Applied Psycholinguistics (2021), **42**, 937–967
doi:10.1017/S0142716421000175

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

From "No, she does" to "Yes, she does": Negation processing in negative yes–no questions by Mandarin speakers of English

Haoruo Zhang[1] and Norbert Vanek[2]*

[1]College of Foreign Languages, University of Shanghai for Science and Technology, Shanghai, China and [2]School of Cultures, Languages and Linguistics, University of Auckland, Auckland, New Zealand
*Corresponding author. Email: norbert.vanek@auckland.ac.nz

(Received 18 January 2020; revised 14 April 2021; accepted 14 April 2021; first published online 10 June 2021)

Abstract

In response to negative yes–no questions (e.g., *Doesn't she like cats?*), typical English answers (*Yes, she does*/*No, she doesn't*) peculiarly vary from those in Mandarin (*No, she does*/*Yes, she doesn't*). What are the processing consequences of these markedly different conventionalized linguistic responses to achieve the same communicative goals? And if English and Mandarin speakers process negative questions differently, to what extent does processing change in Mandarin–English sequential bilinguals? Two experiments addressed these questions. Mandarin–English bilinguals, English and Mandarin monolinguals ($N = 40$/group) were tested in a production experiment (Expt. 1). The task was to formulate answers to positive/negative yes–no questions. The same participants were also tested in a comprehension experiment (Expt. 2), in which they had to answer positive/negative questions with time-measured *yes*/*no* button presses. In both Expt. 1 and Expt. 2, English and Mandarin speakers showed language-specific *yes*/*no* answers to negative questions. Also, in both experiments, English speakers showed a reaction-time advantage over Mandarin speakers in negation conditions. Bilingual's performance was in-between that of the L1 and L2 baseline. These findings are suggestive of language-specific processing of negative questions. They also signal that the ways in which bilinguals process negative questions are susceptible to restructuring driven by the second language.

Keywords: negation processing; yes–no questions; cognitive restructuring in bilinguals; Mandarin; English

Routine answers to negative yes–no questions (e.g., *Doesn't she like cats?*) contrast sharply across languages. In English, the *yes*/*no* part of the answer is typically of the same polarity as the verb in the answer (e.g., *Yes, she does*/*No, she doesn't*). In Mandarin, however, *shi*/*shi de* "yes" and *bu*/*bu shi*/*bu shi de* "no" typically oppose the polarity of the verb (e.g., *No, she does*/*Yes, she doesn't*). The answering system of Mandarin speakers is known as truth-based while the answering system of English speakers is known as polarity-based (Holmberg, 2015). The contrast between the

two systems is argued to arise because languages of the truth-based and the polarity-based systems attach negation in markedly different ways. Truth-based languages like Mandarin structurally attach negation to the statement of the question, *that is*, *Doesn't she like cats?* -> "She **doesn't** like cats." Instead, languages like English more typically attach negation to the polarity of the question, that is, *Doesn't she like cats?* -> "Is it the case or **not** that [she likes cats]." This difference in the attachment of negation in negative questions is linguistically accountable using Holmberg's (2015) distinction between high versus middle negation. It was considered that there are two forms of negative questions in English, one with high negation (e.g., *Doesn't she like cats?*) and the other with middle negation (e.g., *Does she not like cats?*) (Holmberg, 2015; Huddleston & Pullum, 2002). While middle negation is possible in English, it is less typical compared with high negation (we found 3 middle negation tokens and 47 high negation tokens out of a random sample of 50 tokens in the British National Corpus). In contrast, high negation is absent in Mandarin (Holmberg, 2015). For an illustration, a Mandarin question formed via high negation, such as *[Bu [ta [xi huan mao]] ma], that is, [Doesn't [she like cats]], renders the question ungrammatical. Mandarin is structurally limited to middle negation. In high negation, typical in English, the form *n't* is more typically attached to the question [Doesn't [she like cats]] whereas in middle negation, typical in Mandarin, negation is attached to the statement of the question [Ta [bu [xi huan mao]] ma], that is, [Does [she [not [like cats]]]]. As a result, speakers of the truth-based and the polarity-based systems respond to different statements in negative questions. When answering negative questions such as *Doesn't she like cats?*, English speakers typically respond to the positive statement (e.g., *She likes cats*) (Choi, 1991; Holmberg, 2013, 2014, 2015) while Mandarin speakers typically respond to the negative statement (e.g., *She doesn't like cats*) (Holmberg, 2015; Huang and Liao 2007; Lu, 2003).

Given this sharp crosslinguistic contrast between polarity-based and truth-based answering systems, it is surprising that, with a few exceptions, variation in the processing of negation in negative yes–no questions has been little explored. Language-specific attachment of negation also raises the question for bilinguals. Do speakers of both English and Mandarin use both the polarity-based and the truth-based systems depending on the language context? The aim of this study is to investigate the extent to which language influences the processing of negative yes–no questions in English and Mandarin monolingual speakers and Mandarin–English sequential bilinguals.

The differences in the attachment of negation within the truth-based versus the polarity-based systems established the key crosslinguistic contrast. Using these two systems, we next compared the cognitive implications of how differently attached negation is processed by monolingual speakers of a truth-based and a polarity-based language. This comparison serves as a stepping stone in the following section, where findings from studies with bilingual speakers are surveyed to present the state of the art in the field alongside the remaining research gaps that this study aims to fill.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Processing demand variation in the truth-based and the polarity-based systems

Evidence from developmental studies

Using different answering systems can lead to varied cognitive demands. Crosslinguistic evidence from developmental studies suggests that responding to negative questions using the truth-based system is more difficult than using the polarity-based system. Similar to Mandarin speakers, Korean speakers typically use the truth-based system when they answer negative questions (Choi, 1991; Holmberg, 2015) and so do Japanese speakers (Akiyama, 1979, 1992; Holmberg, 2015). Choi (1991) examined early language production in Korean and English children (1;7–3;3) and found that Korean children gave more elaborate answers (e.g., *-Isn't it a bird? -*A *bird.*) instead of *yes*/*no* answers when answering negative questions than English children did. Akiyama (1979, 1992) investigated the answers to negative questions in Japanese and English children (3–6 years old) and observed that Japanese children made significantly more errors in responses to negative questions than English children. Both Choi (1991) and Akiyama (1979, 1992) interpreted the relatively later acquisition of Korean and Japanese systems as evidence that it is more difficult to answer negative questions using the truth-based system than the polarity-based system.

Evidence from research on negation processing

A great number of studies demonstrated that it is more demanding to process negative statements than positive statements (Akiyama et al., 1979; Carpenter & Just, 1975; Clark & Chase, 1972; Dale & Duran, 2011; Dudschig & Kaup, 2018; Fischler et al., 1983; Hasson & Glucksberg, 2006; Kaup et al., 2006; Kaup et al., 2007; Lüdtke et al., 2008; Tian et al., 2010; Tian et al., 2016). For instance, Clark and Chase (1972) examined reaction times (RTs) when participants verified positive/negative English sentences (e.g., *A plus is/isn't above a star*) against picture stimuli. They found that the participants were significantly slower to verify negative than positive sentences, which indicates that it is more difficult to process the negative statement than the positive statement.

The RT result in Clark and Chase (1972) was supported by Fischler et al. (1983) who measured event-related brain potentials (ERPs) of participants when they verified positive/negative English sentences (e.g., *A robin is/is not a tree*). The researchers observed that, around 400 ms after the offset of the stimuli, the brain potentials were significantly larger when a positive statement was false than when it was true. In contrast, the brain potentials were significantly larger when a negative statement was true than when it was false. They argued that the negative statement is not processed immediately like the positive statement, otherwise the N400 effect would have been symmetrical when the participants process positive and negative sentences.

Empirical evidence indicating that it is more difficult to process the negative statement than the positive statement was further supported by more recent mouse-tracking and eye-tracking results. For example, Dale and Duran (2011) tracked participants' computer mouse trajectories when they verified English sentences (e.g., *Elephants are small/not small*). The task was to choose a preferred

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

answer by clicking either true or false. The results showed that there were significantly more abrupt shifts between the response alternates when the participants verified negative sentences compared to positive sentences. The greater number of shifts suggests that there is an increased cognitive demand to process the negative statement.

Similar findings were reported by Tian et al. (2016) using a visual world eye-tracking paradigm. They instructed the participants to listen to positive/negative English sentences (e.g., *Matt has/hasn't shut his dad's window*). While listening to audio stimuli, participants saw pictures representing the positive/negative state of affairs (an open window/a closed window). The researchers observed that it took participants 900 ms longer when hearing negative sentences compared to positive sentences to focus on a matching picture (e.g., an open window for *Matt hasn't shut his dad's window*). Longer RTs when processing negative sentences compared to positive sentences point to an added difficulty when the task is to process the negative statement compared to its positive counterpart. To date, there is no crosslinguistic evidence showing that answering negative questions is more difficult for Mandarin speakers than it is for English speakers, which is what we predict based on available research findings.

Negation processing in bilinguals

Evidence from production experiments

In earlier work, Akiyama (1979) instructed Japanese–English bilingual children and Japanese and English monolingual children (3–6 years old) to answer the same set of negative questions and compared their responses. He found that, when answering negative questions in Japanese, bilinguals' response pattern (e.g., *-Can't you eat a block? -No, I can't*) significantly differed from that of Japanese monolingual children (*-Yes, I can't*). In contrast, when answering English negative questions, bilinguals' response pattern resembled that of English monolingual children (*-No, I can't*). Akiyama interpreted the English-like response pattern of bilinguals as evidence for crosslinguistic influence from English to Japanese.

Crosslinguistic influence between bilinguals' answering systems was also reported in Korean–English bilingual children. Choi (2014) instructed one group of Korean monolingual children (mean age 4;9) and one group of Korean–English bilingual children of the same age to listen to stories and answer corresponding negative questions. The results showed that bilingual children made significantly more errors when they answered negative questions in Korean compared to Korean monolingual children. To illustrate the errors, when the negative question was *Didn't the second pig build his house out of mud?* and the case was that the second pig did not build his house out of mud, bilingual children answered *No, (he didn't)* while *Yes, (he didn't)* was the correct answer following the truth-based (Korean) system. Also, bilingual children made significantly more errors when they answered negative questions in Korean than in English. Comparably to Akiyama (1979), Choi interpreted the production results as evidence for crosslinguistic influence from the English system to the Korean system in Korean–English bilinguals. Considering the findings in Akiyama (1979) and Choi

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

(2014), one can expect Mandarin–English bilinguals in this study to also answer negative questions like English speakers.

Evidence from comprehension experiments

Few studies tested negation processing in bilinguals, with some contradictory findings. Manning et al. (2018) investigated whether L1 and L2 English speakers process negation in the same way or not. The researchers measured ERPs of French[1] learners of English and simultaneous French–English bilinguals when they processed true/false positive/negative sentences (e.g., *The jury found him innocent/guilty because the fire was recognized as intentional/not intentional in court*). In L2 learners, a greater N400 was observed in true-negative sentences (. . . *innocent* . . . *not intentional* . . .) than true-positive sentences (. . . *guilty* . . . *intentional* . . .). However, no greater N400 was found in simultaneous bilinguals. The researchers concluded that there is an additional processing cost for L2 learners compared with native speakers when they process negative sentences.

Unlike Manning et al. (2018), Ćoso and Bogunović (2019) argued that it is not more difficult for bilinguals to process negation in an L2. The researchers instructed Croatian[2] learners of English to verify positive/negative sentences (e.g., *Hearts are not above arrows*) against pictures. They found that Croatian learners showed comparable accuracy rates and response speed when processing negative sentences in L2 English and L1 Croatian. The researchers concluded that Croatian learners processed negation in English and Croatian in a similar manner.

The research gaps this article aims to fill are to what extent language can influence the processing of negation in negative questions in English and Mandarin monolingual speakers and Mandarin–English bilinguals. If English and Mandarin monolingual speakers and Mandarin–English bilinguals do not process negation in the same way, different attachments of negation are expected to manifest themselves as variation in response type and processing speed. Building on previous research on negation processing (e.g., Akiyama, 1979; Choi, 2014; Ćoso & Bogunović, 2019; Manning et al., 2018), we designed a production experiment to measure preferred response types and speech onset times (Expt. 1) and a comprehension experiment to measure response times (Expt. 2). In the production experiment, Mandarin–English bilinguals and English and Mandarin monolinguals were instructed to formulate yes/no responses to negative questions (critical trials) and positive questions (control trials). In the comprehension experiment, the same participants answered positive/negative questions with time-measured *yes*/*no* button presses.

Experiment 1. Formulating yes/no responses

To test the extent to which Mandarin and English monolingual speakers and Mandarin–English bilinguals differ in their answering systems, in Experiment 1 we asked participants to verbalize answers to negative and positive questions and we examined the proportion of their *yes*/*no* answers and their speech onset times.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Participants

Forty English (31 females) and 40 Mandarin (38 females) monolingual speakers and 40 Mandarin learners of English (39 females) took part in this experiment. The English participants (MAGE = 19.4, max. 23, min. 18) were recruited from a university in the UK and the Mandarin participants were recruited from a college in China (MAGE = 20.5 years, max. 22, min. 19). The Mandarin participants were functionally monolingual, with some knowledge of basic English limited to a few phrases. All monolingual participants were right-handed and reported no fluency in any language other than their L1. The bilingual participants were recruited from a university in the UK. Following Athanasopoulos et al. (2011), Park and Ziegler (2014), and Vanek and Selinker (2017), the bilingual participants were asked before the experiments to self-assess their language background through a questionnaire (e.g., AOA, frequency of L2 speaking and writing). Their average age was 22.45 years (max. 26, min. 21). They started to learn English at the age of 9.13 ($SD = 1.98$) on average. Their average score of the Oxford Placement Test (max. 100) was 72.60 ($SD = 7.50$). They had been living in the UK for 2.07 months ($SD = 1.12$) on average at the time of testing. Participants were asked to self-assess the percentage of their speaking/writing in English on a typical day. Their average English speaking was 38.00% ($SD = 15.72$) of the time per day and average English writing was 60.63% ($SD = 27.11$) of the time per day. All bilingual participants were right-handed and reported no fluency in any language other than Mandarin and English.

Materials

The instruments for this task consisted of 24 informational declarative statements (henceforth referred to as "statements"), 48 yes–no questions and 24 distractor filler questions in English and Mandarin.

Statements

The stimulus preparation started with forming 24 statements both in English and in Mandarin (see examples 1 and 2). Twelve out of 24 statements – critical statements – were used to form yes–no questions (see Table 1), and the other 12 statements – filler statements – were used to form distractor filler questions (see examples 3 and 4). All statements were positive, transitive, and declarative. The syntactic structure of all the statements within a language was kept the same (i.e., subject–verb–object–adverb). All English statements were checked for authenticity (i.e., *Do these statements sound natural or not?*) by an English native speaker to ensure that potentially ambiguous or infelicitous statements were excluded from the stimuli. Two Mandarin–English bilinguals then checked all the Mandarin translations from English for authenticity, and only translations for which consensus was reached were included in the experiment.

(1) Example of an English statement:

Mr. Fox stole a roast duck from a farm.
[$_{TP}$ Mr. Fox [$_{VP}$ stole [$_{NP}$ a roast duck] from a farm.]]

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Table 1. Examples of yes–no questions in English and Mandarin for the statement *Mr. Fox stole a roast duck from a farm*

Question type	Condition	Language	Example									
Positive	Positive-same	English	*Did Mr. Fox steal a roast duck from a farm?*									
		Mandarin	*Hu li*	*xian sheng*	*cong*	*nong chang*	*tou le*	*yi zhi*	*kao*	*ya*	*ma?*	
			fox	sir	from	farm	steal	a	roast	duck	Q	
			"Did Mr. Fox steal a roast duck from a farm?"									
Positive	Positive-different	English	*Did Mr. Fox steal a roast chicken from a farm?*									
		Mandarin	*Hu li*	*xian sheng*	*cong*	*nong chang*	*tou le*	*yi zhi*	*kao*	*ji*	*ma?*	
			fox	sir	from	farm	steal	a	roast	chicken	Q	
			"Did Mr. Fox steal a roast chicken from a farm?"									
Negative	Negative-same	English	*Didn't Mr. Fox steal a roast duck from a farm?*									
		Mandarin	*Hu li*	*xian sheng*	*mei*	*cong*	*nong chang*	*tou le*	*yi zhi*	*kao*	*ya*	*ma?*
			fox	sir	Neg	from	farm	steal	a	roast	duck	Q
			"Did Mr. Fox not steal a roast duck from a farm?"									
Negative	Negative-different	English	*Didn't Mr. Fox steal a roast chicken from a farm?*									
		Mandarin	*Hu li*	*xian sheng*	*mei*	*cong*	*nong chang*	*tou le*	*yi zhi*	*kao*	*ji*	*ma?*
			fox	sir	Neg	from	farm	steal	*a*	roast	chicken	Q
			"Did Mr. Fox not steal a roast chicken from a farm?"									

(2) Example of a Mandarin statement:

Hu li	*xian sheng*	*cong*	*nong chang*	*tou le*	*yi zhi*	*kao*	*ya.*
fox	sir	from	farm	steal	a	roast	duck

"Mr. Fox stole a roast duck from a farm."

[$_{TP}$ Hu li xian sheng [$_{VP}$ cong nong chang tou le [$_{NP}$ yi zhi kao ya]]]

Yes–no questions

English yes–no questions were formed from the critical statements by subject–verb inversion, and Mandarin yes–no questions were formed by adding the question marker *ma* at the end of each critical statement. Forty-eight yes–no questions were formed from 12 critical statements. Each question belonged to one of four conditions (see Table 1). The four conditions were classified based on two factors: (a) the polarity of the question (positive/negative) and (b) the sameness of the direct objects/patients in the given statement and the following question (same/different). For English negative questions, we used high negation in the current study, and the motivation for this choice was that it was found to be the most typical form in English. The four conditions were labeled as positive-same, positive-different, negative-same, or negative-different. The 48 yes–no questions were grouped into four lists (A, B, C, and D). For counterbalancing purposes, within each list the yes–no questions came from different statements. Each list (see Appendix for the full set of lists) consisted of 12 yes–no questions, that is, 3 yes–no questions from each of the 4 conditions (3 positive-same + 3 positive-different + 3 negative-same + 3 negative-different). Each participant was randomly assigned to one list. Another level of randomization was that each participant saw the statement-question pairs in a random order.

The question setup (Table 1) conformed to the scope of negation in Mandarin negative yes–no questions with two features in mind. First, the questions were formed considering that "in Chinese, negation normally starts from the word immediately following the negator and stretches to the end of the clause" (Xiao & McEnery, 2008: 299). This is an important feature in the current design because it ensures that negation is not focused on, or limited to, the adjunct (e.g. *from a farm*) but it negates the whole verb phrase (e.g. *stole a roast duck from a farm*). Second, the negator *mei* "not" in Mandarin has aspectual requirements in a sense that it negates bounded events (Ernst, 1995). What these characteristics help to clarify is that the design does not include double-barrelled questions (about stealing A or B vs. whether A or B was stolen from C), where participants would be responding to two propositions, one with negation and one without negation.

Distractor filler questions

Twenty-four distractor filler questions (i.e., other than yes–no questions, see examples 3 and 4) were included to mask the purpose of the experiment and to avoid reactivity in participants. These questions were not part of the analyses. Each filler statement was transformed into two distractor filler questions. As a result, each participant saw 12 critical yes–no questions and 24 distractor filler questions in total.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

The pairs of filler statements and filler questions were randomly distributed among the critical pairs of statements and yes–no questions for each participant.

(1) Example of an English filler statement and a distractor filler question:

Mr. Dog received a letter from his grandpa.
Who received a letter from his grandpa? (Mr. Dog /Mr. Fox)

(2) Example of a Mandarin filler statement and a distractor filler question:

Gou	*xian sheng*	*shou dao le*	*ta*	*ye ye*	*ji lai de*	*xin.*
dog	sir	receive	his	grandpa	send	letter

"Mr. Dog received a letter from his grandpa."

Shei	*shou dao le*	*ta*	*ye ye*	*ji lai de*	*xin?*
who	receive	his	grandpa	send	letter
(gou	*xian sheng/*	*hu li*	*xian sheng)*		
(dog	sir	fox	sir)		

"Who received a letter from his grandpa? (Mr. Dog /Mr. Fox)"

Procedure

Participants were tested individually. They were asked to carefully read the instructions on a computer screen. They were informed that they would see one statement at a time. After each statement, they saw and heard one question. Their task was to read each statement carefully and answer each question aloud (including *yes*/*no* whenever possible), accurately and as quickly as possible. Participants received a brief training (four items). Only after they confirmed that they had understood the task and the procedure, did the experimenter start the computerized test.

During the training and the experimental session (set up in E-Prime 2.0), participants first read one statement displayed on the screen for 8 s. The 8-s interval was kept constant across statements in order to ensure that each participant had the same maximum reading time. Then, a corresponding question followed automatically (Figure 1). Participants were not able to go back to the statement once the question appeared. Their answers were audio-recorded. When a participant did not specifically give a *yes*/*no* answer in their response, the experimenter would remind them about the task (i.e., *Include yes/no in your answer whenever possible*) before moving to the following trial. The bilingual participants were tested in English (i.e., in their L2), the English participants in English, and the Mandarin participants in Mandarin Chinese.

For the analyses, responding with a positive answer (i.e., *yes* in English and *shi (de)* "yes" in Mandarin) was coded 1; a negative answer (i.e., *no* in English and *bu (shi de)* "no" in Mandarin) was coded 0. Answers in each condition were analyzed separately and the mean score for each condition was used to show the proportions of *yes* answers.

Results

Yes/no responses

Figure 2 shows *yes/no* answers to positive/negative questions verbalized by English and Mandarin monolingual speakers and Mandarin–English bilinguals. Our

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

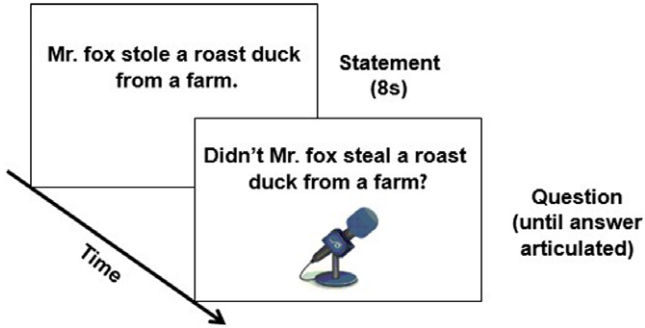

Figure 1. Protocol of the Production Experiment (Expt. 1).

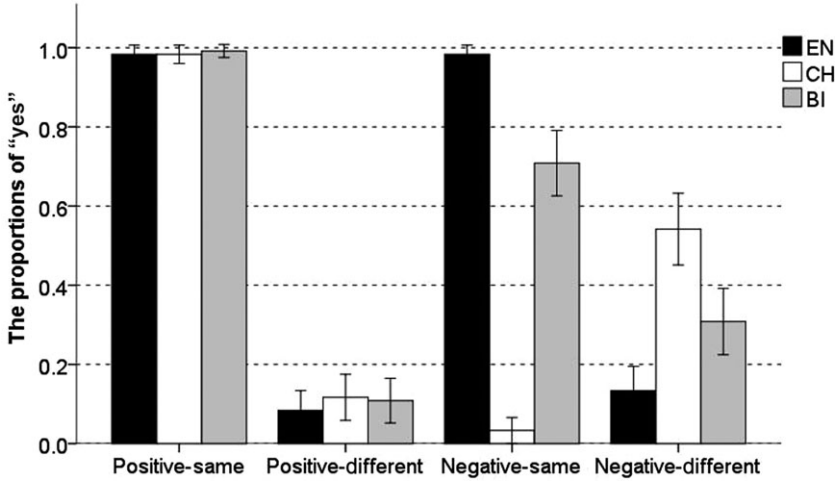

Figure 2. Mean Proportions of *yes* Answers by English and Mandarin Monolingual Speakers and Mandarin–English Bilinguals in the Production Experiment (Expt. 1) (Error Bars = 95% Confidence Interval).

examination started with answers to (critical) negative questions, followed by answers to (control) positive questions. In the critical negative-same condition, English speakers predominantly answered *yes* ($M = 0.98$, $SD = 0.13$) while Mandarin speakers predominantly answered *bu (shi de)* "no" (the proportion of *shi (de)* "yes" $M = 0.03$, $SD = 0.18$). The predominant answer given by bilingual participants was *yes* ($M = 0.71$, $SD = 0.46$) like English speakers. In the critical negative-different condition, English speakers heavily preferred *no* (the proportion of *yes* $M = 0.13$, $SD = 0.34$). Mandarin speakers, however, showed no clear preference for either *shi (de)* "yes" ($M = 0.54$, $SD = 0.50$) or *bu* (*shi de)* "no." In bilingual participants, *no* was preferred over *yes* ($M = 0.31$, $SD = 0.46$) like in English speakers. In the control positive-same condition, English ($M = 0.98$, $SD = 0.13$) and Mandarin speakers ($M = 0.98$, $SD = 0.13$) and bilinguals ($M = 0.99$, $SD = 0.09$)

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

almost exclusively answered *yes*. In the control positive-different condition, English ($M = 0.08$, $SD = 0.28$) and Mandarin speakers ($M = 0.12$, $SD = 0.32$) and bilinguals ($M = 0.11$, $SD = 0.31$) rarely answered *yes* and predominately answered *no*.

To test the effect of language group on the proportion of *yes*/*no* answers in each condition, we built mixed-effects regression models using the lme4 package (Baayen et al., 2008) in the R software (Version 3.5.1; R Development Core Team, 2018). The fixed effect factors were *Group* (English/Chinese/Bilingual), *Question* (positive/negative), and *Sameness* (same/different), the binary dependent variable was *Answer* (yes/no), and the random effect factors were *Participant* and *Item*. The model included all possible random effects (Barr et al., 2013), with random slopes over question by participant and random slopes over question, sameness, and their interaction by item as follows:

```
answer ~ 1 + question * sameness * group +
         (1 + question | participant) +
         (1 + question * sameness | item)
```

The results are shown in Table 2. The full dataset with responses for each participant in each condition, as well as the materials used, is made available to the research community on the project website https://osf.io/x4536/. These steps were taken to promote reproducible and replicable practices in bilingualism research more generally (Bolibaugh et al., 2021), following examples from language-modulated event cognition research with Mandarin–English bilinguals in particular (Tang et al., 2021; Vanek, 2020).

The full model returned two significant three-way interactions, namely between *Group*, *Question*, and *Sameness* for bilinguals versus Mandarin speakers, and also between *Group*, *Question*, and *Sameness* for bilinguals versus English speakers. To further explore the nature of the interactions, we next built a reduced model excluding *Question*. A comparison of the reduced model with the full model showed that the presence of *Question* significantly increased the model fit, $\chi^2(6) = 60.42$, $p < .001$, confirming that participants answered positive and negative questions differently. Then, we proceeded with a forward variable selection and zoomed in on the negative questions, comparing a model including *Group* with a reduced model without *Group* in the data for negative questions only. This comparison confirmed significant contribution of *Group* to the variation in responses to negative questions, $\chi^2(4) = 136.75$, $p < .001$. And as the final statistical step, we ran Tukey-adjusted pairwise comparisons between groups (Table 3) to examine how group responses to specific types of negative questions differ. In the negative-different condition, we found a significant difference between the Chinese and English monolinguals (estimate = 2.35, $SE = 0.40$, Z ratio = 5.84, $p < .001$) and between the bilinguals and Mandarin monolinguals (estimate = −1.16, $SE = 0.35$, Z ratio = −3.32, $p = .003$) as well as between the bilinguals and the English monolinguals (estimate = 1.19, $SE = 0.41$, Z ratio = 2.89, $p = .011$). In the negative-same condition, there was a significant difference between the Mandarin and English monolinguals (estimate = −10.04, $SE = 1.32$, Z ratio = −7.58, $p < .001$) and between the bilinguals and Mandarin monolinguals (estimate = 6.18, $SE = 0.88$, Z ratio = 7.04, $p < .001$) as well as between the bilinguals and the English monolinguals (estimate = −3.86, $SE = 1.14$, Z ratio = −3.38, $p = 0.002$). These results indicate that bilinguals' responses to both types of negative questions were in-between those of English and Mandarin controls.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Table 2. Coefficients from a mixed-effects model fitted to the answers of English-Mandarin bilinguals and English and Mandarin monolingual speakers in the production experiment (Expt. 1)

Fixed effects	Estimate	*SE*	*Z* value	*p*
Intercept	−0.81	0.20	−4.00	< .001**
Question (pos)	−1.80	0.48	−3.76	< .001**
Sameness (same)	1.73	0.31	5.61	< .001**
Group (CH)	0.98	0.27	3.64	< .001**
Group (EN)	−1.07	0.33	−3.22	.001*
Question (pos) × Sameness (same)	6.10	1.29	4.73	< .001**
Question (pos) × Group (CH)	−0.92	0.53	−1.72	.086
Question (pos) × Group (EN)	0.71	0.60	1.19	.232
Sameness (same) × Group (CH)	−5.33	0.61	−8.70	< .001**
Sameness (same) × Group (EN)	4.30	0.83	5.17	< .001**
Question (pos) × Sameness (same) × Group (CH)	4.58	1.56	2.93	.003*
Question (pos) × Sameness (same) × Group (EN)	−4.63	1.67	−2.77	.006*
Random effects	**Variance**			***SD***
Participants (intercept)	0.00			0.00
Question (pos)	0.37			0.61
Item (intercept)	0.03			0.17
Question (pos)	0.86			0.93
Sameness (same)	0.18			0.42
Question (pos) × Sameness (same)	2.62			1.62

Note. *$p < .05$; ** $p < .001$.

Table 3. Tukey-adjusted pairwise comparisons of answers to the negative questions between groups in the production experiment (Expt. 1)

Condition	Group	Estimate	*SE*	*Z* ratio	*p*
Negative-different	English versus Mandarin	2.35	0.40	5.84	< .001**
	Bilingual versus Mandarin	−1.16	0.35	−3.32	.003*
	Bilingual versus English	1.19	0.41	2.89	.011*
Negative-same	English versus Mandarin	−10.04	1.32	−7.58	< .001**
	Bilingual versus Mandarin	6.18	0.88	7.04	< .001**
	Bilingual versus English	−3.86	1.14	−3.38	.002*

Note. *$p < .05$; **$p < .001$.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Speech onset times

We next consider the speech onset times (i.e., RT intervals between the onset of the stimulus and the onset of articulation) from Expt. 1. In order to compare the cognitive demands of the answering systems in each group, RTs were only included in the analyses if an English answer was *yes* for the positive-same and the negative-same conditions in English monolinguals and bilinguals. Analogously, RTs were only included in the analyses if an English answer was *no* for the positive-different and the negative-different conditions. For the bilinguals, the mean proportion of target-like responses was 0.71 in the critical negative-same condition and 0.69 in the critical negative-different condition. For the Mandarin monolinguals, RTs were included in the analyses if the answer was *shi (de)* "yes" for the positive-same and *bu (shi de)* "no" for the positive-different and the negative-same conditions (the mean proportions of the target-like responses was 0.97). This step was taken to ensure crosslinguistic comparability of the RTs of expected answers. Notably, for the negative-different condition, Mandarin answers do not strictly follow the truth-based system, that is, they are not clearly geared toward *shi (de)* "yes", but are distributed similarly as either *shi (de)* "yes" or *bu (shi de)* "no." The analysis of the RTs first considers the *shi (de)* "yes" answers for the negative-different condition because *shi (de)* "yes" in Mandarin speakers is the expected answer if the Mandarin follow the truth-based system. Then we also include the RTs of the nearly 50% "incorrect" answers in the analysis (i.e., we take native Mandarin speakers' responses as the baseline).

There were a few outliers in each group. Following Keating and Jegerski (2015) and Norris (2015), for the English participants, 1 data entries (0.2% of total RTs of expected English answers) were more than 2.5 standard deviations away from the group mean in each condition. These outlier RTs were replaced by the cut-offs (group mean +/− 2.5 *SDs*). For the Mandarin participants, 8 data entries (2.0% of total RTs of expected Mandarin answers) were more than 2.5 standard deviations away from the group mean in each condition. These were also replaced by the cut-offs. For the bilingual participants, there were 19 outliers (4.8% of total RTs of expected answers), and they were replaced by the cut-offs. RTs of the expected answers in each language group for each condition are illustrated in Figure 3.

To test the effect of language group on response speed when participants process positive and negative questions, we built a series of mixed-effects models. We specified Group (English/Chinese/Bilingual), Question (positive/negative), and Sameness (same/different) as fixed effect factors, and Participant and Item as random effect factors. Interactions between the fixed effects factors were also tested. The model included a maximal random effects structure, including random slopes over group by item, and random slopes over question, sameness and their interaction by item: RT ~ 1 + question × sameness × group + (1 + question × sameness | participant) + (1 + question | item). The results are shown in Table 4 (the full dataset with RTs per participant/condition is available on the project website https://osf.io/x4536/).

First, to statistically test whether participants' speech onset times significantly differed in answers to negative and positive questions, we compared a full model including *Question* with a reduced model excluding *Question*. This comparison showed that the model fit was significantly improved with the presence of *Question*, $\chi^2(6) = 52.86$, $p < .001$, confirming that speech onset times to negative questions were significantly slower than to positive questions. Second, we zoomed in on the negative questions and compared a model including *Group* with a reduced model without *Group* in the data for

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

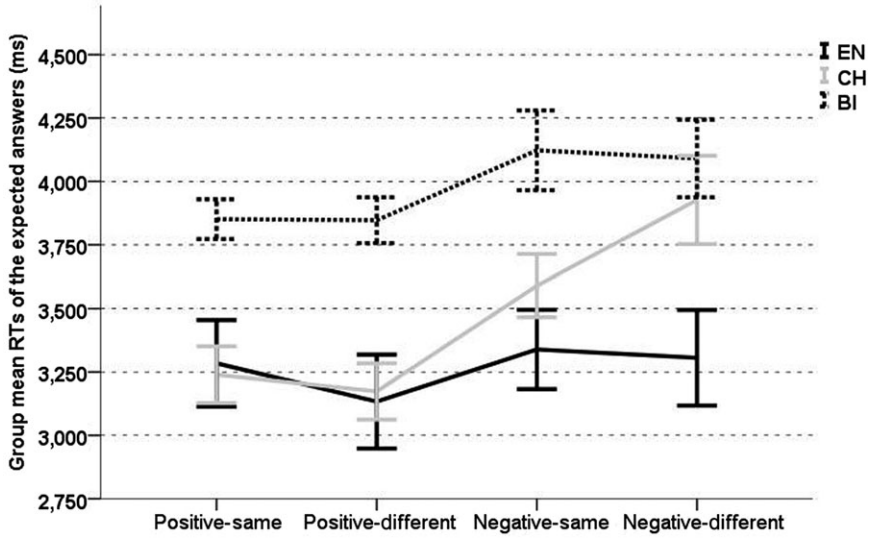

Figure 3. Mean RTs of English and Mandarin Monolingual Speakers and Mandarin–English Bilinguals in the Production Experiment (Expt. 1) (Error Bars = 95% Confidence Interval).

negative questions only. This comparison confirmed a significant between-group difference in speech onset times for responses to negative questions, $\chi^2(4) = 39.38$, $p < .001$. Third, we looked at the nature of the three-way interaction found in the full model (Table 4) more closely by running a Tukey-adjusted pairwise comparisons (Table 5) to examine how speech onset times of responses to specific types of negative questions differ between groups. In the negative-different condition, we found a significant difference between Mandarin and English monolinguals (estimate = 615.94, $SE = 190.64$, t ratio = 3.23, $p = .006$) as well as between bilinguals and English monolinguals (estimate = 847.15, $SE = 158.62$, t ratio = 5.34, $p < .001$), but not between bilinguals and Mandarin monolinguals (estimate = 231.21, $SE = 203.19$, t ratio = 1.14, $p = .495$). In the negative-same condition, there was a significant difference between bilinguals and Mandarin monolinguals (estimate = 536.11, $SE = 179.64$, t ratio = 2.98, $p = .014$) and between bilinguals and English monolinguals (estimate = 793.26, $SE = 135.88$, t ratio = 5.84, $p < .001$) but not between the English and the Mandarin monolinguals (estimate = 257.15, $SE = 164.99$, t ratio = 1.56, $p = .275$).

These results overall show that it took Mandarin speakers longer than English speakers to articulate answers to negative compared to positive questions. The interpretation here is that of relatively greater processing difficulties in answer formulation using a truth-based compared to a polarity-based system. Considering bilinguals' speech onset times, as shown in Figure 3, the differences by condition within the bilingual group exhibit characteristics that are in-between the patterns found for Mandarin monolinguals (i.e., it took bilinguals longer to respond to negative than to positive questions) and English monolinguals (i.e., bilinguals' speech onset times were similar in the two negative conditions).

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Table 4. Coefficients from a mixed-effects model fitted to the speech onset times of English-Mandarin bilinguals and English and Mandarin monolingual speakers in the production experiment (Expt. 1)

Fixed effects	Estimate	*SE*	*t* value	*p*
Intercept	4,128.04	121.98	33.84	< .001**
Question (pos)	−278.35	89.50	−3.11	< .001**
Sameness (same)	−26.74	88.40	−0.30	.763
Group (CH)	−210.15	158.46	−1.33	.187
Group (EN)	−840.51	152.35	−5.52	< .001**
Question (pos) × Sameness (same)	28.48	109.77	0.26	.796
Question (pos) × Group (CH)	−467.95	122.35	−3.82	< .001**
Question (pos) × Group (EN)	146.03	113.63	1.29	.202
Sameness (same) × Group (CH)	−301.74	125.63	−2.40	.018*
Sameness (same) × Group (EN)	75.83	117.86	0.64	.522
Question (pos) × Sameness (same) × Group (CH)	363.76	156.05	2.33	.021*
Question (pos) × Sameness (same) × Group (EN)	47.70	149.38	0.32	.750
Random effects	**Variance**			***SD***
Participants (intercept)	351,608			592.97
Question (pos)	57,058			238.87
Sameness (same)	67,870			260.52
Question (pos) × Sameness (same)	66,160			257.22
Item (intercept)	33,747			183.70
Question (pos)	12,798			113.13
Residual	238,709			488.58

Note. *$p < .05$; **$p < .001$.

Table 5. Tukey-adjusted pairwise comparisons of speech onset times in the negative conditions between groups in the production experiment (Expt. 1)

Condition	Group	Estimate	*SE*	*t* ratio	*p*
Negative-different	English versus Mandarin	615.94	190.64	3.23	.006*
	Bilingual versus Mandarin	231.21	203.19	1.14	.495
	Bilingual versus English	847.15	158.62	5.34	< .001**
Negative-same	English versus Mandarin	257.15	164.99	1.56	.275
	Bilingual versus Mandarin	536.11	179.64	2.98	.014*
	Bilingual versus English	793.26	135.88	5.84	< .001**

Note. *$p < .05$; **$p < .001$.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

As a matter of fact, close to 50% of the RTs were excluded from the analysis (Table 4) in the critical negative-different conditions for the monolingual Mandarin speakers as "unexpected responses," which may make the "expected responses" of Mandarin natives seem rather prescriptive. To avoid prescriptivism, we also ran the same analysis with the "unexpected/incorrect" responses of Mandarin natives included (Table 6), and we found that the results with unexpected responses differed little from the results with expected responses only (Table 4).

There are a few noteworthy observations (Table 7) in addition to the analyses directly responding to the research questions. Mandarin monolingual speakers showed a clear pattern of the kind observed in earlier verification tasks (e.g., Clark & Chase, 1972; Carpenter & Just, 1975), where the equivalent of true-negatives ($M = 3{,}927$, $SD = 704$) takes longer than false-negatives ($M = 3{,}590$, $SD = 676$) (Figure 3). This pattern also holds for the "incorrect" answers in the bilingual group, that is, with longer RTs in the negative-different condition ($M = 5{,}142$, $SD = 1{,}657$) than in the negative-same condition ($M = 4{,}513$, $SD = 1{,}332$). Looking at RTs of the "incorrect" responses of the Mandarin monolinguals, the gap between the negative-different condition ($M = 3{,}712$, $SD = 1{,}296$) and the negative-same condition ($M = 3{,}590$, $SD = 676$) narrows, resembling the pattern of the English speakers (negative-different condition: $M = 3{,}306$, $SD = 970$; negative-same condition: $M = 3{,}339$, $SD = 858$) and that of the bilinguals' target responses (negative-different condition: $M = 4{,}091$, $SD = 702$; negative-same condition: $M = 4{,}123$, $SD = 730$).

Discussion

Tracking the answers to negative questions confirms the robustness of the crosslinguistic contrast assumed. Mandarin speakers typically used a truth-based system while English speakers used a polarity-based system to respond to negative questions. Mandarin speakers predominantly answered *bu (shi de)* "no" in the negative-same condition (e.g., the question was *Didn't he steal a duck?* and the given statement was *He stole a duck*) with the proportion of *no* significantly higher than that of *no* by English speakers. The interpretation here is that Mandarin speakers attach negation to the statement of the question, that is, *Didn't he steal a duck?* -> "He didn't steal a duck." English speakers almost exclusively answered *yes* in the negative-same condition (see the example mentioned before). The interpretation here is that English speakers typically attach negation to the polarity of the question, that is, *Didn't he steal a duck?* -> "Is it the case or not that [he stole a duck]." These interpretations were also corroborated by the results in the negative-different condition (e.g., the question was *Didn't he steal a chicken?* and the given statement was *He stole a duck*), where the proportion of *shi (de)* "yes" given by Mandarin speakers was significantly higher than that of *yes* responded by English speakers. These preferred answers suggest that English speakers are highly likely to respond to the positive statement of a negative question (Choi, 1991; Holmberg, 2013, 2014, 2015). In contrast, answers typical of Mandarin speakers suggest that Mandarin speakers respond to the negative statement of a negative question (Holmberg, 2015; Huang and Liao 2007; Lu, 2003).

Extending evidence for crosslinguistic differences in response type, the crosslinguistic contrasts we observed in speech onset times provide another piece of evidence that the processing of negation in negative questions by English and

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Table 6. Coefficients from a mixed-effects model fitted to the RTs of English–Mandarin bilinguals and English and Mandarin monolingual speakers in the production experiment (Expt. 1) with "incorrect" answers from the Mandarin monolingual speakers in the negative-different condition included

Fixed effects	Estimate	*SE*	*t* value	*p*
Intercept	4,126.34	120.19	34.33	< .001**
Question (pos)	−276.33	88.31	−3.13	.002*
Sameness (same)	−19.36	83.89	−0.23	.818
Group (CH)	−249.79	148.55	−1.68	.095
Group (EN)	−838.07	149.76	−5.60	< .001**
Question (pos) × Sameness (same)	20.72	106.30	0.20	.846
Question (pos) × Group (CH)	−427.20	110.07	−3.88	< .001**
Question (pos) × Group (EN)	143.34	111.22	1.29	.199
Sameness (same) × Group (CH)	−268.96	110.01	−2.44	.016*
Sameness (same) × Group (EN)	67.86	111.46	0.61	.544
Question (pos) × Sameness (same) × Group (CH)	330.49	143.80	2.30	.022*
Question (pos) × Sameness (same) × Group (EN)	55.87	144.57	0.39	.700
Random effects	**Variance**			***SD***
Participants (intercept)	335,038			578.82
Question (pos)	43,448			208.44
Sameness (same)	37,002			192.36
Question (pos) × Sameness (same)	32,914			181.42
Item (intercept)	33,409			182.78
Question (pos)	13,662			116.88
Residual	243,702			493.66

Note. *$p < .05$; **$p < .001$.

Table 7. Response patterns and the corresponding RT differences between the negative-different condition and the negative-same condition in the production experiment (Expt. 1)

	Negative-different		Negative-same		
Response	*M*	*SD*	*M*	*SD*	RT difference
Mandarin correct responses	3,927	704	3,590	676	337
Bilingual incorrect responses	5,142	1,657	4,513	1,332	629
Mandarin incorrect responses	3,712	1,296	3,590	676	122
Bilingual correct responses	4,091	702	4,123	730	32
English responses	3,306	970	3,339	858	33

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Mandarin speakers is language-specific. It took Mandarin speakers significantly longer than English speakers to process negative questions compared to positive questions. A large body of empirical evidence suggests that to process a negative statement is more difficult than to process a positive statement (e.g., Akiyama et al., 1979; Carpenter & Just, 1975; Clark & Chase, 1972; Dale & Duran, 2011; Fischler et al., 1983; Kaup et al., 2006, 2007; Lüdtke et al., 2008; Tian et al., 2010, 2016). Building on this evidence, the relatively longer RT in Mandarin speakers, compared with English speakers, can be explained as a result of attaching negation to the statement of a question and thus Mandarin speakers respond to the negative statement. In contrast, English speakers are more likely to attach negation to the polarity of the question and as a result they respond to the positive statement.

Bilinguals' results indicate a shift from using Mandarin-like responses toward English-like responses when they answered negative questions in English. In response to negative questions, Mandarin–English bilinguals approximated to a polarity-based pattern, which is in line with the findings in Japanese–English and Korean–English bilinguals (Akiyama, 1979; Choi, 2014). To illustrate this, in the negative-same condition, bilinguals predominantly answered *yes* to *Didn't he steal a duck?* (the given statement was *He stole a duck*). This pattern resembled that of English speakers who predominantly answered *yes* to the same combination of statement and yes–no question, but it diverged from Mandarin speakers who typically answered *bu (shi de)* "no." Bilinguals' performance suggests shifts toward the English-like attachment of negation, characterized by attaching negation to the polarity of the question rather than to the statement of the question.

Bilinguals, resembling English speakers, showed shorter speech onset times than Mandarin speakers when they answered negative questions compared to positive questions. The advantage in bilinguals' L2 RT compared to the Chinese controls' RT can be interpreted as an L2-driven change in the processing routines from a truth-based toward a cognitively less demanding polarity-based system. If bilinguals had processed negation following the truth-based system, then one would not expect that in their L2 they would outperform Mandarin monolingual speakers in the processing of negative questions. Following the L2-based system, the results suggest that bilinguals more typically attached negation to the polarity of the question and processed the positive statement. Nonetheless, bilinguals also exhibited overall long speech onset times compared to the English monolinguals, indicating difficulty in processing and formulating responses in the L2.

An unexpected result was the interchangeable use of *yes*/*no* by Mandarin speakers found in the negative-different condition, where *yes* was predicted to be their predominant answer. For an illustration, Mandarin speakers showed a similar preference for *shi (de)* "yes" and *bu (shi de)* "no" when they answered *Didn't he steal a chicken?* (the given statement was *He stole a duck*). Clark and Chase (1972) and Fischler et al. (1983) observed prolonged RTs when participants verified true-negative sentences (e.g., *-He didn't steal a chicken. -True, he didn't.*) and argued for a greater cognitive demand in this condition (see Tian et al. (2016) for a detailed discussion). Given the comparability between sentence verification and answering yes–no questions following the truth-based system, a plausible explanation here is that responding to the negative statement of the question (i.e., using the truth-based system) in the negative-different condition is difficult, and thus processing in

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Mandarin speakers can break down. In this condition, there are two mismatches between the negative statement and the given sentence, namely the polarities (negative vs. positive) and objects (chicken vs. duck). In contrast, English speakers who respond to the positive statement of a negative question (i.e., following the polarity-based system) may not experience such complexity as there is only one mismatch between the positive statement and the given sentence, the objects. The longer speech onset times observed in Mandarin speakers than English speakers in the negative-different condition further support this idea.

Experiment 2. Comprehending yes–no questions

This experiment was designed to further test the extent to which English and Mandarin monolingual speakers and Mandarin–English bilinguals differ, namely in cognitive demands, when processing negative questions. Expt. 1 helped us to establish the robustness of an important crosslinguistic difference – and its implications for bilingual speakers – as it surfaces in production. Language production involves articulation and its monitoring (Levelt, 1989), which are processes adding to the complexity of responding to negative questions. With the rationale to control for possible effects of such added complexity, we designed Expt. 2 focusing on how participants comprehend negative questions in a task where yes–no answers had been pre-formulated for them and their task was to quickly and accurately choose one. This design characteristic is advantageous for RT measures because, unlike in production, it helps us to discount potential effects of differences in articulation speed.

Participants

The same participants as in Expt. 1 were tested in Expt. 2 immediately after Expt. 1.

Materials

Materials of Expt. 2 were identical as those in Expt. 1, that is, 24 statements, 48 yes–no questions as well as 24 distractor filler questions.

Procedure

Participants were tested individually. They were asked to carefully read the instructions on the computer screen. They were informed that they would see one statement first. After each statement, one question and two answer choices (*yes*/*no*) appeared on the computer screen. Their task was to read each statement and the subsequent question carefully and choose their preferred answer as quickly as possible (Figure 4). They were asked to press the "↑" key on the keyboard to choose yes or the "↓" key to choose no. For distractor filler questions, participants chose from the *Mr. Dog*/*Mr. Fox* type alternatives by pressing the "↑" key or the "↓" key corresponding to the position of answer choices (*Mr. Dog*/*Mr. Fox*) displayed on the screen. Participants received a brief training (four items). Only after the participants confirmed that they had understood the task and procedures, did the experimenter start the computerized test.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

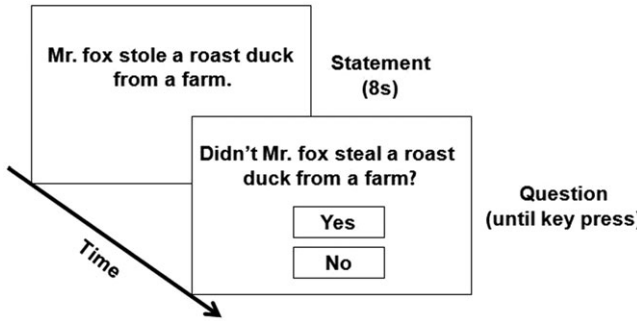

Figure 4. Protocol of the Comprehension Experiment (Expt. 2).

During the training and the experimental session (set up in E-Prime 2.0), participants first read one statement displayed on the screen for 8 s (Figure 3). The 8-s interval was kept constant across statements in order to ensure that each participant had the same maximum reading time. Then, a corresponding question followed automatically. Participants were not able to go back to the statement once the question appeared. When a participant answered a question, that is, once they pressed "↑" or "↓", the computerized task automatically continued with the next trial. Each answer and RT were recorded. The bilingual participants were tested in English, the English participants in English, and the Mandarin participants in Mandarin Chinese.

For the analyses, selecting a positive answer (i.e., an up-pointing arrow "↑" press) was coded 1; a negative answer (i.e., a down-pointing arrow "↓" press) was coded 0. Answers in each condition were analyzed separately and the mean score for each condition was used to show the proportion of *yes* answers.

Results

Responses

We first report the answers in English and Mandarin monolingual speakers and Mandarin–English bilinguals (Figure 5), starting with answers to (critical) negative questions, followed by answers to (control) positive questions. In the critical negative-same condition, English speakers predominantly answered *yes* ($M = 0.98$, $SD = 0.16$) while Mandarin speakers predominantly answered *bu (shi de)* "no" (the proportion of *shi (de)* "yes" $M = 0.07$, $SD = 0.25$). Bilinguals, like English speakers, heavily preferred *yes* ($M = 0.69$, $SD = 0.46$). In the critical negative-different condition, English speakers predominantly answered *no* (the proportion of *yes* $M = 0.07$, $SD = 0.25$) while Mandarin speakers showed no clear preference for either *shi (de)* "yes" ($M = 0.54$, $SD = 0.50$) or *bu* (*shi de)* "no." Bilinguals, like English speakers, heavily preferred *no* (the proportion of *yes* $M = 0.32$, $SD = 0.47$). In contrast, in the control positive-same condition, English ($M = 0.99$, $SD = 0.09$) and Mandarin speakers ($M = 0.99$, $SD = 0.09$) and bilinguals ($M = 0.99$, $SD = 0.09$) almost exclusively answered *yes*. In the control positive-different condition, English ($M = 0.03$, $SD = 0.18$) and Mandarin speakers ($M = 0.05$, $SD = 0.22$) and bilinguals ($M = 0.09$, $SD = 0.29$) rarely answered *yes* and almost exclusively answered *no*. These results replicated the response

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

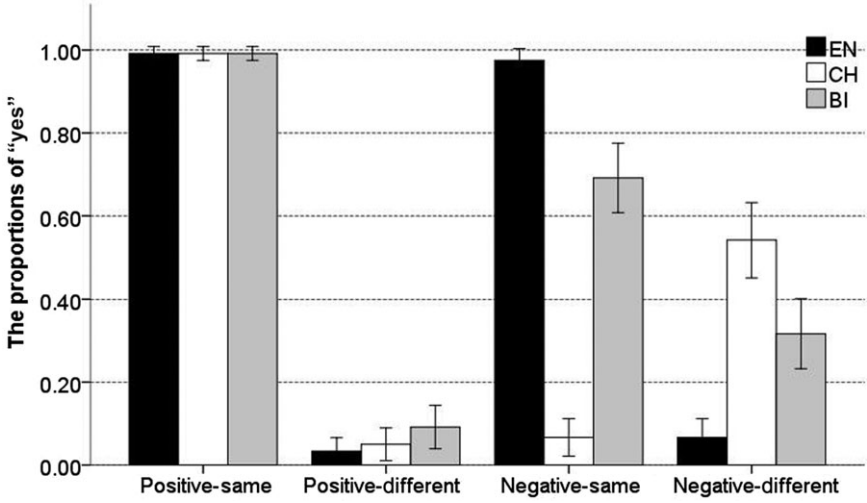

Figure 5. Mean Proportions of *yes* Answers by English and Mandarin Monolingual Speakers and Mandarin–English Bilinguals in the Comprehension Experiment (Expt. 2) (Error Bars = 95% Confidence Interval).

pattern of each language group found in previous Expt. 1 (see Figure 2), indicating crosslinguistic differences in *yes*/*no* answers to negative questions. The use of identical lists of trials across experiments cannot rule out potential practice or familiarity effects. What we observed instead were negligible oscillations and highly comparable patterns of responses in Expt. 1 (Figure 2) and Expt. 2 (Figure 3), pointing to minimal changes in response patterns that could be attributable to increasing stimulus familiarity.

Reaction times

We next consider the RTs from Expt. 2. Ten data entries from the English monolinguals (2.2% of total RTs of expected English answers), 11 data entries from the Mandarin monolinguals (2.7% of total RTs of expected Mandarin answers), and 8 data entries from the bilinguals (2.0% of total RTs of expected answers) were more than 2.5 standard deviations away from the group mean in each condition and they were replaced by the cut-offs (group mean ± 2.5 *SD*s). RTs of the expected answers in each language group for each condition are illustrated in Figure 6.

To test the effect of language group on response speed when participants process positive and negative questions, we built a series of mixed-effects models following the structure in Expt. 1: RT ~ 1 + sameness × group + (1 + question × sameness | participant) + (1 + question | item). The results are shown in Table 8 (the full dataset with RTs for each participant in each condition is available on the project website https://osf.io/x4536/).

We first compared a full model including *Question* with a reduced model excluding *Question*. This comparison showed that the model fit was significantly improved with the presence of *Question*, $\chi^2(6) = 88.20$, $p < .001$, confirming that it took participants significantly longer to answer negative questions than positive questions.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

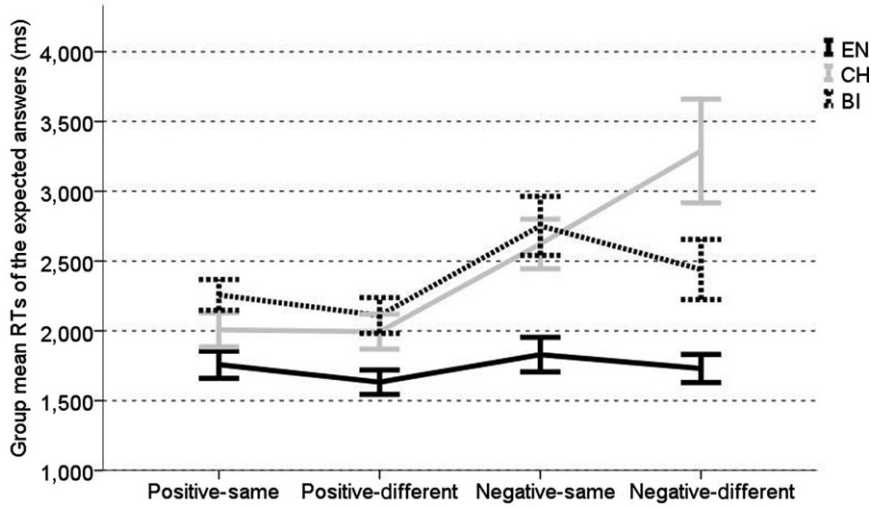

Figure 6. Mean RTs of English and Mandarin Monolingual Speakers and Mandarin–English Bilinguals in the Comprehension Experiment (Expt. 2) (Error Bars = 95% Confidence Interval).

Then we focused on the negative questions and compared a model including *Group* with a reduced model without *Group* in the data for negative questions only. This comparison confirmed that there were significant between-group differences in the processing of negative questions, $\chi^2(4) = 66.87$ $p < .001$. And third, we ran Tukey-adjusted pairwise comparisons to examine more closely how RTs to specific types of negative questions differ between groups. The results are presented in Table 9.

The RT patterns found were comparable to those in previous Expt. 1 (see Figure 2), confirming that, as predicted, Mandarin monolinguals took longer than English monolinguals to answer negative compared to positive questions. Bilinguals were slower than English speakers but faster than Mandarin speakers to respond to negative questions compared with positive question. And while in the negative-same condition bilinguals resembled the Mandarin monolinguals, in the negative-different condition bilinguals' processing was in-between that of English and Mandarin monolingual speakers. It should be noted that presenting the same trials to participants across the two experiments could have led to faster RTs in Expt. 2. Although possible training effects cannot be ruled out, we found the RT patterns in Expt. 1 and Expt. 2 to be highly comparable in terms of the key processing differences between groups.

We also ran the analysis including the RTs of the "incorrect/unexpected" responses of the Mandarin monolinguals in the negative-different condition. As shown in Table 10, the results were similar to those when we only included the expected answers (Table 8). The RTs of "incorrect/unexpected" answers from the Mandarin in the negative-different condition and the bilingual participants in the negative-different condition and the negative-same conditions are shown in Table 11.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Table 8. Coefficients from a mixed-effects model fitted to the RTs of English–Mandarin bilinguals and English and Mandarin monolingual speakers in the comprehension experiment (Expt. 2)

Fixed effects	Estimate	*SE*	*t* value	*p*
Intercept	2,565.26	145.48	17.63	< .001**
Question (pos)	−445.48	129.50	−3.44	< .001**
Sameness (same)	264.39	132.07	2.00	.048*
Group (CH)	783.10	204.93	3.82	< .001**
Group (EN)	−845.19	192.63	−4.39	< .001**
Question (pos) × Sameness (same)	−125.28	147.89	−0.85	.399
Question (pos) × Group (CH)	−903.56	186.91	−4.83	< .001**
Question (pos) × Group (EN)	353.87	173.25	2.04	.044*
Sameness (same) × Group (CH)	−1,027.21	188.60	−5.45	< .001**
Sameness (same) × Group (EN)	−159.66	174.79	−0.91	.364
Question (pos) × Sameness (same) × Group (CH)	898.02	210.39	4.27	< .001**
Question (pos) × Sameness (same) × Group (EN)	153.40	198.03	0.78	.440
Random effects	**Variance**			***SD***
Participants (intercept)	567,059			753.03
Question (pos)	305,339			552.58
Sameness (same)	288,727			537.33
Question (pos) × Sameness (same)	231,138			480.77
Item (intercept)	15,385			124.04
Question (pos)	3,747			61.21
Residual	331,698			575.93

Note. *$p < .05$; **$p < .001$.

Table 9. Tukey-adjusted pairwise comparisons of RTs in the negative conditions between groups in the comprehension experiment (Expt. 2)

Condition	Group	Estimate	*SE*	*t* ratio	*p*
Negative-different	English versus Mandarin	1,626.53	196.98	8.26	< .001**
	Bilingual versus Mandarin	−861.89	214.84	−4.01	< .001**
	Bilingual versus English	764.64	197.43	3.87	< .001**
Negative-same	English versus Mandarin	769.79	153.95	5.00	< .001**
	Bilingual versus Mandarin	223.14	176.83	1.26	.422
	Bilingual versus English	992.94	171.96	5.77	< .001**

Note. *$p < .05$; **$p < .001$.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Table 10. Coefficients from a mixed-effects model fitted to the RTs of English–Mandarin bilinguals and English and Mandarin monolingual speakers in the comprehension experiment (Expt. 2) with "incorrect" answers from the Mandarin monolingual speakers in the negative-different condition included

Fixed effects	Estimate	*SE*	*t* value	*p*
Intercept	2,549.28	133.41	19.11	< .001**
Question (pos)	−428.33	119.59	−3.58	< .001**
Sameness (same)	290.13	125.53	2.31	.023*
Group (CH)	272.81	174.85	1.56	.121
Group (EN)	−826.84	175.71	−4.71	< .001**
Question (pos) × Sameness (same)	−152.30	145.84	−1.04	.298
Question (pos) × Group (CH)	−396.15	157.97	−2.51	.013*
Question (pos) × Group (EN)	333.39	158.74	2.10	.038*
Sameness (same) × Group (CH)	−526.14	164.99	−3.19	.002*
Sameness (same) × Group (EN)	−187.21	165.27	−1.13	.260
Question (pos) × Sameness (same) × Group (CH)	398.33	195.12	2.04	.043*
Question (pos) × Sameness (same) × Group (EN)	182.92	195.26	0.94	.350
Random effects	**Variance**			***SD***
Participants (intercept)	431,916			657.20
Question (pos)	182,508			427.21
Sameness (same)	195,142			441.75
Question (pos) × Sameness (same)	149,120			386.16
Item (intercept)	14,024			118.42
Question (pos)	4,508			67.14
Residual	377,397			614.33

Note. *$p < .05$; **$p < .001$.

Table 11. Response patterns and the corresponding RT differences between the negative-different condition and the negative-same condition in the comprehension experiment (Expt. 2)

	Negative-different		Negative-same		
Response	*M*	*SD*	*M*	*SD*	RT difference
Mandarin correct responses	3,289	1,500	2,622	949	667
Bilingual incorrect responses	3,524	1,004	3,537	1,593	13
Mandarin incorrect responses	2,270	898	2,622	949	352
Bilingual correct responses	2,440	981	2,752	969	312
English responses	1,730	537	1,830	676	100

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

Discussion

Longer RTs found in Mandarin monolinguals than English monolinguals when they responded to negative questions in the comprehension experiment bring further empirical support for language specificity in the processing of negative questions. Other possible explanations for the longer response speed in Mandarin speakers to answer negative questions relate to the frequency of using negative questions and *yes*/*no* answers in Mandarin. It is possible that Mandarin speakers lack familiarity with negative questions in the critical negative-different condition and/or with the alternatives they had to choose from. We conducted brief diagnostic analyses of corpus data using the PKU-CCL-CORPUS (The Modern Chinese Corpus, 2016) and the British National Corpus. From a sample of 50 randomly selected Mandarin negative questions, 22 were in the negative-different condition while that frequency for English negative questions was 37 out of 50. Based on these frequencies, Mandarin speakers may be less familiar with the negative-different condition compared with English speakers. Another possible related reason for longer RTs in Mandarin speakers is that they may be more used to giving echo answers (e.g., -*Does she likes cats*? -*She likes cats*.) (Holmberg, 2015; Li & Cheng, 2008) rather than short *yes*/*no* answers to negative questions. Although Mandarin speakers' lack of familiarity with negative yes–no questions and *yes*/*no* answers to them could explain their longer response time, potential lack of familiarity does not provide an explanation for the different response patterns in English and Mandarin speakers. In this respect, language-specific attachment of negation has greater explanatory power.

Bilinguals, resembling English speakers, exhibited shorter RTs than Mandarin speakers, particularly when they answered negative-different questions. This resemblance to the target response system can be interpreted as an L2-driven shift in the processing routines from a truth-based toward a cognitively less demanding polarity-based system. Using two negation environments (a simpler with one mismatch, i.e., in the negative-same condition; or a more complex with two mismatches, i.e., in the negative-different condition) proved to be a particularly beneficial design feature to demonstrate in finer detail that closer approximation toward the less demanding L2-based processing system emerges when the pragmatic context is more difficult.

However, bilinguals did not exhibit a complete shift to the L2 system; differences were still found between their and the English monolingual speakers' response patterns. Namely, some bilinguals took notably longer to answer negative than positive questions in comparison with English monolingual speakers. Two likely candidates that could drive variation in bilinguals' performance are L2 exposure and L2 proficiency. With the truth-based system as the initial state of L2 learning of English negative questions, bilinguals could be approximating to the target polarity-based response system with increasing L2 exposure and L2 proficiency. Bilinguals' performance in L2 found between that of the two monolingual control groups suggests an intermediate degree of cognitive restructuring (Athanasopoulos, 2011) on the level of negation processing. An ultimate test of cognitive restructuring is when the developing L2 also affected the established L1 (e.g., Brown & Gullberg, 2010; Bylund & Jarvis, 2011). This possibility invites future research to extend the assessment of how much the L1 system may have been impacted, for instance, by means of testing

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

bilinguals on their comprehension and production of negative questions in L1 Mandarin alongside L2 English. A further intriguing question open for future research is what characterizes the developmental trajectory of learners with a less demanding polarity-based L1 (like English) acquiring a cognitively more taxing truth-based L2 system (like Mandarin or Korean or Japanese). Another beneficial addition to the future research would be to control for the potential effect of linguistic environment, as the bilinguals tested in the current study were all living in the UK. Future designs may be suitably extended by recruiting bilinguals living in China to see whether comparable results would emerge.

General discussion and conclusion

This study set out to investigate the yes–no answering systems in English and Mandarin monolingual speakers and Mandarin–English bilinguals by examining their responses and RTs when answering negative questions. We found that the *yes*/*no* answers to negative questions and the slowdowns in English and Mandarin substantially differ. One plausible explanation is that Mandarin speakers typically attach negation to the statement of the question while English speakers typically attach negation to the polarity of the question. Bilinguals' performance between that of the L1 and L2 baselines suggests that L1-specific processing of negation in negative yes–no questions can flexibly approximate to L2-like processing.

This study is an innovative contribution to research on negation processing in the truth-based system. Building on knowledge from studies on other truth-based languages including Japanese and Korean (Akiyama, 1979, 1992; Choi, 1991), RT slowdowns in Mandarin speakers in this study provide direct empirical support for the idea that it is more demanding to process negative questions for speakers of a truth-based language than for English speakers. We attribute the greater difficulty of negation processing in the truth-based system to attaching negation to the statement rather than to the polarity of a question, which is in line with Holmberg's (2015) distinction of high versus middle negations.

This study fills the research gap in Mandarin–English bilinguals' processing of negation, and it supports the idea that negation processing in an L2 is not necessarily more difficult than that in an L1. This is an important addition since in comparison to the great number of studies on negation processing in monolingual speakers (e.g., Akiyama et al., 1979; Carpenter & Just, 1975; Clark & Chase, 1972; Dale & Duran, 2011; Fischler et al., 1983; Kaup et al., 2006, 2007; Lüdtke et al., 2008; Tian et al., 2010, 2016), negation processing in bilinguals has received little attention. While Manning et al. (2018) argued that it is more demanding to process negation in the L2 than in the L1, Ćoso and Bogunović (2019) claimed that negation processing in L2 and L1 is similar. In this study, bilinguals showed an overall English-like performance, which suggests shifts from the Mandarin-like pattern toward the English-like pattern. When comparing bilinguals and English monolinguals, the findings here are in line with Manning et al. (2018) by showing that it took bilinguals longer than English monolinguals to answer negative questions. However, when comparing bilinguals and L1 Mandarin monolinguals, our observations support Ćoso and Bogunović (2019) by showing that bilinguals can be faster to respond to negative questions in English compared with monolinguals

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

to negative questions in Mandarin. Using the less difficult polarity-based system by Mandarin–English bilinguals seems to compensate for the extra difficulty of an L2.

One of the potential limitations of this study relates to between-group differences. It might be that differences between bilinguals' and Mandarin monolinguals' socio-economic and educational background could have influenced some of the results. We recruited Mandarin participants from a vocational college of preschool education while bilingual participants came from a university overseas. Such socio-demographic variation in the sample could possibly be associated with differences in verbal skills and cognitive processing in general, which in turn could be linked to variation in the speed of processing of negative questions. While mixed-effects modeling helped this study to control for within-group variation, future designs will find it advantageous to include psychometric tests to establish that the samples are fully comparable except for their language.

This study considers the contrast between the answering systems to yes–no questions in English and Mandarin speakers from a processing perspective, and it attributes this contrast to language-specific attachment of negation in negative questions. This account can also explain the different cognitive demands suggested in previous studies with English, Japanese, and Korean (Akiyama, 1979, 1992; Choi, 1991), as well as the difficulty for Mandarin learners to acquire the English answering system (Holmberg, 2015). In-between performance of Mandarin–English bilinguals suggests that some still answer English negative questions following the Mandarin system. In L2 grammar instruction, it may be useful to direct learners' attention to the structural and conceptual differences linked to the variation in negation attachment to the statement and to the polarity of the question. Explicit teaching of the differences between the L1 and the L2 has been found beneficial for learning not only L2 vocabulary (Horst et al., 2010) but also structural properties, for instance, about L2 interrogatives (Ammar et al., 2010) and L2 tense and aspect (McManus & Marsden, 2019). As an example, McManus and Marsden (2019) provided explicit instruction and practice to L1 English learners of L2 French on the L1–L2 differences in tense and aspect use, and found that this group of learners performed significantly better in a French language test than the control group who received explicit instruction and practice about L2 French only. Similarly, we envisage a comparative advantage of explicitly directing attention of Mandarin learners of English to the distinction between negation attachment either to the statement or to the polarity of the question via crosslinguistic Mandarin–English comparisons.

Acknowledgements. Research presented in this article builds on Haoruo Zhang's PhD thesis defended at the University of York in 2019. This research was supported by the Chinese Scholarship Council Grant 201608060057 awarded to Haoruo Zhang. Earlier versions of this article benefited from discussions with researchers at the Centre for Research in Language Learning and Use at York. The authors also appreciate feedback from the audience at the Second Language Research Forum 2018 in Montreal.

Notes

1. In Holmberg's classification of answering systems (2015), French (Romance, Indo-European) is neither strictly truth-based nor polarity-based due to a positive particle *oui* "yes" added to a negative question.
2. Based on Holmberg (2015), Croatian (South Slavic, Indo-European) does not follow the truth-based system.

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

References

Akiyama, M. M. (1979). Yes–no answering systems in young children. *Cognitive Psychology*, **11,** 485–504.

Akiyama, M. M. (1992). Cross-linguistic contrasts of verification and answering among children. *Journal of Psycholinguistic Research*, **21**, 67–85.

Akiyama, M. M., Brewer, W. F., & Shoben, E. J. (1979). The yes–no question answering system and statement verification. *Journal of Verbal Learning and Verbal Behavior*, **18**, 365–380.

Ammar A, Lightbown P, & Spada N. (2010). Awareness of L1/L2 differences: Does it matter?. *Language Awareness*, **19**(2), 129–146.

Athanasopoulos, P. (2011). Cognitive restructuring in bilingualism. In A. Pavlenko (Ed.) *Thinking and speaking in two languages* (pp. 29–65). Multilingual Matters.

Athanasopoulos, P., Damjanovic, L., Krajciova, A., & Sasaki, M. (2011). Representation of colour concepts in bilingual cognition: The case of Japanese blues. *Bilingualism: Language and Cognition*, **14**(1), 9–17.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**, 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**, 255–278.

Bolibaugh, C., Vanek, N., & Marsden, E. (2021). Towards a credibility revolution in bilingualism research: Open data and materials as stepping stones to more reproducible and replicable research. *Bilingualism: Language and Cognition* (In press).

Brown, A., & Gullberg, M. (2010). Changes in encoding of PATH of motion in a first language during acquisition of a second language. *Cognitive Linguistics*, **21**(2), 263–286.

Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition*, **14**(1), 47–59.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, **82**, 45–73.

Choi, H. (2014). The answering system to yes–no truth-functional questions in Korean-English bilingual children. *Journal of Psycholinguistic Research*, **43**, 1–21.

Choi, S. (1991). Children's answers to yes–no questions: A developmental study in English, French, and Korean. *Developmental Psychology*, **27**, 407–420.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, **3**, 472–517.

Ćoso, B., & Bogunović, I. (2019). The role of linguistic cues in bilingual negation processing. *International Journal of Bilingualism*, **23**, 21–36.

Dale, R., & Duran, N. D. (2011). The cognitive dynamics of negated sentence verification. *Cognitive Science*, **35**, 983–996.

Dudschig, C., & Kaup, B. (2018). How does "not left" become "right"? Electrophysiological evidence for a dynamic conflict-bound negation processing account. *Journal of Experimental Psychology: Human Perception and Performance*, **44**, 716–728.

Ernst, T. (1995). Negation in mandarin Chinese. *Natural Language & Linguistic Theory*, **13**(4), 665–707.

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry, N. W. Jr. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, **20**, 400–409.

Hasson, U., & Glucksberg, S. (2006). Does understanding negation entail affirmation? An examination of negated metaphors. *Journal of Pragmatics*, **38**, 1015–1032.

Holmberg, A. (2013). The syntax of answers to polar questions in English and Swedish. *Lingua*, **128**, 31–50.

Holmberg, A. (2014). Answering yes/no-questions in English and other languages. *The EFL Journal*, **5**, 23–42

Holmberg, A. (2015). *The syntax of yes and no*. Oxford University Press.

Horst, M., White, J., & Bell, P. (2010). First and second language knowledge in the language classroom. *International Journal of Bilingualism*, **14**(3), 331–349.

Huang, B., & Liao, X. (2007). *Xiandai hanyu [Modern Chinese]* (pp. 1–159). Higher Education Press.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, **38**, 1033–1050.

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly Journal of Experimental Psychology*, **60**, 976–990.

Keating, G. D., & Jegerski, J. (2015). Experimental designs in sentence processing research: A methodological review and user's guide. *Studies in Second Language Acquisition*, **37**, 1–32.

Levelt, W. J. (1989). Speaking: From intention to articulation. Cambridge, MA: The MIT Press.

Li, D., & Cheng, M. (2008). *A practical Chinese grammar for foreigners.* Beijing Language and Culture University Press.

Lu, J. (2003). *Xiandai hanyu yufa yanjiu jiaocheng [A Study Guide of Modern Chinese Grammar]* (pp. 205–266). Beijing, China: Peking University Press.

Lüdtke, J., Friedrich, C. K., De Filippis, M., & Kaup, B. (2008). Event-related potential correlates of negation in a sentence–picture verification paradigm. *Journal of Cognitive Neuroscience*, **20**, 1355–1370.

Manning, G., Sabourin, L., & Farshchi, S. (2018). *Negation Processing in L2 English Speakers.* (Poster Presentation). Second Language Research Forum, Montreal.

McManus, K., & Marsden, E. J. (2019). Signatures of automaticity during practice: Explicit instruction about L1 processing routines can improve L2 grammatical processing. *Applied Psycholinguistics*, 40(1), 205–234.

Norris, J. M. (2015). Statistical significance testing in second language research: Basic problems and suggestions for reform. *Language Learning*, **65**, 97–126.

Park, H. I., & Ziegler, N. (2014). Cognitive shift in the bilingual mind: Spatial concepts in Korean–English bilinguals. *Bilingualism: Language and Cognition*, **17**(2), 410–430.

R Development Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org/

Tang, M., Vanek, N., & Roberts, L. (2021). Crosslinguistic influence on English and Chinese L2 speakers' conceptualization of event series. *International Journal of Bilingualism*, **25**(1), 205–223.

The Modern Chinese Corpus (PKU-CCL-CORPUS). Retrieved July 8, 2016 from http://ccl.pku.edu.cn:8080/ccl_corpus/index.jsp

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account. *Quarterly Journal of Experimental Psychology*, **63**, 2305–2312.

Tian, Y., Ferguson, H., & Breheny, R. (2016). Processing negation without context – why and when we represent the positive argument. *Language, Cognition and Neuroscience*, **31**, 683–698.

Vanek, N. (2020). Changing event categorization in second language users through perceptual learning. *Language Learning*, **70**(2), 309–348.

Vanek, N., & Selinker, L. (2017). Covariation between temporal interlanguage features and nonverbal event categorisation. *IRAL – International Review of Applied Linguistics in Language Teaching*, **55,** 223–243.

Xiao, R. Z., & McEnery, A. M. (2008). Negation in Chinese: A corpus-based study. *Journal of Chinese Linguistics*, **36**(2), 274–330.

Appendix
Yes–No Questions in Expt. 1 and Expt. 2

List A	List B	List C	List D
Positive-same	Positive-different	Negative-same	Negative-different
Mr. Fox stole a roast duck from a farm.			
1. Did Mr. Fox steal a roast duck from a farm?	1. Did Mr. Fox steal a roast chicken from a farm?	1. Didn't Mr. Fox steal a roast duck from a farm?	1. Didn't Mr. Fox steal a roast chicken from a farm?
Mrs. Fox baked a cake for her family.			
2. Did Mrs. Fox bake a cake for her family?	2. Did Mrs. Fox bake some potatoes for her family?	2. Didn't Mrs. Fox bake a cake for her family?	2. Didn't Mrs. Fox bake some potatoes for her family?
Mr. Sheep watched a football game on Friday.			
3. Did Mr. Sheep watch a football game on Friday?	3. Did Mr. Sheep watch a basketball game on Friday?	3. Didn't Mr. Sheep watch a football game on Friday?	3. Didn't Mr. Sheep watch a basketball game on Friday?
Positive-different	Negative-same	Negative-different	Positive-same
Mrs. Sheep read the newspaper after dinner.			
4. Did Mrs. Sheep read the novel after dinner?	4. Didn't Mrs. Sheep read the newspaper after dinner?	4. Didn't Mrs. Sheep read the novel after dinner?	4. Did Mrs. Sheep read the newspaper after dinner?
Mr. Duck lost his watch during his trip to Europe.			
5. Did Mr. Duck lose his scarf in a trip?	5. Didn't Mr. Duck lose his watch in a trip?	5. Didn't Mr. Duck lose his scarf in a trip?	5. Did Mr. Duck lose his watch in a trip?
Mrs. Duck had her piano class on Tuesday.			
6. Did Mrs. Duck have her dance class on Tuesday?	6. Didn't Mrs. Duck have her piano class on Tuesday?	6. Didn't Mrs. Duck have her dance class on Tuesday?	6. Did Mrs. Duck have her piano class on Tuesday?
Negative-same	Negative-different	Positive-same	Positive-different
Mr. Swan gave a necklace to his wife.			
7. Didn't Mr. Swan give a necklace to his wife?	7. Didn't Mr. Swan give a ring to his wife?	7. Did Mr. Swan give a necklace to his wife?	7. Did Mr. Swan give a ring to his wife?
Mrs. Swan cleaned her feathers in a lake.			
8. Didn't Mrs. Swan clean her feathers in a lake?	8. Didn't Mrs. Swan clean her boots in a lake?	8. Did Mrs. Swan clean her feathers in a lake?	8. Did Mrs. Swan clean her boots in a lake?
Mr. Lion sent a box of DVDs to his cousin.			
9. Didn't Mr. Lion send a box of DVDs to his cousin?	9. Didn't Mr. Lion send a box of chocolates to his cousin?	9. Did Mr. Lion send a box of DVDs to his cousin?	9. Did Mr. Lion send a box of chocolates to his cousin?

(*Continued*)

https://doi.org/10.1017/S0142716421000175 Published online by Cambridge University Press

(*Continued*)

List A	List B	List C	List D
Negative-different	Positive-same	Positive-different	Negative-same
Mrs. Lion broke a glass in her kitchen.			
10. Didn't Mrs. Lion break a plate in her kitchen?	10. Did Mrs. Lion break a glass in her kitchen?	10. Did Mrs. Lion break a plate in her kitchen?	10. Didn't Mrs. Lion break a glass in her kitchen?
Mr. Dove opened his gift shop in the city center.			
11. Didn't Mr. Dove open his restaurant in the city center?	11. Did Mr. Dove open gift shop in the city center?	11. Did Mr. Dove open his restaurant in the city center?	11. Didn't Mr. Dove open his gift shop in the city center?
Mrs. Dove bought a coat during the Christmas sale.			
12. Didn't Mrs. Dove buy a hat during the Christmas sale?	12. Did Mrs. Dove buy a coat during the Christmas sale?	12. Did Mrs. Dove buy a hat during the Christmas sale?	12. Didn't Mrs. Dove buy a coat during the Christmas sale?

Cite this article: Zhang, H. and Vanek, N. (2021). From "No, she does" to "Yes, she does": Negation processing in negative yes–no questions by Mandarin speakers of English. *Applied Psycholinguistics* **42**, 937–967. https://doi.org/10.1017/S0142716421000175