for investigating psychophysiological concomitants of language. Different response patterns were found for L1 and L2, and heightened reactivity was found for taboo words in an L1, especially in the auditory modality. The current work also identified a new phenomenon: emotional reactivity to childhood phrases in an L1 but not an L2.

Research on L2 acquisition has traditionally focused on the acquisition of linguistic systems (phonology, morphology, grammar, lexicon). Reports of differences in the emotional experience associated with language are a reminder that more than the cognitive realm must be included when considering how learning an L1 differs from learning an L2. Our study thus contributes to a growing body of work on differences in emotional responsitivity to speakers' different languages (e.g., Anooshian & Hertel, 1994; Ayçiçeği & Harris, in press; Bond & Lai, 1986; Dewaele & Pavlenko, 2002; Gonzalez–Reigosa, 1976; Pavlenko, 1999, 2002).

Taboo words and reprimands elicited stronger physiological responses, as measured by skin conductance, in speakers' L1 than did equivalent words in their L2, which was learned after the age of 12. Our results explain, in part, the relative comfort and ease experienced by bilingual speakers when saying taboo

words in their L2: the terms, although semantically equivalent, are not as physiologically arousing as the L1 terms. The strong reaction to emotionally laden childhood reprimands in the L1 but not the L2 suggests that a physiological difference between the L1 and later languages is not limited to taboo words. An L1 learned early in life may be represented in an individual in a particularly intimate and pervasive way.

APPENDIX A

ENGLISH AND TURKISH STIMULI

Aversive words

anger	öfke
cancer	kanser
crime	suç
cruel	zalim
danger	tehlike
death	ölüm
disease	hastalık
fight	kavga
grave	mezar
kill	öldürmek
murder	katletmek
pain	ağrı
poison	zehir
sick	hasta
slavery	kölelik
war	savaş

Neutral words

box	kutu
branch	şube
chair	sandalye
column	kolon
door	kapı
envelope	zarf
finger	parmak
foot	ayak
job	iş
name	isim
number	sayı
part	kısım
street	cadde

table	masa
tree	ağaç
window	pencere

Positive words

bride	gelin
father	baba
freedom	özgürlük
friend	arkadaş
fruit	meyva
happy	mutluluk
home	ev
honey	canım
joy	sevinç
kiss	öpücük
laugh	gülmek
love	aşk
mother	anne
mountain	dağ
smile	gülümsemek
sunset	gündoğuşu

Reprimands (not all are translation equivalents)

Don't do that!	Yapma!
Go to your room!	Yıkıl karşımdan! (translation: I don't want to see you!)
No!	Hayır
Shame on you!	Seni utanmaz!
Shut up!	Kes sesini!
Stop that!	Dur!
I hate you!	Senden nefret ediyorum!

Taboo words (not all are translation equivalents)

asshole	sevişmek
bitch	kahpe
breast	meme
oral sex	masturbasyon
shit	gerdek
raped	tecavüz
pee	fuhuş
vagina	bekaret
whore	kilot

ACKNOWLEDGMENTS

Bruce Mehler from NeuroDyne Medical Corporation provided expertise regarding electrodermal recording. We thank Hiram Brownell, Jean-Marc Dewaele, Wayne Dinn, and Alison Morris for helpful comments on a previous draft.

REFERENCES

Altarriba, J., & Rivera–Santiago, 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., & Hertel, T. P. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion, 8,* 503–514.

Ayçiçeği, A., & Harris, C. L. (in press). Bilinguals' recall and recognition of emotion words. *Cognition and Emotion*.

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., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language, 44,* 235–249.

Bloom, L., & Beckwith, R. (1989). Talking with feeling: Integrating affective and linguistic expression in early language development. *Cognition and Emotion, 3,* 315–342.

Bond, M. H., & Lai, T. (1986). Embarrassment and code-switching into a second language. *Journal of Social Psychology, 126,* 179–186.

Bostwick, C. M. (1996). Individual differences in young children's emotion knowledge and emotional expressiveness: The role of the family. *Dissertation Abstracts International: Section B. The Physical Sciences and Engineering, 56,* 5193.

Channouf, A., & Rouibah, A. (1997). Suboptimal familiar faces exposure and electrodermal reactions. *Annuario de Psicologia, 74,* 85–97.

Cohen, J. D., MacWhinney, B., Flatt, M. R., & 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.

Crane, L. D., Dieker, R. J., & Brown, C. T. (1970). The physiological response to the communication modes: Reading, listening, writing, speaking, and evaluating. *Journal of Communication, 20,* 231–240.

Dawson, M. E., Schell, A. M., & Filion, D. L. (1990). The electrodermal system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of psychophysiology: Physical, social, and inferential elements* (pp. 295–324). Cambridge: Cambridge University Press.

Dewaele, J.-M., & Pavlenko, A. (2002). Emotional vocabulary in interlanguage. *Language Learning, 52,* 263–322.

Dewaele, J.-M., & Regan, V. (2001). The use of colloquial words in advanced French interlanguage. *EUROSLA Yearbook, 1,* 51–67.

Dinn, W. M., & Harris, C. L. (1999, August). *Orbitofrontal dysfunction: A neurobehavioral continuum*. Paper presented at the APA 107th Annual Convention, Boston.

Dinn, W. M., & Harris, C. L. (2000). Neurocognitive function in antisocial personality disorder. *Psychiatry Research, 97,* 173–190.

Durst, U. (2001). Why Germans don't feel anger: In J. Harkins, & A. Wierzbicka (Eds.), *Emotions in crosslinguistic perspective.* New York: Mouton de Gruyter.

Ellis, H. D., Quayle, A. H., & Young, A. W. (1999). The emotional impact of faces (but not names): Face specific changes in skin conductance responses to familiar and unfamiliar people. *Current Psychology: Developmental, Learning, Personality, Social, 18,* 88–97.

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.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Gleason, J. B. (1985). Language and socialization. In F. S. Kessel (Ed.), *The development of language and language researchers: Essays in Honor of Roger Brown.* Hillsdale, NJ: Erlbaum.

Gonzalez–Regiosa, F. (1976). The anxiety arousing effect of taboo words in bilinguals. In C. D. Spielberger & R. Diaz–Guerrero (Eds.), *Cross-cultural anxiety* (pp. 89–105). Washington, DC: Hemisphere.

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders.* Philadelphia, PA: Lea & Febiger.

Goz, Ö. (in press). *The word frequency dictionary of written Turkish.* Istanbul: Turkish Language Institute.

Gray, S. G., Hughes, H. H., & Schneider, L. J. (1982). Physiologial 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.

Gumperz, J. J., & Hernandez, E. (1971). *Bilingualism, bidialectalism, and classroom interaction in language and social groups.* Palo Alto, CA: Stanford University Press.

Hill, A. B., & Kemp–Wheeler, S. M. (1989). The influence of context on lexical decision times for emotionally aversive words. *Current Psychology: Research and Reviews, 8,* 219–227.

Hugdahl, K. (1995). *Psychophysiology: The mind–body perspective.* Cambridge, MA: Harvard University Press.

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.

Jacobs, A. (1955). Formation of new associations to words selected on the basis of reaction-time-GSR combinations. *Journal of Abnormal and Social Psychology, 51,* 371–377.

Javier, R. A. (1989). Linguistic considerations in the treatment of bilinguals. *Journal of Psychoanalytic Psychology, 6,* 87–96.

Javier, R. A., & Marcos, L. R. (1989). The role of stress on the language-independence and code-switching phenomena. *Journal of Psycholinguistic Research, 18,* 449–472.

Jay, T. B. (1992). *Cursing in America*. Philadelphia, PA: John Benjamins.

LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science, 9,* 490–493.

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. (2002). *The taboo Stroop and related effects of emotion on memory and attention.* Paper presented at the Annual Meeting of the Psychonomics Society, Kansas City, MO.

Manning, S. K., & 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., Richards, A., & Eysenck, M. (1989). Interpretation of homophones related to threat in anxiety states. *Journal of Abnormal Psychology, 98,* 31–34.

Matthews, A., & MacLeod,C. (1985). Discrimination of threat cues without awareness in anxiety states. *Journal of Abnormal Psychology, 95,* 131–138.

McGinnies, E. (1949). Emotionality of perceptual defense. *Psychological Review, 56,* 244–251.

Nothman, F. H. (1962). The influence of response conditions on recognition thresholds for taboo words. *Journal of Abnormal and Social Psychology, 51,* 427–433.

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., & 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.

Sechrest, L., Flores, L., & Arellano, L. (1968). Language and social interaction in bilingual culture. *Journal of Social Psychology, 78,* 215–224.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.

Tranel, D., Fowles, D. C., & Damasio, A. R. (1985). Electrodermal discrimination of familiar and unfamiliar faces: A methodology. *Psychophysiology, 22,* 403–408.

Vanderploeg, R. D., Brown, W. S., & Marsh, J. T. (1987). Judgments of emotion in words and faces: ERP correlates. *International Journal of Psychophysiology, 5,* 193–205.

Van Lancker, D., & Cummings, J. (1999). Expletives: Neurolinguistics and neurobehavioral perspectives on swearing. *Brain Research Reviews, 31,* 83–104.

Weber–Fox, C. M., & Neville, H. J. (1999). Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis.* Mahwah, NJ: Erlbaum.

Wierzbicka, A. (1999). *Emotions across languages and cultures: Diversity and universals.* Cambridge: Cambridge University Press.

Wischner, G. J., & Gladis, M. T. (1969). Rate of formation of associations by schizophrenic and normal subjects as a function of stimulus or response placement of aversive components. *Psychological Reports, 25,* 599–605.

Zajonc, R. B. (1962). Response suppression in perceptual defense. *Journal of Experimental Psychology, 64,* 206–214.