# International Journal of Bilingualism

**Physiological reactivity to emotional phrases in Mandarin––English bilinguals**
Catherine L. Caldwell-Harris, Jimmy Tong, Winvy Lung and Sinlan Poo

The online version of this article can be found at:

Published by:

**$SAGE**

Additional services and information for *International Journal of Bilingualism* can be found at:

**Email Alerts:** http://ijb.sagepub.com/cgi/alerts

**Subscriptions:** http://ijb.sagepub.com/subscriptions

**Reprints:** http://www.sagepub.com/journalsReprints.nav

**Permissions:** http://www.sagepub.com/journalsPermissions.nav

**Citations:** http://ijb.sagepub.com/content/15/3/329.refs.html

>> Version of Record - Aug 12, 2011

OnlineFirst Version of Record - Dec 17, 2010

What is This?

# Physiological reactivity to emotional phrases in Mandarin–English bilinguals

## Catherine L. Caldwell-Harris, Jimmy Tong, Winvy Lung, and Sinlan Poo
Boston University, USA

## Abstract

Chinese–English bilinguals residing in the US were interviewed about their experience of using emotional expressions. They judged L1-Mandarin expressions as feeling stronger than L2-English expressions. Respondents nonetheless preferred to express their emotions in English, citing more relaxed social constraints in English-speaking environments. Electrodermal monitoring was conducted on a similar sample of Chinese–English bilinguals in order to determine how physiological reactivity corresponds to self-reports. For those who had both good Mandarin and good self-reported English ability, English and Mandarin emotional expressions elicited similar magnitude skin conductance responses (SCRs), with the exception of the category of endearments (e.g., *Thank you, I miss you, I love you*), where larger SCRs occurred for English expressions. Given cross-cultural reports that English-speaking societies encourage more open expression of positive emotion than do Chinese cultures, hearing English endearments may have led to easy retrieval of personal situations with strong emotional resonances; these memories then resulted in increased affect and increased SCRs. However, ratings of the emotional intensity of endearments were similar in the two languages, thus conflicting with the SCR findings. Additionally, English childhood reprimands were rated as less intense than L1-Mandarin reprimands, consistent with other studies showing that childhood reprimands are felt to be more intense in the native language. Future work will be needed to understand the conditions under which physiological responses differ from self-report.

## Keywords

cross-cultural differences, electrodermal monitoring, emotion, multilingualism

Bilinguals frequently report experiencing reduced emotion when using their second language, or when using their less proficient language (Altarriba & Santiago-Rivera, 1994; Bond & Lai, 1986; Dewaele, 2004, 2006, 2008; Gonzalez-Reigosa, 1976; Marian & Kaushanskaya, 2004; Pavlenko, 1998, 2002, 2005; Schrauf, 2000). Cultural variation in emotional expression has also been widely studied (Elfenbein & Ambady, 2002; Goddard, 2002; Matsumoto, 2001; Tsai & Levenson, 1997;

**Corresponding author:**
Catherine L. Caldwell-Harris, Department of Psychology, Boston University, 64 Cummington St., Boston, MA 02215, USA.
Email: charris@bu.edu

Yamaguchi, 2001). One of the findings is that frequent verbal declarations are normative in individualistic cultures, while emotional restraint is emphasized in collectivist cultures (Markus & Kitayama, 1991; Matsumoto, Yoo, & Nakagawa, 2008; Mesquite, 2001). Less well understood is how cultural display rules and attitudes about emotional expression interact with bilingualism and language learning history (but see discussion of the case of *I love you* in Dewaele, 2008). Differences between self-reported emotionality and physiological arousal are also not well understood. The current article attempts to fill these gaps by interviewing Mandarin–English bilinguals and analyzing electrodermal responses (skin conductance) while bilinguals listened to emotional phrases in their first and second language. In the following section, we briefly review prior work on self-reported emotional resonances and physiological reactivity while bilingual speakers listened to emotional expressions.

## Prior findings on reactions to emotional phrases

For Turkish immigrants and sojourners living in Boston, skin conductance responses (SCRs) were reduced for emotional words and phrases heard or read in the second language, English (Harris, Aycicegi, & Gleason, 2003). This finding confirmed long-standing reports that a second language, especially when it is of lower proficiency, elicits fewer emotional resonances than does a first language. Turks studying English as a foreign language in Istanbul also had reduced SCRs for English, especially English childhood reprimands (Caldwell-Harris & Aycicegi-Dinn, 2009). This replication minimizes the objections that the increased arousal to the native language reported in Harris et al. (2003) was caused by nostalgia for the mother tongue, or caused by novelty of hearing Turkish in an American university laboratory.

A first language may be experienced as more emotional than the second language because the emotional regulation system co-evolved with language, as suggested by Bloom and Beckwith (1989) in their account of socialization during first language acquisition. However, a second or later-learned language is often experienced as the most emotional language if it has become the more proficient language through long use and immersion (Dewaele, 2006; Schrauf & Durazo-Arvizu, 2006). In Dewaele and Pavlenko's (2001–3) web questionnaire, respondents downgraded the illocutionary strength of L1 taboo/swear words, anger, and the expression *I love you* if L1 was no longer the dominant language (Dewaele, 2004, 2006, 2008). For each of these types of emotional expressions, multiple factors influenced how multilinguals evaluated the emotional force of these expressions in their different languages. The most consistently relevant were, in order of strength of effect: language dominance, proficiency, usage frequency, age of acquisition, and acquisition context (naturalistic vs. instructed).

Theorists have repeatedly argued that contexts of learning will influence how adult speakers perceive the emotional resonances of their languages (Altarriba, 2006; Dewaele, 2004; Pavlenko, 2005). Harris's (2004) skin conductance study with bilinguals who varied in age of acquisition was broadly consistent with these prior reports. Electrodermal responses were highly similar for Spanish–English bilinguals' two languages, for those who grew up in the US from early childhood with Spanish spoken at home. Bilinguals who had immigrated to the US during their teen years or later had higher SCRs to Spanish reprimands (Harris, 2004) but not to other stimuli. The generalization thus appears to be that proficiency and contexts of use are more important in determining psychophysiological reaction than is age of acquisition (Harris, Gleason, & Aycicegi, 2006). Note that for speakers in that study, proficiency and age of acquisition were highly correlated, as they are for most immigrants (Birdsong & Molis, 2001; Jia, Aaronson, & Wu, 2002).

The first language may be frequently experienced as emotional because early childhood includes many emotionally based learning events, such as learning from family members. According to the 'emotional contexts of learning hypothesis' (Harris et al., 2006), language comes to be experienced as emotional to the extent that it is learned and used in emotional contexts. The Spanish–English bilinguals who immigrated in their teens experienced peer socialization in American high schools, which plausibly facilitated acquiring the emotional resonance of English taboo phrases. It is thus not surprising that they responded to English taboo phrases in a similar manner as their monolingual English peers despite English being their second language.

## Pilot study on Mandarin–English bilingual speakers

When a pattern of results is obtained across a number of cultural groups, one can feel more secure in identifying broad psychological mechanisms, rather than mechanisms specific to a particular language or culture. We chose to study native speakers of Mandarin because Chinese language and culture differ along a number of dimensions from Latin American and Turkish language and culture. Unlike these other two cultural groups, Mandarin shares no cognates with English, and has a different writing system, phonology, and grammar. Chinese culture also has a different set of sociocultural expectations about verbal expression, compared to US norms (Potter, 1988).

A pilot study consisting of semi-structured interviews about learning history and emotional responsiveness was carried out with students at Boston University ($N = 23$) who self-identified as fluent in Mandarin. Most were international students from China and Taiwan who had arrived in the US for college, or who had immigrated with their families in middle childhood or later. Questions included language spoken, language learning history, current use of both languages, perceived dominance, and self-ratings of proficiency, similar to those in the LEAP, a self-report instrument (Marian, Blumenfeld, & Kaushanskaya, 2007). We also asked questions modeled after those used in the Bilingualism and Emotion Questionnaire (BEQ; Dewaele & Pavlenko, 2001–3), about which language elicits greater emotional feelings. Questions included which language feels more emotional, which language feels stronger when saying or hearing *I love you*, using or hearing swear words, and hearing or expressing anger. Participants were additionally asked to name which language they preferred to use for expressing anger and which for exchanging confidences and intimate expressions (expressing positive emotions). Interviewers took notes on a clip-board, but explored areas of interest that emerged during discussion. Of the 23 respondents, six had immigrated to the US by age 5 (or had learned English from a native-speaking parent overseas). The remaining 17 can be considered late learners and immigrated to the US at age 12 or later.

Two striking phenomena emerged during the interviews. One of these was reported codeswitching to say *I love you* in English while speaking Mandarin, as in the following example:

| | |
|---|---|
| Interviewer: | How often to you speak Mandarin currently? |
| Student: | About an hour a day, on the phone to my mother. |
| Interviewer: | And the whole hour is in Mandarin? |
| Student: | Yes, except at the end, we say love ya. |

A second unexpected discovery during the interviews was that the question *Which of your two languages feels more emotional to you?* sometimes elicited *English* even for international students who identified Mandarin as their dominant language, and spoke accented, non-colloquial English. In contrast, the Turkish and Spanish speakers discussed above, and Russian immigrants to the US studied by Caldwell-Harris, Staroselsky, Smashnaya, and Vasilyeva (2011) always identified either

their first language or their dominant language as their most emotional language, defined as the language that elicits the strongest emotional feelings during use, consistent with self-reports in Dewaele (2006, 2008).

Follow-up questions asked respondents to explain why a language felt more emotional. In prior studies, during debriefing, Turks most frequently explained that Turkish felt most emotional to them because it was their first language, the language of their family, and their language of greatest expertise. Spanish and Russian speakers also said this, but additionally commented that Spanish (or Russian) felt more emotional to them because Latin (or Russian) culture was so passionate and/or emotionally expressive. Mandarin speakers who identified Mandarin as their most emotional language did mention first language, proficiency, and life-long experience as reasons. The 25 per cent of interviewees who initially identified English as their most emotional language gave explanations such as*: It just is*, *You can say more* and *We don't say I love you in Mandarin*.

These responses suggest that respondents associated English with the cultural permission to speak freely about emotional topics. American culture emphasizes the importance of verbal expression of emotion (Goddard, 2002). Follow-up questions generally confirmed that some respondents had interpreted our question, *Which language feels more emotional?*, to mean *In which language do you feel at ease expressing your emotions?* Respondents tended to mention either that there are fewer cultural restrictions on English emotional expressiveness, or that the Chinese translation equivalent, *Wo ai ni*, is a more serious or consequential expression (see Table 1). The initial misunderstanding of our question during interviewing is striking because it indicates that for this population, a strong or automatic inference about the phrase *Which language feels more emotional?* concerns emotional display rules, not one's own subjective impression of emotional resonances.

After clarifying for participants that we meant subjective impression of emotional resonances, respondents were asked to name which language felt more emotional, or they could say that both languages felt equally strong, in response to the following six questions:

> *Which language feels stronger when you say I love you? When you hear I love you?*
> *Which language feels stronger when you express anger? When you hear angry statements?*
> *Which language feels stronger when you use swear words? When you hear swear words?*

**Table 1.** Responses by late learners of English (*N* = 17) to *How does it feel to say or hear 'I love you' vs. 'Wo ai ni'?*

*Statements asserting fewer cultural restrictions on English emotional expression*
If you want to date a girl, you would feel shy to say *I love you* in Mandarin, but in English it's OK.
It's easier to say things in English.
*I love you* doesn't mean a lot in English because people say it too much.

*Statements that 'Wo ai ni' is culturally restricted, or is a more serious or consequential expression*
We don't say *I love you* in Mandarin
In English, there are several levels of *I love you* but in Mandarin, you would really mean it. [Explained: could say 'I love you' to express affection to a friend or as a humorous compliment.]
*I love you* sounds weird in Mandarin [respondent had trouble expressing why].
You don't usually say *I love you* in Mandarin but it's becoming more open now.
In Mandarin it means a lot more.
My parents told me that I would say *Wo ai ni* to one person in my life.
You can't say *I love you* in Mandarin because it is too strong.

**Table 2.** Judging strength of emotional utterances and preference for expression (pilot study)

| | Percentage of late learners of English ($N = 17$) | | | Percentage of early learners of English ($N = 6$) | | |
|---|---|---|---|---|---|---|
| | Mandarin | English | Both | Mandarin | English | Both |
| Judged expressions to have stronger emotional resonances in … | 53% | 6% | 41% | 17% | 17% | 67% |
| Language preference for expressing anger | 23% | 59% | 18% | 67% | 33% | 0% |
| Language preference for expressing confidences/ intimacies | 18% | 64% | 18% | 17% | 50% | 33% |

Interviewees often reported that they had never said *Wo ai ni* to anyone, or had very little experience with swear words in one or both languages. We nevertheless asked them still to choose which language would be stronger if they were to hear or utter these expressions.

To enable generalizations to be drawn from the pilot study, participants were categorized as being late learners of English (immigrating after age 12) or early learners (immigrated or otherwise acquired English by age 5). The late learners of English reported themselves as having native or near-native proficiency in Mandarin and moderate proficiency in English, while the early learners of English reported the reverse. Reports of usage patterns followed proficiency and age of acquisition. Late learners of English reported using always Mandarin with family, and frequently using it with friends. Early learners of English reported using Mandarin frequently with family and only sometimes with friends.

We discuss first the pattern revealed by the 17 late learners of English. About half (53%) of the late learners judged emotional expressions to have stronger emotional resonances in Mandarin, or judged emotional resonances to be equal; few (6%) said English expressions felt stronger (see Table 2). However, when reporting which language they preferred for expressing emotion, the majority of the late learners preferred English (59% preferred English for expressing anger, while 65% preferred English for expressing intimacies). This difference in perceived emotional force vs. language preference is reminiscent of the discussion by Dewaele (2004) about what it means when bi- and multilinguals prefer to swear in their more emotional or less emotional language (where the more emotional language is usually either the first language or the dominant language). Swearing is accompanied by arousal and/or anxiety (Gonzalez-Reigosa, 1976). If these emotions are valued, then speakers prefer to swear in their more emotional language. Respondents who find anxiety/ arousal aversive report preferring the less emotional language for swearing. The pilot data reported here thus extend Dewaele's insight to the case of expressing intimacies, confidences, and anger, among Chinese–English bilinguals who immigrated to the US between age 12 and 18 (see also Bond & Lai, 1986).

Two-thirds of the six early learners judged expressions to be equally strong in both languages. They had less preference for using English for emotional expression (especially anger) than did the late learners. This suggests that they may not understand the Chinese cultural prohibitions against emotional expression. However, the four participants who said they preferred to express anger in Mandarin judged English expressions to feel stronger than Mandarin. They may actually be adhering to Chinese cultural conventions by preferring to express anger in their less emotional language.

These provocative possibilities deserve to be studied in future work with a larger sample of early Mandarin–English bilingual speakers.

## *Pilot study on emotional expressiveness: discussion*

Mandarin-bilingual speakers who immigrated to the US in the teen years reported in spontaneous comments that they found it easier to express emotion in their second and less proficient language, English. Many frequently named English as their preferred language for expressing anger and sharing intimacies/confidences. One participant mentioned code-switching to use the phrase *Love ya* as a formulaic ending for a phone call to a parent.

These comments echo reports from other studies of emotional expression. Dewaele (2008) commented on the case of Veronica Zhengdao Ye (2004), a Chinese scholar who immigrated to Australia in the 1990s. Ye wrote, 'We do not place so much emphasis on verbal expression of love and affection, because they can evaporate quickly. For a Chinese, love and affection are embodied in care and concern, in doing what we believe are good things for the other party' (Ye, 2004, p. 140). In contrast, frequent verbal declarations of love are the norm in western cultures, especially North America (Matsumoto, 2001; Wilkins & Gareis, 2006). The media depict family members showering each other with *I love you* and dating websites discuss how soon is too early for love declarations.

It is well understood that display rules differ between Chinese and western cultures (Markus & Kitayama, 1991). Matsumoto (2001; Matsumoto et al., 2008) has discussed how Chinese culture emphasizes group harmony and subordination of individual desires. Minimizing emotional displays is thus valued as part of minimizing self-assertion. Another explanation for different display rules is whether cultures are tight or loose about constraining individual choice in a given situation (Triandis & Gelfand, 1998). Hall's (1976) description of cultures as high context or low context is also relevant. Chinese culture is classically considered a high context culture, meaning that interactants who know each other well assume shared context, which reduces the necessity for fully transparent verbal statements. This is consistent with reports by Beijing students interviewed by Caldwell-Harris and Kronrod (in preparation), who reported that declarations of love to a family member would be odd or unacceptable, or strange because an overt declaration would undermine the assumption of steadfast love.

In a anthropological study of traditional Chinese village life, Potter (1988) argued that the primary cross-cultural difference between China and the more individualistic western countries is not display rules or valuing or not valuing expressiveness, but how interpersonal relationships are conceptualized. Expressing love to family members (particularly parent–child) is discouraged in traditional Chinese culture because such declarations cast family members as equal in status, and falsely imply that warm feelings are the basis for the relationship. In traditional Chinese culture, family structure is a societal given; family relationships and obligations are not negotiable. An assertion of love would undermine the social forces of filial and parental duty that transcend daily emotional feelings.

## Physiological study: reactions to a first or second language

In our pilot study, immigrants from Chinese-speaking countries reported Mandarin expressions elicited more powerful emotional resonances than English expressions, but nevertheless preferred to use English to express emotions. This pattern of preferences could have implications for the research program of measuring physiological reactions to emotional expressions in a first or second language. In order to make hypotheses, it is useful to review what

types of emotional expressions lead to skin conductance responses (SCRs); noting that this psychophysiological technique has not been used with language stimuli until recently (outside of taboo word research).

SCRs can be broadly understood as sensitivity to (1) environment threat or relevance, or (2) internal effort/arousal (Boucstein, 1992). Threatening stimuli engage the flight or fight system, which dumps adrenaline into the blood stream and leads to the sweating that increases the transient conductivity of the skin. Stimuli that simply say *Pay attention! I'm relevant to your concerns!* can also engage an alerting system that leads to measurable SCRs. Internally generated arousal, such as fearful or anxious thoughts, also results in SCRs (e.g., Gray, Hughes, & Schneider, 1982). The following are specific examples.

*Conditioned associations.*    Conditioned reactions can be rapidly developed by pairing any stimulus, such as a word or phrase, with electric shock (Dawson, Schell, & Filion, 2000). Taboo phrases are believed to reliably elicit large SCRs via conditioned associations (Manning & Melchiori, 1974; McGinnies, 1949). The explanation is that taboo phrases were previously paired with threatening environmental messages (or punishment for use), and after repeated pairings, the phrase itself became capable of eliciting a fear response.

*Re-experiencing autobiographical memories.*    The Turkish–English bilinguals in the Harris et al. (2003) study had large SCRs to childhood reprimands in Turkish. One respondent reported during debriefing, *I could hear my aunt saying those phrases in my mind.* We can't know if reprimands uttered by the aunt had become conditioned associations with threat. The phrase may have elicited a memory which was connected to fear of punishment. Re-experiencing these memories may have been sufficient to elicit the SCRs. Or simply, the phrases that elicited childhood memories may have been more self-relevant than other phrases, as described below.

*Alerting due to relevance.*    When a few familiar faces were interposed in a series of unfamiliar faces in an electrodermal study by Tranel, Fowles, and Damasio (1985), the familiar faces elicited elevated SCRs. The familiar faces were not threatening, but were merely more relevant than surrounding stimuli.

*Cognitive effort during problem-solving.*    After a laboratory foreign-vocabulary learning task, Dahlen and Caldwell-Harris (2009) recorded skin conductance while learners matched sentences containing the newly learned foreign words to pictures. Incorrect trials elicited larger SCRs than did correct trials. The authors suggested that effort to identify the correct answer, or anxiety associated with uncertainty, may have increased skin conductance.

*Anxiety about performance, especially for speaking.*    When Turks who acquired English as a foreign language read aloud prepared true or false statements (e.g., *My favorite sport is baseball*), SCRs were higher in L2-English (Caldwell-Harris & Aycicegi-Dinn, 2009). Because this occurred for both true and false statements, the elevated SCRs probably indicated anxiety about English speaking ability.

To summarize, language that is associated with prior threatening experiences or with expectation of censure, or language that is relevant, or language that elicits cognitive effort, or language that leads to anxiety about speaking a less proficient language, may all lead to heightened SCRs.

We envision two possible scenarios on how hearing emotional language will be experienced by Chinese–English bilinguals such as the participants in the pilot study. We focus on predictions for

endearments because expressing *I love you* was the case where respondents made the strongest distinction between their two languages.

1. *Predict elevated SCRs for endearments (and possibly other Mandarin emotional expressions), compared to English expressions.*
   Rationale: SCRs to endearments will be stronger in Mandarin than in English because at least some of these expressions have a taboo-like status, due to cultural display rules restricting use.
2. *Predict reduced SCRs for Mandarin emotional expressions, compared to English expressions.*
   Rationale: If emotional expressions, especially endearments, are less common in Chinese cultural environments, then fewer easily retrievable autobiographical memories will be associated with them. English phrases will be associated with more memories, and these will lead to a quicker flood of emotional reactions, leading to heightened SCRs in English and/ or reduced SCRs for Mandarin.

## Method

*Changes from prior methodology.*    Electrodermal activity was recorded while bilinguals listened to and rated emotional and neutral phrases. Two changes from the prior studies (Caldwell-Harris & Aycicegi-Dinn, 2009; Harris, 2004; Harris et al., 2003) were implemented. In prior studies, neutral stimuli were single words selected to be relatively low in affect, according to norms in Toglia and Battig (1978). The emotional stimuli were phrases (e.g., *Shame on you!*). Neutral single words lack strong emotional associations and are less efficient at eliciting autobiographical memories than are emotional phrases. However, a drawback of single words is speakers' difficulty in identifying them with no surrounding context, particularly in Mandarin, given the high rate of homophony in Mandarin. We thus decided to embed neutral words in short phrases/sentences. Because all of the emotional expressions were highly familiar, and many were collocations, we tried wherever possible to create neutral phrases/sentences that were familiar or even formulaic collocations, such as *Take a taxi* or *Please sit down.*

In prior studies, participants were simply asked to rate the word or phrase for pleasantness (Harris, 2004; Harris et al., 2003) or emotional intensity (Caldwell-Harris & Aycicegi-Dinn, 2009). Participants were instead instructed to think of a situation when this phrase was used, and to rate the emotional intensity of that situation. The motivation for this change was to minimize reliance on rote cultural knowledge of the meaning of the phrase. Our hope was that we would elicit participants' personal emotional reaction. As we review later, an unexpected consequence of these instructions may have been to make the task more difficult, which could increase skin conductance due to effort in identifying a suitable autobiographical memory.

*Decisions about participant inclusion.*    Participants were recruited via sign-up sheets asking, 'Are you fluent in Mandarin?' in the Psychology Department of Boston University and via email to the Chinese Student Association. In bilingualism research, it is common to set precise conditions for inclusion. We decided to recruit a naturalistic sample, meaning that all students who identified as Mandarin speakers and who wished to be in the study were included. This naturalistic sample allowed us to discover what range of language-learning histories exists among college students who regard themselves as Mandarin–English bilinguals. Because we did not restrict enrollment in the study, we expect that language and cultural diversity in our sample will be similar to what exists among other Mandarin-speaking students attending universities in the northeastern US.

It is well known that many Mandarin speakers grow up exposed to another Chinese language in the home or community. Cantonese is widely spoken in Hong Kong although Mandarin is taught in schools. Many participants from Taiwan speak Taiwanese at home and when participating in the community (when shopping, chatting with neighbors). Students from mainland China also frequently speak a home language (such as Fujian) but learn Mandarin at school. While preparing for this study with the pilot interviews, we debated whether to include students exposed to a home/community dialect in addition to Mandarin. We worried that exclusion would restrict the pool of potential participants and would reduce our ability to generalize results to the wider group of Chinese–English speakers, and thus we decided against exclusion.

## Participants

Participants were 64 native Mandarin speakers (19 males, 45 females, aged 18–28); 37 per cent had grown up exposed to another Chinese language.

*Learning history questionnaire.*   Participants filled out a language history questionnaire that asked for details regarding age of arrival to an English-speaking country (AoAr), age of exposure to English, age of exposure to Mandarin, age of exposure (if any) to another Chinese language, percentage of time currently using Mandarin at home, and percentage of time currently using Mandarin with friends at school. Participants additionally reported their verbal SAT or TOEFL (Test of English as Foreign Language) score and rated their proficiency on a 7-point scale, with 7 corresponding to a native speaker, 6 near-native, 5 very good, 5 good, 4 fair, 3 not good, 2 poor, 1 little. Administering objective proficiency tests may be preferable to self-reported standardized test scores and self-ratings. However, recent papers have demonstrated strong correlations between self-ratings and standard proficiency tests (Hakuta, Bialystok, & Wiley, 2003; for discussion see Delgado, Guerrero, Goggin, & Ellis, 1999). Separate ratings were obtained for understanding, speaking, reading, and writing, but the mean of these ratings was used as the overall proficiency rating. One reason to include ratings for reading and writing is that speakers typically rated themselves as poorer at these than at understanding and speaking. Including these measures thus yields values that allow more discrimination between near-native speakers and medium-to-good speakers. For example, some students lived in a Mandarin-speaking country until the early teen years, but still self-rated themselves as having less than native-speaker Mandarin proficiency. This resulted because they reported non-native ability in reading and writing, due to attending English-language school throughout childhood.

*Characterizing the sample.*   Age of arrival to an English-speaking country (usually the US) ranged from birth to age 28. The distribution of age of arrival (AoAr) had two peaks in our sample. Twelve individuals (18%) arrived between birth and age 3, and reported native English speaking ability and less-than-native ability in Mandarin. A second peak of AoAr occurred in the teen years, with 17 (25%) of the sample arriving between 14 and 17. These individuals generally reported having native or near-native abilities in Mandarin and less-than-native English speaking ability.

Table 3 summarizes participants' learning history variables, with separate means provided for participants who grew up with and without exposure to a second Chinese language in addition to Mandarin and English. It is noteworthy that trilinguals were indistinguishable from bilinguals on proficiency and usage variables with the exception of current use of Mandarin at home with family. Trilinguals reported less than half as much Mandarin home-use as bilinguals.

**Table 3.** Learning history of participants in skin conductance study

|  | Trilingual (N = 24) | Mandarin, English (N = 40) | Statistical difference |
|---|---|---|---|
| Gender | 75% female | 67% female | n.s. |
| Age | 19.6 (18–28)[a] | 19.2 (18–28) | n.s. |
| Age of exposure to Mandarin | 2.7 (0–8) | 0.4 (0–4)[b] | $t(61) = 4.0, p < .001$ |
| Age of exposure to English | 6.9 (0–15) | 8.2 years old | n.s. |
| Age at arrival in the US | 12.7 (1–20) | 11.3 (0–24) | n.s. |
| Length of stay in the US | 7.3 (0–19) | 7.8 (0–23) | n.s. |
| Mandarin proficiency (1–7 scale)[c] | 5.7 (3–7) | 5.8 (3–7) | n.s. |
| English proficiency (1–7 scale)[c] | 5.7 (4–7) | 5.7 (4–7) | n.s. |
| English verbal SAT | 576 | 584 | n.s. |
| English TOEFL | 315 | 275 | n.s. |
| Current use of Mandarin with friends (%)[d] | 13% (0–90) | 17% (0–95) | n.s. |
| Current use of Mandarin at home (%)[e] | 29% (0–100) | 74% (0–100) | $t(1.61) = 4.9, p < .001$ |
| Which other Chinese language learned | Cantonese N = 14 Taiwanese N = 7 Fujian N = 2 | Not applicable |  |
| Identified other Chinese language as their dominant language | 33% (8 out of 24 trilinguals) |  |  |

Notes: [a] Range provided in parentheses. [b] Four bilingual participants learned Mandarin after English between age 1 and 4. [c] Average of self-ratings for proficiency in speaking, listening comprehension, reading, and writing. [d] Current use at university with friends. [e] At home with family

The learning history variables in Table 3 were used as indicator variables for an exploratory factor analysis which yielded three factors (note that gender and age were excluded as they were uncorrelated with other factors, and length of stay was redundant with age of arrival, with $r = -.94$). The three factors jointly explained 70 per cent of the variance in the indicator variables and were interpreted based on loadings on the indicator variables as representing (1) dominance in another Chinese language, (2) English ability, and (3) Mandarin ability and current use (see Table 4). The factor analysis reveals which learning history variables group together and thus complements the group comparisons in Table 3. Identifying another Chinese language as one's dominant language patterned with later Mandarin exposure and low use of Mandarin at home. Early exposure to English (measured by age of arrival in the US and age of exposure if living overseas) grouped together with English proficiency (correlation was $r = -.46$). Mandarin proficiency patterned with current use of Mandarin with friends at school ($r = .36$).

## Design and materials

*Selection of phrases and norming.* An initial list of phrases was generated by creating Mandarin translations of English items used in the prior Spanish study (Harris, 2004) and Turkish study (Harris et al., 2003). It was decided to use the same phrase categories from the Spanish study: neutral (*Take a taxi*), endearments (*I love you*), insults (*You are so fat*), reprimands (*That's not nice*), and taboo phrases (*She's a slut*) (other examples in each category appear in Table 5). Phrases were scrutinized and discussed by a bilingual panel consisting of three of the co-authors (C. C.-H., a native English speaker, W. L., a native of Hong Kong with fluency in Cantonese and Mandarin, and S. P., a native of Taiwan) and two additional Mandarin-speaking assistants. Phrases judged to be specific to a geographical region were discarded. As phrases from the pilot study were discussed and discarded,

**Table 4.** Factor analysis of questionnaire data: variance explained and factor loadings

| | Orthogonally rotated factors | | |
| --- | --- | --- | --- |
| | 1<br>Learned another<br>Chinese language | 2<br>English acquisition<br>and ability | 3<br>Mandarin current<br>use and ability |
| Variance explained (%) | 29% | 26% | 14% |
| *Factor loadings on variables:* | | | |
| Age of arrival in the US | .26 | **.74** | .20 |
| Age of English exposure | −.28 | **.84** | −.10 |
| Age of Mandarin exposure | **.82** | −.01 | .11 |
| English proficiency | −.02 | −**.68** | −.25 |
| Mandarin proficiency | .00 | .42 | **.63** |
| Other Chinese language<br>is dominant | **.78** | .19 | −.19 |
| Current use of Mandarin<br>at home | −**.84** | .16 | .08 |
| Current use of Mandarin<br>with friends | −.11 | .03 | **.90** |

*Note:* Values in bold type indicate high factor loadings (.6 cutoff).

**Table 5.** Examples of English and Mandarin stimuli in each emotional phrase category

| Emotion category | English | Rating intensity[a] | Mandarin[b] | Rating intensity |
| --- | --- | --- | --- | --- |
| Endearments | I love you | 5.5 | 我愛你 | 6.1 |
| N = 8 | I miss you | 5.2 | 我想你 | 5.5 |
| | Thank you so much | 3.9 | 謝謝你 | 3.9 |
| Insults | Get lost | 4.6 | 你給我滾 | 6.3 |
| N = 8 | You idiot | 5.2 | 你這個大白癡 | 5.6 |
| | Don't be a jerk | 5.7 | 你這個無賴 | 4.9 |
| Reprimands | Shame on you | 4.3 | 真丟臉 | 4.9 |
| N = 8 | Behave yourself | 4.6 | 太過分了 | 4.9 |
| | Stop that | 5.1 | 停了 | 5.2 |
| Taboo | He's an asshole | 5.1 | 他是個王八蛋 | 5.2 |
| N = 8 | He screwed your mother | 6.0 | 幹你娘 | 5.8 |
| | She's a bitch | 5.6 | 她是個賤人 | 5.8 |
| Neutral | I have shoes | 2.6 | 我有鞋子 | 1.3 |
| N = 15 | I bought toys | 2.2 | 我買了玩具 | 4.1 |
| | The wooden desk | 1.8 | 一張木桌子 | 1.1 |
| Filler | The white envelope | 1.2 | 拿你的外套 | 1.7 |
| N = 45 | | | | |

*Notes:* [a] Ratings are from the electrodermal monitoring session and are for the indicated item. Average ratings for categories appear in Figure 1. [b] Traditional Chinese script was shown to raters and used by our panel when evaluating items.

the panel asked for suggestions from friends and co-workers. Over several meetings the panel discussed their intuitions about strength of emotional meaning and usage contexts. Because translation equivalents seldom had similar emotional intensity or familiarity, the English phrases selected for inclusion are not simple translations of the Mandarin stimuli. The panel decided to find, in each language, a similar number of phrases that fit the designated category label (e.g., insults), such that the items

in a given category had overall comparable familiarity, emotional force, and colloquial usage to phrases in the other language. The final list contained eight items in each emotion category (total *N* = 32), and 15 neutral items that were evaluated by the panel to be collocations or familiar phrases. In addition, 45 filler phrases were included which had similar length, but did not repeat content words used in other trials. Their purpose was to decrease the proportion of highly emotional stimuli, given the guidelines of Dawson et al. (2000) that emotional stimuli be fewer than one-third of all trials in order to elicit SCRs for emotional stimuli. Phrases were generally short sentences, with English phrases having a mean length of 3.4 words (range 2–5), and Mandarin phrases consisting of an average of 5.2 characters (range 2–10) in written form (note that a Mandarin character is a single syllable not a word; words in Mandarin are usually two-character compounds). Number of words/ characters did not vary by phrase category in either language. Note that for norming (see below), monolingual raters evaluated the written forms, but participants in electrodermal recording sessions heard audio sequences only. All stimuli are available from the authors.

*Monolingual ratings.*    Written versions (in traditional Chinese script) of the Mandarin emotional expressions were rated for their emotional intensity by 52 female high school students, age 16–18, attending Ching-mei Girls High School in Taipei (obtained during a visit to Taipei by S. P.). Our monolingual English sample consisted of 45 Boston University students, of whom 37 were female,[1] who rated the English emotional expressions for emotional intensity following the same instructions listed in the 'Apparatus and procedure' section later.

Phrases were rated on a 1–7 scale, with 7 being most emotional. Raters were also asked to check a box if a phrase was odd or unfamiliar. We did not include neutral items in the paper-and-pencil rating study because raters would then produce two clumps of ratings; with the neutral phrases in the bottom one-third or one-half, and the emotional phrases occupying the top one-third or one-half. Excluding neutral items allowed the entire 7-point scale to be used for rating emotionality, thus enhancing between-item sensitivity. The mean ratings for Mandarin items in the three emotional categories were judged to have similar emotional intensity, with endearments, insults/taboo, and reprimands having mean ratings of 4.1, 4.1, and 4.2 respectively (means are in the middle of the scale because neutral items were not rated). Ratings varied by emotional category for American college students, who provided ratings of 5.1, 4.25, and 3.5 for the endearments, insults/taboo, and reprimands respectively. The higher emotional-intensity ratings for endearments may fit with the importance of positive emotions for North Americans (Goddard, 2002). Low ratings for reprimands is consistent with American college students' self-reports that that reprimands are not perceived as very anxiety-provoking, and with prior low SCRs for childhood reprimands in monolingual English speakers (Harris, 2004). To provide a basis for bilingual assessment, it would have been ideal to obtain phrases which received highly similar ratings by monolinguals in each language. Our bilingual panel scrutinized items after receiving these ratings and agreed that better items do not exist in the two languages, i.e., there are not widely familiar endearments in Mandarin that would plausibly lead to higher intensity ratings for a set of Chinese-speaking raters; nor do childhood reprimands exist in English which are both highly familiar and likely to lead to higher ratings by the American monolinguals. Although we were not able to obtain a perfectly matched set of stimuli, our norming task did establish that the phrases were at least broadly of similar intensity (with category means in the 3.5–5.0 rating range) and most importantly, that no students indicated that any phrases were odd or unfamiliar.

*Audio recording.*    Due to the difficulty of finding Mandarin–English bilingual speakers with no accent in either language, a decision was made to follow the method in Harris et al. (2003) of having each language recorded by a native speaker of that language. Mandarin phrases were recorded

by a Taiwanese female who had recently moved to Boston and had prior professional broadcasting experience. English phrases were recorded by an American female who had professional speaking experience. Instructions for recording were to read each phrase with intonation appropriate for the phrase. Native-speaking judges rated the phrases for tonal appropriateness. Phrases with insufficient or excessive affect (or uneven volume) were rerecorded.

## Apparatus and procedure

Participants heard auditory clips of the Mandarin and English phrases via a computer loudspeaker, presented in random order. A 2 × 5 within-subjects design was used: two levels of language (Mandarin and English) and five categories of emotionality. Participants rated the emotionality of each phrase on a 1–7 scale, with 7 being the most emotional, by pressing an appropriate key on a standard keyboard.

The specific instructions for rating emotional intensity were intended to elicit autobiographical memories:

> *You will rate phrases for emotional intensity. Do this by trying to think of a real situation that happened to you in which the phrase was used. Then rate how emotional that use of the phrase was for you. (If you can't think of an emotional situation, it is okay to give the phrase a low emotional intensity rating.)*

In prior work in our laboratory, we had found that asking for ratings of *emotional intensity* was more easily understood than asking about *emotional arousal*, as the latter term led some participants to infer that a high rating would require a bodily response or feeling of physical alertness. Our instructions left open the possibility that participants could have translated the phrase into a language different from the presentation language, and then imagined a situation in which the translated phrase was used. However, our impressions from interacting with participants and from discussion during debriefing were that participants understood that the term 'the phrase' meant the actual linguistic expression, not the general meaning. For example, one participant commented, *For the Chinese phrases, I could hear my grandmother saying them.* Future rating studies should be explicit that memories are to be retrieved in the language of the phrase. It would also be useful to conduct a separate study to determine if some categories of emotional experience are more easily retrieved in Mandarin vs. English, as studies of this type do exist for other languages (e.g., Altarriba & Bauer, 2004; Altarriba, Bauer, & Benvenuto, 1999; Marian & Neisser, 2000; Schrauf, 2000; Schrauf & Durazo-Arvizu, 2006).

Stimuli were presented on a computer using PsyScope, experimental control software developed by Cohen, MacWhinney, Flatt, and Provost (1993). Electrodermal activity was recorded using the Davicon C2A Custom Skin Conductance Monitor (NeuroDyne Medical Corporation). Gold-plated electrodes (sensors) were attached to the tip of the first and second fingers of the dominant hand. A 10-second recording interval began coincident with stimulus onset. The Davicon Psychophysiological Assessment Software subtracts the base point 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 (Hugdahl, 1995; Lykken, 1998). To correct for differences in range and to increase the homogeneity of variance in the data, each participant's distribution of micromhos was converted to standard scores (mean of 100, standard deviation of 15), following the recommendations of Ben-Shakhar (1985) to convert raw SCRs to standard scores (mean of 100).
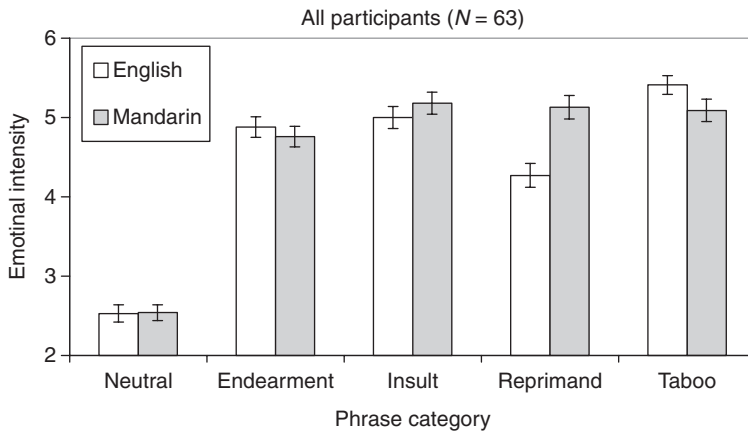
**Figure 1.** Emotional intensity ratings (1–7 scale) provided by participants during electrodermal monitoring. Ratings did not vary by any learning history factors

After recording was complete, sensors were removed from the fingertips and participants were again asked to listen to all items, this time judging whether any items were odd, semantically/grammatically wrong, recorded with incorrect tone, or deviated from the participants' own understanding of standard colloquial usage. Because items were rated as problematic generally by only one or two participants, all items were retained. However, electrodermal data for problematic items (for a specific participant) were excluded from analysis.

## Results

*Preliminary analysis.*    Two-way ANCOVAs (language × 5 categories of emotion stimuli) were conducted on emotional intensity ratings and on skin conductance amplitudes. The factors identified during factor analysis were entered as covariates. Due to equipment failure, SCR data could not be used for one participant, leaving 63 participants to be analyzed.

*Emotional intensity ratings of phrases.*    A strong language × emotion interaction was obtained, $F(4,244) = 13.7$, $p < .0001$. Post-hoc comparisons revealed that Mandarin reprimands received higher emotional intensity ratings than English reprimands, $p < .001$, while English taboo phrases received slightly higher ratings than the corresponding Mandarin items, $p < .02$). As shown in Figure 1, items in the other categories received highly comparable ratings. None of the learning history factors influenced emotional intensity ratings.

Perceiving childhood reprimands from the first language to be emotionally more powerful than those in the second language is intuitively plausible, because of conditioned associations and the emotion-laden context of parents' use of reprimands (as suggested by the contexts of learning hypothesis). Documenting these rating differences is still an important finding. The phrases are mixed in with other emotional phrases and are not identified as childhood reprimands to listeners, nor were research participants ever informed of a category called childhood reprimands.

*Skin conductance responses.*    Not surprisingly, the main effect of emotional phrase category was significant, $F(4,244) = 3.1$, $p < .02$, indicating that SCRs differed by emotional categories. The

**Table 6.** Learning history of participants for two groups displayed in Figures 2 and 3

|  | Balanced bilinguals (N = 46) | Low Mandarin proficiency and use (N = 17) | Statistical difference |
|---|---|---|---|
| Gender | 75% female | 52% female | n.s. |
| Age | 19.2 (18–28)[a] | 19.7 (18–28) | n.s. |
| Age of exposure to Mandarin | 1.3 (0–8) | 1.0 (0–6)[b] | n.s. |
| Age of exposure to English | 7.9 (0–15) | 7.3 (0–15) | n.s. |
| Age at arrival in the US | 11.8 (0–28) | 11.4 (0–21) | n.s. |
| Length of stay in the US | 7.4 (0–19) | 8.0 (0–19) | n.s. |
| Mandarin proficiency (1–7 scale)[b] | 6.3 (4–7) | 4.2 (2–6) | $t(61) = 8.5, p < .001$ |
| English proficiency (1–7 scale)[b] | 5.5 (1–7) | 6.0 (3–7) | n.s. |
| English verbal SAT | 563 | 635 | n.s |
| English TOEFL | 301 | 283 | n.s. |
| Current use of Mandarin with friends (%)[c] | 20% (0–95) | 3% (0–10) | $t(61) = 3.0, p < .01$ |
| Current use of Mandarin at home (%)[d] | 60% (0–100) | 47% (0–100) | n.s. |
| Trilingual | 32% | 52% | n.s. (Fisher exact test, $p = .07$) |
| Other Chinese language is dominant | 9% | 24% | n.s. |

*Notes:* [a] Range provided in parentheses. [b] Average of self-ratings for proficiency in speaking, listening comprehension, reading, and writing. [c] Current use at university with friends. [d] At home with family.

factor representing Mandarin ability/use (which loads on self-reported proficiency and current Mandarin use, see Table 4) was part of a significant three-way interaction with language and emotional phrase category. This interaction occurred because a different response pattern appeared for the three-quarters of the participants with the highest values on the Mandarin ability/use factor compared to those with the lowest Mandarin ability/use. Specifically, those with high Mandarin ability/use had larger SCRs to the English endearments than to the Mandarin endearments. Those with only fair Mandarin ability and low Mandarin current use (and English generally better than Mandarin), showed the opposite: higher SCRs to Mandarin than to English endearments. With Mandarin ability/use factor dichotomized to a high (N = 46) and low (N = 17) group, the three-way ANOVA was significant at $F(4,244) = 2.5$, $p < .05$. Note that the two groups differed primarily on Mandarin ability and current use (see Table 6 for learning history and statistical tests between groups).

Because the low ability/use pattern is an exceptional one and may reflect low familiarity with some items (as explained in the next section), we first explain the more common pattern shown by the high Mandarin ability/use group.

Earlier, we summarized reasons why a set of stimuli could elicit higher SCRs compared to other stimuli, including cognitive effort. Participants' task was to rate the emotional intensity they remembered experiencing on a prior use of the phrase (see section on procedure). This task could have required more effort for some stimuli than others. To guide interpretation, we discuss each condition's reaction times (RTs) together with SCRs.

Figure 2 depicts the pattern of elevated SCRs to English endearments which held for the majority of participants. We first list two reasons that are unlikely to have been operative, followed by
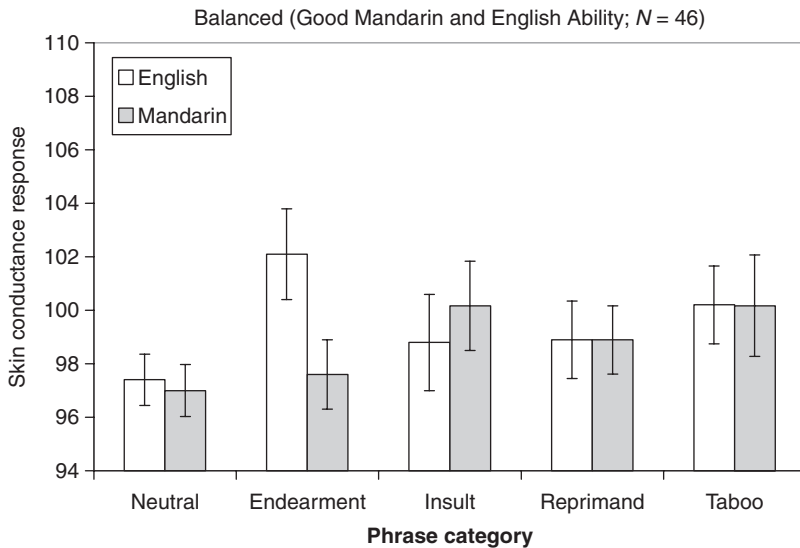
**Figure 2.** SCRs are standard scores for SCR amplitude (mean of 100). The Mandarin dominance factor interacted with emotional category and language. Endearments elicited higher SCRs in those participants with good Mandarin and English ability

two that are more plausible for explaining higher SCRs to English endearments compared to Mandarin endearments.

   We argue that the following are *less* plausible explanations:

- *Conditioned associations between stimulus and threat/punishment.* Implausible because the English endearments would have been largely experienced in positive settings.
- *Cognitive effort during problem-solving.* The instructions to think of a situation in which you heard or used this phrase could have been taxing or anxiety-provoking. Or, trying to evaluate the emotional intensity of the memory could have been difficult. We set this aside as implausible based on rating times. The mean time to rate the English endearments was not longer than for other phrase types. They averaged 4 seconds, the same rating time as the other English phrase types, except for English reprimands, which required 4.2 seconds. Indeed, rating times for the 10 categories (2 languages × 5 phrase types) were relatively uniform, ranging from 3.8 to 4.2 seconds, with no significant differences in rating times, all $ps > .5$.

Could elevated SCRs for English endearments indicate that the endearments were novel, surprising, or unfamiliar? The 46 individuals who were categorized as high Mandarin ability/use also had relatively good English proficiency (5.6 on 7-point scale, with 7 meaning native speaker). The elevated SCRs for English endearments were not disproportionately found among the participants who had late arrival in the US and self-reported only fair English ability. The English endearments are highly familiar, even over-learned phrases to this group of bilinguals.

   We argue that the following are *more* plausible explanations:

English Superior to Mandarin (*N* = 17)
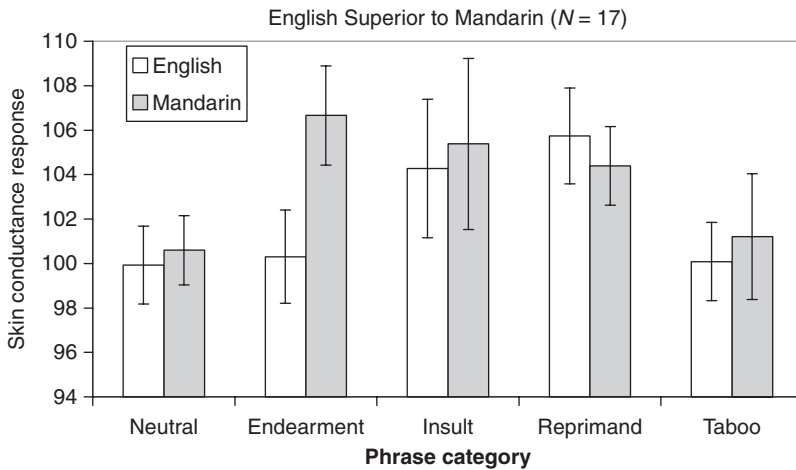
**Figure 3.** A subset of participants reported low Mandarin current use and only medium proficiency ability. Their high response to Mandarin endearments may reflect difficulty in retrieving instances of using the endearments and/or difficulty rating emotional intensity

- *Re-experiencing autobiographical memories*. English endearments may have elicited more autobiographical memories. In the pilot study, interviewees said they felt more at ease expressing emotions in English and some said they had used at least some of these expressions more frequently in English (particularly *I love you*) than the corresponding expressions in Mandarin. This makes it possible that the instructions to think of a situation in which you heard or used this phrase easily elicited relevant memories, which were then accompanied by emotional arousal appropriate to the meaning of the phrase. Mandarin endearments may have elicited fewer autobiographical memories, because of cultural restriction on use, with the result that less emotion was experienced during the memory-retrieval and rating task.
- *Alerting due to relevance*. It is possible that the English endearments could be viewed as more self-relevant because listeners could more easily think of saying the phrase to a friend (or hearing it said in their presence) in their current life living in the US, while the Mandarin phrases may have not elicited a recent personal memory and may have instead elicited a more abstract conceptualization of the meaning of the phrase. However, note that word and phrase frequency does not typically influence skin conductance (Harris et al., 2003). The reason is that SCRs are mainly sensitive to threat and emotional arousal, not the processing fluency that is a hallmark of highly frequent stimuli.

*Low Mandarin ability/use participants.* Figure 3 depicts the pattern for the participants with the lowest Mandarin ability/use: heightened SCRs to the Mandarin endearments, with no other between-language differences.

The heightened reactivity to Mandarin endearments among participants with the lowest Mandarin ability/use suggests that the elevated SCRs observed for Mandarin endearments stems from problem-solving and effort, either to retrieve a memory of a situation or to evaluate it. How words and phrases prompt retrieval of meaning and/or a specific situation is known to vary with

proficiency and language use for bilingual speakers. It is well known that ease of retrieval of a word or phrase and retrieval of autobiographical memories varies for bilingual speakers (e.g., Altarriba & Bauer, 2004; Altarriba et al., 1999). This group reported using Mandarin at home only 5 per cent of their time and had a mean self-rating of 4.0 (7-point scale with 7 = native speaker; compare to values in Table 3). Low use of Mandarin resulted from either identifying another Chinese language as their dominant language, or early arrival in the US.

Reaction times partially support the proposal that infrequent use of Mandarin made retrieving and/or evaluating situations containing the endearments difficult. A 2 × 5 ANOVA on RTs revealed a significant interaction of phrase type and language, $F(4,64) = 4.2$, $p < .001$. Fastest rating times occurred for English, especially endearments and reprimands, 4 seconds each. The longest RTs occurred for Mandarin (specifically endearments, taboo expressions, and neutral expressions, 4.2 s, 4.4 s, and 4.5 s, respectively). A drawback to the effortful explanation of elevated SCRs for Mandarin endearments is that taboo rating times were also long (4.5 s), but Mandarin taboo expressions did not have the same elevated SCRs as the Mandarin endearments. Puzzles thus remain in this subset of participants, but we tentatively conclude that retrieval/evaluation effort combined with low proficiency to produce a particular mix of elevated SCRs and RTs.

*Discrepancies between emotional intensity ratings and SCRs.*    Ratings of the emotional intensity for the English reprimands were lower than for other categories, but no SCR differences were found between English and Mandarin childhood reprimands (compare Figure 3 with Figures 1 and 2). This is puzzling, and is also inconsistent with three prior studies which found decreased SCRs for L2 childhood reprimands (Caldwell-Harris & Aycicegi-Dinn, 2009; Harris, 2004; Harris et al., 2003). (See also a discussion of discrepancies in Havelka & Eilola, in press.)

In those prior tasks, participants rated stimuli for pleasantness or emotional intensity, and could thus tap their cultural and semantic knowledge of the conventional meaning of expressions. These instructions would then be minimally effortful, allowing electrodermal monitoring to be sensitive to automatic emotional associations elicited by the language content. In the current study, effort associated with autobiographical memory retrieval may have increased SCRs in a manner that would cancel out or obscure the expected pattern of reduced SCRs for L2 childhood reprimands. Rating times suggest that evaluating the English reprimands may have been more difficult (4.3 s), than Mandarin reprimands (3.8 s), a statistically significant difference at $p < .01$. An extra half second to rate a phrase indicates more problem-solving, which could bring more effort-associated arousal, and this could elevate SCRs that would have been low due to lack of emotional arousal associated with the meaning of the phrase.

It is worth comparing the pattern of SCRs vs. ratings found in prior studies. Turkish immigrants and sojourners in Boston rated Turkish and English reprimands as having identical unpleasantness values (Harris et al., 2003), as did Spanish–English bilinguals (Harris, 2004). Turks studying English as a foreign language in Istanbul were given yet a different rating task: to rate words and phrases for emotional intensity (without instructions about retrieving a memory and rating the memory). In this study, ratings followed the same pattern as SCRs: L2 reprimands were rated as less emotional than L1 reprimands.

Recommendations can be made from these cross-experiment comparisons. The rating/evaluation task should be made as simple as possible when psychophysiological measures are the primary interest, to allow automatic emotional associations to be measured without interference from task difficulty. This is consistent with the practice of using simple tasks during neuroimaging (Cabeza & Kingstone, 2006). However, when ratings themselves are of interest, then questions about

personal emotional feelings including autobiographical memories may be the most sensitive to language-learning history.

*Could bilinguals' emotional reactions differ because cross-language stimuli were not equivalent?* In an ideal research study, stimuli in the two languages are rated by monolingual speakers and shown to have similar emotional intensity and familiarity. If emotional intensity ratings of these stimuli by bilingual speakers are then found to vary by emotional category and language, this variation licenses inferences about how bilingualism (e.g., learning history factors such as first vs. second language, age of acquisition, proficiency) influences perceived emotionality of specific categories of emotional expression. In contrast, if differences in monolingual ratings are obtained, this may require a return to stimulus selection, to find matched stimuli. But cross-linguistic differences in ratings may be interesting in their own right, indicating cultural differences in how emotional expressions are perceived. We argue that the current study reveals both cultural differences and effects of bilingualism, and we try to briefly justify this with the following discussion.

As we noted when describing stimuli selection, the reprimand and endearments were not perfectly matched according to monolingual emotional intensity norms. English monolingual speakers judged the English reprimands to have lower emotional intensity than did the monolingual Mandarin raters. The bilingual raters agreed with this (see Table 2). Is this because English childhood reprimands are actually less intense than Mandarin reprimands, either in their content or their cultural references? Or did we choose English reprimands that are less intense? Alternatively, the meaning of the expressions could be of similar intensity, but if Chinese parenting is more authoritarian (e.g., Chao, 1994), then phrases may be imbued in memory with more admonishing resonances. Evidence that childhood reprimands elicit low intensity responses in American culture comes from a study of American monolinguals (described in Harris, 2004), where SCRs elicited by reprimands were as low as SCRs elicited by neutral stimuli. The American monolingual speakers in that electrodermal monitoring study reported little distress when hearing childhood reprimands; a few participants mentioned their parents didn't use them much, and some participants even laughed when hearing the reprimands. Interestingly, age of immigration determined the presence of a 'reprimand' effect in an electrodermal monitoring study of Spanish–English bilinguals. Spanish L1, English L2 college student who immigrated early or were born in the US had low reactivity to both Spanish and English reprimands, making them resemble their English monolingual peers. In contrast, bilinguals who lived in Latin America until the teen years did have elevated responses to L1-Spanish reprimands (Harris, 2004). This suggests that reprimands may accrue more emotional force when childhood is experienced in Latin America compared to North America. While we cannot rule out language-history factors (proficiency, usage) as an explanation, this pattern of findings is at least suggestive of an effect of cultural child-rearing practices on perceived emotional intensity. We submit the current findings of a reprimand effect in ratings of Mandarin as likely also reflecting both L1/L2 status and cultural factors.

American monolingual students also rated the endearments as higher in emotional intensity (5.1 on 1–7 scale) than did the Mandarin monolingual students (4.1). The bilinguals rated the endearments similarly in the two languages, but SCRs were higher for English (setting aside the low Mandarin proficiency/use group). The SCRs are thus consistent with the monolingual ratings. As in the case of the reprimands, it is difficult to evaluate whether this means that English endearments are generally more intense than Mandarin endearments, or if cultural factors (greater encouragement of positive emotional expression; Goddard, 2002) imbue English endearments with more intensity.

## Conclusions and future research

We reported results of interviews with a small set of Mandarin–English bilinguals and psycho-physiological monitoring of reactions to emotional expressions with a larger sample. Bilinguals usually prefer to use their more proficient/more emotional language (Dewaele, 2008; Pavlenko, 2005). Exceptions have been generally idiosyncratic to an individual's situation and learning history. However, our interviews revealed a pattern that may transcend individual history and reflect cultural factors. Many Mandarin–English bilinguals speakers preferred to express anger, taboo phrases, and confidences/intimacies in L2-English. Interviewees attributed this to the greater social constraints in Chinese culture to minimize emotional expression (giving explanation like, *It's harder to say things in Chinese*). Future research could determine more fully how cultural conventions about emotional expression interact with learning history factors such as age of acquisition and cultural and linguistic proficiency.

Continuing the method used in prior psychophysiological monitoring studies, international students from Chinese-speaking countries and Chinese-American students listened to (and rated) emotional and neutral phrases in Mandarin and English while electrodermal activity was monitored. Ratings of emotional intensity showed a pattern found in prior work: childhood reprimands heard in L2-English were rated as less emotional than comparable expressions presented in Mandarin. This is consistent with the emotional contexts of learning hypothesis, which states that language comes to be experienced as emotional when it is learned or used in an emotional context (Harris et al., 2006). The context of childhood socialization with parents may be the most emotional context (Bloom & Beckwith, 1989), and thus it makes sense that childhood reprimands in English would not be experienced as emotional.

Respondents' self-reports of their proficiency revealed that, averaged across the 64 participants, English and Mandarin proficiencies were frequently similar (between very good and near-native). This is consistent with the average age of arrival in the US (12.7 years), given prior findings that immigrating between ages 10 and 14 leads to high levels of proficiency in both languages (Caldwell-Harris et al., 2011; Jia et al., 2002). Given these similar proficiencies when averaging across participants, it is thus not surprising that physiological reactivity elicited by English and Mandarin emotional phrases was highly similar. This replicates the finding in Harris (2004) that emotional expressions elicited highly similar SCRs for Spanish–English balanced bilinguals.

A between-language difference in SCRs occurred for the category of endearments, where SCRs varied by Mandarin proficiency/use. For the participants who had good Mandarin proficiency and good English proficiency, English endearments elicited elevated SCRs. For the participants who reported low Mandarin proficiency and very little current use, it was Mandarin endearments that elicited elevated SCRs. This unexpected pattern is a reminder of the difficulties of studying physiological reactions to emotional language, because of the large number of factors influencing the emotional arousal system. In addition to automatic emotional resonances associated with particular expressions, novelty, surprise, familiarity, relevance to the self, and cognitive effort in retrieval and rating can all influence physiological arousal. We took into account the effortful rating task required of participants (*think of a situation when this phrase was used, and rate the emotional intensity of that situation*), and explained the observed pattern as follows:

1. For the low Mandarin proficiency/use group, low use of Mandarin endearments could have made retrieving and/or evaluating an episodic memory difficult. This difficulty resulted in cognitive effort which increased skin conductance.

2. For the high proficiency group, a cognitive effort explanation for the high SCRs to English endearments is implausible because rating times were fast for the English endearments. Instead, this may be a true emotional resonances effect, given that Mandarin–English bilinguals said it was easy for them to express their emotions in English.

These admittedly speculative explanations highlight how many aspects of emotional reactivity remain to be empirically explored and confirmed. Theorists who want to better understand when and why language has emotional resonances could investigate questions such as the following.

## Does ease of retrieval of emotional memories influence impressions of strong emotionality?

Emotional expressions may 'feel more emotional' if they rapidly elicit episodic memories of emotional events. This hypothesis could be tested by timing how long participants require to retrieve a memory of using an expression such as *I miss you* and the rated emotional strength of the retrieved memory. If skin conductance is also monitored, then a finding of rapid retrieval and high skin conductance would support inference #2 above.

The current study identified discrepancies between ratings and psychophysiology. Emotional intensity ratings were not influenced by variables related to learning history (such as age of arrival, usage frequency, report of language dominance), but SCRs to the endearments did vary by Mandarin proficiency and current usage. Future work could examine what factors influence discrepancies between ratings and SCRs. Instructions to rate the pleasantness of a word or phrase will measure knowledge of the semantic denotation and connotations. Learning factors such as proficiency and current use plausibly influence this semantic knowledge, but what level of proficiency is necessary to produce ratings similar to those of monolinguals? This could be investigated by correlating simple judgments of conventional meaning with learning history factors. The current study showed that ratings of childhood reprimands and taboo items differed for L1 and L2, consistent with a role for learning and use. Are there other cases where learning and use factors influence perceived emotionality? Researchers could explore learning a new language with a new romantic partner, rearing children after immigrating to a new country, or arguing with (or receiving praise from) bilingual co-workers.

## Do we know enough about emotion to understand its physiological correlates – especially for bilingual speakers from different cultures?

Interviews have the drawback of being hard to quantify, and self-report surveys have the drawback of being hard to believe (at least by some critics: why would people tell the truth; results are an undecipherable melange of social desirability and the respondents' naive theories). If only we could bypass self-report in order to access true emotion! On this view, physical responses such as neuroimaging or autonomic system measurements (skin conductance, heart rate) can be seen as inherently superior because they bypass self-report.

The current study illustrates how measuring physiology is not the panacea to emotion researchers' woes. We noted the interpretive difficulties that emerged when a relatively simple task used in prior work (rate emotional phrases) was extended to be psychologically richer (rate the intensity of a situation when a phrase was used), and when the population studied was one in which the two languages (in this case, Mandarin and English) are used in cultures with different emotional display rules (Matsumoto, 2001). Nor is this a simple problem of the autonomic nervous system being a less direct route to brain activity than neuroimaging. These labor-intensive and hugely expensive

techniques require very simple tasks. Neuroimaging data are notoriously hard to interpret, as suggested by recent claims that the past decade of findings in social neuroscience is methodologically flawed (Vul, Harris, Winkielman, & Pashler, 2009).

At present, the best approach is convergent methodology: continuing to develop creative experimental tasks drawing on the full theoretical and empirical resources of cross-cultural, cross-linguistic, sociolinguistic, psycholinguistic, and neurolinguistic fields of study, to develop a richer understanding of how language is used and experienced by monolingual, bilingual, and multilingual speakers from around the world.

## Funding

## Note

1.  The eight male students did not differ in their ratings from the 37 female students (higher proportion of females reflects gender ratio in introduction psychology classes).

## References

Altarriba, J. (2006). Cognitive approaches to the study of emotion-laden and emotion words in monolingual and bilingual memory. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 232–256). Clevedon: Multilingual Matters.

Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *American Journal of Psychology, 117*, 389–410.

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context-availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, and Computers, 31*, 578–602.

Altarriba, J., & 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.

Ben-Shakhar, G. (1985). Standardization within individuals: A simple method of neutralizing individual differences in skin conductance. *Psychophysiology, 22*, 292–299.

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.

Boucstein, W. (1992). *Electrodermal activity*. New York: Plenum.

Cabeza, R., & Kingstone, A. (2006). *Handbook of functional neuroimaging of cognition* (2nd ed.). Cambridge, MA: MIT Press.

Caldwell-Harris, C. L., & Aycicegi-Dinn, A. (2009). Emotion and lying in a non-native language. *International Journal of Psychophysiology, 71*, 193–204.

Caldwell-Harris, C. L., & Kronrod, A. (in preparation). Say less, do more: Expressing positive emotions in Chinese culture.

Caldwell-Harris, C. L., Staroselsky, M., Smashnaya, S., & Vasilyeva, N. (2011). Emotional resonances of bilinguals' two languages vary with age of arrival: The Russian–English bilingual experience in the U.S. In P. Wilson (Ed.), *Perspectives on emotion* (Lodz Studies in Language 20). Frankfurt am Main: Peter Lang.

Chao, R. K. (1994). Beyond parental control and authoritarian parenting style: Understanding Chinese parenting through the cultural notion of training. *Child Development, 65*, 1111–1119.

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.

Dahlen, K., & Caldwell-Harris, C. L. (2009). Rehearsal and aptitude in foreign vocabulary learning. Unpublished manuscript.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2000). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 200–223). New York: Cambridge University Press.

Delgado, P., Guerrero, G., Goggin, J. P., & Ellis, B. B. (1999). Self-assessment of linguistic skills by bilingual Hispanics. *Hispanic Journal of Behavioral Sciences, 21*, 31–46.

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

Dewaele, J.-M. (2006). Expressing anger in multiple languages. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 118–151). Clevedon: Multilingual Matters.

Dewaele, J.-M. (2008). The emotional weight of I love you in multilinguals' languages. *Journal of Pragmatics, 40*, 1753–1780.

Dewaele, J.-M., & Pavlenko, A. (2001–3). Web questionnaire bilingualism and emotions. University of London.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*, 203–235.

Goddard, C. (2002). Explicating emotions across cultures: A semantic approach. In S. R. Fussel (Ed.), *The verbal communication of emotions: Interdisciplinary perspectives* (pp. 19–54). Mahwah, NJ: Lawrence Erlbaum.

Gonzalez-Reigosa, 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.

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.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science, 14*, 31–38.

Hall, E. T. (1976). *Beyond culture*. New York: Doubleday

Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development, 25*, 223–247.

Harris, C. L., Aycicegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first than in a second language. *Applied Psycholinguistics, 4*, 561–578.

Harris, C. L., Gleason, J. B., & Aycicegi, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation*. Clevedon: Multilingual Matters.

Havelka, J., & Eilola, T. M. (in press). Behavioural and physiological responses to the emotional and taboo stroop tasks in native and non-native speakers of English. *International Journal of Bilingualism*.

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

Jia, G., Aaronson, D., & Wu, Y. (2002). Long-term language attainment of bilingual immigrants: Predictive variables and language group differences. *Applied Psycholinguistics, 23*, 599–621.

Lykken, D. (1998). *A tremor in the blood: Uses and abuses of the lie detector* (2nd ed.). New York: Perseus.

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.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research, 50*, 940–967.

Marian, V., & Kaushanskaya, M. (2004). Language-mediated self-construal and emotion in bicultural bilinguals. *Journal of Memory and Language, 51*, 190–201.

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General, 129*, 361–368.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224–253.

Matsumoto, D. (2001). Culture and emotion. In D. Matsumoto (Ed.), *The handbook of culture and psychology* (pp. 171–194). New York: Oxford University Press.

Matsumoto, D., Yoo, S. H., & Nakagawa, S. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology, 94*, 925–937.

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

Mesquite, B. (2001). Emotions in collectivist and individualist contexts. *Journal of Personality and Social Psychology, 80*, 68–74.

Pavlenko, A. (1998). Second language learning by adults: Testimonies of bilingual writers. *Issues in Applied Linguistics, 9*, 3–19.

Pavlenko, A. (2002). Bilingualism and emotions. *Multilingua, 21*, 45–78.

Pavlenko, A. (2005). *Emotions and multilingualism*. Cambridge, MA: Cambridge University Press.

Potter, S. (1988). The cultural construction of emotion in rural Chinese social life. *Ethos, 16*, 181–208.

Schrauf, R. W. (2000). Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture and Psychology, 6*, 387–417.

Schrauf, R. W., & Durazo-Arvizu, R. (2006). Bilingual autobiographical memory and emotion: Theory and methods. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 284–311). Clevedon: Multilingual Matters.

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

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

Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology, 74*, 118–128.

Tsai, J. L., & Levenson, R. W. (1997). Cultural influences on emotional responding: Chinese American and European American dating couples during interpersonal conflict. *Journal of Cross-cultural Psychology, 28*, 600–625.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4*, 274–290.

Wilkins, R., & Gareirs, E. (2006). Emotion expression and the locution 'I love you': A cross cultural study. *International Journal of Intercultural Relations, 30*, 51–75.

Yamaguchi, S. (2001). Culture and control orientations. In D. Matsumoto (Ed.), *The handbook of culture and psychology*. New York: Oxford University Press.

Ye, V. Z. (2004). La double vie de Veronica: Reflections on my life as a Chinese migrant in Australia. *Life Writing, 1*, 133–146.