認知科学 第 28 巻 第 2 号 (2021) pp. 271–281 https://doi.org/10.11225/cs.2021.001
Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, Vol. 28, No. 2, pp. 271–281

Research Paper

Foreign language side effect when inner language is suspected to accompany thinking: Lowered thinking ability in daily verbal communication

Yohtaro Takano* Takashi Yagyu

The University of Tokyo

Foreign language side effect (FoLSE) refers to a temporary decline of thinking ability while nonproficient foreign language is being used. This decline is produced by stronger interference between thinking and a heavier cognitive load of foreign language processing. Although FoLSE was shown to occur in laboratories, it may not occur in daily verbal communication when thinking is accompanied by inner language. The reason is as follows: In general, the more similar two concurrent cognitive tasks, the stronger their mutual interference. Inner language is usually experienced as effortless native language. When outer language used in verbal communication is native language, it is more similar to inner native language. Therefore, outer native language is expected to produce stronger interference with inner native language involved in thinking, and thus could produce a larger reduction in thinking performance. This larger reduction due to outer native language may cancel out the reduction due to outer foreign language (FoLSE). To examine this possibility, Japanese college students performed verbal and thinking tasks concurrently in two dual-task experiments. The verbal task was presented in either Japanese (native language) or English (foreign language). The thinking task was always presented in Japanese (native language). Past empirical studies strongly suggested that inner language should be evoked in the employed thinking tasks (i.e., validity judgment on categorical syllogism and intelligence test problems that loaded heavily on verbal factors of intelligence). The results revealed that performance in the thinking task was lower when the verbal task was presented in the foreign language. This means that FoLSE was stronger than the interference between the inner and outer native language. It follows that FoLSE is likely to occur in daily verbal communication as well even when it is accompanied by inner language.

Keywords: foreign language, thinking, bilingualism, inner language, dual-task

Received 31 August 2020; Accepted 9 January 2021

1. Introduction

Foreign language side effect (FoLSE) is temporary decline of thinking ability while foreign language is being used (Takano & Noda, 1993, 1995).[1)] Here, *foreign language* refers to a non-native language, in which its user is not as proficient as in his/her own native language. This research investigates whether FoLSE manifests itself when thinking is accompanied by inner language as often presumed in daily verbal communication.

1.1 Theoretical Prediction of the Foreign Language Side Effect (FoLSE)

Takano and Noda (1993) explained the reason for the occurrence of FoLSE as follows: Daily verbal communication typically consists of linguistic processing and thinking. On the one hand, *linguistic processing* refers to the information processing of verbal input and output: for example, phonetic analysis, morphemic analysis, syntactic analysis, prosodic analysis, and so on in language comprehension; linguistic planning, sentence generation, articulation, self-monitoring, and so on in language production. The

Yohtaro Takano is now at Meiji University and Takashi Yagyu is now at Tokyo Future University

∗ Corresponding author. E-mail: takano@L.u-tokyo.ac.jp

1) Takano and Noda (1993, 1995) referred to the same phenomenon as *foreign language effect.* However, *effect* often carries a positive connotation, whereas this phenomenon is negative. In this article, therefore, *effect* is replaced by *side effect*, which usually bears a negative connotation. This renaming will help distinguish this phenomenon from recently discussed reduction of decision biases due to foreign language, which is also called *foreign language effect* (see Discussion).

difference between foreign and native language is primarily concerned with this linguistic processing. On the other hand, *thinking* refers to semantic information processing that does not directly deal with verbal input and output: for example, inference, detection of relationship between a speaker's current and former utterances, examination of logical or factual validity of an utterance, and so on in listening; real time construction of semantic structures in an ongoing speech, retrieval of related knowledge, interpretation of a listener's facial expression, and so on in speaking.[2)]

In daily verbal communication, linguistic processing and thinking are usually performed in parallel: We think while listening; we speak while thinking. If linguistic processing and thinking are not performed concurrently and thinking is delayed until completion of linguistic processing, many undesirable consequences are expected: For example, information needed for thinking may disappear from working memory; a long silence for thinking may be interpreted as a sign of inability to refute or even a sign of a slow-witted mind; and a long silence will be certainly interrupted by another speaker's utterance (for implicit rules of turn taking in conversation, see Sacks et al., 1974). In addition, ample experimental evidence has established that inference is performed at the same time of listening or reading in language comprehension (e.g., Bransford et al., 1972; Johnson et al., 1973; McKoon & Ratcliff, 1981, 1992).

Meanwhile, it is well known that two demanding cognitive tasks interfere with each other when they are conducted concurrently (e.g., Broadbent, 1958; Norman & Bobrow, 1975; Treisman, 1969). In daily verbal communication where linguistic processing and thinking are conducted in parallel, therefore, performance of either or both can be impaired by the mutual interference. However, thinking tends to be sacrificed in most cases because appropriate linguistic processing is usually a prerequisite for appropriate verbal communication: If linguistic processing is neglected in language comprehension, a speaker's utterance will become irrelevant to a communication partner's preceding utterances; in language production, a speaker will emit meaningless sounds or no sound at all.

It is also well known that intensive practice of cognitive tasks leads to decrease in the interference (e.g., Hirst et al., 1980; LaBerge, 1981; Schneider et al., 1984; Spelke et al., 1976; Underwood, 1974). In general, native language has been practiced intensively throughout a lifetime, and thus its linguistic processing will cause a relatively small amount of interference with concurrent thinking. In contrast, foreign language defined above has been typically practiced for a much shorter period of time, and thus its linguistic processing will cause a relatively large amount of interference. When foreign language is used, therefore, thinking performance will be lower than when native language is used. This lower performance is FoLSE.[3)] It is a specific case of the general principle that harder cognitive tasks produce stronger mutual interference.

FoLSE can be also explained in terms of the resource theory of attention (Kahneman, 1973) or working memory (Just & Carpenter, 1992) as follows: In general, a cognitive task needs a certain amount of cognitive resource whose total amount is limited. When two or more demanding cognitive tasks are performed concurrently, the amount of resource is not sufficient to perform all those tasks, and thus the performance of some of them has to decline. However, practice of a cognitive task reduces the amount of resource needed for its performance. Usually, foreign language is practiced much less than native language and thus needs much more resource. When thinking is performed concurrently with linguistic processing, therefore, performance of thinking will be lower when foreign language is used than when native language is used. This lower thinking performance is FoLSE. It has to be noted, however, that FoLSE does not depend on the resource theory or any other particular theory. As shown in the next section, FoLSE is a robust experimental fact that has to be explained by any valid theory of cognition.

FoLSE has practical implications: It has been shown that lower proficiency in foreign language leads to lower performance in critical thinking (Floyd, 2011; Lun et al., 2010; Manalo & Sheppard, 2016) and lower probability of spontaneous use of diagrams when students are writing explanations of what they have learned (Manalo & Uesaka, 2012). It is reasonable to consider that FoLSE is at least partly responsible for these negative effects of foreign language.

Meanwhile, a number of studies have shown that bilingualism tends to improve intellectual ability (e.g., Bialystok et al., 2012; Hakuta & Diaz, 1985; Peal & Lambert, 1962; Ricciardelli, 1992). Although FoLSE may appear to contradict these findings at first glance,

2) It is not necessary for the subsequent arguments to assume that a clear line can be drawn between linguistic processing and thinking. Although a borderline between them may be somewhat obscure, it is obvious that typically conceived *thinking* is not equivalent to the above defined linguistic processing such as phonemic analysis and articulation.

3) The lowered performance is virtually synonymous with lowered ability in this case because under strong interference it is impossible to raise performance up to the normal level by mere effort.

there is no real contradiction because the improvement of intellectual ability in bilinguals is a long-term effect, whereas FoLSE is a short-term effect that occurs only while a foreign language is actually being used. One of the reasons that the intellectual ability of a bilingual is improved in the long run may be that he/she strives to overcome immediate difficulty of thinking (i.e., FoLSE) while struggling with a less proficient language.

1.2 Empirical Confirmation of FoLSE

Although FoLSE is thus predicted theoretically, empirical confirmation is indispensable for two reasons: First, thinking and linguistic processing may not be demanding enough to actually produce mutual interference. Second, little or no interference is generated when two cognitive tasks are distinct enough from each other (e.g., Allport, 1980: Navon & Gopher, 1979; Treisman & Davies, 1973). The possibility that thinking and linguistic processing are distinct enough cannot be precluded on an a priori ground. If no interference is generated between thinking and linguistic processing, FoLSE will not take place. To examine whether FoLSE can be observed in reality, Takano and Noda (1993) conducted two dual-task experiments, which required participants to perform a thinking task and a verbal task at the same time. This dual-task experimentally simulated daily verbal communication, in which thinking is performed concurrently with linguistic processing.

In the thinking task, no foreign language was used, whereas either native or foreign language was used in the verbal task. If the native language and foreign language conditions had been set in the thinking task, it would have been impossible to discriminate whether lower thinking performance in the foreign language condition, if any, was produced by lowered thinking ability (i.e., FoLSE) or foreign language difficulty per se. This was why no foreign language was used in the thinking task while the native language and foreign language conditions were set in the concurrent verbal task.

In one of their experiments, the thinking task was addition of two-digit numbers, while the verbal task was to answer general knowledge questions (e.g., “Is a lion an animal that lives in water?”). The performance in the thinking task was significantly lower when the verbal task was presented in the foreign language than when it was presented in the native language. FoLSE was thus demonstrated to occur actually. The results were essentially identical for both native speakers of Japanese and those of English.

It has to be stressed that no foreign language was used at all in the thinking task. This means that the decline of thinking performance in the foreign language condition cannot be interpreted to directly reflect well-known foreign language difficulty (for example, it was *not* the case that the participants made more errors because they could not understand the thinking task problems presented in the difficult foreign language). Rather, the performance decline was an indirect consequence of the foreign language difficulty through its interference with thinking (i.e., FoLSE).

FoLSE should not be confounded with foreign language difficulty per se: The latter is difficulty in linguistic processing, whereas the former is difficulty in thinking. Although the foreign language difficulty is widely recognized, FoLSE usually goes unnoticed, and the difficulty accompanying foreign language tends to be wholly attributed to the well-known difficulty in the linguistic processing of foreign language.

In the other experiment, Takano and Noda (1993) demonstrated that FoLSE could occur in non-verbal thinking as well: As a thinking task, they employed those intelligence test problems that had been shown to have heavy loadings on non-verbal spatial factors in factor-analytic studies of intelligence tests (e.g., card rotations test problems, maze-tracing speed test problems, and so on).

Takano and Noda (1995) further demonstrated that the magnitude of FoLSE was modulated by similarity between native and foreign language. The overall difficulty of learning a foreign language tends to be larger when it is dissimilar to a native language than when it is similar (Odlin, 1989).[4] If the time spent in learning a foreign language is comparable, accordingly, those whose native language is dissimilar to that foreign language will be less proficient than those whose native language is similar. It follows that FoLSE will be larger for those whose native language is dissimilar because FoLSE results from proficiency difference between foreign and native language.

Takano and Noda (1995) tested this reasoning in two dual-task experiments. In the first experiment, they compared native speakers of Japanese and those of German while employing English as a common foreign language. It was predicted that FoLSE would be larger for the Japanese speakers because English is much more dissimilar to Japanese than to German. The results confirmed this prediction. In the sec-

4) Although language transfer can be either facilitatory or inhibitory for individual linguistic components (Flege, 1987, 1991; Navarra et al., 2005; Selinker et al., 1975), it is facilitatory for the acquisition of a second language as a whole.

ond experiment, they compared native speakers of English and those of Korean while employing Japanese as a common foreign language. It was predicted that FoLSE would be larger for the English speakers because Japanese is much more dissimilar to English than to Korean. Again, the results confirmed this prediction.

Tyler (2001) also conducted a dual-task experiment to see if FoLSE could be observed for non-native speakers of English who had stayed in Australia for more than 10 years. As a verbal task, he aurally presented Bransford and Johnson's (1972) *Washing Text*, which is very difficult to comprehend when its title (i.e., *washing*) is not provided. He found that performance of a concurrent thinking task (i.e., verification of single-digit calculations) was lower for non-native than native speakers.

1.3 Linguistic Similarity Effect Due to Inner Language

Although the occurrence of FoLSE has been thus demonstrated in laboratories, it is sometimes questioned whether it occurs in daily verbal communication as well. One such criticism focuses on the presence of inner language that may accompany thinking. In daily verbal communication, thinking develops mainly on the basis of verbal input. Accordingly, it is likely that thinking is accompanied by continuing linguistic processing. This linguistic processing, if any, should take the form of inner language because thinking aloud is rarely observed. Here, the term *inner language* is used in its ordinary meaning, namely, covert linguistic processing without vocalization. Although the presence of inner language is disputable because it is not directly observable, its presence cannot be denied confidently for the same reason; in addition, most of us often experience it intuitively.

The criticism goes on as follows: Unless a special conscious effort is made, inner language tends to be native language because it is easier to use. When native language is also used as outer language in verbal communication, therefore, the similarity between inner and outer language will be larger; when foreign language is used as outer language, the similarity will be smaller. It is well documented that similar cognitive tasks tend to produce larger interference with each other (e.g., Brooks, 1968; Eagan & Chein, 2012; Marsh et al., 2009; Pashler, 1994; Shah & Miyake, 1996; Stroop, 1935; Treisman, 1964; Turner & Eagle, 1989; Vasilev et al., 2019). It follows that when native language is used as outer language, its interference with concurrent thinking will be stronger because outer native language is similar to inner native language involved in thinking; when foreign language is used as outer language, its interference will be weaker because outer foreign language is less similar to inner native language. Accordingly, it is predicted that the decline of thinking ability will be larger when native language is used than when foreign language is used in verbal communication. In what follows, this differential interference will be referred to as *linguistic similarity effect*, which is a special case of the above general similarity effect between cognitive tasks.

The predictions of the linguistic similarity effect and FoLSE are exactly opposite: When foreign language is used in verbal communication, the linguistic similarity effect predicts weaker interference, whereas FoLSE predicts stronger interference. If the linguistic similarity effect is potent enough, therefore, FoLSE will be cancelled out by the linguistic similarity effect and the decline of thinking ability will not be manifested. If the linguistic similarity effect is too potent, FoLSE will be overridden and enhanced thinking ability will be manifested when foreign language is used.

The previous laboratory experiments were inadequate to examine this possibility because their thinking tasks involved minimal inner language, if any: Takano and Noda (1993) deliberately excluded linguistic processing in one of their experiments by employing spatial problems that loaded heavily on non-verbal factors of intelligence. Although inner language might be involved in the calculation tasks (Takano & Noda, 1993, 1995; Tyler, 2001), its cognitive load should have been minimal because the presumed inner language is supposed to have been mere retrieval of routine phrases (e.g., "One plus one equals two"). Generally, a smaller cognitive load tends to produce weaker interference (e.g., Lavie et al., 2004; Norman & Bobrow, 1975). FoLSE might have been observed in the calculation tasks because the cognitive load of their inner language was minimal and thus the linguistic similarity effect was too weak to offset FoLSE.

2. Experiment 1

The purpose of Experiment 1 was to examine whether FoLSE would occur when substantial inner language is presumed to accompany thinking. This experiment inherited Takano and Noda's (1993, 1995) basic experimental framework: concurrent performance of a thinking task which was given only in native language, and a verbal task which was given in either native or foreign language. In the present experiment, the thinking task was validity judgment about a categorical syllogism, in which two premises and a

conclusion were presented verbally in native language. An example of the syllogisms was as follows: "All ships are yachts. No motorboats are yachts. Therefore, some motorboats are not ships."[5] To make a validity judgment, it is argued that linguistic processing is performed internally without vocalization on verbal input (e.g., Polk & Newell, 1995; Wetherick & Gilhooly, 1990). In this thinking task, moreover, thinking proceeds on the basis of verbal input as in daily verbal communication. If it is presumed that daily verbal communication is accompanied by substantial inner language, there is no reasonable basis to deny that this thinking task is also accompanied by substantial inner language, although it is not directly observable.

In this experiment, what is assumed about the inner language is confined to the following two properties: First, it is not accompanied by overt vocalization. Second, it is linguistic processing of native language. The purpose of this experiment is neither to show the presence of inner language nor to investigate its properties, but to see whether FoLSE is observed when inner language is presumed to accompany thinking.

2.1 Method

2.1.1 Design

In a within-participants design, the independent variable had three levels: Control condition, in which the thinking task alone was performed without the verbal task; Native condition, in which the concurrent verbal task was imposed in native language; and Foreign condition, in which the verbal task was imposed in foreign language. It has to be noted that the names of the conditions, Native and Foreign, originated from the kinds of language used in the verbal task, not in the thinking task which was always presented in native language. The dependent variables were numbers of correct answers in the thinking and verbal tasks.

2.1.2 Participants

As unbalanced bilinguals, 24 native speakers of Japanese (graduates and undergraduates: age, $M = 22.5$ years, $SD = 3.1$) served as paid volunteers. They had learned English as a foreign language in high schools and universities for 8 to 12 years; none of them had stayed in English-speaking countries for more than one year.

2.1.3 Materials

In the thinking task, two premises and a conclusion of a categorical syllogism were printed in the native language on a page of a booklet. Three booklets each carrying 42 syllogisms were prepared; none of the syllogisms were repeated between the booklets. The assignment of the three booklets to the three conditions was randomly changed between participants. Three additional booklets each carrying 15 syllogisms were also prepared for practice trials.

The verbal task was sentence verification. Two lists of sentences were prepared in both Japanese and English, respectively. When one of the lists was presented in Japanese to a particular participant, the other was presented in English to the same participant; the assignment of the lists to the languages was changed randomly between participants. One list was composed of 36 complex sentences (e.g., "A horse is an animal that lives in water."), which were aurally presented every 8 sec by a tape recorder: "Yes" was a correct answer for half of the sentences; "No" for the other half. The lists were tape-recorded in Japanese by a female native speaker of Japanese; in English by a female native speaker of English. The recorded tape first presented the direction, "Ready? Start," and ended with the direction, "Stop"; the interval between these two directions was 296 sec. In one list, the total net time spent in presenting the 36 sentences was about 124 sec for the Japanese version and about 126 sec for the English version; in the other list, about 123 sec for the Japanese version and 124 sec for the English version. Two similar lists each consisting of 6 sentences were additionally prepared in both Japanese and English to be used in practice trials. The recorded lists were identical to those used by Takano and Noda (1993, Experiment 2).

2.1.4 Procedures

Each participant was tested individually in an experimental session consisting of an instruction, three conditions (i.e., Control, Native, and Foreign), introspective reports, and questions about English learning. In the thinking task, the participant was asked to make as many validity judgments as possible within a fixed period of time (i.e., 296 sec) by writing down either a circle (for valid) or an X (for invalid) for each syllogism; both speed and accuracy were equally stressed. When a syllogism was too difficult, the participant was allowed to skip it without spending too much time. In the verbal task, the participant answered by saying "Yes" or "No" for each sentence; the participant was asked to say, "I don't know," when he/she did not know the answer or failed to understand the presented sentence. Absence of an answer was regarded as "I don't know." The participant was allowed to answer in either

5) Although its first premise and conclusion may not agree with our common knowledge, its logical inference is valid.

of the two languages that first came to mind, irrespective of the language used to present the sentences. The dual-task (Native or Foreign) condition was composed of a practice trial of 56 sec and a test trial of 296 sec. A test trial proceeded as follows: Upon hearing "Start," the participant started to make validity judgments. The first sentence was presented 8 sec later. After emitting a response to this sentence, the participant could concentrate on the thinking task until the next sentence was presented. The participant terminated the thinking task with the direction, "Stop." The participant was asked to perform both thinking and verbal tasks with equal efforts. The single-task (Control) condition was also composed of a practice and a test trial, in which the participant was asked to make as many validity judgments as possible within the same fixed period of time as in the dual-task conditions. The order of the three test conditions was counterbalanced between participants.

2.2 Results

In the thinking task, the participants made 17.5 correct validity judgments on the average ($SD = 6.1$) in Control condition, 13.5 ($SD = 4.8$) in Native condition, and 11.0 ($SD = 4.5$) in Foreign condition. It should be stressed again that no foreign language was used in the thinking task. As discussed previously (1.2. Empirical confirmation of FoLSE), therefore, the lower performance in Foreign condition reflects FoLSE, not foreign language processing difficulty per se.

The numbers of the correct validity judgments were converted to *interference rate* by the following equation as in Takano and Noda (1993):

$$I = (S - D)/S \times 100(\%) \qquad (1)$$

where S denotes the number of correct validity judgments in the single-task (Control) condition, and D denotes that in the dual-task (Native or Foreign) condition. This index represents the magnitude of the interference by the verbal task, while taking as a baseline the performance in the single-task condition where no verbal task was imposed: In other words, this index represents what proportion of the single-task performance was reduced by imposing the concurrent verbal task in either the native or foreign language. If FoLSE was present, the interference rate should be larger in Foreign than in Native condition. The mean interference rates shown in Figure 1 agree with this prediction. In Figure 1, the difference in the interference rate between Foreign and Native conditions reflects the magnitude of FoLSE.[6] The mean interference rate was significantly larger in Foreign than Native condition, Wilcoxon test $z = 2.555$, $p = .011$, effect size $r = .522$. This means the occurrence of FoLSE.

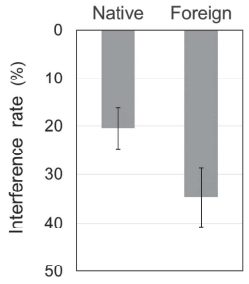

Fig. 1 Mean interference rates with standard errors in Experiment 1.

In the verbal task, the mean number of correct answers was 35.0 ($SD = 1.0$) in Native condition, and 27.5 ($SD = 4.0$) in Foreign condition. This difference was also significant, Wilcoxon test, $z = 4.286$, $p < .0001$, effect size $r = .875$. This confirms that the precondition of FoLSE was satisfied in that the present participants were unbalanced bilinguals who were less proficient in the foreign language (English) than in the native language (Japanese).

3. Experiment 2

Experiment 2 attempted to replicate the findings in Experiment 1 with a different kind of thinking task that had been suggested to evoke inner language by past empirical studies. More specifically, Experiment 2 employed as its thinking task those intelligence test problems that had been shown to have heavy loadings on verbal factors in factor-analytic studies of intelligence (Ekstrom et al., 1976; Eysenck, 1979; Okuno, 1969a, 1969b; Osaka & Okuno, 1956). In one of those problems, for example, words in an incomplete sentence were presented in a scrambled order, and the participant had to mentally rearrange those words to choose between five alternatives a word that would complete the sentence. This problem requires thinking to closely cooperate with linguistic processing, which

6) Although the difference in the interference rate between Native and Foreign conditions was 14.4% in this experiment, it does not represent the general magnitude of FoLSE. The 296 sec in the test trial was not fully occupied by the verbal task because the net time spent in presenting its sentences was 123 to 126 sec and a response ("Yes" or "No") did not take much time. Thus, the participants could concentrate on the thinking task alone for a substantial portion of the given time. If the sentences had been presented in a more rapid pace, the magnitude of FoLSE should have been much larger. The magnitude of FoLSE will be largely affected by such situational factors as time pressure, difficulty of linguistic and/or non-linguistic processing, and so on.

should be inner language because the problem has to be solved silently.

3.1 Method

The method of Experiment 2 was identical to that of Experiment 1 with the following exceptions.

3.1.1 Participants

As unbalanced bilinguals, 24 native speakers of Japanese (graduates and undergraduates: age, $M = 22.4$, $SD = 3.2$) served as paid volunteers. None of them participated in Experiment 1. They had learned English as a foreign language in high schools and universities for 8 to 12 years; none of them had stayed in English-speaking countries for more than one year.

3.1.2 Materials

The verbal task was identical to that in Experiment 1. The thinking task was composed of the intelligence test problems that were designed to assess verbal intelligence. These problems were always presented in native language in both Native and Foreign conditions as in Experiment 1. The problems were adopted from the following three intelligence tests: Kyoto University NX15 Intelligence Test (Osaka & Umemoto, 1984); University of Tokyo A-S Intelligence Test (Yoda et al., 1959); New Tanaka A Intelligence Test (Tanaka & Sakakibara, 1949), which is a Japanese version of Binet test. To obtain a sufficient number of problems, new problems of the same kinds were newly prepared. Three lists of 45 problems each were prepared to be used in the three conditions (Control, Native, and Foreign), respectively. The assignment of these three lists to the three conditions was counterbalanced between participants. Three lists of 12 problems each were also prepared to be used in practice. All the problems required participants to choose between four to five alternatives. Each problem and its numbered alternatives were presented on a computer display.

3.1.3 Procedures

The participant was given practice trials for Control, Native, and Foreign conditions in this order. After all the practice trials were completed, the test trials were conducted in an order counterbalanced between participants. In the thinking task, the participant answered a problem by pressing a number key corresponding to a chosen alternative on the numerical keypad.

3.2 Results

In the thinking task, participants made 20.5 correct answers on the average ($SD = 5.5$) in Control condition, 15.2 ($SD = 3.4$) in Native condition, and 11.7 ($SD = 3.7$) in Foreign condition. The mean interference rate in Foreign condition was larger than that in Native condition (see Figure 2). The difference between these means represents the magnitude of FoLSE. The difference between Foreign and Native conditions was significant, Wilcoxon test, $z = 2.996$, $p = .003$, effect size $r = .612$. This attests to the presence of FoLSE. In the verbal task, the mean number of correct answers was 34.2 ($SD = 1.4$) in Native condition, and 24.9 ($SD = 6.0$) in Foreign condition. The difference between these means was also significant, Wilcoxon test $z = 4.286$, $p < .0001$, effect size $r = .875$. Again, the precondition for FoLSE was satisfied.

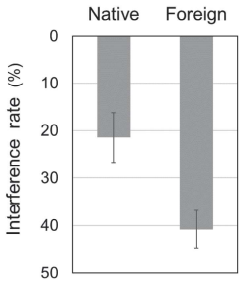

Fig. 2 Mean interference rates with standard errors in Experiment 2.

4. Discussion

The purpose of the present study was to investigate whether FoLSE would be observed when inner language was presumed to accompany thinking as in daily verbal communication. It was predicted that FoLSE would not be observed if the stronger interference between inner and outer native language (i.e., the linguistic similarity effect) was potent enough.

Although inner language is not directly observable by definition, there is no critical difference between daily verbal communication and the present thinking tasks in the possibility of evoking inner language. The presence of inner language is usually presumed in daily verbal communication for two reasons: First, most of materials to be thought of are provided verbally. Second, we intuitively experience inner language. The above thinking tasks satisfied both of these conditions: First, they were presented verbally. Second, the participants reported that they experienced inner language in their introspective reports. In addition, the empirical studies strongly suggest that inner language should be produced in the above thinking tasks (Ekstrom et al., 1979; Okuno, 1969a, 1969b; Osaka & Okuno, 1956; Polk & Newell, 1995; Wetherick

& Gilhooly, 1990). If inner language is presumed at all in daily verbal communication, there is no reason to doubt the presence of inner language in the above thinking tasks.

It is reasonable to assume that the presumed inner language was the native language for two reasons: First, the thinking tasks were always presented in the native language. Second, the effortless native language was advantageous in answering the problems in the thinking tasks. If the inner language was the native language, the interference between the verbal and thinking tasks should have been stronger when the verbal task was presented in the native language (i.e., the linguistic similarity effect). In actuality, however, the performance in the thinking task was lower when the verbal task was presented in the foreign language. This means that FoLSE was stronger than the linguistic similarity effect in these experiments. Accordingly, it is likely that FoLSE is present in daily verbal communication as well even when it involves substantial inner language.

It is sometimes pointed out that the semantic contents of the verbal task and those of the thinking task are unrelated to each other in the dual-task experiments, whereas the semantic contents of thinking and those of verbal input and output are usually closely related in daily verbal communication. However, this difference does not mean non-occurrence of FoLSE in daily verbal communication: When verbal input or output is similar to thinking in semantic contents, this similarity stays the same whether native or foreign language is used in the verbal input or output. Therefore, FoLSE will not be reduced or cancelled out by this similarity in semantic contents. In fact, FoLSE was observed in the above two experiments although the semantic contents of the verbal input and those of thinking were closely related in their thinking tasks.

In a different experimental paradigm, moreover, FoLSE was also shown to occur when thinking was closely related to verbal input. Morishima (2013) measured reading time of successively presented English sentences to see if inconsistency between two sentences could be detected. Although native speakers of English detected the inconsistency when the two sentences were separated by an intervening sentence, non-native speakers detected it only when the two sentences were adjacent to each other with no intervening sentence. This means that thinking directly related to verbal input (i.e., detection of inconsistency between input sentences) was harder when the verbal input was given in the foreign language. It is thus strongly suggested that FoLSE could occur when thinking is closely related to verbal input as well.

Meanwhile, it was recently found that decision biases were mitigated when foreign language was used (e.g., Costa et al., 2014; Costa et al., 2017; Keysar et al., 2012): For instance, the framing effect (Kahneman & Tversky, 1979) disappeared when the problem was presented in foreign language (Keysar et al., 2012). Keysar et al. (2012) named this phenomenon *foreign language effect*. At least four explanations have been proposed for this phenomenon (see Costa et al., 2017; Geipel et al., 2015; Keysar et al., 2012): For instance, foreign language may trigger a deliberate mode of thinking, which replaces an ordinary intuitive mode of thinking that leads to decision biases.

At first glance, the foreign language effect and FoLSE may appear to be incompatible because the foreign language effect is positive, whereas FoLSE is negative. Does foreign language promote thinking or hinder it? In actuality, however, they are fully compatible because they need different conditions for occurrence. The foreign language effect needs at least two conditions: First, an ordinary intuitive mode of thinking has to result in a bias. When this mode of thinking gives a correct answer as in simple addition, which was used in some of the dual-task experiments (Takano & Noda, 1993, 1995; Tyler, 2001), the debiasing effect (i.e., the foreign language effect) is not expected. Second, more importantly, the foreign language effect was observed when thinking could be performed after reading a problem that was presented in foreign language. In this situation, the assumed deliberate mode of thinking was not interfered by linguistic processing of foreign language, and thus could lead to a normative answer. When thinking has to be conducted while foreign language is being processed, in contrast, thinking is interfered more strongly by linguistic processing of foreign language than that of native language. This results in declined performance of thinking (i.e., FoLSE), even if it is performed in a deliberate mode.

A study that may appear to contradict this latter argument is Díaz-Lago and Matute (2019) who demonstrated that foreign language reduced the causality bias (i.e., recognition of a causal relation that does not exist objectively: Matute et al., 2015; Vadillo et al., 2016). Their participants rated causality after learning contingency between a fictitious drug and a fictitious disease for 40 successively presented fictitious patients. In this experiment, thinking and linguistic processing were performed concurrently because participants had to learn contingency (i.e., thinking) while reading verbally presented information (i.e., linguistic processing). However, the linguistic processing required

in this experiment was minimal because three simple sentences were repeated 40 times (e.g., "The patient overcame the crisis": see their Figure 1) with slight changes as to whether the drug was taken and whether the patient recovered. When linguistic processing of foreign language is easy, it is not expected to interfere with concurrent thinking. Although thinking was performed in parallel with foreign language processing, therefore, it is not surprising that FoLSE was not observed in their experiment.

In many instances of daily verbal communication such as conversation, academic discussion, and business or diplomatic negotiation, thinking has to be performed in parallel with substantial linguistic processing of foreign language. In these settings, FoLSE is expected to be prevalent. It is true that FoLSE will not occur in a routine daily conversation such as "How are you doing?" "I'm doing good," because no interference with concurrent thinking is expected theoretically when linguistic processing of foreign language is not demanding enough. When it is demanding enough, however, the currently available empirical evidence strongly suggests that thinking ability will be certainly lowered in daily verbal communication as well.

FoLSE has a variety of important social implications in real life. One of the implications is underestimation of a foreigner's intelligence. Suppose, for example, that we have a foreign student whose native language we do not understand. We are routinely engaged in intuitive assessment of intelligence of other people in daily life (Sternberg et al., 1981). When we communicate with that foreign student in our own native language, we may try to ignore his/her anomalous verbal expressions (e.g., childish wordings, grammatical errors, and so on), and try to base our intuitive assessment of intelligence on the contents of what he/she says, instead of how he/she says, because most of us have some direct experience about foreign language difficulty. Only when the contents of his/her speech reveal some deficits (e.g., erroneous logical reasoning, internal contradiction, lack of original ideas, and so on), might we feel justified in inferring low intelligence. Although this strategy appears to be reasonable, it may be grossly misleading in actuality because it takes into account only the direct effects of foreign language difficulty. In reality, the contents of his/her speech may be degraded due to FoLSE. When the foreign student is using his/her own native language, his/her speech may reveal much less deficits in contents and he/she may look much more intelligent. Even academic estimation of intelligence could be biased by FoLSE: It was believed, for example, that Asian students are inferior to Western students in critical thinking although Asian students were tested in English, a foreign language for them (see Floyd, 2011; Lun et al., 2010; Manalo & Sheppard, 2016). The underestimation of a foreigner's intelligence may well produce a variety of undesirable consequences (e.g., prejudice on the part of a native speaker and antipathy on the part of a foreigner). If we are alert to FoLSE, we may be able to refrain from making hasty underestimation of a foreigner's intelligence.

Acknowledgment

This research was supported in part by a grant from Grant-in-Aid for Scientific Research, Ministry of Education, Science, and Culture, 15300085, and a grant from Murata Science Foundation. Portions of this research were presented at the 67th annual convention of the Japanese Psychological Association, September 2003, Tokyo; and the 12th International Conference on Language and Social Psychology, June 2010, Brisbane. The authors are very grateful to Ko-ichi Kishimoto, Yoko Hirai, and Kaneko-Shobo.

References

Allport, D. A. (1980). Attention and performance. In G. Claxton (Ed.), *New directions in cognitive psychology* (pp. 112–153). Routledge & Kegan Paul.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, *16* (4), 240–250. https://doi.org/10.1016/j.tics.2012.03.001

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, *3* (2), 193–209. https://doi.org/10.1016/0010-0285(72)90003-5

Broadbent, D. (1958). *Perception and communication*. Pergamon.

Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, *22* (5), 349–368. https://doi.org/10.1037/h0082775

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *Plos One*, *9* (4), e94842. https://doi.org/10.1371/journal.pone.0094842

Costa, A., Vives, M. L., & Corey, J. D. (2017). On language processing shaping decision making. *Current Directions in Psychological Science*, *26* (2), 146–151. https://doi.org/10.1177/096372141668026

Díaz-Lago, M., & Matute, H. (2019). Thinking in a foreign language reduces the causality bias. *Quarterly Journal of Experimental Psychology*, *72* (1), 41–51. https://doi.org/10.1177/1747021818755326

Eagan, D. E., & Chein, J. M. (2012). Overlap of phonetic features as a determinant of between-stream phonological similarity effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38* (2), 473–

481. https://doi.org/10.1037/a0025368

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Educational Testing Service.

Eysenck, H. J. (1979). *The structure and measurement of intelligence*. Springer-Verlag. https://doi.org/10.1007/BF00365371

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics*, *8* (2), 162–177. https://doi.org/10.1093/applin/8.2.162

Flege, J. E. (1991). The interlingual identification of Spanish and English vowels: Orthographic evidence. *Quarterly Journal of Experimental Psychology*, *43* (3), 701–731. https://doi.org/10.1080/14640749108400993

Floyd, C. B. (2011). Critical thinking in a second language. *Higher Education Research and Development*, *30* (3), 289–302. https://doi.org/10.1080/07294360.2010.501076

Geipel, J., Hadjichristidis, C., & Surian, L. (2015). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology, 59*, 8–17. https://doi.org/10.1016/j.jesp.2015.02.001

Hakuta, K., & Diaz, R. M. (1985). The relationship between degree of bilingualism and cognitive ability: A critical discussion and some new longitudinal data. In K. E. Nelson (Ed.), *Children's language* (Vol. 5) (pp. 319–344). Erlbaum.

Hirst, W., Spelke, E. S., Reaves, C. C., Caharack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. *Journal of Experimental Psychology: General*, *109* (1), 98–117. https://doi.org/10.1037/0096-3445.109.1.98

Johnson, M. K., Bransford, J. D., & Solomon, S. K. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology*, *98* (1), 203–220. https://doi.org/10.1037/h0034290

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99* (1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47* (2), 263–291. https://doi.org/10.2307/1914185

Keysar, B., Hayakawa, S. L., & An, S. G. (2012). The foreign-language effect: Thinking in a foreign tongue reduces decision biases. *Psychological Science*, *23* (6), 661–668. https://doi.org/10.1177/0956797611432178

LaBerge, D. (1981). Automatic information processing: A review. In J. Long, & A. D. Baddeley (Eds.), *Attention and performance, IX* (pp. 173–186). Erlbaum.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133* (3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lun, V. M. C., Fischer, R., & Ward, C. (2010). Exploring cultural differences in critical thinking: Is it about my thinking style or the language I speak? *Learning and Individual Differences*, *20* (6), 604–616. https://doi.org/10.1016/j.lindif.2010.07.001

Manalo, E., & Sheppard, C. (2016). How might language affect critical thinking performance? *Thinking Skills and Creativity*, *21*, 41–49. https://doi.org/10.1016/j.tsc.2016.05.005

Manalo, E., & Uesaka, Y. (2012). Elucidating the mechanism of spontaneous diagram use in explanations: How cognitive processing of text and diagrammatic representations is influenced by individual and task-related factors. In P. Cox, B. Plimmer, & P. Rodgers (Eds.), *Diagrammatic representation and inference*, Vol. 7352 of *Lecture Notes in Computer Science* (pp. 35–50). Springer. https://doi.org/10.1007/978-3-642-31223-6_9

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition*, *110* (1), 23–38. https://doi.org/10.1016/j.cognition.2008.08.003

Matute, H., Yarritu, I., & Vadillo, M. A. (2011). Illusions of causality at the heart of pseudoscience. *British Journal of Psychology*, *102* (3), 392–405. https://doi.org/10.1348/000712610X532210

McKoon, G., & Ratcliff, R. (1981). The comprehension processes and memory structures involved in instrumental inference. *Journal of Verbal Learning and Verbal Behavior*, *20* (6), 671–682. https://doi.org/10.1016/S0022-5371(81)90238-3

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99* (3), 440–466. https://doi.org/10.1037/0033-295X.99.3.440

Morishima, Y. (2013). Allocation of limited cognitive resources during text comprehension in a second language. *Discourse Processes*, *50* (8), 577–597. https://doi.org/10.1080/0163853X.2013.846964

Navarra, J., Sebastián-Gallés, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: New evidence from an implicit measure. *Journal of Experimental Psychology: Human Perception and Performance*, *31* (5), 912–918. https://doi.org/10.1037/0096-1523.31.5.912

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86* (3), 214–255. https://doi.org/10.1037/0033-295X.86.3.214

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7* (1), 44–64. https://doi.org/10.1016/0010-0285(75)90004-3

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524537

Okuno, S. (1969a). A factorial study on the differentiation of intelligence (I). *Bulletin of the Faculty of Education, Yamanashi University*, *4*, 164–194.

Okuno, S. (1969b). A factorial study on the differentiation of intelligence (II). *Memories of the Faculty of Liberal Arts and Education, Yamanashi University, 20*, 148–155.

Osaka, R., & Okuno, S. (1956). Some analytical studies of mental abilities. *Kyoto University Research Studies in Education, 2*, 177–194.

Osaka, R., & Umemoto, T. (1984). *Kyoto University new NX15 intelligence test* (2nd ed.). Taisei.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116* (2), 220–

244. https://doi.org/10.1037/0033-2909.116.2.220

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, *76* (27), 1–23. https://doi.org/10.1037/h0093840

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, *102* (3), 533–566. https://doi.org/10.1037/0033-295X.102.3.533

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of Psycholinguistic Research, 21*, 301–316. https://doi.org/10.1007/bf01067515

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematic for the organization of turn-taking for conversation. *Language*, *50* (4), 696–735. https://doi.org/10.2307/412243

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman, & D. R. Davies (Eds.), *Varieties of attention* (pp. 1–27). Academic Press.

Selinker, L., Swain, M., & Dumas, G. (1975). The interlanguage hypothesis extended to children. *Language Learning*, *25* (1), 139–152. https://doi.org/10.1111/j.1467-1770.1975.tb00114.x

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125* (1), 4–27. https://doi.org/10.1037/0096-3445.125.1.4

Spelke, E. S., Hirst, W. C., & Neisser, U. (1976). Skills of divided attention. *Cognition*, *4* (3), 215–230. https://doi.org/10.1016/0010-0277(76)90018-4

Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, *41* (1), 37–55. https://doi.org/10.1037/0022-3514.41.1.37

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18* (6), 643–662. https://doi.org/10.1037/h0054651

Takano, Y., & Noda, A. (1993). A temporary decline of thinking ability during foreign language processing. *Journal of Cross-Cultural Psychology*, *24* (4), 445–462. https://doi.org/10.1177/0022022193244005

Takano, Y., & Noda, A. (1995). Interlanguage dissimilarity enhances the decline of thinking ability during foreign language processing. *Language Learning*, *45* (4), 657–681. https://doi.org/10.1111/j.1467-1770.1995.tb00457.x

Tanaka, K., & Sakakibara, K. (1949). *New Tanaka A intelligence test.* Kaneko-Shobo.

Treisman, A. M. (1964). Verbal cues, language, and meaning in selective attention. *American Journal of Psychology*, *77* (2), 206–219. https://doi.org/10.2307/1420127

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76* (3), 282–299. https://doi.org/10.1037/h0027242

Treisman, A. M., & Davies, A. (1973). Divided attention to ear and eyes. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 101–117). Academic Press. https://doi.org/10.1093/acprof:osobl/9780199734337.003.0005

Turner, M. L., & Eagle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28* (2), 127–154. http://dx.doi.org/10.1016/0749-596X(89)90040-5

Tyler, M. C. (2001). Resource consumption as a function of topic knowledge in nonnative and native comprehension. *Language Learning*, *51* (2), 257–280. https://doi.org/10.1111/1467-9922.00155

Underwood, G. (1974). Moray vs. the rest: The effects of extended practice on shadowing. *Quarterly Journal of Experimental Psychology*, *26* (3), 368–373. https://doi.org/10.1080/14640747408400426

Vadillo, M. A., Blanco, F., Yarritu, I., & Matute, H. (2016). Single- and dual-process models of biased contingency detection. *Experimental Psychology, 63*, 3–19. https://doi.org/10.1027/1618-3169/a000309

Vasilev, M. R., Liversedge, S. P., Rowan, D., Kirkby, J. A., & Angele, B. (2019). Reading is disrupted by intelligible background speech: Evidence from eye-tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *45* (11), 1484–1512. https://doi.org/10.1037/xhp0000680

Wetherick, N., & Gilhooly, K. (1990). Syllogistic reasoning: Effects of premise order. In K. Gilhooly, M. T. G. Keane, R. Logie, & G. Erdos (Eds.), *Lines of thinking: Reflections on the psychology of thought* (Vol. 1), *Representation, reasoning, analogy and decision making* (pp. 99–108). Wiley.

Yoda, A., Sawada, K., Miki, Y., & Hidano, T. (1959). *University of Tokyo A-S intelligence test*. Tokyo-Shinri.

Yohtaro Takano（Member）

Professor emeritus at University of Tokyo. Ph.D. at Cornell University in 1985. Lecturer at University of Virginia, associate professor at Waseda University, and professor at University of Tokyo. Now at Institute for Service Innovation Studies, Meiji University. He has been studying the foreign language side effect, linguistic relativity, mirror reversal, mental rotation, and individualism/collectivism among others.

Takashi Yagyu

Associate professor at School of Child Psychology, Tokyo Future University. He is a member of the World Association for Infant Mental Health and the International Association of Early Childhood Education. He is interested in the environmental composition that may affect formation of motivation for learning or mental and physical development of infants and children. He completed coursework of the doctoral program in psychology at Graduate School of Humanities and Sociology, University of Tokyo in 2005.