EMOTION RECOGNITION ABILITY IN ENGLISH AMONG L1 AND LX USERS OF ENGLISH[1]

Pernelle Lorette and Jean-Marc Dewaele
Birkbeck, University of London

Abstract

This article focuses on individual differences in emotion recognition ability among 356 first language (L1) and 564 foreign language (LX) users of English. Recognizing emotions can be particularly challenging in LX contexts. Depending on their linguistic profile, individuals may interpret input very differently, and LX learners and users have been found to perform significantly worse than native control groups (Rintell 1984) in tests of emotion recognition ability. In the present article, we investigate the effect of three independent variables, namely, L1 versus LX status, proficiency in English, and cultural background, on emotion recognition ability. We used an online survey in which participants had to identify the emotion portrayed by a native English-speaking actress in six audiovisual clips. Despite LX users having lower proficiency scores, English-L1 users and LX users' emotion recognition ability scores were broadly similar. A significant positive relationship was found between LX proficiency and emotion recognition ability. A similar but only marginally significant relationship emerged among L1 users. A significant effect of L1 culture was found on emotion recognition ability scores, with Asian LX users scoring significantly lower than European LX users. It thus seems that audiovisual input allows advanced LX users to recognize emotions in LX as well as L1 users. That said, LX proficiency and L1 culture do have an effect on emotion recognition ability.

Keywords: Emotion recognition, L1 versus LX users, proficiency, culture

1. Introduction

The present-day world is characterized by high-speed means of transport, widely accessible means of communication and frequent migration. Due to growing globalization, many people communicate with interlocutors who do not share the same first language (L1) (Stavans & Hoffman 2014).[2] The ability to infer the emotional state of one's interlocutor is critical to interpret the content of the interlocutor's utterances. The proposition expressed by a speaker could be interpreted very differently depending

[1] Pre-print version of the paper published in (2015) *International Journal of Language and Culture,* 2 (1): 62–86. DOI 10.1075/ijolc.2.1.03lor

[2] 'L1' refers here to any native language acquired before the age of three and 'LX' is used as a generic term for any language other than the L1, acquired after the age of three (Dewaele 2013; 2005).

on the speaker's affective orientation regarding this proposition (Ochs & Schieffelin 1989). However, as Rintell (1984) pointed out, there seems to exist an implicit social rule prohibiting the direct expression of one's emotional state in certain contexts. Thus, most people tend to unveil relatively indirect cues to their emotional state, rather than describing it directly--especially in the case of negative emotions.

Although every healthy person appears to have some emotion recognition ability, there are individual differences in this ability. This seems to be even more pronounced when people communicate in a language that is not their L1 (Briggs 1970). Both the expression of one's own emotional state and the comprehension of others' emotional states appear to be less intuitive when the communication occurs in a foreign language (LX) than when the communication occurs in an L1. The question is thus whether recognizing emotions in an LX is actually more difficult than in the L1, and if so, what could account for this added difficulty. The value of this type of research is more than purely academic. As Thompson and Balkwill (2006: 421) nicely summarize:

> Understanding how emotion judgments are guided by physical properties of stimuli and cultural norms has valuable implications for cross-cultural communication in many domains, including business, education, and conflict resolution.

The present paper focuses on some factors that have been linked to recognizing emotions, namely the status of L1 versus LX user, LX users' proficiency in the language in which the emotion is expressed, and the participants' L1 culture.

2. Literature review

2.1. Definition of emotion

Dewaele (2006) pointed out that while everybody has an intuitive understanding of what emotions are, their sheer complexity makes them difficult to define. Unsurprisingly, neuroscientists focus on the physiological aspects of emotion (Lindquist, Wager, Bliss-Moreau, Kober, & Barrett 2012), while applied linguists are more interested in the way multilingual individuals communicate their emotions in their different languages (Dewaele 2013, 2015; Pavlenko 2005). A universally accepted definition of emotion does not exist (Pavlenko 2008), but there is a broad agreement that emotions are related to "social systems, psychological systems and biological systems" (Averill 1982: 19) and that "emotions are psychological events that emerge out of more basic psychological operations that are not specific to emotion" (Lindquist et al. 2012: 123). We adopt Keltner and Shiota's (2003: 89) definition because of the focus on channels through which emotion is communicated:

> An emotion is a universal, functional reaction to an external stimulus event, temporally integrating physiological, cognitive, phenomenological, and behavioral channels to facilitate a fitness-enhancing, environment-shaping response to the current situation.

As emotions integrate several channels, it is possible to gather emotional information conveyed by these different channels.

2.2. Channels providing emotional cues

Following Burns and Beier's (1973) pioneering work, channels providing emotional cues can be categorized into three categories: verbal, (nonverbal) vocal, and visual. Verbal cues are conveyed via the actual content of language. Vocal cues might be provided by pitch, rhythm, timbre, speaking rate or intensity, while visual cues might for instance be gleaned from facial expression or body language. As proposed earlier, the identification and interpretation of emotional cues might be more challenging for LX users. First, the extraction of emotional cues amid all the input a recognizer receives during an interaction might require extra attention. Moreover, some bits of information conveyed in the LX by various channels might not be equally relevant in the listener's L1 language/culture. Furthermore, the interpretation of these emotional cues might also be language/culture specific to a certain extent. Several studies have investigated the universal vs. language-/culture-specific character of information conveyed by the different channels. We will review some of them in the following subsections.

Visual channel

The bulk of research into the visual channel has focused on facial expression. Various studies have shown that facial expression conveys crucial emotional information (see, e.g., Mesquita & Frijda 1992). There has been a long-lasting debate about the universal versus culture-specific character of facial emotion recognition ability. Researchers working in Ekman's (1992) tradition have attempted to demonstrate that the facial recognition of so-called 'basic' emotions, i.e., emotions that are assumed to be the result of biology rather than culture, is universal. This position has been strongly criticized by researchers who found an in-group advantage for facial emotion recognition ability, meaning that people from the same culture tend to recognize each other's facial expression better than people from different cultures (see Elfenbein & Ambady 2002 for a review). More recent research suggests that the truth might fall somewhere in between: there is both universality and culture-specificity in recognizing emotions (Matsumoto 2009; Ożańska-Ponikwia 2013).

However, facial expression alone might not be sufficient to infer someone's emotional state, especially in the case of secondary emotions (Sauter 2010). People's general behavior might also provide some visual cues to their emotional state. Again, one's behavior seems to be the product of both (universal) biology and (specific)

culture. Differences in emotional behaviors might, for instance, result from differences in culture-specific "behavior repertoires" (Mesquita & Frijda 1992). These repertoires are derived from culture-specific models and expectations of what is appropriate in certain social circumstances and are learned in childhood. In the Balinese culture, for instance, it is common to fall asleep as a reaction to unfamiliar or frightening events (Bateson & Mead 1942, in Mesquita & Frijda 1992: 196). Westerners may misinterpret this reaction, given the lack of such an association between event and reaction in the West. This relates to cultural differences in behavior regulation of universal repertoires, i.e., Ekman and Friesen's (1969) so-called "cultural display rules." These rules, acquired early in life, limit the range of possible behaviors to the restricted collection of behavior expressions that are appropriate within a particular culture. A striking example of cultural display rules might be found in the community of the Utku Eskimos, who refrain from crying--even young children are instructed not to cry (Briggs 1970). As Irvine (1982: 37) explained:

> To assess affect, […] one must compare the speaker's actual display with what kind of behavior is expectable and appropriate from that kind of speaker in that kind of situation. A single tear from speaker A in situation A could mean as much (in terms of imputable emotion) as noisy weeping from speaker B.

Difficulty in interpreting emotion via the visual channel is not just a problem at an individual level; it can have serious economic consequences. Tombs, Russell-Bennett & Ashkanasy (2014) analyzed the accuracy of 153 service providers' perceptions of emotional expressions of anger, fear, shame, and happiness of customers from different cultural backgrounds in a laboratory setting. Participants viewed video vignettes of customers complaining (without audio) and had to assess the emotional state of the customers portrayed in the video. Participants in culturally mismatched dyads were more likely to misread anger, happiness and shame expressed by dissatisfied customers. Happiness was misread in both culturally matched and mismatched dyads. Anger was recognized by the Anglo but not by the Confucian Asian customers. Anglo service providers also misread both shame and happiness in Confucian Asian customers. The authors point out that emotion recognition ability is crucial for effective functioning of service employee teams in culturally diverse workplaces.

Vocal channel

In line with research into the visual channel, evidence points towards the cross-cultural aspect of vocal emotion recognition ability as well as to a certain in-group advantage.

Sauter, Eisner, Ekman, and Scott (2010) conducted a study in which they compared the recognition of nonverbal emotional vocalizations from 29 Himba

participants--belonging to a seminomadic community in Namibia--and 26 English-speaking European participants. While cross-cultural vocal recognition was found for basic emotions, some secondary emotions were not recognized across cultures. The authors pointed out that the non-cross-culturally recognized emotions were positive emotions. They propose an evolutionary account for this in-group advantage: positive emotions are assumed to enhance social cohesion within a group, it is thus not aimed at non-group members. However, other authors (see Thompson & Balkwill 2006: 420 for an overview) have suggested that the recognition of negative emotions such as sadness or fear might also enhance group cohesion, as sadness signals a need for help, protection or support and fear signals a threat to the other member of the community. It should also be noted that in Sauter et al.'s study the secondary emotions were exclusively positive and that both types of emotions might thus have been confounded.

Zhu's (2013) comparative research on Chinese and Dutch participants also suggests an in-group advantage for positive emotions, while some vocal (prosodic) cues to negative emotions such as anger or sadness could be recognized equally well by participants of different cultures. According to Zhu, this supports the hypothesis that negative emotions, which signal danger, must be recognized by all humans, irrespective of their linguistic or cultural background.

Scherer, Banse and Wallbott (2001) conducted another important study about vocal emotion recognition ability involving participants with various cultural and linguistic backgrounds. The 428 participants had to judge the emotion(s) conveyed in meaningless multilanguage sentences, i.e., artificial sentences constructed on the basis of syllables coming from different languages, pronounced by four German actors. Female recognizers were found to have a significantly higher emotion recognition ability than male recognizers. Moreover, country of origin yielded a main effect. The German participants had the best recognition rate, followed by the French-speaking Swiss, who are likely to be familiar with German prosody since German is the second official language of Switzerland. Next came the British, Dutch and American participants, i.e., speakers of a Germanic language, then the Italian, French and Spanish participants, i.e., speakers of a Romance language. Lastly were the Indonesian participants, i.e., speakers of a language supposed to be the most distant from German among the languages included in the study. These results suggest an effect of linguistic and/or cultural distance on (vocal) emotion recognition ability in an LX. The authors did acknowledge that the design of their study did not allow them to draw strong conclusions about linguistic distance.

Thompson and Balkwill (2006) conducted a study in which 20 L1 English listeners had to judge the emotions conveyed in semantically-neutral utterances produced by English, German, Chinese, Japanese and Tagalog speakers. Although the recognition rate for each emotion--joy, sadness, anger and fear--was above chance level (25%) in all languages (mean correct answers = 70%), the L1 English listeners

recognized the emotions expressed by English speakers better than by speakers of other languages. The authors argue that their data provide evidence that cultural distance might affect emotion recognition ability, since the least accurate answers appeared for Chinese and Japanese speakers (54% and 59% correct respectively). However, Pell, Monetta, Paulmann, and Kotz (2009) challenged this claim, arguing that Thompson and Balkwill's data did not strongly support an effect of linguistic and/or cultural distance on emotion recognition ability, since the listeners were not significantly more accurate in their emotion identifications in German (67.5% correct) than in Tagalog (72.2% correct). Furthermore, Pell and colleagues' results did not reveal any relationship between linguistic distance and emotion recognition ability.

Verbal channel

Obviously, one can also rely on the actual content of an utterance to infer the interlocutor's emotional state. However, one's understanding of terms, especially abstract ones, can (slightly) differ from someone else's understanding, particularly in multilingual contexts. As Jakobson (1959) originally argued, there is no perfect equivalence between a 'code-unit' in language A and its translation in language B--an idea supported by Pavlenko (1999). Jakobson argues for a distinction between three levels in each lexicalized concept:

> - *a lexical component*: a word form with its phonological and morphophonological properties (…);
> - *a semantic component*: explicitly available information which relates the word to other words, idioms and conventionalized expressions in the language; it is characterized by polysemy (…);
> - *a conceptual component*: non-linguistic multimodal information, which includes imagery, schemas, motor programs, and auditory, tactile and somatosensory representations, based on experiential world knowledge (…) (Pavlenko 1999: 211--212).

In later work, Pavlenko (2008: 15) defined emotion concepts as:

> prototypical scripts, formed as a result of repeated experiences that involve causal antecedents, appraisals, physiological reactions, consequences, and means of regulation and display. These concepts are embedded in larger systems of beliefs about psychological and social processes, often viewed as cognitive models, folk theories of mind or ethno-psychologies.

Accordingly, an emotion word in language A and its 'translation equivalent' in language B will activate different conceptual representations. Therefore, an LX user might be confused when interpreting verbal emotional cues in the LX, being unfamiliar

with the LX emotion concept. Pavlenko (2008) differentiates three categories of emotion words. She defines 'emotion words' as "words that directly refer to particular affective states (…) or processes" (Pavlenko 2008: 148), e.g., *happy*, and *to worry*. Contrariwise, 'emotion-related words' do not directly refer to particular emotions but describe behaviors related to them, e.g., *tears*, and *to scream*. Finally, 'emotion-laden words' express or elicit emotions from the interlocutor without directly referring to any emotion, e.g., *loser*, and *cancer*. As Pavlenko (2014) points out, the level of differentiation of emotions differ between various languages. For example, German has three terms corresponding to the English *anger* while Pintupi has only one word for both *shame* and *embarrassment* (Myers 1979). This might cause misunderstandings of emotion words for LX users, and thus confusion in emotion recognition.

Predominance of a particular channel?

Depending on the situation, some channels might provide more relevant emotional information than others. In their pioneering study, Burns and Beier (1973) demonstrated the predominance of visual cues over vocal cues for gathering emotional information. However, some findings suggest that the dominance of a particular channel might be mediated by the emotion to be recognized, as well as by the recognizer's native language and/or culture (e.g., Paulmann & Pell 2011; Riviello, Esposito, Chetouani, & Cohen 2011). Research has shown that members of individualist cultures strive to be self-sufficient and autonomous, and tend to express their emotions freely and frequently. In contrast, members of collectivist cultures, with interdependent selves, show much more emotional restraint (Markus & Kitayama 1991). Cultural differences also seem to affect strategies used in the recognition of the emotional state of an interlocutor. Tanaka, Koizumi, Imai, Hiramatsu, Hiramoto, and de Gelder (2010) found that cultural background determined how much attention participants paid to the voice and the face of interlocutors in their perception of emotion. Japanese participants paid more attention to vocal processing in the multisensory perception of emotion, while Dutch participants focused more on facial expression (2010: 1259).

Different channels can also provide incongruent information, making emotion recognition trickier. Mehrabian and Wiener (1967) found that when confronted with incongruent vocal and verbal cues, people generally rely on vocal rather than verbal cues. For instance, if *scram*--a word with a negative valence--was expressed with a positive tone of voice, the total communication was generally interpreted as positive. More recently, Collignon and colleagues' (2008) research on fear and disgust led to some unexpected results. Their 32 participants generally relied more on visual cues, which the authors interpreted as a confirmation of visual dominance in emotional processing. However, participants focused more on the vocal channel when the quality of the visual stimuli was diminished.

This shows that individuals are able to rely on different channels depending on the quality and the congruence of the information conveyed by these channels. It thus

seems that in cases where the verbal channel does not convey crucial cues, recognizers will shift their attention to other channels, such as visual or vocal channels, which most people find more difficult to control.

The effect of L1 user versus LX user on emotion recognition ability

Rintell (1984) focused on LX users' verbal-vocal emotion recognition ability. Participants included 19 L1 speakers and 127 LX learners/users of English (with either Arabic, Chinese or Spanish as their L1). They had to identify the emotion expressed by one of the conversational partners in 11 recordings in which both basic and secondary emotions were included "to a noticeable degree" (p. 258). While age and gender did not appear to have an effect on emotion recognition ability, Rintell found that the control group of L1 English-speaking Americans listening to the tape recording of the conversation recognized the emotions significantly better than the LX users. Moreover, LX users with better English proficiency had better emotion recognition ability than less proficient LX users. Besides, when comparing the three biggest language groups represented among the LX users, Chinese speakers scored significantly lower than Spanish and Arabic speakers. Rintell interpreted this as an effect of cultural distance.

Graham, Hamblin, and Feldstein (2001) conducted a similar study, also dealing with LX users' verbal-vocal emotion recognition ability. The same monologue was presented by L1 English actors, each time with a different emotion, including both basic and secondary emotions. The participants, comprising 85 L1 American-English speakers, 45 Japanese and 38 Spanish LX users of English, had to identify the emotion portrayed in each recording. As in Rintell's (1984) study, L1 speakers outperformed LX users (with an average rate of correct identification of 59% across all conditions compared to 42% for Spanish and 38% for Japanese LX users). Graham and colleagues' results also seem to point to a linguistic and/or cultural distance effect, as the confusion patterns of English speakers were more similar to those of Spanish-native LX users than those of Japanese-native LX users of English. However, in contrast to Rintell's findings, the difference between the Spanish and Japanese LX users' emotion recognition ability was not significantly different. Further, Rintell's proficiency effect was not replicated in this study. The authors interpreted this as an indication that the ability to identify and interpret vocal emotional cues in LX speech might not be acquired without extensive natural exposure to the LX in a native context. However, Zhu's (2013) results contradicted this last hypothesis, as the 20 Dutch-native LX Chinese users in her study had a significantly better vocal emotion recognition ability in Chinese (54% correct) than the 20 Chinese L1 speakers (46% correct), even though the LX users had not been extensively exposed to a Chinese-speaking environment.

The results of a study conducted by Dromey, Silveira, and Sandor (2005) suggest that being a native speaker of a language does not guarantee higher scores on emotion recognition ability in that language. In this study, 32 polyglot L1 English speakers performed marginally better than 57 LX English users, but 53 monoglot L1 English

speakers did not outperform LX users in detecting affective prosody at the single word level. The authors postulate that the learning of an additional language might develop "additional sensitivity to certain aspects of speech that carries over to native language tasks" (2005: 356). The authors do admit that education might have been a confounding factor, as the polyglots were more highly educated. In all three groups, female participants outperformed their male peers.

Implications for the present study

In summary, L1 and LX users might be confronted with a range of emotional cues provided by several channels. In the present study, we will examine how they deal with situations in which verbal, vocal, and visual information is available simultaneously. The rationale behind this focus on multisensory perception is that previous studies investigating nativeness, proficiency and culture have focused on situations in which emotional cues were integrated in only one or two channels at a time. However, emotion recognition is not always facilitated by simultaneous delivery of information across three channels—think, for instance, of incongruence between verbal cues and vocal or visual cues. As people are simultaneously confronted with verbal, vocal and visual information in most day-to-day situations, we have chosen to examine such situations to strengthen the ecological validity of this study.

3. Research questions

The literature review reveals contrasting findings and shows that important questions remain unanswered. To our knowledge, few studies have focused on individual differences in the recognition of basic emotions by L1 and LX users when the visual, vocal, and verbal channels are simultaneously available. Therefore, we will address the following research questions:

1. Are there differences between *L1 and LX users of English* in their ability to recognize a (basic) emotion conveyed by a L1 British English speaker?
2. Is *proficiency in English* linked to the ability to recognize a (basic) emotion conveyed by a L1 British English speaker?
3. Is a participant's *L1 culture* linked to their ability to recognize a (basic) emotion conveyed by a L1 British English speaker?

4. Methodology

4.1. Participants

A total of 920 participants (687 females, 233 males) participated in the study. Complete demographic information is presented in Table 1.[3]

---

[3] A preliminary analysis revealed no interesting effects for gender and age of participants, so these two independent variables were left out of the present study.

Table 1. *Participants' demographics (n = 920).*

	L1 users of English (n = 356)		LX users of English (n = 564)	
	Mean (SD)	Range	Mean (SD)	Range
Age	33.3 (16.8)	11--82	30.5 (11.1)	15--70
Proficiency (%) (as measured by a lexical test--see next section)	94.6 (5.7)	80--100	83.4 (11.3)	45--100

The relatively high score for English proficiency among LX users is not surprising considering that the questionnaire was formulated in English, thus requiring sufficient proficiency to fill it in. However, the proficiency scores of the L1 users were significantly higher than those of the LX users, with a smaller standard deviation (see Table 1).

English L1 users mainly originated from the United Kingdom (UK) and Ireland (n = 167), followed by North America (n = 78). The largest group of LX users were Slovenian (n = 120), followed by Belgian (n = 89) and Dutch (n = 70). One hundred and one participants reported more than one nationality. Including only the single-nationality participants, we divided the sample into groups according to nationality. Considering only those groups forming clusters with a substantial number of participants, we will analyze the results of participants from Britain and Ireland (n = 160 L1 English speakers + 2 LX English users), North America (n = 84 L1 + 2 LX), continental Europe (n = 21 L1 + 441 LX), and Asia (n = 27 L1 + 48 LX).[4] We refer to these clusters as participants' 'culture'. We are fully aware that this categorization cannot be watertight, as "cultures are portable schemas of interpretation of actions and events that people have acquired through primary socialization and which change over time as people migrate or enter into contact with people who have been socialized differently" (Kramsch 2015: 638). In other words, we assume that participants in a cluster may have been through comparable primary socialization, while admitting that they may have been through secondary socialization, which could have diluted their original culture and made them culturally hybrid.

4.2. Instrument

An online survey was developed and pilot tested twice with 20 participants. This led to the reformulation of some instructions. The research design and questionnaire obtained approval from the Ethics Committee of the School of Social Science, History and Philosophy at Birkbeck, University of London.

[4] Participants from Australia and New Zealand, from Central and South America, from Africa and from the Arab World were not included because of very small sample sizes.

Firstly, participants had to answer some questions about their social and linguistic background: gender, age, nationality (or nationalities), actual country of residence, L1(s) and information about the acquisition, use and self-rated proficiency in LX(s) if applicable.

Secondly, participants were presented with six audiovisual stimuli, each between 30 and 55 seconds long (see the Appendix for a transcription). To ensure some homogeneity amongst the video clips and to have maximum control over the stimuli, we constructed videos especially for this study instead of using natural material from existing films or sitcoms. In each video, a 43-year-old professional actress improvised a short sketch displaying a particular emotion.[5] Audiovisual stimuli have been chosen to elicit data since they correspond to typical day-to-day situations in which one must rely on simultaneous visual, verbal and vocal cues to gather emotional information. As the research questions required controlled stimuli, we chose to involve a professional actress for the recording of the stimuli. This actress had an interesting linguistic background for our purpose: her accent was not strongly associated with a particular region, as she was born in Canada, was brought up bilingually in Latvian and English, then settled in London. She defined herself as a currently monolingual English speaker using a British Received Pronunciation (RP) accent with shades of London. The choice of an actress rather than an actor was based on the finding that females' emotional states are generally better recognized than males' emotional states (Matsumoto et al. 2000; Scherer, Banse, & Wallbott 2001). The actress was provided with brief example scenarios related to each basic emotion--happiness, sadness, anger, fear, surprise, and disgust. She could choose to improvise small situations based on these scenarios or based on her own ideas. This absence of strict guidelines was meant to strengthen the authenticity of her play. Just as in real life, participants should not always focus on the same channel to the same extent: in some videos the verbal channel was the most informative, whereas in others cues were particularly conveyed via the vocal and/or visual channels. The URLs of the 6 recordings are included in the Appendix.

The third part of the survey consisted of the English version of the LexTALE, a 60-item lexical test developed by Lemhöfer and Broersma (2012). Participants had to decide whether each item was an existing (British) English word or not. The test has been shown to reliably measure the lexical proficiency of learners with different cultural and linguistic backgrounds (Lemhöfer & Broersma 2012). Moreover, it also gives a good indication of overall English proficiency, at least for intermediate and advanced English learners (Lemhöfer & Broersma 2012).

4.3. Procedure

We used snowball sampling to recruit participants; in other words, the call for participation in the survey was forwarded to friends and colleagues around the world,

[5] Karen Glossop from the Wishbone Theatre in Camden.

and put on various social media and mailing lists, including Linguistlist. Time on task was not limited. About 150 incomplete submissions were deleted.

5. Results

We first calculated the participants' emotion recognition ability score by assigning one point for each video in which the intended emotion had been identified correctly. Each participant could thus score between 0 and 6. The average emotion recognition ability score was 4 (see Figure 1).

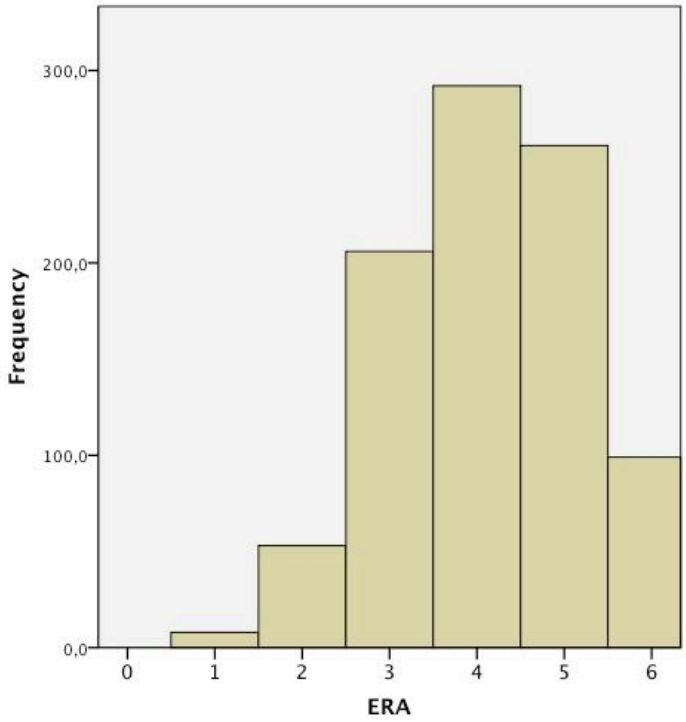

Figure 1. *Bar plot showing the number of participants who obtained a specific emotion recognition ability score (n = 920). No participant recognized less than one emotion.*

The best recognized emotion was disgust (95% correct), followed by fear (80%), surprise (67%), happiness (66%), sadness (54%) and anger (51%) (see Table 2).

Table 2. *Response patterns for each stimulus in order of appearance in the survey (n = 920).*

	Responses							
	Disgust	Anger	Happiness	Fear	Surprise	Sadness	Neutral Emotion	No Idea
Intended emotion								
Disgust	**95,11%**	0,22%	0,11%	0,98%	0,22%	1,30%	1,74%	0,33%
Anger	1,20%	**51,20%**	6,74%	1,09%	3,26%	5,98%	28,15%	2,39%
Happiness	0,54%	11,74%	**66,63%**	2,07%	2,61%	0,33%	9,24%	6,85%
Fear	0,11%	0,33%	0,00%	**79,57%**	0,00%	19,02%	0,87%	0,11%
Surprise	0,11%	0,11%	32,50%	0,00%	**66,74%**	0,22%	0,11%	0,22%
Sadness	0,33%	6,52%	0,87%	2,07%	0,98%	**53,80%**	32,93%	2,50%

The visual channel of the stimuli depicting disgust was quite obvious, which might account for its high recognition rate. Sadness and anger might have been recognized less easily because they were arguably acted more subtly than the other emotions--see the high rate of 'neutral emotion' responses-. Further, these were the only two stimuli in which the verbal channel was very similar: the last video clip depicted sadness, but the actress used almost the same monologue as in the second video, which depicted anger. Therefore, we analyzed the data several times: once considering all six emotions, once ignoring the responses for the video depicting disgust and once ignoring the responses for the video depicting sadness. These permutations had no effect on the results, which convinced us that the internal consistency for the six emotions is sufficient. In the following sections, we present the analysis of emotion recognition ability scores including all six emotions.

Since the data are not normally distributed (Kolmogorov-Smirnov $D(356) = .18$, $p < .001$ for the L1 users, and $D(564) = .17$, $p < .001$ for the LX users), a (nonparametric) Mann-Whitney U test was conducted with *group* (L1 users versus LX users of English) as the independent variable, and emotion recognition ability score as the dependent variable. No significant difference was found between the groups ($p = ns$). In both groups, participants recognized four emotions on average and half of the participants recognized three to five emotions. The range is somewhat broader for the LX users (one to six emotions correctly recognized) than for the L1 users (two to six emotions correctly recognized), but the longer lower whiskers for the LX users is due to only 8 LX users who recognized fewer than two emotions (see Figure 2). These 8 LX users did not stand out by a particular characteristic measured in our study that could account for their low emotion recognition ability score. Their average proficiency score was not particularly low, ranging from 54% to 96% (mean = 80, SD = 15). Moreover, their origin does not seem to relate to this low emotion recognition ability score. Three

of the 8 LX users came from a non-English speaking European country, four came from an Asian country, and one did not provide information about their origin. All began the acquisition of English between ages 3 and 15. The context of acquisition and the frequency of use of English do not seem to account for the low emotion recognition ability scores of these participants. Five of them had not spent any time in an English-speaking country, one had spent half a year, one had spent 6 years and one had spent 40 years in an English-speaking country.

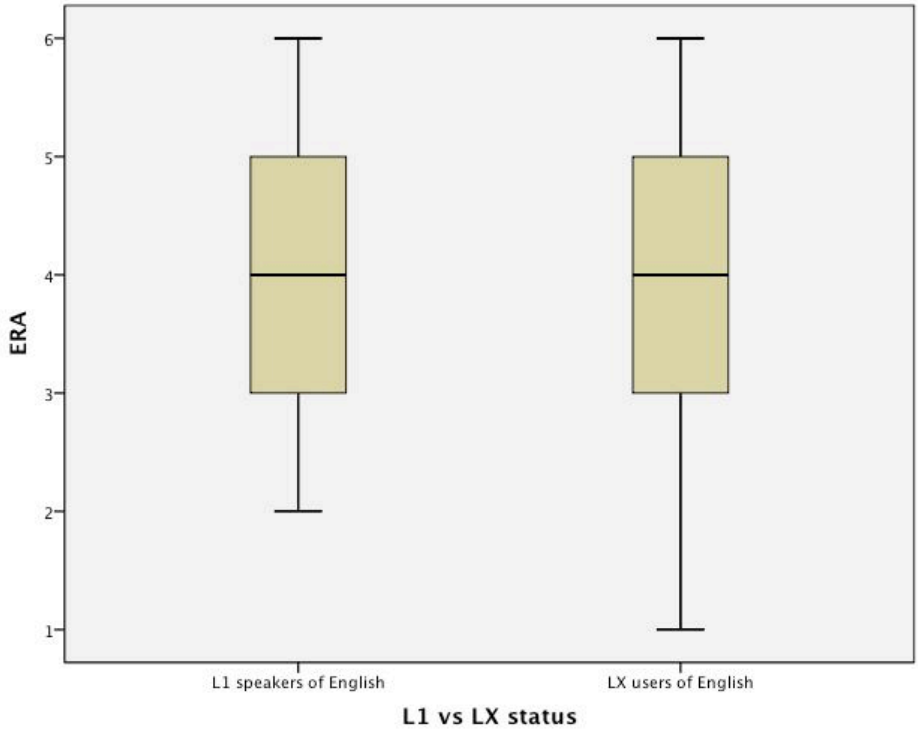

Figure 2. *Boxplot showing emotion recognition ability scores broken down by group (n = 920).*

5.1. Proficiency

Two separate Spearman's correlation analyses between emotion recognition ability score and proficiency in the L1 and LX groups showed a significant positive relationship for the LX users ($N = 564$, $\rho = .09$, $p < .04$) (see Figure 3) and a marginally significant positive relationship for the L1 users ($N = 356$, $\rho = .10$, $p < .055$) (see Figure 4). In other words, more proficient LX users were more likely to have higher emotion recognition ability scores, and the same relationship -though only marginal--existed for L1 users.

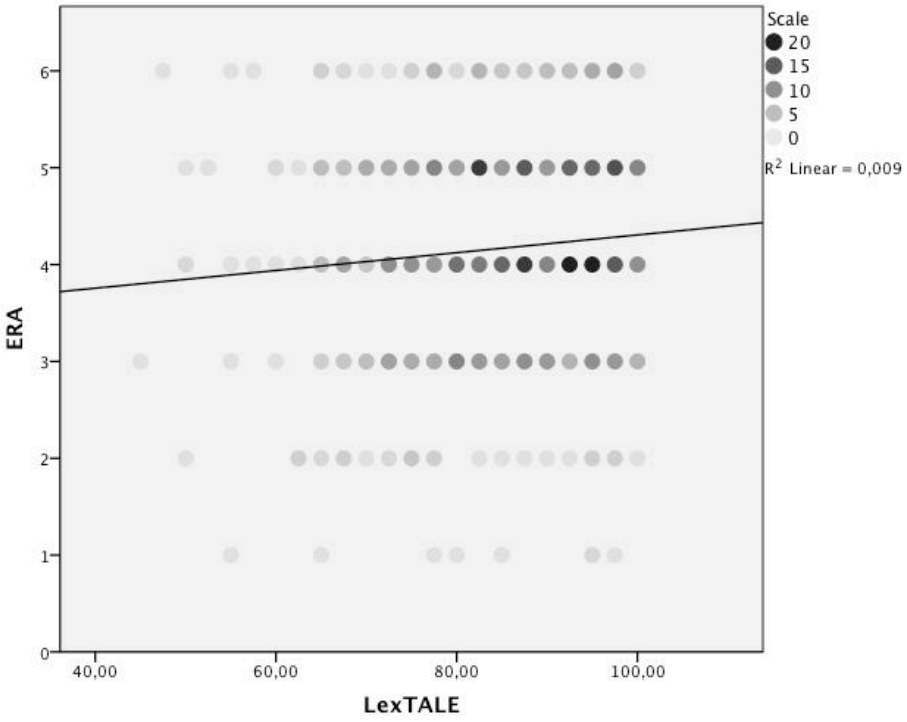

Figure 3. *Scatterplot of LX users' LexTALE scores (max. 100) and emotion recognition ability scores with fit line (n = 564).*

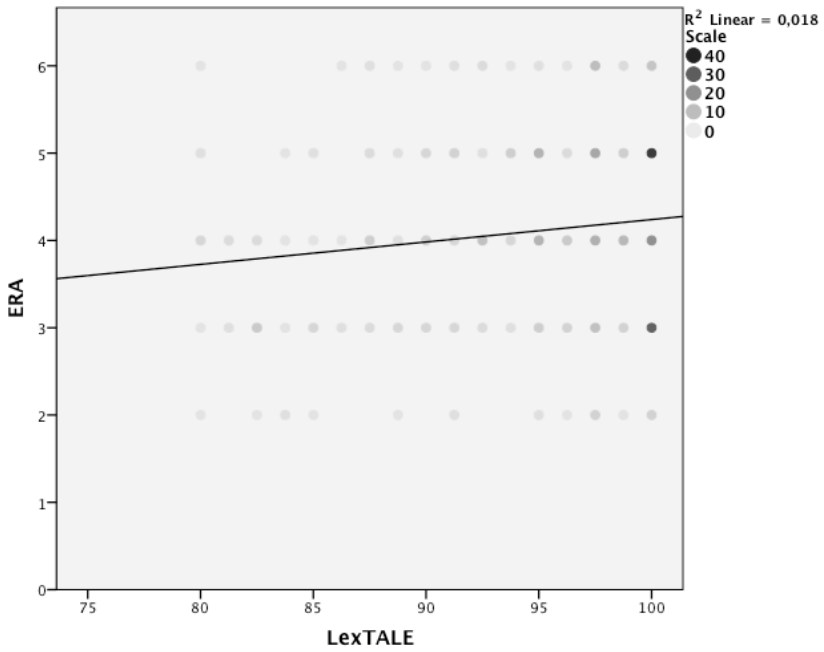

Figure 4. *Scatterplot of L1 users' LexTALE scores (max. 100) and emotion recognition ability scores with fit line (n = 356).*

L1 culture

A Kruskal-Wallis test with Emotion recognition ability score as the dependent variable and L1 Culture as the independent variable turned out to be significant ($\chi^2 = 30.7$, $p < .001$). A series of post-hoc Mann Whitney tests indicated that Asian participants experienced significantly more difficulty in recognizing the intended emotions than British or Irish ($U = 4049$, $r = -.28$, $p < .001$), North American ($U = 2125.5$, $r = .30$, $p < .011$) or continental European ($U = 10658$, $r = -.24$, $p < .001$) participants (see Figure 5).

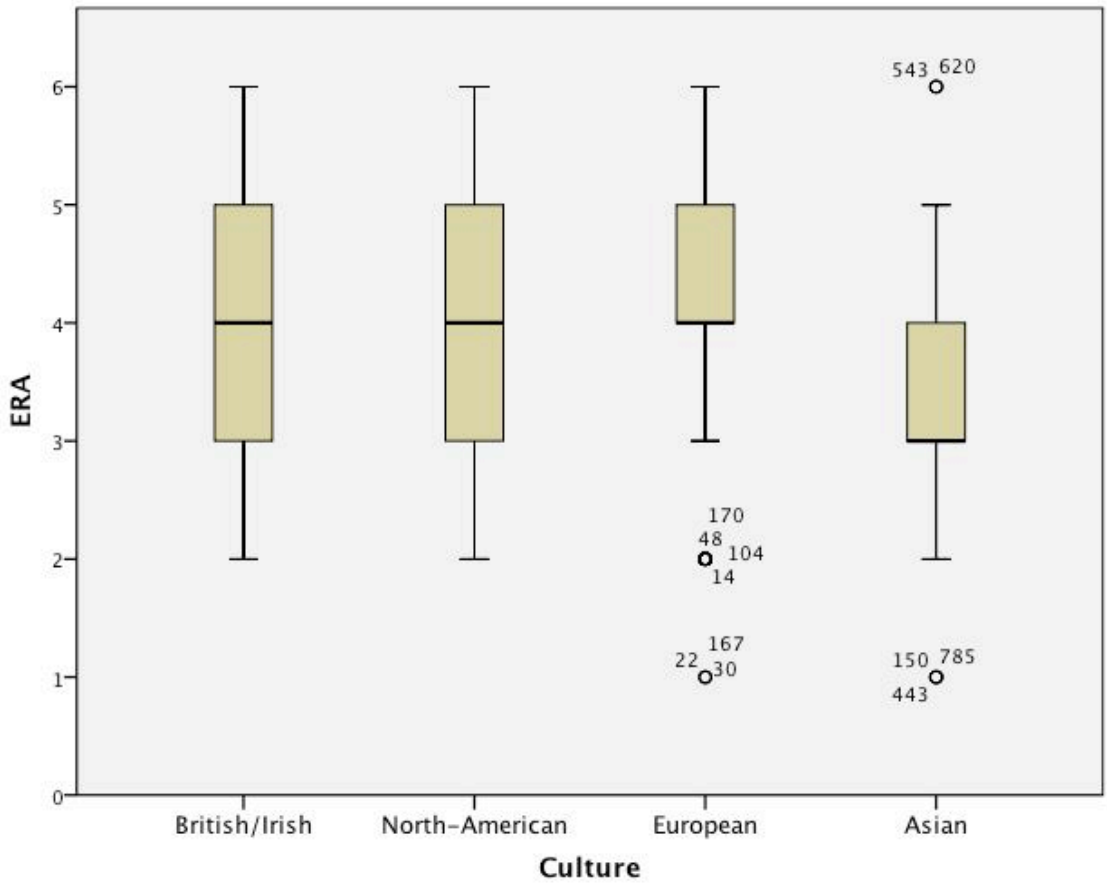

Figure 5. *Boxplot showing the emotion recognition ability scores for British/Irish, North-American, Continental European and Asian participants (n = 785).*

6. Discussion

Our findings revealed mostly similar scores for L1 and LX users' emotion recognition ability, which is surprising given previous findings (e.g., Graham et al. 2001; Rintell 1984). It is important to note that Rintell (1984) and Graham and colleagues (2001) investigated both basic and secondary emotions, while we focused solely on basic emotions. A number of possible reasons can be imagined for the different results of the current study. Firstly, Dromey et al.'s (2005) results suggest that the number of spoken languages seems to have a stronger effect rather than L1 versus LX status in determining emotion recognition ability. Our LX users were more multilingual than the L1 users (an average of 3.6 versus 2.9 spoken languages respectively). However, a quick additional Spearman correlation analysis showed no relationship between number of languages known and emotion recognition ability scores ($N = 919$, $\rho = .02$, $p = ns$). Secondly, our LX users sample was mainly made up

of authentic English LX *users*, while Rintell's (1984) and Graham and colleagues' (2001) samples consisted entirely of English LX *learners* enrolled in relatively short intensive English as a Foreign Language (EFL) courses. It is thus very likely that our LX users had more experience in authentic interactions in the LX. Indeed, more than two-thirds of our LX users sample had already had more than one year of contact with L1 users of English, and more than half of the LX users had already lived for at least one year in an English-speaking country. Thirdly, one could argue that the emotions in our stimuli might have been somewhat too easily recognizable, which could have neutralized the expected difference between both groups. However, if that was the case, one would expect most participants to be at ceiling, which is not true for our data (see Figure 1).

Furthermore, a relationship was found between *emotion recognition ability score* and *proficiency*. This suggests that (at least) once an upper-intermediate proficiency level has been reached, an increase in proficiency is linked to an increase in emotion recognition ability. Unfortunately, the design of the present study does not allow us to draw conclusions about LX users with beginner and lower-intermediate proficiency levels. However, we hypothesize that the correlation between emotion recognition ability score and proficiency would be higher had we had more participants with lower proficiency in English. The relationship between these two variables might possibly be of a logarithmic type, with a high(er) correlation among beginners and lower-intermediate LX users, which then weakens as the proficiency level increases. This hypothesis is strengthened by the discovery of a similar--though only marginally significant--relationship between emotion recognition ability score and proficiency among L1 users. This suggests that even though L1 users were significantly more proficient than LX users and that their scores were more narrowly distributed around the (higher) mean, those L1 users who scored higher on the lexical recognition test also tended to do better at recognizing emotions. It is possible that in both groups, those with a higher verbal aptitude (see Abrahamsson & Hyltenstam, 2008), or those with higher levels of emotion perception, which is one facet of trait emotional intelligence (see Petrides, Pita & Kokkinaki, 2007), had less difficulty in recognizing emotions in audiovisual stimuli. The significant difference in proficiency but not in emotion recognition ability between L1 and LX users, and the relationship between proficiency and emotion recognition ability, suggest that the LX users were able to stretch themselves in identifying the emotions presented to them. They may have relied more on visual and vocal cues in order to compensate for their lower lexical knowledge. It may be that LX users were also more highly motivated than L1 users in doing this challenging task well, to show off their ability to recognize emotions in English.

Regarding the effect of L1 culture, our results indicate a trend towards a lower emotion recognition ability (in English) for participants from Asian culture compared to other cultures. This finding echoes the observations by Zhu (2013) about Chinese and

Dutch participants and by Tanaka et al. (2010) about differences between Japanese and Dutch participants in terms of the strategies they used to recognize the emotional state of an interlocutor. This might relate to differences in affective socialization in the East and West (Wang & Ross 2005, see Pavlenko 2014: 268). It is possible that the European LX users of English performed better in emotion recognition because of greater familiarity with how Europeans express emotions. In order to rule out proficiency as the cause of the lower emotion recognition ability scores of the Asian LX users, we carried out a Mann-Whitney U test, comparing the LexTALE scores of the European and Asian LX users of English. No difference was found between the groups ($p = ns$), meaning they were equally proficient in English.

7. Limitations and perspectives for further research

This study has a number of limitations that need to be acknowledged. Firstly, all stimuli have been generated by an actress. Although she is a professional, acted emotions might not always reflect natural, unconsciously encoded emotions. Moreover, with a single actress some of the observed patterns might result from her unique encoding and/or portrayal of emotions (Wallbott & Scherer 1986). Finally, it could be argued that the portrayed emotions were not entirely 'pure'--if such a thing exists.

Secondly, one drawback of the stimuli is that in some video clips the actress acted the emotional event itself, while in others she reported a past emotional event. However, it is not uncommon to experience the reactivation of an emotion while recounting a particular event associated with this emotion. When reporting a past emotional event, the actress made sure to depict a particular emotion, as if she was re-experiencing it while telling the particular event. Therefore, we assume that this limitation had a minimal effect on the results.

Another potential limitation of our instrument is that all the participants viewed the stimuli in the same order. However, precisely because it was the case for all the participants, we postulate that a possible influence of the order would affect the responses of the participants of each group in the same direction. We therefore assume that the relationships we have found between emotion recognition ability and the independent variables are not an artefact of the methodology. Moreover, we have consciously not drawn any conclusion about the better--or worse--recognition of any particular emotion among the cohort.

We also pointed out that the grouping of participants in L1 culture groups was only a rough categorization. We are aware that the lumping together of different nationalities to form bigger entities is inevitably a generalization, dictated mainly by statistical concerns. Finer-grained, more homogeneous categories--assuming a sufficiently large number of participants--could yield richer results.

Lastly, the patterns discovered in our data are limited to the recognition of basic emotions displayed in a 'prototypical' context--without irony, without any attempt to conceal the emotion, and with little or no incongruence across the different channels.

In further research, it would thus be interesting to investigate whether the inclusion of secondary emotions in the stimuli could lead to more striking differences (1) between L1 and LX users, (2) between high- and low-proficiency LX users, (3) between L1 users from various cultures, and (4) between LX users from various cultures.

8. Conclusion

According to our findings, (advanced) LX users of English can generally recognize basic emotions in English speech as accurately as L1 users of English despite lower levels of proficiency. Although LX users' proficiency in English is related to emotion recognition ability in English, these findings suggest that the threshold for the successful recognition of basic emotions in an LX is probably lower than has been assumed so far: one does not need to be fully proficient in order to be able to accurately infer the emotional state of an interlocutor, at least in the case of basic emotions (see Zhu 2013). Interestingly, English L1 users with high proficiency scores tended to score higher on emotion recognition ability, which could hint at variation on some underlying linguistic or psychological dimension. Our findings suggest that a link exists between cultural distance and emotion recognition ability. The Asian participants stood out as having a significantly lower ability to recognize the actress's emotional state than other groups, despite having English proficiency levels comparable to those of Continental European participants.

To sum up, although emotion recognition ability scores for emotions in English were not significantly different for L1 and LX users, we did find an effect of LX proficiency and L1 culture. Emotion recognition ability in multilingual environments merits further investigation, as it is crucial for successful intercultural communication in today's globalized world (Tombs et al. 2014).

Acknowledgment

We would like to thank the reviewers for their excellent comments on a previous version of this paper and our participants for having agreed to fill out our questionnaire.

9. References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481--509.

Averill, J. (1982). *Anger and Aggression: An Essay on Emotion*. New York: Springer-Verlag.

Briggs, J. L. (1970). *Never in anger: Portrait of an Eskimo family*. Cambridge, MA: Harvard University Press.

Burns, K. L., & Beier, E. G. (1973). Significance of vocal and visual channels in the decoding of emotional meaning. *The Journal of Communication, 23*, 118--130.

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research, 1242*, 126--135.

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. (2013). *Emotions in Multiple Languages* (2nd ed.). Basingstoke: Palgrave Macmillan.

Dewaele, J.-M. (2015). Culture and emotional language. In F. Sharifian (Ed.), *The Routledge Handbook of Language and Culture* (pp. 357--370). Oxford: Routledge.

Dromey, C., Silveira, J., & Sandor, P. (2005). Recognition of affective prosody by speakers of English as a first or foreign language. *Speech Communication, 47*, 351--359.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6,* 169--200.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*, 49--98.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*(2), 203--235.

Graham, C. R., Hamblin, A. W., & Feldstein, S. (2001). Recognition of emotion in English voices by speakers of Japanese, Spanish and English. *International Review of Applied Linguistics in Language Teaching, 39*, 19--37.

Irvine, J. T. (1982). Language and affect: Some cross-cultural issues. In H. Byrnes (Ed.), *Contemporary Perceptions of Language: Interdisciplinary Dimensions* (pp. 31--47). Washington, D.C: GU Press.

Jakobson, R. (1959). On linguistic aspects of translation. In R. A. Brower (Ed.), *On Translation* (pp. 232--239). Cambridge Mass.: Harvard University Press.

Keltner, D., & Shiota, M. N. (2003). New displays and new emotions: A commentary on Rozin and Cohen (2003). *Emotion, 3*, 86--91.

Kramsch, C. (2015). Language and culture in foreign language learning. In F. Sharifian (Ed.), *The Routledge Handbook of Language and Culture* (pp. 634--644). Oxford: Routledge.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44*(2), 325--343.

Lindquist, K. A., Wager, T. D., Bliss-Moreau, E., Kober, H., & Barrett, L. F. (2012). What are emotions and how are they created in the brain? *Brain and Behavioral Sciences*, *35*, 175--184.

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

Matsumoto, D. (2009). *Facial Expressions are Not Universal? The study and its flaws…* Retrieved from http://www.humintell.com/2009/09/facial-expressions-are-not-universal-the-study-and-its-flaws.

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., Yrizarry, N., Loewinger, S., Uchida, H., Yee, A., Amo, L., & Goh, A. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior, 24*(3), 179--209.

Mehrabian, A., & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of Personality and Social Psychology, 6*, 108--114.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin, 112*(2), 179--204.

Myers, F. (1979). Emotions and the self: A theory of personhood and political order among Pintupi Aborigines. *Ethos*, 343--370.

Ochs, E., & Schieffelin, B. (1989). Language has a heart. In E. Ochs (Ed.), *The Pragmatics of Affect. A Special Issue of Text* (pp. 7--25). Berlin: Mouton de Gruyter.

Ożańska-Ponikwia, K. (2013). *Emotions from a Bilingual Point of View: Personality and Emotional Intelligence in Relation to Perception and Expression of Emotions in the L1 and L2*. Newcastle: Cambridge Scholars Publishing.

Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion, 35*, 192--201.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition, 2*(3), 209--230.

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

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition, 11*(2), 147--164.

Pavlenko, A. (2014). *The Bilingual Mind and What It Tells Us about Language and Thought*. Cambridge: Cambridge University Press.

Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing emotions in a foreign language. *Journal of Nonverbal Behavior, 33*, 107--120.

Petrides, K. V., Pita, R., & Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology, 98,* 273--289.

Rintell, E. (1984). But how did you FEEL about that? The leaner's perception of emotion in speech. *Applied Linguistics, 5*(3), 255--264.

Riviello, M. T., Esposito, A., Chetouani, M., & Cohen, D. (2011). Inferring emotional information from vocal and visual cues: A cross-cultural comparison. *2nd IEEE International Conference on Cognitive Infocommunications (CogInfoCom),* 1--4.

Sauter, D. (2010). More than happy: The need for disentangling positive emotions. *Current Directions in Psychological Science, 19*(1), 36--40.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences of the United States of America, 107*(6), 2408--2412.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76--92.

Stavans, A., & Hoffmann, C. (2014). *Multilingualism.* Cambridge: Cambridge University Press.

Tanaka, A., Koizumi, A., Imai, H., Hiramatsu, S., Hiramoto, E., & de Gelder, B. (2010). I feel your voice. Cultural differences in the multisensory perception of emotion. *Psychological Sciences*, *21*, 1259--62.

Thompson, W. F., & Balkwill, L.-L. (2006). Decoding speech prosody in five languages. *Semiotica, 158*, 407--424.

Tombs, A. G., Russell-Bennett, R., & Ashkanasy, N. M. (2014). Recognising emotional expressions of complaining customers: A cross-cultural study. *European Journal of Marketing*, *48*(7/8).

Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology, 51*(4), 690--699.

Wang, Q., & Ross, M. (2005). What we remember and what we tell: The effects of culture and self-priming on memory representations and narratives. *Memory, 13*(6), 594--606.

Zhu, Y. (2013). *Expression and Recognition of Emotion in Native and Foreign Speech. The Case of Mandarin and Dutch*. Doctoral dissertation, Leiden University.

10. Appendix

The stimuli can be consulted on the following links. They have been transcribed below, in order of appearance in the survey. The instruction inviting the participants to choose a particular emotion was formulated as such:

"Please choose the answer that you think applies most to the video above. The main emotion expressed by the speaker is: anger - happiness - surprise - disgust - sadness - fear - neutral emotion - no idea."

1) Disgust: http://www.youtube.com/embed/rH6evcth9Vs

So, Jerry, you wanted to discuss the proposal that I put together for the two separate groups. You, you, you've got something... Kind of... No, no, it's not... It's sort of there. No, it's still there. It's now dripping down a little bit. Maybe if you use a napkin somewhere that you could wipe it with.

2) Anger: http://www.youtube.com/embed/8VcoNbk3HVE

Yesterday, I went to see my mother-in-law. It was actually her birthday the day before yesterday, but I couldn't go because I had a business meeting. And I bought her a very nice bunch of flowers. Very nice. And when I got there, she said, "What is this about?" And I said, "Well, it is your birthday, Maria. Happy birthday!" And she said, "It's not my birthday; it was my birthday yesterday." So anyway, I really hope she liked the flowers.

3) Happiness: http://www.youtube.com/embed/x1S3IzTmf6A

So, I went to my Pilates class after a really long time of absence of a few weeks, which you start to really notice if you haven't been. But the teacher is absolutely amazing. What she's really into is torturing us, basically. And she, she wants you to work really, really hard. And she says, "Oh, when I'm coming in a... You know, if I am in a bad mood, if I see you there and I can hear you groaning a little bit, and gasping and running out of breath, then I think "Brilliant, I'm really getting them to do some good work."

4) Fear: http://www.youtube.com/embed/T_5uBEYC8Wc

So, I've got quite bad back pain, and it's been like that for about three weeks. It's really on my right side. And I suppose what I want to know is what... what it... you know, because I've tried doing some stretching but they haven't... haven't really worked at all. And I just kind of wondered whether you could tell me if you could exclude some things that it could be. It's just that I know that one of the indications is some kind of... I know this sounds stupid but... some indications of... And I know I'm probably fine but... some indications of... of... of... certain kinds of cancers can be... to do with back pain. And that's kind of when... I don't know, if you could just, kind of, eliminate it, that would be really helpful.

5) Surprise: http://www.youtube.com/embed/rHuCJ6rojzE

So this is like a really beautiful restaurant. It's just really, really nice, and... I just, you know, kind of... Oh my god! Really? Yes, okay!

6) Sadness: http://www.youtube.com/embed/B-k3ivqrVDw

So, yesterday, I went to see my mother in law. It was actually her birthday the day before that and I actually couldn't go. I was, you know, away working. So, I went

the following day. And I bought her some flowers, and gave her the bouquet. And she was asking me why I bought her some flowers. And I said, "Well, because it is your birthday, Maria." And she said, "No, it isn't; it was my birthday the day before." So, yeah, well anyway, I really hope she liked the flowers.

Authors' addresses
Jean-Marc Dewaele
Department of Applied Linguistics and Communication
Birkbeck, University of London
26 Russell Square
London WC1B 5DT
UK
j.dewaele@bbk.ac.uk

Pernelle Lorette
Rue de la Maladrée, 17
6181 Gouy-lez-Piéton
Belgium
pernellelorette@hotmail.com

About the authors
Jean-Marc Dewaele is Professor of Applied Linguistics and Multilingualism at Birkbeck, University of London. He has published widely on individual differences in psycholinguistic, sociolinguistic, pragmatic, psychological and emotional aspects of Second Language Acquisition and Multilingualism. He is author of the monograph *Emotions in Multiple Languages* (2010). He is Vice-President of the International Association of Multilingualism, Convenor of the Research Network Multilingualism within AILA, and former president of the European Second Language Association. He became General Editor of the *International Journal of Bilingual Education and Bilingualism* in 2013.
Pernelle Lorette studied Germanic languages and literature at University of Namur, Belgium and graduated in 2014 from the research master in Linguistics at Utrecht University, the Netherlands. She has carried out research projects on factors influencing second language acquisition/use. She is currently starting her PhD project at Birkbeck, University of London.