BIROn - Birkbeck Institutional Research Online

The relationship between bi/multilingualism, nativeness, proficiency and multimodal emotion recognition ability

Pernelle Lorette and Jean-Marc Dewaele

Abstract

Aims and Objectives: The primary aim of this study is to investigate the relationship between visual–vocal–verbal emotion recognition ability (ERA) and multilingualism – that is, both bilingual first language (L1) acquisition and the level of multilingualism operationalised as the number of languages one can speak. Besides these two factors, we also consider nativeness and proficiency as possible predictors of ERA. *Methodology:* One-thousand-two-hundred-and-twenty participants completed a survey online consisting of a sociobiographical background questionnaire, an English lexical test and an emotion recognition test including six stimuli. For each of the six audiovisual recordings, participants had to indicate which emotion they thought the L1 English speaker was conveying – happiness, sadness, anger, (positive) surprise, fear, disgust or no/neutral emotion. *Data and Analysis:* An individual ERA score was calculated for each participant. Correlations between ERA and the different variables were computed – including interactions – and significant correlations were fed into a linear regression model. *Findings:* The number of spoken languages was unrelated to ERA in our sample. The data revealed an interaction between BFLA and nativeness: bi/multilingually raised English second or foreign language (LX) users outperformed monolingually raised LX users, but bi/multilingually raised L1 users of English scored lower than monolingually raised L1 users. Proficiency was significantly related to ERA. *Originality:* This study points to a bilingual advantage in emotion recognition in English for participants with specific linguistic profiles. Participants who grew up with two languages from birth had an advantage if it did not include English. The advantage seemed to be cancelled out among bi/multilingually raised English L1 users, possibly due to interferences from their other L1(s) or L1 culture(s). *Significance:* This study contributes to the scarce literature on bilingual advantage in the affective domain and offers a nuanced view on bilingualism and ERA.

Keywords : Bilingualism, emotion recognition, proficiency, bilingual advantage

1. Introduction

The putative 'bilingual advantage' has been heavily debated among scholars and their findings have been widely reported in the media. Studies investigating linguistic and cognitive outcomes and meta-analyses of those studies seem to draw divergent conclusions, some of them supporting the bilingual advantage hypothesis (BAH) (e.g. Adesope, Lavin, Thompson, & Ungerleider, 2010) and some of them not supporting it or even suggesting a disadvantage on some dimensions (e.g. Lehtonen et al., 2018). In the present article, we move away from the linguistic and/or cognitive focus dominating the literature so far and investigate whether bi/multilingualism might give a socio-emotional advantage. Specifically, we focus on the ability to perceive how one's interlocutor is feeling. This ability is a crucial aspect of communication, as we interpret the content of our interlocutors' messages differently depending on our interpretation of their emotional state. Yet, some people struggle with this. Several possible factors have been investigated (Lorette & Dewaele, 2015). In this study, we aim to investigate factors possibly linked to emotion recognition ability (ERA) in the simultaneous presence of visual, vocal and verbal cues. We consider bilingual first language acquisition (BFLA) and the level of multilingualism – operationalised as the number of languages one can speak – as well as two factors that have been extensively investigated, namely nativeness[1] and proficiency.

2. Literature review

2.1. Definition of emotion

The concept of emotion is conceptualised in different ways depending on the theoretical paradigm. The basic or discrete approach, famously represented by Paul Ekman, regards an emotion as a discrete entity that arises automatically in response to a particular trigger in the environment and brings about specific behavioural and physical reactions (e.g. Ekman, 1992). The well-known 'Basic Six' (Keltner & Cordaro, 2017) – referring to the six emotions of happiness, sadness, sur- prise, disgust, fear and anger – are assumed to be expressed and experienced in similar ways across culture and thus to be universally recognisable, be it from facial (e.g. Ekman, Sorenson, & Friesen, 1969) or from vocal manifestations (Sauter, Eisner, Ekman, & Scott, 2010). On the opposite side of the spectrum of emotion theories, the integrative approach considers emotions to be domain-general constructions of the mind based on sensory input and past experiences, which are organ- ised around the dimensions of valence – that is, how pleasant or unpleasant one is feeling – and arousal – that is, how agitated or calm one is feeling. These dimensions are assumed to be the only universal features of an emotion (Russell, 1995), whereas the classification of emotions into dis- crete categories is based on one's previous experiences. In other words, according to the integrative approach, the perception of a particular

emotion instance is dependent on every individual sociali- sation pathway.

The aim of the present multilingualism study is not to contribute to the debate on the nature of emotions in psychology, but rather to investigate linguistic factors that might affect the perception of emotions in multimodal communication. Therefore, we adopt the definition of emotion pro- posed by Keltner and Shiota (2003, p. 89) that emphasises the different modalities in which emo- tional information is conveyed:

> 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.

2.2. Multimodal ERA

In a typical daily-life situation, information to infer the emotional state of one's interlocutor can be simultaneously gleaned from the verbal, the vocal and the visual channels (Burns & Beier, 1973), as well as from the context (e.g. Barrett, 2017). The content of the message of one's interlocutor might provide verbal cues to their emotional state; their rhythm, pitch, timbre, speaking rate and intensity might convey vocal cues; visual cues might be gleaned from their facial expression or body language. The emotion decoder has to process this stream of information, extract the relevant cues and interpret them according to the context in which they are being delivered. However, the relevance and the appropriate interpretation of cues might differ from one linguistic and/or cultural group to another (e.g. Irvine, 1982). In the following section, we review studies that investigate cross-linguistic and/or cross-cultural ERA, touching upon the long-lasting debate about the univer- sality or language/culture-specificity of emotional information integrated in the verbal, vocal and visual channels.

2.3. ERA across languages and cultures

Cross-linguistic communication entails the use of linguistic codes that are, for at least one of the interlocutors, different from the ones they have been confronted with from birth. As translated terms are never perfectly equivalent and might be linked to (slightly) different conceptual repre- sentations (Pavlenko, 1999), the interpretation of verbal cues to one's emotional state might lead to extra difficulties in the case of cross-linguistic (emotional) communication.

However, the bulk of research on cross-linguistic and/or cross-cultural ERA has focused on non- verbal communication. Facial expression, which is considered to be an important source of emotional information (e.g. Mesquita & Frijda, 1992), has received much attention in emotion recognition research. Ekman and his team have been very productive

in this area, mostly using static stimuli – that is, posed photographs of actors portraying an emotion – and asking participants to choose which of the 'Basic Six' is conveyed in each picture. The main finding of this line of research, repeatedly replicated, is assumed to support the hypothesis that basic emotions can be universally recognised from facial expression (e.g. Ekman et al., 1969). At the same time, results also suggest that people from the same cultural group recognise each other's emotions more accurately than people from dif- ferent cultures (Elfenbein & Ambady, 2002). In the Ekmanian framework, this in-group advantage is accounted for by the so-called 'cultural display rules'. These rules are acquired from childhood on and modulate the range of innate behavioural patterns by restricting one's behaviours to only those that are appropriate in a particular culture (Ekman & Friesen, 1969; Matsumoto & Ekman, 1989). Despite this account for variability, Ekmanian studies have been criticised for their methodology, which might conceivably bias their findings towards supporting universality. Not only has the use of static, prototypical stimuli arguably lacking ecological validity received criticism (Russell, 1995, 2015), but also the response format used in most studies of the Ekmanian tradition has been chal- lenged. Participants are typically asked to pick one of the six proposed emotion labels, without having the opportunity to indicate if they did perceive another emotion or no emotion at all. As demonstrated in a study by Gendron, Roberson, van der Vyver, and Barrett (2014) among American and Himba participants, the use of this response format might bias the results and conceal important cross-cul- tural differences in the facial perception of emotions.

Beside visual input, vocal input can also provide valuable emotional information, which might likewise be subject to linguistic and/or cultural variability. As Anolli, Wang, Mantovani, and De Toni's (2008) study comparing Chinese and Italian speakers suggests, members of different cul- tures encode emotions vocally in different ways. Thus, it seems reasonable to expect cultural dif- ferences in vocal emotion perception.

Sauter et al. (2010) found both cross-cultural similarities and differences in Himba and English speakers' recognition of emotional vocalisations, such as laughter or screams. Participants were asked to listen to an emotional story. They were then presented with two emotional vocalisations, each one from a different emotion category – expressed by either Himba or English speakers – and had to pick the vocalisation that conveyed the same emotion as in the story they had just heard. The so-called basic emotions could be recognised by both groups irrespective of the linguistic/cultural background of the speakers who vocalised the emotions. However, the other emotions included in the study – which were all positive ones – were not recognised cross-culturally. The ones expressed by Himba speakers were more reliably recognised by Himba participants and the ones expressed by English speakers were more reliably recognised by English participants, supporting the in- group advantage hypothesis.

Whereas Sauter and colleagues investigated two groups of participants with no experience of each other's language and/or culture, Bak (2016) focused on one group of participants speaking two languages. She investigated the vocal emotion perception of Polish first language (L1) learners of English with no or little authentic exposure to L1 English culture. In the first task, participants – who were English major students at university – had to rate the valence and arousal conveyed in low-pass filtered English speech – that is, a technique used to make the word unintelligible while retaining prosodic information. In the second task, participants had to categorise the emotion con- veyed in each stimuli as 'happiness', 'sadness' or other. Bak found a surprising effect of English proficiency, with the low-proficient outperforming the high-proficient participants in their vocal ERA. The author explains this results following heuristics principles (e.g. Gigerenzer, 2001): the high-proficient second or foreign language (LX) users, who had been exposed to English for a longer period, had a better knowledge of the language; they were thus trying to integrate more information into their emotion processing to make an 'optimal' inference, and hence the greater probability of making an error. Bak's (2016) findings are out of line with previous research, such as Rintell's (1984) study of vocal–verbal ERA among L1 English users and LX English learners with Arabic, Chinese or Spanish backgrounds. In this study, participants heard 11 recordings in English. After each recording, they had to choose one of the 11 emotional labels corresponding to the emotional state of the speaker. The strongest effect was found for LX proficiency, with the intermediate and advanced learners being better able to recognise the intended emotions than the beginners. A possible explanation for these divergent findings might rely in the nature of the LX participants in both studies. Bak's participants are language 'specialists', as they are majoring in English at university, while Rintell's participants are typical learners enrolled in an English as a foreign language (EFL) course. Following Bak's (2016) supposition, the 'language specialists' might have been more affected than Rintell's 'typical EFL learners' by the impeding processing load of trying to integrate as much information as possible. Moreover, an important distinction between both studies is that Bak focused on purely vocal ERA, while Rintell investigated vocal– verbal ERA. It is reasonable to expect that language proficiency is especially related to the decod- ing of verbal cues and hence positively related to verbal ERA in particular. This speculation finds support in Graham, Hamblin, and Feldstein's (2001) study into vocal ERA, in which eight different emotion portrayals of the same monologue were presented to L1 and LX users of English. Contrary to Bak's (2016) and Rintell's (1984) findings, LX proficiency did not appear to be significantly related to ERA in this study – in which only vocal cues were available.

2.4.   Bilingualism

2.4.1. The bilingual cognitive advantage

Bi/multilingualism has been linked to a number of cognitive out- comes, comparing how monolinguals and bi/multilinguals perform on different cognitive tasks. Some findings seem to support the so-called BAH, pointing towards positive relationships between bi/multilingualism and cognitive outcomes. Adesope et al. (2010) conducted a meta-analysis of data from 63 studies that all included participants equally or almost equally proficient in two lan- guages and interpreted their findings as supportive of the BAH. Although they acknowledged that the reviewed studies vary in their findings and effect sizes – with some studies even showing nega- tive correlations between bilingualism and cognitive measures – they concluded that bilingualism is generally positively linked to metalinguistic and metacognitive awareness, abstract and sym- bolic representations, attentional control and problem solving.

Other studies, however, challenge this putative bilingual advantage. In a recent meta-analysis project reviewing 152 studies that include early and late bilingual adults, Lehtonen and colleagues (2018) argue that their findings do not support the BAH, as only very small effect sizes were found for some of the investigated cognitive measures – of the order of 1% of explained variance – and those effects even disappeared when controlling for outliers and publication biases. Eventually, their analysis yielded no effect of bilingualism for inhibitory control, monitoring, shifting or working memory. A small effect remained for attention, but pointing towards a *dis*advantage for bilinguals.

2.4.2. Bi/multilinguals' ERA

Besides research linking bi/multilingualism to variables that are usually regarded as belonging to the domain of cognition, a few studies have also investigated the link between bi/multilingualism and affective variables, although this research area is much scarcer.

Dromey, Silveira, and Sandor (2005) probably conducted the first study investigating the link between bi/multilingualism and emotion recognition. In their study, an English-speaking actor read words in either a neutral or an angry tone of voice. The L1 monolingual, L1 multilingual and LX multilingual users of English had to choose the emotion label corresponding to the actor's tone of voice. The L1 multilinguals outperformed the LX multilinguals, but, against expectations, the L1 monolinguals did not. As the authors argue, '[t]he learning of a second language may have helped the [L1 multilinguals] develop additional perceptual skills in decoding speech emotion in their native language' (Dromey et al., 2005, p. 351).

Recently, Alqarni and Dewaele (2018) compared bilinguals' and monolinguals' ability to perceive emotions in English. Their findings – in line with Dromey and colleagues' (2005) results – offer tenta- tive support of what they call a 'bilingual emotional advantage', with Arabic LX users of English outperforming English monolinguals in their ability to perceive emotions in English. The authors hypothesise that this result could be linked to bilinguals' broader emotional lexical space – as they have acquired emotion words from two languages and therefore have a broader knowledge of various emo- tion concepts. Following Barrett's argumentation (2017) in Alqarni and Dewaele (2018), this would in turn enable their brain to make a wider range of emotional predictions, and hence broader emotional perceptions. The authors also speculate that bilinguals' exposure to two cultures might strengthen socio-pragmatic skills, which can be beneficial to emotion communication across languages.

The different findings in the studies reviewed above might be linked to the different inclusion criteria regarding the 'bilingual' participants. Adesope and colleagues (2010) included studies with participants reporting to be (almost) equally proficient in two languages. Although no information is provided about context nor age of acquisition of these languages, the authors mention that 'par- ticipants who were learning second languages were not regarded as bilinguals but rather as second language learners [...]' and '[...] were excluded from this meta-analysis' (pp. 211–212). Lehtonen and colleagues (2018), on the contrary, included studies with both early and late bilinguals, controlling for a possible effect of age of acquisition on the cognitive outcomes variables, but did not find any effect. Finally, Alqarni and Dewaele's (2018) study included only late bilinguals.

In the present study, we examine a possible bilingual advantage distinguishing between partici- pants who grew up as bilinguals from birth (BFLA) and participants who became bi/multilinguals after the age of 3. Moreover, we investigate whether the knowledge of more languages constitutes an advantage in ERA – regardless of proficiency level, context of acquisition or age of acquisition.

3. Research questions and hypotheses

In this study, we address the following questions and hypotheses.

1. Is BFLA linked with an increased ability to recognise emotions expressed in English via the verbal, vocal and visual channels?

Similar to other cognitive and affective skills found to be better developed in bilinguals, we hypothesise that ERA is positively linked to BFLA, meaning that individuals who acquired more than one language before age 3 are better able to recognise emotions expressed in English than monolingually raised individuals.

2. Is the knowledge of more languages linked with an increased ability to recognise emotions expressed in English via the verbal, vocal and visual channels?

Speaking several languages might help one to develop a particular sensitivity to paralinguistic aspects of speech, which might support the ability to recognise emotions in general. Hence, we hypothesise that the more languages one can speak, the better one will be at recognising emotions.

3. Is nativeness in English linked to an increased ability to recognise emotions expressed in English via the verbal, vocal and visual channels?

Following previous findings with similar methodology (Lorette & Dewaele, 2015), we hypoth- esise that L1 users will not outperform LX users of English. This might seem surprising at first glance, but findings suggest that when LX users can simultaneously rely on visual, verbal and vocal cues, they might be able to accurately recognise emotions as well as L1 users – probably using more (or more diverse types of) cues than L1 users.

4. Is proficiency in English linked with an increased ability to recognise emotions expressed in English via the verbal, vocal and visual channels?

We hypothesise that the more one is proficient in English, the better one will be at recognising emotions in English, since more proficient participants might be able to interpret verbal cues more accurately.

4. **Methodology**

4.1. Participants

One-thousand-two-hundred-and-twenty participants (837 females, 383 males) with a mean age of 34 years (*SD = 14*) took part in this study. Five-hundred-and-sixty-four are LX users of English, while 656 are L1 users. The L1 users of English were mainly from the UK and Ireland (*n* = 317), followed by North America (*n* = 229). The LX users mainly originated from Slovenia (*n* = 120), Belgium (*n* = 89) and The Netherlands (*n* = 70). The participants' mean English proficiency score – based on a lexical test described in the *Instrument* section – is 83% (*SD* = 11.3) for the LX users and 91% (*SD* = 11.2) for the L1 users. Moreover, 207 participants acquired more than one language before age 3 (henceforth named BFLA participants), while 1013 participants were not exposed to more than one language at least until the age of 3 (henceforth named non-BFLA participants). Regarding the number of languages participants can speak – regardless of proficiency level, age of acquisition or context of acquisition – the LX users know 3.6 languages on average (*SD* = 1.3) and the L1 users know 2.3 languages on average (*SD* = 1.3).

4.2. Instrument

Data was collected through an online survey consisting of a sociobiographical questionnaire, an emotion recognition test with six audiovisual stimuli and a lexical test.

In the first part of the survey, participants were asked about their gender, age, nationality, actual country of residence, L1(s) and possible LX(s) – and for each LX, the age and context of acquisi- tion, length and frequency of use with L1 and LX speakers, self-rated proficiency and length of residence in a country in which this LX is an official language.

Next, six audiovisual stimuli (30–55 seconds long) were presented to the participants – see the Appendix for transcriptions. The rationale for the use of dynamic audiovisual stimuli relies on previ- ous research showing that these are ecologically more valid than static stimuli deprived of context (Alves, 2013; Elfenbein & Ambady, 2002). In each stimulus, an actress with slightly London-coloured English Received Pronunciation improvised a daily-life situation in which she intended to convey one of the six so-called 'basic' emotions – that is, happiness, sadness, surprise, disgust, anger or fear. A female actress has been chosen as females' emotions tend to be decoded more easily than those of males (e.g. Matsumoto et al., 2000). Note that, as valence might affect emotion recognition (Sauter et al., 2010; Zhu, 2013), the instance of surprise included in this study is a positively valenced one in order to avoid happiness being the only positive emotion. The improvised situations were based on either written scenarios created by the authors or on the actress' own ideas – in agreement with the authors – in order to foster the authenticity of her play. Cues to her emotional state can be gathered from both the verbal and the non-verbal channels – that is, her facial expression, tone of voice, etc. The stimuli were presented via YouTube videos embedded in the survey. After each stimulus, partici- pants were required to choose one of the six emotion labels that reflected the main emotion conveyed in the recording. Crucially, they were also able to choose 'no emotion' or 'no idea' in order to limit the bias of a forced-choice response format (Russell, 1994).

Finally, the survey included the English version of the LexTALE (Lemhöfer & Broersma, 2012), a 60-item lexical test in which participants have to indicate, for each item, whether it is an existing English word or not. The LexTALE only takes a few minutes to complete and has been shown to reliably measure the lexical proficiency of LX users with different social and linguistic backgrounds and to be a good indicator of overall English proficiency (Lemhöfer & Broersma, 2012).

4.3. Data collection and analysis

The online survey was advertised on social media, public mailing lists and via emails to colleagues and acquaintances, who were in turn asked to spread it in their own network.

Thanks to this snow- ball sampling method using different kinds of media, participants with different social and linguistic backgrounds could be reached. The 919 participants reached via this method took part voluntarily, without time limit and were informed that they could withdraw at any time. In order to reach a more balanced sample between L1 and LX users, 301 additional L1 users of English were recruited via a digital agency that constituted a panel of volunteers who were remunerated to complete the survey.

An individual ERA score was calculated for each participant. For each stimulus, a point was added to the participant's ERA score if the emotion indicated by the participant corresponded to the emotion intended by the actress. As there were six stimuli, each participant's ERA score can range between 0 and 6. The data was analysed in SPSS software.

5. Results

As shown in Figure 1, the majority of the participants identified the intended emotion in four out of six stimuli (mean *ERA score* = 3.9, *SD* = 1.2).
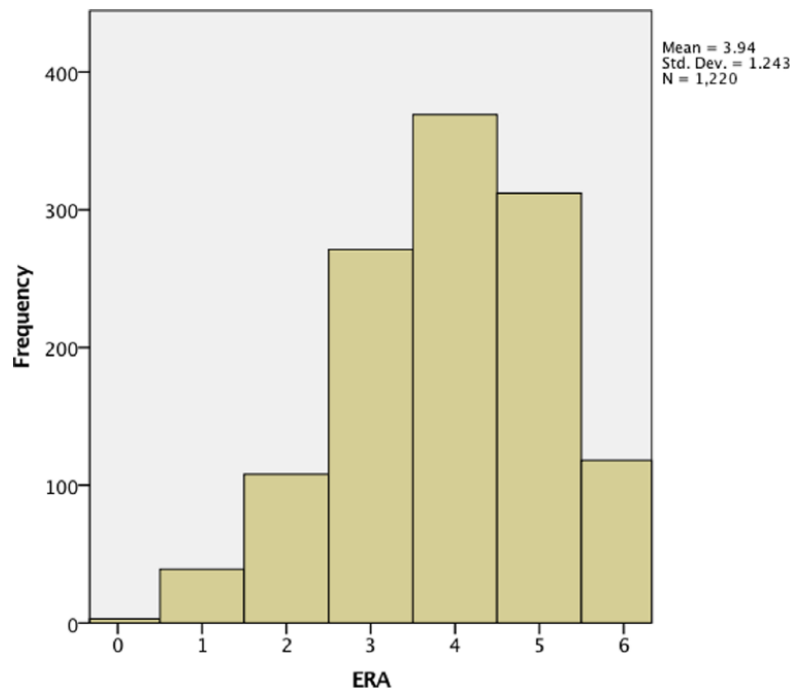

Mean = 3.94
Std. Dev. = 1.243
N = 1,220

Figure 1. *Bar plot showing the overall frequency of responses (N = 1220).*

The relationship between the response variable *ERA score* (henceforth *ERA*) and the explana- tory variables included in this study has been examined by means of bivariate correlations. Therefore, the categorical variables *BFLA* and *Nativeness* have been dummy coded and the interac- tion variables *Nativeness × BFLA, Proficiency × BFLA, Number of*

languages × *BFLA*, *Number of languages* × *Nativeness* and *Number of languages* × *Proficiency* have been created.[2]

ERA significantly correlates with the variables *Nativeness* (ρ = −.146, *p* < .000*), Proficiency* (ρ = .136, *p* < .000) and *Number of languages* (ρ = .153, *p* < .000). Regarding interactions, *ERA* significantly correlates with *Nativeness* × *BFLA* (ρ = −.121, *p* < .000), *Proficiency* × *BFLA* (ρ = .056, *p* = .049) and *Number of languages* × *Nativeness* (ρ = .106, *p* < .000).

The independent variables correlating with *ERA* were then fed into a linear regression model in order to assess the predicting value of each factor for the outcome variable *ERA* while controlling for the effects of the other independent variables. As the assumptions of homoscedasticity and normality of residuals were not met, the bootstrapping procedure was implemented with 95% biased-corrected accelerated confidence intervals (Bca CIs) based on 1000 bootstrap samples.

The resulting five-predictor regression model accounts for 10.2% of the variance in *ERA* ($F(6, 1213) = 22.98$, *p* < .000).

Model		B	Bootstrap[a]				
			Bias	Std. Error	Sig. (2–tailed)	BCa 95% Confidence Interval	
						Lower	Upper
1	(Constant)	4.231	−.003	.056	.001	4.126	4.329
	Nativeness	−.420	.005	.091	.001	−.592	−.232
	Proficiency	.024	5.101E−5	.004	.001	.016	.032
	NrLang	.039	.002	.036	.288	−.035	.113
	NatxBFLA	−.342	.002	.124	.009	−.573	−.094
	ProfxBFLA	.011	1.165E−5	.008	.182	−.005	.028
	NrLangxNat	.059	−.001	.052	.260	−.033	.158

Table 1. *Summary of the five-predictor regression model with emotion recognition ability as the outcome variable (n = 1220).*

As shown in Table 1, the two factors *Nativeness* (*p* < .001, 95% CI [−.592, −.232]) and *Proficiency* (*p* < .001, 95% CI [.017, .031]), as well as the interaction term *Nativeness* × *BFLA* (*p* < .009, 95% CI [−.573, −.094]), are significant predictors of *ERA* in this model. The initial (non- bootstrapped) standardised coefficients suggest that *Proficiency* (*Beta* = .231) is the best predictor of *ERA* in this model.[3] The more proficient participants are in English, the more likely they are to correctly identify intended emotions expressed visually, vocally and verbally in English. Although the main effect of *Nativeness* with a negative *B* (and, in the non-bootstrapped model, a negative *Beta* of −.169) surprisingly suggest that English LX users outperform L1 users in their ability to recognise emotions in English, this relationship is best understood taking the interaction *Nativeness* × *BFLA*

into account. To clarify this interaction, a *t*-test with *ERA* as the dependent variable and *BFLA* as the independent variable was run separately for LX users and for L1 users – see Figure 2. Among LX users, BFLA participants (*mean ERA = 4.46*) obtained significantly higher ERA scores than non-BLFA participants (*mean ERA = 4.11, t(562) = –2.40, p < .017*). In contrast, non-BFLA L1 users (*mean ERA = 4.19*) obtained significantly higher ERA scores than BFLA L1 users (*mean ERA = 3.88, t(353) = 2.40, p < .017*). Thus, BFLA seems to present an advantage for emotion rec- ognition among LX users of English, but seems to be disadvantageous for L1 users.
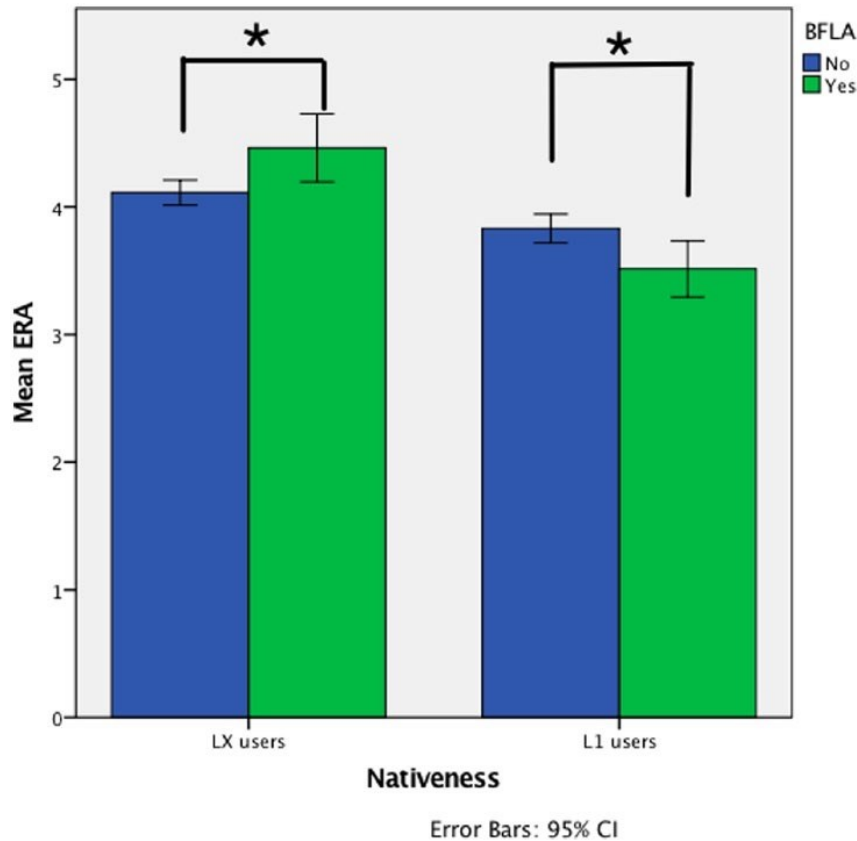

Figure 2. *Barplot showing bilingual first language acquisition (BFLA) versus non-BFLA participants' emotion recognition ability (ERA) among second or foreign language users and first language users (n = 919, *p < .05).*

6.   Discussion and limitations

Our study supports previous results regarding the positive relationship between English profi- ciency and ERA in English when visual, vocal and verbal cues are simultaneously available to the emotion decoder (Lorette & Dewaele, 2015). Proficiency is most likely to particularly relate to the interpretation of verbal cues – as suggested by the proficiency effect found in Rintell's (1984) study into verbal–vocal ERA but not found in Graham

and colleagues' (2001) study into vocal ERA. Despite the availability of non-verbal cues, proficiency appeared to be a significant predictor of ERA in our study. Therefore, we might speculate that participants have relied on verbal cues to a great extent despite the presence of non-verbal cues. This suggests that the verbal channel is an important source of emotional cues for L1 and LX users even when non-verbal cues are also avail- able (Lorette & Dewaele, 2018).

Concerning BFLA, the equivocal picture arising from the literature regarding a bilingual (dis) advantage in the cognitive domain is reflected in our findings on the affective level. Being raised bi/multilingually appears to be advantageous for ERA in English among LX users of English, but disadvantageous for ERA in English among individuals raised in both English *and* one (or several) additional language(s). This can be interpreted as a partial support for the bilingual emotional advantage (Alqarni & Dewaele, 2018), although this general, language-independent advantage for recognising emotions might be cancelled out by 'negative transfers' or interferences that bi/multi- lingual L1 users might experience between English and their other L1(s). Individuals raised bi/ multilingually are confronted not only with different codes to express themselves, but also with different language- and/or culture-specific norms and models about appropriate behaviour in a particular context. Specifically in the case of emotion communication, being confronted with 'emotional scripts' – as Wierzbicka (1999) calls them – from different languages and/or cultures might bring about some confusion for bi/multilingually raised individuals regarding the way emo- tions are communicated in a particular language and/or culture.

Regarding a more general relationship between bi/multilingualism and ERA, our analyses did not confirm the hypothesis that knowing more languages is linked to an increased ERA. Possibly, the results might have been different if we had distinguished between the total number of languages one can speak and the number of additional languages one has *learned*. Acquiring several languages from birth and learning additional foreign languages later in life arguably involve different processes and might thus relate to ERA in different ways. We might hypothesise that people confronted with a par- ticular language/culture from birth on have internalised the emotional scripts of this language/culture to a greater extent than people confronted with this language/culture later in life; the former being thus more likely to use and perceive emotions accurately in different contexts than the latter. Besides age of acquisition, the context of acquisition might also play a role. People learning a language in a formal context probably have less authentic contact with this language/culture, and are thus likely to be less familiar with the emotional scripts of this language/culture, compared to people acquiring this language in a naturalistic context (Dewaele, 2013). Consequently, late and formal LX learners might have a more restricted array of emotional scripts to their disposal, thus being more restricted in the number of emotional predictions that they can construct in a particular context. Furthermore, unlike acquiring a L1, learning an additional language later in life seems to be related to the

development of specific attentional skills (MacWhinney, 2005). Language learners have to shift their attention to particular cues that are less salient and/or relevant in their L1 but are very salient and informative in that particular language – for instance in the case of a non-tonal L1 speaker learning a tonal LX. Moreover, during the learning process of a foreign language, LX users are confronted with profi- ciency gaps, and thus have to turn to other, non-verbal cues to try to interpret the input. Hypothetically, learning additional languages might foster a particular sensitivity to non-verbal cues, which might be advantageous for emotion recognition. Next, the learning of an additional language has been related to increased metalinguistic and metapragmatic awareness (Sanz, 2013), which might also be benefi- cial in the case of cross-linguistic and/or cross-cultural emotion recognition (Alqarni & Dewaele, 2018). Therefore, we hypothesise that it is not the *total* number of languages one *knows*, but rather the number of *additional* languages one has *learned*, that is positively related to ERA – a hypothesis that deserves to be tested in future research.

Finally, it is important to acknowledge the rather modest proportion of variance explained by our model. This might be due to a few important factors that have repeatedly been shown to cor- relate with ERA, but that we did not intend to investigate in the present study as we intentionally focused on less-widely investigated variables. Firstly, participants' cultural backgrounds have been shown to be related to ERA (e.g. Elfenbein & Ambady, 2002; Gendron et al., 2014; Lorette & Dewaele, 2015). Secondly, trait Emotional Intelligence has been linked to ERA in numerous stud- ies (e.g. Alqarni & Dewaele, 2018; Dewaele, Lorette, & Petrides, in press; Petrides & Furnham, 2003). Future studies could include both widely and less-widely investigated factors in their mod- els. The inclusion of more factors in a study requires larger samples but it would allow the develop- ment of a more accurate, fine-grained picture of the complex phenomena of ERA.

7. Conclusion

The primary aim of this study was to investigate the link between bi- and multilingualism and emo- tion perception when visual, vocal and verbal cues are available. Our results did not support a general emotional bilingual advantage in the case of multilingualism as operationalised as the mere number of languages one can speak, regardless of proficiency, age or context of acquisition. Nevertheless, considering the particular case of BFLA, our findings suggest that simultaneous acquisition of two languages from birth conferred an advantage for emotion perception among LX users of English. However, this advantage seems to be overridden by other phenomena among English L1 users, which we hypothesise to be related to interferences from the L1 users' other L1(s) or L1 culture(s). Finally, we found a positive relationship between English proficiency and visual–vocal–verbal ERA in English, demonstrating the effect of proficiency even when non-verbal

cues are available.

Notes

1. We would like to emphasise that we do not intend to use the term 'nativeness' to support a clear-cut dichotomy between 'native' and 'non-native' speakers. Instead, we advocate the use of 'first language (L1) users' and 'second or foreign (LX) users', which convey less negative connotation and account for the heterogeneity of both groups (Dewaele, 2018).

2. The variables *Proficiency* and *Number of languages* have been centred before creating the interaction terms to reduce multicollinearity.

3. We are aware that the coefficients of the non-bootstrapped model have to be interpreted carefully since the assumptions of homoscedasticity and normality were not met. However, as the non-bootstrapped and the bootstrapped models give similar results, it is reasonable to consider the standardised coefficients as being informative.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*(2), 207–245.

Alqarni, N., & Dewaele, J.-M. (2018). A bilingual emotional advantage? An investigation into the effects of psychological factors in emotion perception in Arabic and in English of Arabic-English bilinguals and Arabic/English monolinguals. Unpublished manuscript.

Alves, N. T. (2013). Recognition of static and dynamic facial expressions: A study review. *Estudios de Psicologia (Natal)*, *18*(1), 125–130.

Anolli, L., Wang, L., Mantovani, F., & De Toni, A. (2008). The voice of emotion in Chinese and Italian young adults. *Journal of Cross-Cultural Psychology*, *39*(5), 565–598.

Bak, H. (2016). *Emotional prosody processing for non-native English speakers: Towards an integrative emo- tion paradigm*. Cham: Springer.

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. New York: Houghton Mifflin Harcourt.

Burns, K. L., & Beier, E. G. (1973). Significance of vocal and visual channels in the decoding of

emotional meaning. *Journal of Communication*, *23*(1), 118–130.

Dewaele, J.-M. (2013). *Emotions in multiple languages*. 2nd edition. Basingstoke: Palgrave-MacMillan. Dewaele, J.-M. (2018). Why the dichotomy 'L1 versus LX user' is better than 'Native versus Non-native speaker'. *Applied Linguistics*, *39*(2), 236–240.

Dewaele, J.-M., Lorette, P., & Petrides, K. V. (in press). The effects of linguistic proficiency, trait emotional intelligence and cultural background on emotion recognition by English native speakers. In L. Alba Juez & L. Mackenzie (Eds.), *Emotion in discourse*. Amsterdam: Benjamins.

Dromey, C., Silveira, J., & Sandor, P. (2005). Recognition of affective prosody by speakers of English as a first or foreign language. *Speech Communication*, *47*(3), 351–359.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3–4), 169–200.

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

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion.

Science, *164*, 86–88. Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion*, *14*(2), 251–262.

Gigerenzer, G. (2001). Decision making: Nonrational theories. *Elsevier Science*, *5*, 3304–3309.

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*(1), 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, DC: Georgetown University Press.

Keltner, D., & Cordaro, D. T. (2017). Understanding multimodal emotional expressions. Recent advances in basic emotion theory. In J.-M. Fernández-Dols & J. A. Russell (Eds.), *The science of facial expression*. New York: Oxford University Press.

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

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review.

Psychological Bulletin, *144*(4), 394–425.

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.

Lorette, P., & Dewaele, J.-M. (2015). Emotion recognition ability in English among L1 and LX users of English. *International Journal of Language and Culture*, *2*(1), 62–86.

Lorette, P., & Dewaele, J.-M. (2018). Emotion recognition ability across different modalities: the role of language status (L1/LX), proficiency and cultural background. *Applied Linguistics Review*.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion*, *13*(2), 143–157.

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., & Yee, 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.

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

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

Petrides, K. V., & Furnham, A. (2003). Trait emotional intelligence: Behavioural validation in two studies of emotion recognition and reactivity to mood induction. *European Journal of Personality*, *17*(1), 39–57.

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

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–141.

Russell, J. A. (1995). Facial expressions of emotion: What lies beyond minimal universality? *Psychological Bulletin*, *118*(3), 379–391.

Russell, J. A. (2015). *Moving on from the basic emotion theory of facial expression*. Retrieved June 6, 2017 from http://emotionresearcher.com/moving-on-from-the-basic-emotion-theory-of-facial-expressions/

Sanz, C. (2013). Multilingualism and metalinguistic awareness. In C. Chapelle (Ed.), *The

Encyclopedia of applied linguistics. London: Blackwell Publishing Ltd.

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*, *107*(6), 2408–2412.

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

Zhu, Y. (2013). Which is the best listener group?: Perception of Chinese emotional prosody by Chinese natives, naïve Dutch listeners and Dutch L2 learners of Chinese. *Dutch Journal of Applied Linguistics*, *2*(2), 170–183.

Appendix

The stimuli can be consulted on the following links. They have been transcribed below, in order of apparition in the survey.

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

Lorette and Dewaele 15 *birthday, Maria. Happy birthday!' And she said: 'It's not my birthday, it was my birthday yester-*

day.' 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 can kind of just 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 flow- ers. 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.