frontiers | Frontiers in Language Sciences

TYPE Original Research
PUBLISHED 08 January 2024
DOI 10.3389/flang.2023.1293673

Check for updates

OPEN ACCESS

EDITED BY
Yan Jing Wu,
Ningbo University, China

REVIEWED BY
Xuemei Chen,
University of Electronic Science and Technology of China, China
Katy Borodkin,
Tel Aviv University, Israel

*CORRESPONDENCE
Wenwen Yang
✉ wny20rxq@bangor.ac.uk

RECEIVED 13 September 2023
ACCEPTED 08 December 2023
PUBLISHED 08 January 2024

CITATION
Yang W, Rauwolf P, Frances C, Molina-Nieto O, Duñabeitia JA and Thierry G (2024) Chinese-English bilinguals prefer being truthful in the native language.
Front. Lang. Sci. 2:1293673.
doi: 10.3389/flang.2023.1293673

Chinese-English bilinguals prefer being truthful in the native language

Wenwen Yang[1]*, Paul Rauwolf[1], Candice Frances[2], Olivia Molina-Nieto[1], Jon Andoni Duñabeitia[3] and Guillaume Thierry[1,4]

[1]School of Psychology and Sport Science, Bangor University, Bangor, United Kingdom, [2]Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, [3]Nebrija Research Center in Cognition, Nebrija University, Madrid, Spain, [4]Faculty of English, Adam Mickiewicz University, Poznań, Poland

Bilinguals make decisions differently across languages likely due to reduced affect in the foreign language, but very little is known about language use in relation to deception. Here, we tested whether late Chinese-English bilinguals prefer to lie in the foreign language when betting against a virtual opponent. In each trial, participants freely announced bets in Chinese or English depending on whether they had drawn a coin or not. Results showed that bilinguals preferred using their native language, Chinese, over their foreign language, English, when being truthful—namely, announcing a coin when they had one. Even though participants did not choose English more to lie than to tell the truth, our results can be interpreted as a tendency to lie more in English when their behavior in the truth condition is considered the baseline. Participants also switched between languages more often after telling the truth than lying, and after telling the truth they switched more to Chinese. These results provide the first empirical evidence for strategic language use in bilinguals.

KEYWORDS

strategic language use (SLU), foreign language effect (FLE), lying, native/foreign language, decision-making

1 Introduction

Depending on the context in which they find themselves, bilingual individuals who use two languages on a regular basis, switch between languages more or less frequently. Some bilinguals who operate in a multilingual environment (e.g., a bilingual community, such as Montreal, or a school for interpreters) tend to switch between their languages more frequently than those who live in a more linguistically uniform environment (e.g., a monolingual community or at home). For instance, bilinguals spontaneously switch between their languages in a picture naming task with minimal external constraints (De Bruin et al., 2018, 2020). This language-switching behavior brings up an interesting question: If people find themselves in a situation where they want to deceive their interlocutor and they can freely choose between languages, will they be more likely to lie in their native or foreign language? And if they lie, do they use languages strategically?

The foreign language effect (FLE), a concept introduced by Keysar et al. (2012), refers to the fact that the language context affects decision-making in bilinguals and multilinguals. In their seminal study, Keysar et al. (2012) showed that bilinguals weigh the positive and negative consequences of situations differently in their foreign and their native language, since reasoning in a foreign language can reduce decision-making biases. The authors asked participants to make critical decisions in emotionally loaded contexts such as the

"Asian disease" problem (Tversky and Kahneman, 1985) and an equal-odds betting task presented through the medium of either their native or the foreign language. They found that participants tended to act more rationally when operating in a foreign language.

Evidence in support of the FLE mostly comes from the decision-making and moral judgement literature (Keysar et al., 2012; Costa et al., 2014; Gao et al., 2015; Geipel et al., 2015; Hadjichristidis et al., 2015; Corey et al., 2017). When encountering scenarios that involve risk perception or moral dilemmas in different language contexts, bilinguals' ultimate decision is language dependent. For example, Corey et al. (2017) found that people were more likely to give an affirmative answer in the classic "would you sacrifice a person to save five others" question when they used a foreign language than when the scenario was presented in the native language. Costa et al. (2014) found that bilinguals' intuitive bias was reduced when problems are presented in a foreign language in three different contexts focusing on loss aversion, economic decision-making, and decision-making under risk and uncertainty, which they interpreted as the manifestation of a robust FLE. Geipel et al. (2015) found that bilinguals tended to be more lenient in their moral judgments of acts violating moral norms, such as consensual incest when they made these judgements in a foreign as compared to the native language. Hadjichristidis et al. (2015) found that the use of a foreign language can reduce disproportionate risk perception relative to the native language, bringing a potential benefit to bilinguals who, for instance, have irrational anxieties, such as refusing to travel by plane for safety reasons. Gao et al. (2015) engaged Chinese-English bilinguals in a game of bets. Participants received feedback on each trial, either positive or negative, depending on the outcome of each bet and either in Chinese (their native language) or English (the foreign language). They found that participants tended to take more risks when receiving positive feedback in Chinese than when receiving equivalent feedback in English. On first approach, this result seems inconsistent with those of Keysar et al. (2012), who found a reduction of the negative bias in the foreign language. However, the results of the two studies are in fact convergent if one considers that bilinguals are more likely to accept negative outcomes in a foreign language (Keysar et al., 2012) whereas they are prompted to take more risks after receiving positive feedback in their native language (Gao et al., 2015).

A likely source for the FLE is that bilinguals tend to have different emotional reactions in their two languages (Pavlenko, 2004, 2005, 2012; Puntoni et al., 2009; Wu and Thierry, 2012; Caldwell-Harris, 2015; Hsu et al., 2015; Jończyk et al., 2016; Gao et al., 2020). For example, Puntoni et al. (2009) found that marketing slogans are perceived as more emotional when text is written in the consumers' native language rather than in a foreign one. Gao et al. (2020) showed that bilinguals who were asked to affectively rate self-critical statements found the statements less unpleasant when they were expressed in a foreign language. Studies characterizing explicit behavior regarding emotion and bilingualism based on surveys and questionnaires (Pavlenko, 2004, 2005, 2008, see 2012 for a review; Dewaele, 2004a,b, 2008a,b, 2010) have repeatedly shown that bilinguals tend to feel more detached from both negative and positive feelings evoked by language when they operate in a foreign language. For instance, Dewaele (2004a,b) showed that multilinguals have attenuated affective responses to swear words and taboo words in the languages they acquired subsequently to their native language (see Pavlenko, 2012, for a review). In the same vein, other studies have shown reduced psychophysiological reactions to negative stimuli in the foreign than in the native language, as suggested by pupil dilation or galvanic skin responses (see Iacozza et al., 2017; García-Palacios et al., 2018). Some electrophysiological studies have also shown that the response to words conveying negative emotions such as taboo words is reduced in the second relative to the native language (Wu and Thierry, 2012; Jończyk et al., 2016). For instance, Wu and Thierry (2012) showed that negative words in a foreign language tend to activate translation equivalents in the native language to a lesser extent than positive or neutral words. Convergingly, Jończyk et al. (2016) investigated N400 modulations, which indexes semantic processing depth, in a sentence reading task featuring positively and negatively valenced sentences ending in an expected or unexpected fashion. They found that N400 modulations were weaker in the foreign language (English) than the native language (Polish). Frances et al. (2020b) found that the effects of emotionality were not modulated by language when Spanish-English participants listened to neutral or positive descriptions about two countries in either Spanish or English (Frances et al., 2020a). Also, Hsu et al. (2015) showed that bilinguals report having a stronger emotional experience when reading emotion-laden texts in their native language, which elicited brain activations distinct from those elicited by the foreign language.

There is substantial research on decision-making and moral judgement in relation to language of operation in bilinguals as well as on the cognitive bases or lying and lying-emotion relationships (e.g., Buller and Burgoon, 1996; Abe et al., 2007; Bizzi et al., 2009; Dong et al., 2010; Garrett et al., 2016). Nevertheless, the links between language use and lying remain relatively unexplored. This can be considered surprising given the wide impact that could be expected from research on language use and deception in domains such as the law, employment, and advertising. One explanation for the paucity of studies in this field is that encouraging participants to vary both their lying and language choice behavior is arguably very difficult Simply instructing participants to lie in some experimental conditions and manipulating the language of operation at the same time is likely to foster unrealistic behavior based on metacognitive evaluation, which bears little resemblance to natural trends. Studies that have investigated lying in a bilingual context have led to highly inconsistent findings, some showing more likelihood of lying in L2, others showing greater ease of lying in L1, and yet others, showing no significant differences between language contexts (Caldwell-Harris and Ayçiçegi-Dinn, 2009; Duñabeitia and Costa, 2015; Gai and Puntoni, 2018; Suchotzki and Gamer, 2018; Bereby-Meyer et al., 2020).

Caldwell-Harris and Ayçiçegi-Dinn (2009), for instance, asked participants to rate the emotional intensity of words and phrases, varying in emotional content and language. Participants exhibited reduced skin conductance responses (SCRs) when listening to emotional phrases in their second language (English) as compared to their first language (Turkish-their L1.). In a second task, they asked Turkish-English bilinguals to read truthful and deceptive statements aloud in both the native and the foreign language and rate the way they felt when uttering each statement. Participants reported a stronger negative emotional impact of the Turkish lie than the English lie condition. However, whilst greater skin conductance responses (SCRs) were elicited by untruthful statements than truthful ones, it was English lies rather than Turkish ones that elicited the strongest SCR responses. Together, results showed that lies tend to evoke larger SCRs than truthful statements, and irrespective of truth status, SCRs tend to be higher in the L2 than in the L1. They concluded that two factors come into play when bilinguals lie in their two languages: (i) increased arousal when lying compared to telling the truth and (ii) anxiety when bilinguals have to speak in their L2.

Duñabeitia and Costa (2015) asked Spanish-English bilinguals to produce descriptions of one of three animals in a picture, either truthfully or untruthfully, and failed to find an interaction between language of operation and truthfulness, although reaction times were longer in the foreign than the native language, and for lying as compared to telling the truth. Bereby-Meyer et al. (2020) invited a large cohort of bilinguals to roll a die three times and report the first number drawn. Participants lied less in the foreign than the native language, which the authors interpreted as resulting from the fact that decision-making in a foreign language is less intuitive and thus leads to longer deliberations. Suchotzki and Gamer (2018) asked German-English bilinguals private questions that were neutral or affectively challenging in both the native and the foreign language, and they found a smaller difference in reaction times between lying and truth-telling in the foreign language. Gai and Puntoni (2018) engaged three different groups of bilinguals with different languages in a spot-the-difference task in which they were rewarded based on the number of differences reported between two images. Since the number of differences was not verified, they could lie and inflate their score. The authors found that the foreign language affected the tendency to lie differently depending on the magnitude of the lie: foreign language use decreased the rate of minor lies whilst increasing the rate of major lies. Despite the inconsistencies listed above, a recurrent theme is that of attenuated emotional responses in the foreign language (Geipel et al., 2015; Hayakawa et al., 2017). Therefore, we hypothesed that it may be easier for bilinguals to lie in their foreign language (Keysar et al., 2012; Costa et al., 2014; Hayakawa et al., 2017). However, we had no clear prediction for truths.

Here, rather than focussing on how contextual effects influence bilinguals' intention to lie, we aim to investigate the language choices made by bilinguals placed in a lie-incentivizing situation. Our research questions are the following: How does the decisional situation in which bilinguals find themselves influence language choice? And, in particular, does the incentive to lie in some situations lead to strategic language use (SLU)? Considering that lying requires more cognitive resources than telling the truth and that the foreign language is less emotionally salient (see above), do bilinguals show a preference for the foreign language when they intend to lie?

In other words, in the current study, we focus on SLU as a kind of "reverse FLE," that is how bilinguals' intention to tell the truth or to lie determines the choice of language rather than how the language of operation constrains the likelihood of telling the truth or lying. Language use is thus considered a measured (dependent) rather than a manipulated (independent) variable.

To investigate SLU, we designed an online coin-drawing game and invited Chinese-English bilinguals to freely choose between their native language (Chinese) or the foreign language (English) to announce to their opponent whether or not they had drawn a coin in each trial. Each trial of the game started with a coin draw, the outcome of which was Coin Draw (a coin has been drawn) or No Coin Draw (no coin has been drawn; see Table 1).

After the outcome of a draw is displayed on the screen, participants decide whether to announce a coin (Bet) or not (Drop). If participants had drawn a coin and chose to announce one, they were thus telling the truth; whereas if they chose to announce a coin when they had not drawn one, they were lying. The announcement was made to an "Artificial Intelligence" agent (AI) who, in each trial, made the decision to accept or reject the participants' offer. This experimental context placed participants implicitly in a strategic context whereby they were likely to try and figure out how the AI made decisions, even though the mechanism behind the "AI" was in fact very simple.[1] They were thus incentivized to respond with self-serving lies, in a context where language choice did not seem critical, so that we could study their tendency to choose a particular language when lying or telling the truth. In addition, we were able to consider switching behavior from one trial to the next, to investigate whether participants would change language strategically. Given that switching behavior is exhibited based on acceptance of the last bet, it might be bilinguals perceive the language choice as crucial.

Our hypotheses stem from the idea that emotional concern is reduced in the second language (e.g., Wu and Thierry, 2012; Jończyk et al., 2016). Since lying is usually associated with greater emotional challenge, and since decision-making is less influenced by emotion in the L2, lying could be less demanding in the second language, or at least, bilinguals might feel less guilty about lying in their L2, leading them to choose their second language more often when they have to lie. We thus hypothesized that Chinese-English bilinguals would choose the foreign language (English) over the native language when they had elected to lie about the outcome of a draw (i.e., announcing that they have a coin when they did not draw one). As for announcements of truth (announcing the possession of a coin when they had one), we had no prediction regarding the choice of language. Thus, we predicted a higher count of foreign language selection for "bet" trials without a coin (lies) and no difference for "bet" trials with a coin (truths).

Reaction times (RTs) should show a consistent pattern: longer RTs for native language choices when lying relative to foreign language choices, and no expected differences for truth trials (longer RTs would thus be expected when participants (1) have to lie or (2) are emotionally challenged, meaning that their RTs would be longer when lying in the native language).

We further investigated the occurrence of language switches between consecutive trials and also the direction of switching

1 Participants' answers at debriefing support that, as anticipated, most of them failed to understand how the AI worked. On a scale of 1–5 (1 being "no" and 5 being "certain") rating how confident they were that they had guessed what the AI was doing, they had an average rating of 2.26 ± 1.06. This suggests that the great majority of them did not resolve the AI as being a simple algorithm.

TABLE 1 Possible coin draw outcomes, conditions, and AI decisions.

Draw	Decision	Condition	AI decision	Score	Probability
Coin	Bet	Truth	Accept	+1	0.6
			Reject	−1	0.4
	Drop	Drop	–	−1	–
No Coin	Bet	Lie	Accept	+2	0.9 × 2 → 0.1
			Reject	−2	0.1 × 2 → 0.9
	Drop	Drop	–	0	–

(Chinese-to-English or English-to-Chinese) when it happened. By examining when and how language switches occurred, we sought to tease apart overall strategic effects (that is, what language is preferred in a given lie/truth condition) from language priming effects (how language use depends on previous language choices independent of the current lie/truth condition) likely to occur from one trial to the next, given that language exposure was mixed throughout the experiment. However, for this exploratory part of the analysis, we did not formulate a hypothesis as to the direction of the effects. The structure of the experiment and its main hypotheses have been pre-registered on aspredicted.com[2] (https://aspredicted.org/R7X_THY).

2 Methods

2.1 Participants

Sixty-four Chinese-English bilingual speakers participated in this study. We excluded 9 participants who were in one of the following situations: misunderstanding of the game rules (or failure to follow instructions) or aborted experiment (i.e., incomplete datasets). We did not include in the analyses data from 13 participants who did not vary their response as a function of language (i.e., who responded using the same key throughout the experiment, regardless of the language associated with it) and one participant who failed to show any form of strategic playing (e.g., betting all the time).[3] The number of participants was based on a heuristic of testing at least 40 participants, as stated in the pre-registration. The analyses were thus conducted on 41 individuals [27 females; Mean age = 25, SD = 3.15; 35 right-handed and six ambidextrous users, assessed by Edinburgh Handedness Inventory (EHI, Oldfield, 1971)]. Participants included in the analyses had college-level education and they had received an average of 17.7 ± 2.3 years of education.

TABLE 2 Chinese-English bilinguals' L2 background.

Measure	Mean	SD
Age of L2 acquisition	9.1	2.8
Length of L2 learning	15.2	3.5
Daily Chinese usage (%)	76.2	14.7
Daily English usage (%)	21.7	13.7

Participants' language proficiency was assessed using the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). All participants had exposure to Chinese from birth and low-to-average levels of daily exposure to English at the time of testing (46% ± 10% on a scale of 0%-almost never to 100%-always). They self-reported their language proficiency in terms of listening comprehension, speaking, and reading (0-none, 5-adequate, 10-perfect; Chinese, $M_{\text{listening}} = 9.51$, SD = 0.60; $M_{\text{speaking}} = 9.15$, SD = 0.85; $M_{\text{reading}} = 9.66$, SD = 0.57; English, $M_{\text{listening}} = 6.68$, SD = 1.66; $M_{\text{speaking}} = 5.76$, SD = 1.61; $M_{\text{reading}} = 7.29$, SD = 1.63). Participants had low to medium proficiency in English, making them unbalanced Chinese-English bilinguals who mostly used and considered English a foreign language. Participants' L2 background is summarized in Table 2.

All participants self-reported normal or corrected-to-normal vision, no hearing impairments, and no learning or language disabilities. They were recruited from mainland China and received payment for their participation. Participants' motivation was heightened by informing them that the top three high scorers in the game would get a payment bonus after the end of the data collection. They gave written consent to participate in the study. The experiment was approved by the ethics committee of the

2 Note that the pre-registration also mentioned two other participant groups (Spanish-English and Welsh-English bilinguals) for whom the data collection could not be completed within the timeframe stated in the pre-registration form. We did not have any prediction regarding between-group comparisons and the completion of data collection in the Chinese-English bilingual group enabled us to test our core hypothesis of strategic language use in a context of lying.

3 There was a slight divergence between the exclusion criteria used in the analyses and those specified in the pre-registration form, due to an unexpectedly large proportion of participants who elected to respond with the same key throughout the experiment. We decided to exclude these participants after data collection as it was not possible to know whether they used the same key to bet throughout the experiment because (1) they actually did not want to be strategic about language choice, (2) they wanted to minimize effort, or (3) they assumed that language was irrelevant in the game based on the design. The use of both keys does not necessarily imply strategic language use, but at least it ensured that participants were not just minimizing effort or ignoring the language manipulation. This was the only exclusion criterion that differed between pre-registration and current analysis.

School of Human and Behavioral Sciences at Bangor University (authorisation number: 2021-16892).

2.2 Materials

Two sets of written instructions were prepared, one in Mandarin Chinese and one in English, alongside screen captures of the coin drawing game to illustrate the different situations in which participants could find themselves when playing. The main experiment screen appeared like a game app (see Figure 1) with a circle which could contain a golden coin or not, depending on the draw of the trial. The choices of "offers" written in Chinese and English were each presented next to the image of a keyboard key (F or J). The image of the golden coin was presented between the two keys as well as the image of a spacebar, a display of the current trial's score, and the current overall score.

2.3 Experimental design

The rules of the game were the following: When participants drew a coin shown in a circle at the top of the screen (Coin Draw), they were instructed to bet by pressing either the F or J key on the keyboard, depending on the language in which they wanted to make the offer. They were explicitly told that they should bet when they had a coin (Truth condition). The AI opponent then "made a decision" to either accept or reject their bet. If the AI accepted the bet, the score for that trial was +1; if the bet was rejected by the AI, the score for this trial was −1 (Table 1). Since they were instructed to bet when they drew a coin, participants automatically scored −1 if they pressed the spacebar to "drop" (i.e., not bet). In the case where participants did not draw a coin (No Coin Draw), they could freely choose to bet or drop by pressing the spacebar. If they chose to bet without a coin (Lie condition), they could also freely choose the language in which they made the offer by pressing the F or J key. The AI opponent then "made a decision" to accept or reject the bet. If the AI accepted the bet, the score for that trial was +2 (since it involved the risk of lying); if the bet was rejected by the AI, the score was −2 (since the lie was "detected"). However, since betting involved a lie, dropping by pressing the spacebar incurred no loss of points.

Importantly, the language options for betting were randomly assigned to the F and J key on a trial-by-trial basis, such that neither key was more associated with one language more than the other across the experimental session. This was a way of making it particularly effortful for participants to systematically choose the same language to make their offers (i.e., behave in a monolingual way) whilst at the same time ensuring that pressing always the same key would fail to reveal any language preference. However, the issue of language choice was never announced or described explicitly in the instructions. Note that the decision to drop (spacebar) involved no language choice.

Unbeknownst to the participants, the "AI" agent making decision was a very simple probability-based algorithm. In the truth condition (Coin Draw), the AI accepted bets at a fixed rate of 60%, in order to produce an overall increase in score for Coin Draw trials (motivation to play). In the lie condition (No Coin Draw), the AI accepted bets at a rate of 90% for two successive lies and the acceptance rate fell to 10% for the third lie in a row. This change in probability prevented participants from systematically choosing to bet in No Coin Draw trials, which would amount to non-strategic playing. By focusing on understanding the AI behavior and finding the best strategy to gain points, we steered participants' attention away from the language choice they needed to make in every Bet trial. The acceptance rate reverted to 90% only after participants dropped a bet after two consecutive lies, the rate staying at 10% if they kept betting without a coin. Note that the AI made no decision when participants chose to drop (spacebar).

2.4 Procedure

After completing and signing the consent form, participants either completed the questionnaires (LEAP-Q and EHI) on Qualtrics or started the experiment (the game) with order randomized between participants. The experiment was programmed in OpenSesame (Mathôt et al., 2012) and run in JATOS (Lange et al., 2015), and consisted of a practice session with 10 trials and a test session with 160 trials.

In order to place participants in a bilingual context, the experiment started with written instructions presented paragraph-by-paragraph with English and Mandarin Chinese alternating from one paragraph to the next. The language of the first paragraph was chosen at random for each participant. Participants were informed that their goal was to score as many points as they could during the game. Then, screen captures illustrating all possible conditions of the game (six cases in total) were presented, with explanatory text and arrows in the participants' native or foreign language to familiarize them with the game before the session started. Then participants completed 10 practice trials before the overall score was reset and data collection started.

In each trial, a golden coin flickered inside a circle presented at the top of the screen for ∼480 ms. Participants then were given the result of the coin draw for the current trial: either they saw a coin (Coin Draw) or they did not see a coin (no Coin Draw). They then either announced a coin (bet) or drop for this trial. If announcing a Coin Draw, they could press the "F" or "J" key on the keyboard to choose the language in which they wanted to announce the coin (我有硬币—"I have a coin"). Coin and No Coin draws were equally probable (80 trials of each type) throughout the session and appeared in randomized order. The association between the response keys and language was also equiprobable and randomized independently of coin draws. There was no time limit for participants to make a decision. After their response, if they had bet, they were informed of the AI's decision through an icon (shaking hands for acceptance and a stop hand for rejection) and saw a display of the current trial's score before they proceeded to the next trial.

After the experiment was completed, participants were asked to answer questions in English regarding their thoughts about the AI's behavior and possible relationships between their choice of language use and their decisions (translations of answers in Chinese

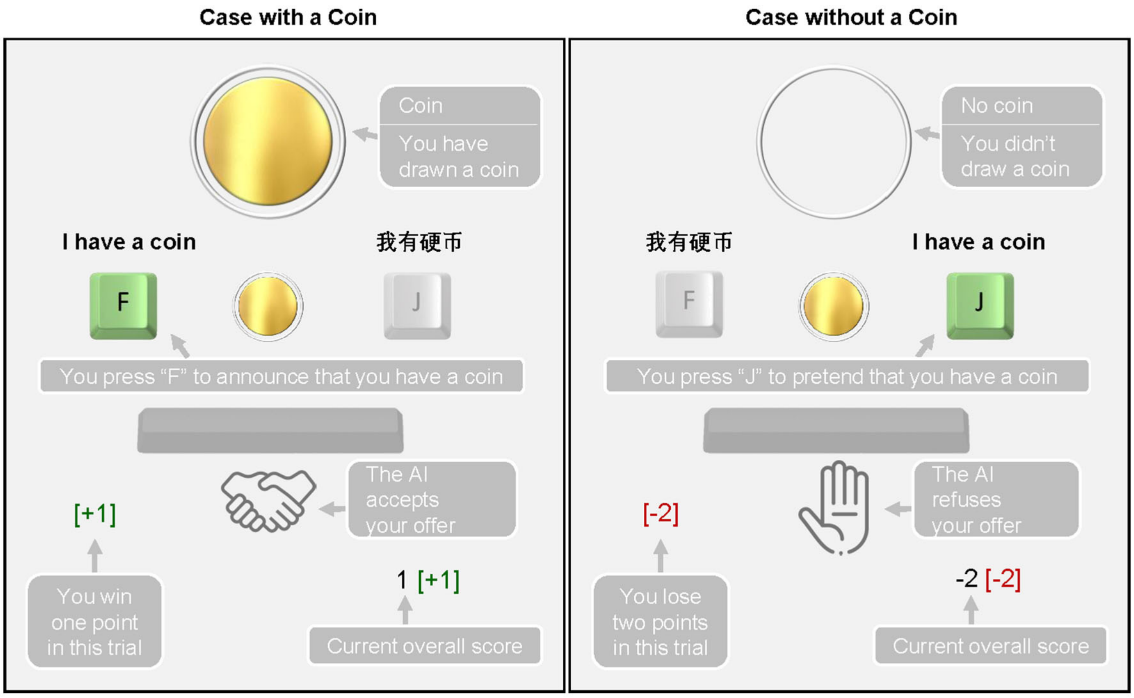

FIGURE 1
Examples of instructions screens of the coin drawing game [**(left)** case when a coin is drawn, **(right)** case when no coin is drawn]. Note that the commentaries in gray rounded boxes appear only during the pre-session instructions.

are available on the Open Science Framework: https://tinyurl.com/4nc4saan).

2.5 Analysis

In what follows, unless otherwise specified, all data analyses concern bet trials, that is, trials in which participants announced to the AI that they had a coin, irrespective of whether or not they did draw one. Indeed, by design, drop trials do not convey any information regarding language choice. Therefore, the condition in which participants drew a coin and bet is conceptually equivalent to a "truth" trial; and the condition in which they had no coin and bet is conceptually equivalent to a "lie" trial. First, we conducted an ANOVA looking at the effects of Truth Value and language use on counts of occurrence (i.e., how many trials of each kind were counted) in order to determine whether one language was used more than the other overall (main effect of language), and whether, as expected, participants engaged in more truth than lie decisions (main effect of Truth Value), exploring also any interactions between these two factors. The dependent variables were counts of occurrence (language choice) and reaction times (RTs). As stated in the pre-registration of this study, trials in which RTs were too short to be plausible (<200 ms) or too long to be representative (>5,000 ms) were discarded before analysis.

Beyond this first analysis, we investigated language switching behavior based on proportions normalized within each of the truth and lie conditions considered separately, taking into account the language of the previous trial. We elected to do this to correct for the main effect of Truth Value expected in the first analysis above, which was driven by the structure of the experiment: participants had to bet when they had a coin and could choose to drop when they did not, meaning that the count of lie trials had to be necessarily lower than that of truth trials. Normalizing data within each of the truth and lie conditions separately was expected to eliminate the difference between conditions driven by the experimental design and discarded the possibility that any interactions between Truth Value and other independent variables would be artificially driven by the main effect of Truth Value. We made no predictions regarding the proportion of language switches between conditions or the direction of language switching as this was the exploratory part of the analysis. Instructions, compiled data, and statistical output are available on the Open Science Framework (https://tinyurl.com/4nc4saan). The analyses regarding language switching behavior were exploratory and not pre-registered.

3 Results

3.1 Truth value and language use

We first analyzed counts of announcements in either of the two languages and the corresponding reaction times (RTs) in bet trials. Overall, participants did not choose English over Chinese or *vice versa* when announcing coins, since there was no main effect of Language, $F_{(1,40)} = 2.03$, $p = 0.16$, $\eta_p^2 = 0.005$. There was, however, a main effect of Truth Value, $F_{(1,40)} = 146.75$, $p < 0.001$, $\eta_p^2 = 0.786$, with greater counts of truth than lie announcements. There was also a significant interaction between Language and Truth Value,

$F_{(1,40)} = 11.06$, $p = 0.002$, $\eta_p^2 = 0.217$, such that more truths were announced in Chinese than in English, $t_{(40)} = 2.78$, $p_{bonferroni} = 0.050$, whereas no such difference was found for lies, $t_{(40)} = -1.38$, $p_{bonferroni} = 1$ (see Figure 2).

A control analysis looking at side of response ("F" and "J" key) in relation to language found no main effect or interaction ($ps > 0.1$). As regards RTs, there was only a main effect of Truth Value on response latency, $F_{(1,40)} = 14.29$, $p < 0.001$, $\eta_p^2 = 0.263$, such that language selection for truth announcements was faster than language selection for lies. No other main effect or interaction was significant ($ps > 0.1$).

We then analyzed language switching behavior in our participants, that is, their tendency to change language from one bet to the next. Amongst the different situations in which participants could find themselves, we chose to explore the cases in which they had bet in the previous trial, whether they had lied or not, whether the AI had accepted or rejected their offer, and the cases in which participants had previously declined to bet (drop trials). In all cases, we studied the relationship between Truth Value and AI Decision in relation to (a) switching behavior (switch/no switch) and (b) switching direction in switching trials (English-to-Chinese or Chinese-to-English).

3.2 Truth value, AI decision, and switching occurrence (all bet analysis)

A repeated measures ANOVA was conducted to investigate the relationship between Truth Value (truth, lie), AI Decision in the previous trial (accept, reject), and Switching Occurrence (switch, no switch). There was a main effect of AI Decision,[4] such that more bets were accepted than rejected overall, $F_{(1,40)} = 16.88$, $p < 0.001$, $\eta_p^2 = 0.297$. We also found an interaction between AI Decision and Switching Occurrence, $F_{(1,40)} = 13.72$, $p < 0.001$, $\eta_p^2 = 0.255$, such that participants were less likely to switch between languages when the AI accepted their offer as compared to when the AI refused it (see Figure 3). There was no other significant main effect or interaction in this analysis ($ps > 0.1$).

As regards RTs, there was a main effect of Truth Value, $F_{(1,40)} = 8.47$, $p = 0.006$, $\eta_p^2 = 0.175$, such that participants took longer RTs to lie than tell the truth. We also found a main effect of AI Decision, $F_{(1,40)} = 15.23$, $p < 0.001$, $\eta_p^2 = 0.276$, showing that participants were slower when they had been accepted by the AI in the previous round than when they had been rejected. There was no other main effect and interaction ($ps > 0.1$).

4 It should be noted the main effect of AI Decision was a programming artifact, likely driven by the instruction to bet when participants had a coin, and indeed, in cases where they had drawn a coin and announced it, the probability of acceptance on the part of the AI agent was fixed at 60% against 40% rejection. However, for the lie trials it happens that participants overall got more accepted than rejected also, which means that overall, participants managed to extract some information regarding AI operation (albeit implicitly because they could not provide a metacognitive account of this at debrief).

3.3 Truth value, AI decision, and switching direction (switch trials only)

A repeated-measures ANOVA was conducted to investigate the relationship between Truth Value (Truth, Lie), AI Decision in the previous trial (accept, reject), and Switching Direction (Chinese to English, English to Chinese). We found an interaction between Truth Value and Switching Direction, $F_{(1,40)} = 6.68$, $p = 0.014$, $\eta_p^2 = 0.143$, such that participants were more likely to switch from English to Chinese when they had told the truth in the previous trial, $t_{(40)} = 2.88$, $p_{bonferroni} = 0.038$ (see Figure 4). There was no other significant main effect or interaction in this analysis ($ps > 0.1$).

As regards RTs, there was a main effect of Truth Value, $F_{(1,40)} = 4.93$, $p = 0.006$, $\eta_p^2 = 0.123$, such that participants took longer RTs to lie than tell the truth. We also found a main effect of AI Decision, $F_{(1,40)} = 5.27$, $p = 0.028$, $\eta_p^2 = 0.131$, showing that participants were slower when they had been accepted by the AI in the previous round than when they had been rejected. There was no other main effect or interaction ($ps > 0.1$).

3.4 Language switching after a trial in which participants did not bet (drop trials)

A paired samples *t*-test showed that bilingual participants tended not to switch between their languages more after a drop trial than after a bet trial ($p > 0.05$). No significant difference in frequency was found between switch directions, and there no significant difference in RTs between switch directions (all $ps > 0.05$).

3.5 Metacognitive assessment of SLU by participants

At the debriefing, we asked participants to explain if they had been able to work out the process underlying AI decisions and whether they felt that they had used one language over the other in relation to betting. None of the participants was able to extract or describe in a meaningful way how they thought the AI operated. One participant had the intuition that the experiment concerned language over and above decision-making and betting, but even that participant, like all others, did not report overtly preferring one language over the other, when telling a lie or telling the truth (see Supplementary Table 1). The data from this participant was included in the analyses but we reran the analyses after excluding the corresponding dataset and the results remained qualitatively unchanged.

4 Discussion

Here, we investigated strategic language use (SLU) as a kind of "reverse FLE," that is, bilinguals' tendency to preferentially use one of their languages when they are incentivized to lie in an extremely simplified—but tightly controlled—language-choice

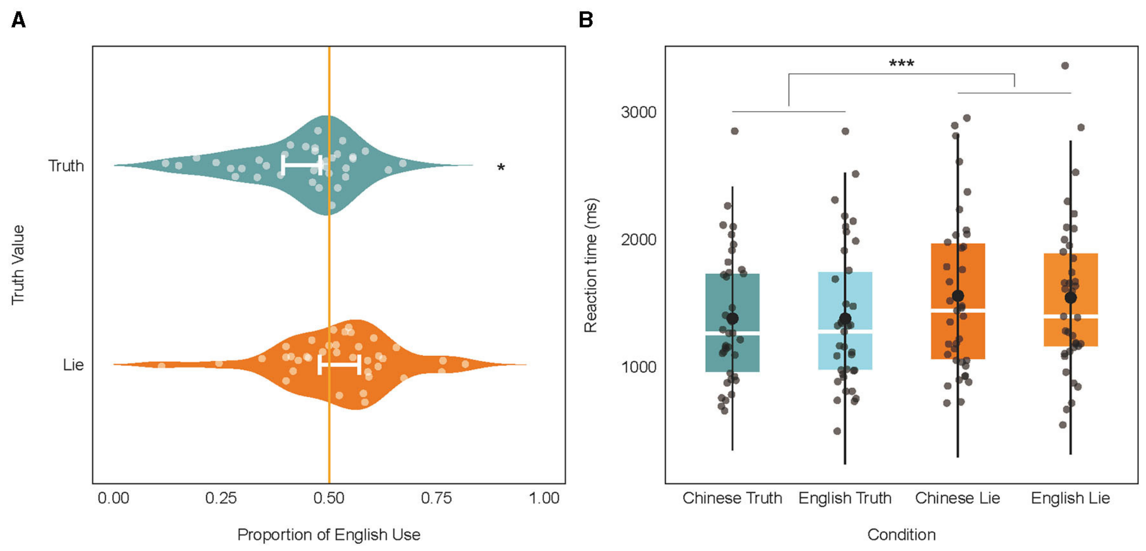

FIGURE 2
Language use in truth and lie conditions. **(A)** Violin plots of proportions of language use (0 is 100% Chinese and 1 is 100% English). **(B)** Boxplot of reaction times for each language x truth value combination. Dots represent data for each participant. Error bars in violin plots depict 95% CIs. *$p \leq 0.05$, ***$p \leq 0.001$.

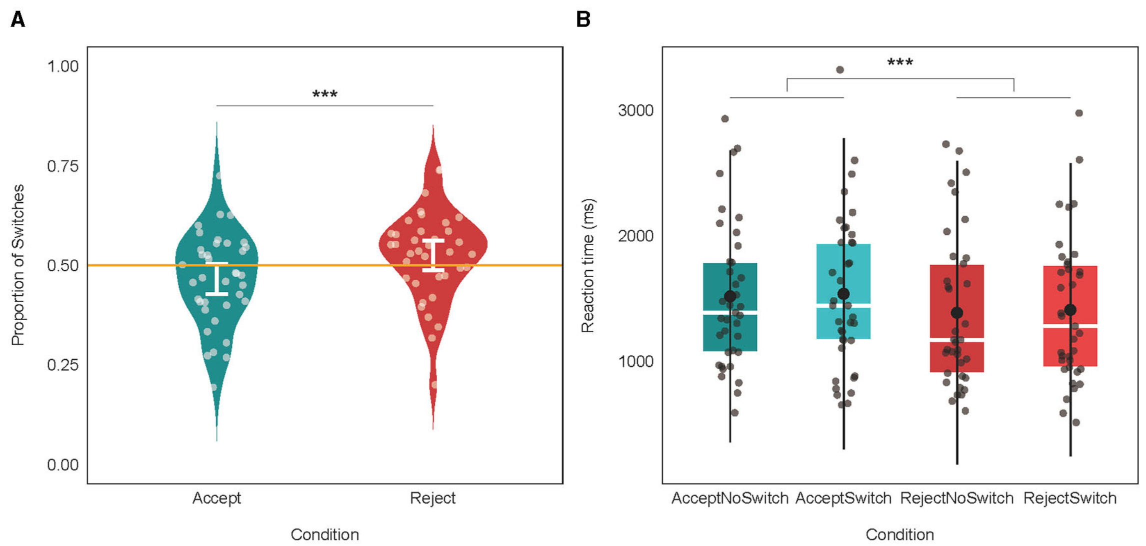

FIGURE 3
Switching occurrence when participant bets were accepted and rejected in truth and lie conditions respectively. **(A)** Violin plots of proportions of switches (0 is no switch and 1 is 100% switch). **(B)** Boxplot of reaction times for each AI decision by switch combination. Dots represent data for each participant. Error bars in violin plots depict 95% Cis. ***$p \leq 0.001$.

paradigm. Participants did not show the expected preference for English (the foreign language) over Chinese (their native language) when they lied, and their RTs did not differ between languages. However, they displayed a preference for Chinese over English when telling the truth, and bilinguals tended to switch less from one language to the other when their offer had been accepted rather than rejected. In addition, participants were more likely to switch from English-to-Chinese after a truth-telling trial and reciprocally, from Chinese-to-English, after a lie trial (both marginal effects). When bilinguals chose not to play in the previous trial (after a drop trial), they did not have a preferred switching direction. It is noteworthy that participants were likely unaware of the core manipulation of language use in the experiment since they were unable to report what was being tested apart from the decision process itself in the betting game. Moreover, at debrief, participants were surprised to be asked about a possible link between their language use and their being truthful or deceitful, suggesting that participants were not consciously aware of the rationale underlying the experiment.

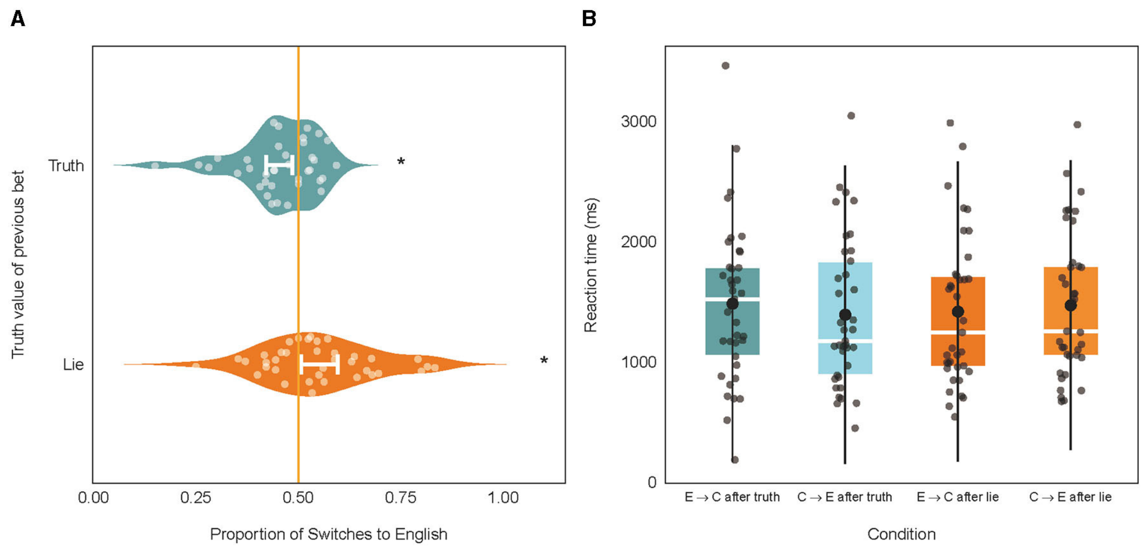

FIGURE 4
Switching direction in relation to the nature of the previous bet (Truth or Lie). **(A)** Violin plots of proportions of switches to English (0 is 100% switch to Chinese and 1 is 100% switch to English). **(B)** Boxplot of reaction times for each truth value in the previous trial and switch direction combination. Dots represent data for each participant. Error bars in violin plots depict 95% CIs. *$p \leq 0.05$.

We interpret the greater use of Chinese over English when participants made truth announcements as a counterpoint of the core hypothesis underlying this study rather than a negative result. Indeed, this result is consistent with the hypothesis that bilinguals use their languages strategically when trying to appear convincing in the eyes of an opponent (in this case an AI agent). In fact, if we consider participants' response pattern in their L1 Chinese as the baseline, the relative proportion of lies in English was greater. There is no reason why the difference between languages should numerically manifest in the lie condition alone as it is essential to compare language choice across conditions, and indeed we found an interaction between Language and Truth Value. Furthermore, this result is consistent with previous studies that have reported attenuated emotional resonance of the foreign language compared with the native language (Caldwell-Harris and Ayçiçegi-Dinn, 2009; Keysar et al., 2012; Wu and Thierry, 2012; Jończyk et al., 2016, 2019; Gao et al., 2020), since telling the truth likely triggers the need to appear sincere and trustworthy. In addition, it is not entirely surprising that the majority of truth announcements were made in the bilinguals' native language, if we consider the cooperative principle of Grice's Maxim of Quality (Grice, 1989). According to this principle, a speaker naturally avoids saying anything that they consider to be false, and it follows that we typically tend to believe that others' statements are true, leading to the classic truth bias effect (DePaulo et al., 1997). Statistically, despite the relative pervasiveness of lies in daily communication, truth statements are overwhelmingly more frequent, making us assume that most statements that we hear are true (O'Sullivan, 2003). Hence, it could be inferred that stating the truth is the default mode of communication, and that this default might concern the native language more than the foreign language, given that it is the most commonly used language in imbalanced bilinguals.

Moreover, an interesting idea is that language choice serves as a way to (implicitly) signal honesty, and indeed, such a hypothesis is consistent with our results: Bilinguals prefer telling the truth in Chinese, because the native language is implicitly associated with telling the truth. As suggested by an anonymous reviewer, language choice when lying might be influenced by one's concept of the self, to which honesty is likely to be an important contributing factor (Strohminger and Nichols, 2014). Therefore, one could have predicted that lying in the native or the second language depends on whether or not lying is perceived as desirable (Gai and Puntoni, 2021). Since it is generally assumed that people tell the truth most of the time (DePaulo et al., 1996; Levine, 2014) and given that they dominantly operate in the native language, lying in the native language is likely to be considered unacceptable and might therefore be incompatible with one's self-concept (Alempaki et al., 2021). There may thus be a relationship between honesty and native language use, which could lead to the prediction that when bilinguals tell the truth, they are more likely to use the native language. However, it must be noted that the previous references concern the likelihood of lying when participants are in a given language context, rather than choosing the language in which to operate in a lying context.

As for the main effect of lying on RTs, participants' faster RTs for truths than lies is a classical RT deception effect (see Suchotzki et al., 2017 for a review). It may thus be that longer RTs in the lie condition have blurred the language selection patterns in this condition, whereas shorter RTs in the truth condition may have helped reveal language choice patterns.

An alternative hypothesis which we did not consider here is that using the foreign language is more cognitively demanding. Some studies have shown that increasing the cognitive load in decision making experiments tends to increase the proportion of truth decisions (Gilbert et al., 1990). Thus, we could have predicted that bilingual participants would be more likely to choose their native language to lie, given that it would be more difficult to do so in the second language. However, we took great care in the current experiment to minimize cognitive load differences between languages since language choice was reduced to its simplest expression of choosing between two highly repeated short statements (i.e., "I have a coin"/"我有硬币"). Furthermore, the results are inconsistent with such a hypothesis given the lack of any RT differences between languages either in the lie or the truth condition.

It should be noted at this point that the main effect of Truth Value showing a greater count of truth trials than lie trials was expected and stemmed directly from the experimental design, since participants were instructed to tell the truth in 50% of the trials but could choose to lie or drop in the other 50% of trials. Thus, since they were instructed to do so, participants almost always announced a coin when they had one (i.e., a truth trial), with the incidence of drop trials being very low (Mean = 2.5 ± 1.4 out of 80 trials). When they did not draw a coin, participants could choose to announce a coin (lie) or drop, which necessarily drove the number of lie trials down (Mean = 26.9 ± 14.5 out of 80 trials). The critical observation, however, is the interaction between Truth Value and Language use discussed above. It is possible that the language difference was only found on truth trials because they were more frequent than lie trials, and thus a difference of language use in the lie condition may also have been present but undetected due to low statistical power.

Regarding language switching behavior, bilinguals tended to switch less between languages after the AI accepted their offer than when their offer was rejected. This might relate to the "hot hand effect" (Gilovich et al., 1985; Ayton and Fischer, 2004), according to which participants may misconstrue an acceptance decision as a sign that they are more likely to win again in the next trial, in which case they tend to stick with their previous language choice. When the AI rejects their offer in the previous trial, however, participants would be more likely to change "strategy" which may lead to a greater likelihood of language switching.

As regards switching direction, we found that participants were more likely to switch from English to Chinese when they had told the truth in the previous trial. This could be considered as a delayed decision priming effect on language irrespective of the decision to tell the truth or lie in the next trial. In other words, we could consider that telling the truth in a given trial primes Chinese usage in the next (Brodeur and Lupker, 1994). Consistent with our original hypothesis, we could expect that when participants lied in a given trial, they would be primed to use English in the next trial. However, this was not confirmed by the data analysis. This result is consistent with Grice's Maxims (Grice, 1989), if we consider using the native language and telling the truth as the default operation mode. When this default mode of operation is disrupted by external factors (the requirement to lie strategically), the cooperative principle of the Maxim of Quality is violated and thus participants do not use either language preferentially to lie. Recall that we also found main effects of Truth Value and AI Decision in the subcase of switch trials and these effects were broadly aligned with the overall effects found in the "all bet analysis" suggesting that these main effects are stable: Participants are generally slower to respond when they lie.

It must be kept in mind that the results obtained likely reflect extraneous variables which were not manipulated in this study, such as the age of acquisition (e.g., Tremblay, 2006; Ferré et al., 2018), frequency of use (e.g., Kroll and Stewart, 1994) and exposure (e.g., Tremblay, 2006), affective relationship to the foreign language (e.g., Eilola and Havelka, 2011), cultural effects, etc. Participants in this study were unbalanced late bilinguals living in China (that is, immersed in Chinese culture) with much higher proficiency in Chinese than English. They started learning English from the age of nine and used English for about a fifth of their daily life. Although we recruited Chinese-English bilinguals with a CET-4, or CET-6 score of 450, or IELTS score of 6 or higher and assessed their fluency in the native and foreign language based on self-reports (LEAP-Q), fluency measures may have lacked precision, since the timing of English tests was not controlled. In addition, it is noteworthy that although self-reported tests such as LEAP-Q provide valuable insights, they may not always align precisely with objective assessments and should be considered alongside standardized test results when evaluating participants' proficiency. All these factors could have contributed to differences in language use, although an exploratory test of relationships between such predictors and proportion of Chinese and English use in the truth and lie conditions, respectively, failed to show any significant relationship (see Supplementary Table 2, possibly because the range of language use was limited in our sample).

Finally, the current study contributes to a nascent theoretical understanding of strategic language use in bilinguals, showing that individuals who can communicate in more than one language do not use their languages interchangeably and solely based on the language in which they are being addressed. This helps us grasp the complexity of language selection, especially when emotions come into play. Furthermore, the practical implications of our research extend to the realm of law enforcement and criminal justice where deception is frequently encountered. For instance, a court of law may want to require the use of native language in particular conditions to better support the investigation of criminal cases and ensure greater fairness and transparency.

5 Conclusion and future directions

Overall, the current study aimed to investigate Chinese-English bilinguals' strategic language use (SLU) in a situation that incentivized lying behavior. For the first time, we establish that Chinese-English bilinguals tend to use their native and foreign languages strategically when making true and false statements. Specifically, bilinguals preferred to use Chinese over English to tell the truth and were primed to use Chinese after telling the truth. But we can interpret that they chose English over Chinese to tell lies if we consider their behavior when telling the truth as a baseline. The results suggest the existence of a reliable link between truth value and language choice in bilinguals, especially in the context of a demanding game of bets in which language choice is irrelevant (no language-directed instruction). Future research may explore whether strategic language choice can be observed overall in the case of lying under different conditions or in a different population of bilinguals and examine likely cultural factors at play, with implications for the generalization of the results to other cultural settings. Future investigations will also hopefully tackle the nature of interactions between feedback and SLU (e.g., Gao et al., 2015).

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://tinyurl.com/4nc4saan.

Ethics statement

The studies involving humans were approved by School of Human and Behavioural Science, Bangor University, Bangor, United Kingdom. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

WY: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. PR: Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – review & editing. CF: Conceptualization, Investigation, Methodology, Project administration, Software, Validation, Writing – review & editing. OM-N: Conceptualization, Investigation, Methodology, Project administration, Validation, Writing – review & editing. JD: Conceptualization, Investigation, Methodology, Project administration, Validation, Writing – review & editing. GT: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. WY was supported by the Chinese Scholarship Council (202006070017). GT was supported by the Polish National Agency for Academic Exchange (NAWA) under the NAWA Chair Programme (PPN/PRO/2020/1/00006).

In memoriam

In memoriam of Albert Costa.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/flang.2023.1293673/full#supplementary-material

References

Abe, N., Suzuki, M., Mori, E., Itoh, M., and Fujii, T. (2007). Deceiving others: distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *J. Cogn. Neurosci.* 19, 287–295. doi: 10.1162/jocn.2007.19.2.287

Alempaki, D., Dogan, G., and Yang, Y. (2021). Lying in a foreign language? *J. Econ. Behav. Organ.* 185, 946–961. doi: 10.1016/j.jebo.2020.10.027

Ayton, P., and Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness? *Mem. Cognit.* 32, 1369–1378. doi: 10.3758/BF03206327

Bereby-Meyer, Y., Hayakawa, S., Shalvi, S., Corey, J. D., Costa, A., Keysar, B. (2020). Honesty speaks a second language. *Top. Cogn. Sci.* 12, 632–643. doi: 10.1111/tops.12360

Bizzi, E., Hyman, S. E., Raichle, M. E., Kanwisher, N., Phelps, E. A., Morse, S. J., et al. (2009). *Using Imaging to Identify Deceit: Scientific and Ethical Questions*. Cambridge, MA: American Academy of Arts and Sciences.

Brodeur, D. A., and Lupker, S. J. (1994). Investigating the effects of multiple primes: an analysis of theoretical mechanisms. *Psychol. Res.* 57, 1–14. doi: 10.1007/BF00452990

Buller, D. B., and Burgoon, J. K. (1996). "Emotional expression in the deception process," in *Handbook of Communication and Emotion*, eds P. A. Andersen, and L. K. Guerrero (Cambridge: Academic Press), 381–402. doi: 10.1016/B978-012057770-5/50016-3

Caldwell-Harris, C. L. (2015). Emotionality differences between a native and foreign language: implications for everyday life. *Curr. Dir. Psychol. Sci.* 24, 214–219. doi: 10.1177/0963721414566268

Caldwell-Harris, C. L., and Ayçiçegi-Dinn, A. (2009). Emotion and lying in a non-native language. *Int. J. Psychophysiol.* 71, 193–204. doi: 10.1016/j.ijpsycho.2008.09.006

Corey, J. D., Hayakawa, S., Foucart, A., Aparici, M., Botella, J., Costa, A., et al. (2017). Our moral choices are foreign to us. *J. Exp. Psychol.: Learn. Mem. Cogn.* 43, 1109. doi: 10.1037/xlm0000356

Costa, A., Foucart, A., Arnon, I., Aparici, M., and Apesteguia, J. (2014). "Piensa" twice: on the foreign language effect in decision making. *Cognition* 130, 236–254. doi: 10.1016/j.cognition.2013.11.010

De Bruin, A., Samuel, A. G., and Duñabeitia, J. A. (2018). Voluntary language switching: when and why do bilinguals switch between their languages? *J. Mem. Lang.* 103, 28–43. doi: 10.1016/j.jml.2018.07.005

De Bruin, A., Samuel, A. G., and Duñabeitia, J. A. (2020). Examining bilingual language switching across the lifespan in cued and voluntary switching contexts. *J. Exp. Psychol. Hum. Percept. Perform.* 46, 759. doi: 10.1037/xhp0000746

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., and Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Pers. Soc. Psychol. Rev.* 1, 346–357. doi: 10.1207/s15327957pspr0104_5

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., and Epstein, J. A. (1996). Lying in everyday life. *J. Pers. Soc. Psychol.* 70, 979–995. doi: 10.1037/0022-3514.70.5.979

Dewaele, J. (2010). *Emotions in Multiple Languages*. Basingstoke: Palgrave-Macmillan. doi: 10.1057/9780230289505

Dewaele, J.-M. (2004a). Blistering barnacles! What language do multilinguals swear in? *Estud. Sociolingüíst.* 5, 83–106. doi: 10.1558/sols.v5i1.83

Dewaele, J.-M. (2004b). The emotional force of swearwords and taboo words in the speech of multilinguals. *J. Multiling. Multicult. Dev.* 25, 204–222. doi: 10.1080/01434630408666529

Dewaele, J.-M. (2008a). "The effect of type of acquisition context on perception and self-reported use of swearwords in L2, L3. L4 and L5," in *Investigations in Instructed Second Language Acquisition*, eds A. Housen, and M. Pierrard (Berlin: De Gruyter Mouton), 531–560. doi: 10.1515/9783110197372.4.531

Dewaele, J.-M. (2008b). The emotional weight of I love you in multilinguals' languages. *J. Pragmat.* 40, 1753–1780. doi: 10.1016/j.pragma.2008.03.002

Dong, G., Wu, H., and Lu, Q. (2010). Attempting to hide our real thoughts: electrophysiological evidence from truthful and deceptive responses during evaluation. *Neurosci. Lett.* 479, 1–5. doi: 10.1016/j.neulet.2010.05.014

Duñabeitia, J. A., and Costa, A. (2015). Lying in a native and foreign language. *Psychon. Bull. Rev.* 22, 1124–1129. doi: 10.3758/s13423-014-0781-4

Eilola, T. M., and Havelka, J. (2011). Behavioural and physiological responses to the emotional and taboo Stroop tasks in native and non-native speakers of English. *Int. J. Biling.* 15, 353–369. doi: 10.1177/1367006910379263

Ferré, P., Anglada-Tort, M., and Guasch, M. (2018). Processing of emotional words in bilinguals: testing the effects of word concreteness, task type and language status. *Second Lang. Res.* 34, 371–394. doi: 10.1177/0267658317744008

Frances, C., De Bruin, A., and Duñabeitia, J. A. (2020a). The effects of language and emotionality of stimuli on vocabulary learning. *PLoS ONE* 15, e0240252. doi: 10.1371/journal.pone.0240252

Frances, C., De Bruin, A., and Duñabeitia, J. A. (2020b). The influence of emotional and foreign language context in content learning. *Stud. Second Lang. Acquis.* 42, 891–903. doi: 10.1017/S027226311900072X

Gai, J., and Puntoni, S. (2018). Major or minor: when foreign language increases versus decreases cheating. *ACR North Am. Adv.* 46, 547–548.

Gai, P. J., and Puntoni, S. (2021). Language and consumer dishonesty: a self-diagnosticity theory. *J. Consum. Res.* 48, 333–351. doi: 10.1093/jcr/ucab001

Gao, S., Luo, L., and Gou, T. (2020). Criticism in a foreign language hurts less. *Cogn. Emot.* 34, 822–830. doi: 10.1080/02699931.2019.1668751

Gao, S., Zika, O., Rogers, R. D., and Thierry, G. (2015). Second language feedback abolishes the "hot hand" effect during even-probability gambling. *J. Neurosci.* 35, 5983–5989. doi: 10.1523/JNEUROSCI.3622-14.2015

García-Palacios, A., Costa, A., Castilla, D., Río, E., Casaponsa, A., Duñabeitia, J. A., et al. (2018). The effect of foreign language in fear acquisition. *Sci. Rep.* 8, 1157. doi: 10.1038/s41598-018-19352-8

Garrett, N., Lazzaro, S. C., Ariely, D., and Sharot, T. (2016). The brain adapts to dishonesty. *Nat. Neurosci.* 19, 1727–1732. doi: 10.1038/nn.4426

Geipel, J., Hadjichristidis, C., and Surian, L. (2015). The foreign language effect on moral judgment: the role of emotions and norms. *PLoS ONE* 10, e0131529. doi: 10.1371/journal.pone.0131529

Gilbert, D. T., Krull, D. S., and Malone, P. S. (1990). Unbelieving the unbelievable: some problems in the rejection of false information. *J. Pers. Soc. Psychol.* 59, 601–613. doi: 10.1037/0022-3514.59.4.601

Gilovich, T., Vallone, R., and Tversky, A. (1985). The hot hand in basketball: on the misperception of random sequences. *Cogn. Psychol.* 17, 295–314. doi: 10.1016/0010-0285(85)90010-6

Grice, P. (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Hadjichristidis, C., Geipel, J., and Savadori, L. (2015). The effect of foreign language in judgments of risk and benefit: the role of affect. *J. Exp. Psychol. Appl.* 21, 117. doi: 10.1037/xap0000044

Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., and Keysar, B. (2017). Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychol. Sci.* 28, 1387–1397. doi: 10.1177/0956797617720944

Hsu, C.-T., Jacobs, A. M., and Conrad, M. (2015). Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. *Cortex* 63, 282–295. doi: 10.1016/j.cortex.2014.09.002

Iacozza, S., Costa, A., and Duñabeitia, J. A. (2017). What do your eyes reveal about your foreign language? Reading emotional sentences in a native and foreign language. *PLoS ONE* 12, e0186027. doi: 10.1371/journal.pone.0186027

Jończyk, R., Boutonnet, B., Musiał, K., Hoemann, K., and Thierry, G. (2016). The bilingual brain turns a blind eye to negative statements in the second language. *Cogn. Affect. Behav. Neurosci.* 16, 527–540. doi: 10.3758/s13415-016-0411-x

Jończyk, R., Korolczuk, I., Balatsou, E., and Thierry, G. (2019). Keep calm and carry on: electrophysiological evaluation of emotional anticipation in the second language. *Soc. Cogn. Affect. Neurosci.* 14, 885–898. doi: 10.1093/scan/nsz066

Keysar, B., Hayakawa, S. L., and An, S. G. (2012). The foreign-language effect: thinking in a foreign tongue reduces decision biases. *Psychol. Sci.* 23, 661–668. doi: 10.1177/0956797611432178

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008

Lange, K., Kühn, S., and Filevich, E. (2015). "Just Another Tool for Online Studies"(JATOS): an easy solution for setup and management of web servers supporting online studies. *PLoS ONE* 10, e0130834. doi: 10.1371/journal.pone.0130834

Levine, T. R. (2014). Truth-default theory (TDT) a theory of human deception and deception detection. *J. Lang. Soc. Psychol.* 33, 378–392. doi: 10.1177/0261927X14535916

Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–067. doi: 10.1044/1092-4388(2007/067)

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* 44, 314–324. doi: 10.3758/s13428-011-0168-7

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: the boy-who-cried-wolf effect. *Pers. Soc. Psychol. Bull.* 29, 1316–1327. doi: 10.1177/0146167203254610

Pavlenko, A. (2004). 'Stop doing that, ia komu skazala!': language choice and emotions in parent - child communication. *J. Multiling. Multicult. Dev.* 25, 179–203. doi: 10.1080/01434630408666528

Pavlenko, A. (2005). *Emotions and Multilingualism*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511584305

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Biling.: Lang. Cogn.* 11, 147–164. doi: 10.1017/S1366728908003283

Pavlenko, A. (2012). Affective processing in bilingual speakers: disembodied cognition? *Int. J. Psychol.* 47, 405–428. doi: 10.1080/00207594.2012.743665

Puntoni, S., De Langhe, B., and Van Osselaer, S. M. (2009). Bilingualism and the emotional intensity of advertising language. *J. Consum. Res.* 35, 1012–1025. doi: 10.1086/595022

Strohminger, N., and Nichols, S. (2014). The essential moral self. *Cognition* 131, 159–171. doi: 10.1016/j.cognition.2013.12.005

Suchotzki, K., and Gamer, M. (2018). The language of lies: behavioral and autonomic costs of lying in a native compared to a foreign language. *J. Exp. Psychol. Gen.* 147, 734. doi: 10.1037/xge0000437

Suchotzki, K., Verschuere, B., Van Bockstaele, B., Ben-Shakhar, G., and Crombez, G. (2017). Lying takes time: a meta-analysis on reaction time measures of deception. *Psychol. Bull.* 143, 428. doi: 10.1037/bul0000087

Tremblay, M.-C. (2006). Cross-linguistic influence in third language acquisition: the role of L2 proficiency and L2 exposure. *Clo/opl* 34, 109–119.

Tversky, A., and Kahneman, D. (1985). "The framing of decisions and the psychology of choice," in *Behavioral Decision Making*, ed. G. Wright (Berlin: Springer), 25–41. doi: 10.1007/978-1-4613-2391-4_2

Wu, Y. J., and Thierry, G. (2012). Unconscious translation during incidental foreign language processing. *Neuroimage* 59, 3468–3473. doi: 10.1016/j.neuroimage.2011.11.049