Journal of Pan-Pacific Association of Applied Linguistics, 24(1), 107-131.
https:// doi.org/10.25256/PAAL.24.1.6

The Effects of Speaking Activities on Brain Blood Flow: An NIRS Study*

Junko Negishi, Hiroyuki Kanzaki, Yuko Yamada, Masumi Murakami, Erika Ozawa and Yoshiyuki Nakamura**

Negishi, J., Kanzaki, H., Yamada, Y., Murakami, M., Ozawa, E., & Nakamura, Y. (2020). The effects of speaking activities on brain blood flow: An NIRS study. *Journal of Pan-Pacific Association of Applied Linguistics, 24*(1), 107-131.

This study explored appropriate speaking activities for speakers of different proficiency levels by measuring the changes in their brain blood flow with near-infrared spectroscopy (NIRS). Fifty adult speakers and learners of English of various proficiency levels participated in the study by undertaking four types of task: a single-speaker picture description, interview, and two paired interaction activities on an easy and a difficult topic. By objectively analyzing the collected data, the authors revealed which speaking activity facilitated brain activity, which speaking activity was suitable for which level of learner, and from which area of the brain it was possible to obtain activation data. In short, advanced language learners need to tackle more difficult tasks, while easy, patterned tasks such as the picture description and interview are more appropriate for elementary level learners. As for intermediate learners, their language learning is likely to be promoted through most of the activity types, except for the interactions on difficult topics, which were shown to decrease their brain blood flow.

Keywords: speaking, interaction, brain science, NIRS

---

* We wish to acknowledge our deep appreciation to Dr. Yoshiyuki Nomura for statistical processing and Professor Hiroshi Tomonari for his kind advice.

** First/Corresponding Author: **Junko Negishi**, Associate Professor, Department of English and American Language and Literature, School of Literature, Tsurumi University.

Co-Authors: **Hiroyuki Kanzaki** Associate Professor, Department of Orthodontics, School of Dental Medicine, Tsurumi University; **Yuko Yamada**, Research associate, Department of Orthodontics, School of Dental Medicine, Tsurumi University; **Masumi Murakami**, Clinical Fellow, Department of Orthodontics, School of Dental Medicine, Tsurumi University; **Erika Ozawa**, Clinical Fellow, Department of Orthodontics, School of Dental Medicine, Tsurumi University; **Yoshiyuki Nakamura**, Professor Emeritus, Department of Orthodontics, School of Dental Medicine, Tsurumi University

1 Introduction

In Japan, the new Courses of Study have been sequentially implemented since April of 2020. In the foreign language unit, "speaking," which is one of the traditional four skills, has been divided into two sub-categories, namely, "(spoken) production" and (spoken) interaction," following the Common European Framework of Reference (CEFR) classification (Council of Europe, 2001). Interaction is a communication activity between multiple speakers and requires language learners to not only speak but also understand what the interlocutors say.

Recently, Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT) administered a status survey and test of the four key English skills to third-year junior high school (JHS Ss) and high school (HS Ss) students (MEXT, 2017a, 2017b). Results demonstrated that 100% of the 20,000 JHS Ss and 87.2% of the 10,000 HS Ss surveyed had reached only the A1 level of the CEFR (7.1% and 18.8% received zero scores, respectively). The results were not nearly as good as MEXT had expected and revealed that HS Ss had much less experience with speaking activities than JH Ss: 63.3% of the JHS Ss answered that "they participated in speeches and presentations in English," compared with only 36.9% of the HS Ss. In response to the results, MEXT has called for continued efforts from educational institutions to practice productive language activities, including speaking, in order to cultivate productive and interactive skills.

The demand for improving English speaking skills has been emphasized since the government began to promote "the action plan to cultivate 'Japanese people who can use English'" (MEXT, 2003). However, research on speaking, specifically on interaction, is scarce compared with research on other language skills. Second language acquisition (SLA) theory has been advocated from learners and teachers' subjective, cognitive rules of thumb (Oishi, 2006); nevertheless, various kinds of equipment to measure brain actions have been developed since 1990s, which has enabled researchers to conduct SLA research from an objective, brain science perspective. In the present study, with the goal of generating more objective data, an experiment was carried out using near-infrared spectroscopy (NIRS), which is capable of measuring the hemoglobin concentration change in a brain. The study explores the effects of different speaking activities and speakers' proficiency level on brain blood flow, and it proposes appropriate speaking activities for different levels of L2 English proficiency.

2 Literature Review

2.1 Brain science and NIRS

The brain, or cerebrum, controls our mind and behavior. It consists of four lobes: the frontal lobe, temporal lobe, parietal lobe, and occipital lobe. The frontal lobe, which plays the most advanced role, is considered to be central to language processing, thought, cognitive functions, and short-term and working memory. One of the important speech areas, Broca's area, is located in the frontal lobe. The temporal lobe is associated with language comprehension, visual memories, and emotion, and houses the hippocampus in its interior side. Wernicke's area, another area related to speech, is located in the posterior part of the temporal lobe. The parietal lobe is also responsible for processing language. Lastly, the occipital lobe processes visual information. Recent studies have indicated that language processing is an interworking of various parts of the brain over neural networks (c.f. Yokoyama, 2010), although this study focuses on the prefrontal cortex, as described below.

As aforementioned, the development of various types of medical devices has provided exponential advancement in the functional analyses of the brain and has made it possible to objectively study brain functions. NIRS is non-invasive (c.f. Yamashita, Maki, Watanabe, & Koizumi, 1998) and portable, which enables researchers to choose a desired location for measurement and does not restrict movement of the subject. As such, NIRS is considered to be a suitable device for measuring the activated level of the brain under normal conditions. Based on the theory that brain activity increases local blood flow, NIRS devices measure and image changes in hemoglobin concentration caused by brain activity; as hemoglobin concentration increases in the local blood, which absorbs near-infrared light, the detected light decreases. The data obtained from NIRS are levels of oxygenated hemoglobin (Oxy-Hb), deoxygenated hemoglobin (Deoxy-Hb), and total hemoglobin (Total-Hb= Oxy-Hb+Deoxy-Hb).

There are various types of NIRS, including some with many channels (e.g., 56 channels) for the entire brain and two channels for the forehead. For the present study, a smaller sized 16-channel NIRS was employed. In terms of the reliability of NIRS, Zhang et al. (2011) reported that fNIRS produced highly reliable results, using the cluster as the minimal analytical unit. Schecklmann, Ehlis, Plichta, and Fallgatter (2008) also reported that group and cluster analyses have sufficient test-retest reliability.

2.2 Second language acquisition from a brain science viewpoint

Research using NIRS has mainly been conducted in the fields of engineering and the medical sciences. The former has focused on the technology's function and applied it in real-life circumstances, while the latter has studied NIRS' use for medical purposes. However, NIRS research related to language and brain science is scarce. In the past, studies concerned with language acquisition have generally obtained learners' responses using a cognitive approach (c.f. Ellis, 1995; Krashen, 1977). SLA research from a brain science perspective has mainly focused on specifying the brain functions used when a certain language stimulus was given to a subject and identifying the activated area of the brain.

Studies measuring activated brain conditions when "producing a language" have been conducted by Kubota et al. (2005), using English as the first language, and by Hatta et al. (2009), using Japanese as the first language. These studies set meaning category tasks in which subjects were given a cue and uttered a word that emerged in their mind while researchers measured blood concentration changes in the frontal lobe. Thereafter, the changes caused by different tasks were identified by analyzing the participants' blood flow. Kikuchi and Taura (2011) conducted research to investigate the critical period hypothesis on what they called, "speaking tasks accompanied by cognitive activity," in which they administered bilingual stroop tests (color and word meaning mismatches, such as where the word "RED" is written in blue, but the subject must say "blue" while looking at the word "RED"). The results showed differences in blood flow based on when the subjects started learning English (before/after six years old). These studies merely focused on utterances of a recalled word, and no research has been conducted in which Japanese learners of English speak or interact in English for a certain length of time while their brain blood flow is measured using NIRS.

Does NIRS only detect brain activity when a stimulus is given? Is it considered impossible to measure the condition of the brain where multiple speakers undertake a task on a continuous basis? There have been some relevant studies in the field of engineering and the medical sciences; for example, Funane et al. (2011) used NIRS to observe multiple subjects undertaking a collaborative task from an engineering perspective. In the medical sciences, Shindo, Kitamura, Tachibana, and Someya (2013) observed two participants interacting in Japanese for a certain time period to determine whether an interlocutor's gender caused social anxiety and, consequently, increased/decreased brain blood flow. In the experiment, the two speakers first vocalized five Japanese vowels "a, i, u, e, o" in rotation to determine a baseline blood flow, and, thereafter, they had a conversation in Japanese on a provided topic, the Great East Japan Earthquake, for 240 seconds. During the first 120 seconds, participant A interviewed participant B, and during the second half, they took turns. The results indicated that brain activity in the frontal lobe during the conversation was negatively correlated with an indicator of cooperative behavior and that there was a gender difference in anxiety levels. These studies suggest the feasibility of utilizing NIRS for measuring brain blood flow while multiple speakers interact in their second language.

2.3 Brain blood flow and English proficiency level

According to Sakai (2009), the brain does not actively work when a learner begins to study a second language. Along with acquisition progress, the learner consciously learns the language by exploiting knowledge, and brain activity proportionately increases. However, further acquisition automates the language, and brain activity becomes "economical" (Figure 1).

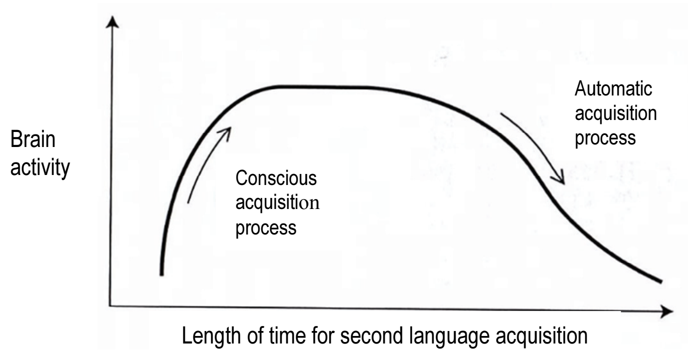


Figure 1. Second language acquisition process and brain activity (Sakai, 2009, p. 142)

Oishi (2006), who took a practical approach to SLA and brain science, revealed differences in brain activity when observing Japanese learners of English. She undertook various experiments using NIRS with learners of different English proficiency levels to investigate their receptive skills (i.e., reading and listening). In accordance with learners' proficiency levels, from novice to advanced, she noted four activation states:

(1) non-activated state (a state where the subject abandons the given task as it is too difficult)
(2) excessively activated state (a state where too much load is placed on the wider brain area for the subject to execute the given task)
(3) selectively activated state (a state where the necessary brain area is used to execute the given task)
(4) automated activated state (automated processing is executed as per native speakers, which causes a less/non-activated state)

(Oishi, 2006, p. 190)

These four states represent what Sakai (2009) advocates in Figure 1. Taura (2016a) reported that those with studying abroad experience participated in translation tasks and used less brain energy than those who had never studied abroad, based on measuring blood flow in the prefrontal lobe using fNIRS. The brain state of students who had studied abroad, thus, seems to be close to Oishi's automated state (4). These studies indicate that an automated activation state can occur not only in native speakers but also in second language learners.

What does an activated blood flow mean? Oishi (2006) assigned an English listening and reading task of varying difficulties to 16 university students. She observed the results of the comprehensive test and change in brain blood flow using NIRS. The results indicated that the elementary/intermediate and advanced level learner groups demonstrated a smaller amount of change in brain blood flow while completing easier tasks. Oishi mentioned that higher brain activation is required for difficult tasks. By contrast, when the degree of difficulty is lower, higher brain activation is not necessary and brain activation approaches an automated processing state. In other words, more difficult tasks increase a learner's brain blood flow.

The present study measured the blood flow change in the prefrontal lobe only, rather than in the whole brain. Is it possible to measure only the prefrontal lobe instead of all language areas located in other parts of the brain? Oishi (2006) mentioned that the activation level was lower in the prefrontal lobe than in the left side of the brain; however, she confirmed that the blood flow volume in the prefrontal lobe also increased significantly during listening and reading activities. Taura (2016b), meanwhile, compared two devices—a 28-channel fNIRS (OMM-3000) and portable 16-channel fNIRS (LIGHTNIRS)—and found a high correlation between the two devices' measure of Oxy-Hb change in Broca's area when subjects undertook language tasks, but a low correlation in Deoxy-Hb and Total-Hb in other areas. Similarly, Taura (2016c) compared two devices—a large 28-channel fNIRS (OMM-3000) and a simplified 2-channel fNIRS (Pocket NIRS); for the 2-channel NIRS, Taura placed one channel on the left forehead and the other on the right forehead. Taura administered bilingual stroop tests to eight subjects and revealed that the two devices were found to be correlated. Although the two experiments did not use speaking tasks, it appears possible to obtain related outcomes by measuring activation levels in the prefrontal lobe. Oishi's (2006) studies focused on receptive tasks such as listening and reading instead of productive tasks, including speaking; conversely, the present study attempted to identify to which proficiency level learners executed tasks with activated blood flow when they engaged in one of the productive activities (i.e., speaking).

2.4 The three types of speaking activities

Examples of speaking activities that are often used for second language learners and examinations are 1) single-speaker talk, 2) interviews, and 3) multiple speaker interaction. These are listed in Table 1, which also highlights their advantages and disadvantages when a teacher is evaluating learners.

First, when a single speaker speaks, such as in a picture description or storytelling task (monologue), the advantage is that there is no influence from the interlocutor(s). Furthermore, teachers can use specific sentence patterns or vocabulary intentionally to control what learners do or say. The learners can practice this type of speaking activity on their own. However, such tasks are unnatural as there is no interaction or "negotiation of meaning"[1] that usually occurs in real interactions.

Table 1. Advantages and Disadvantages of the Three Types of Speaking Activities

Type	Advantages	Disadvantages	Research
Single speaker talk	No influence from interlocutor(s). Can use specific sentence patterns and vocabulary.	No negotiation of meaning that is commonly seen in real interactions.	c.f. Young, 2000
Interview	More realistic than monologue. Can use specific sentence patterns and vocabulary.	Only answer the questions. Difficult to use negotiation of meaning. Asymmetric power balance.	c.f. Johnson & Tyler, 1998
Interaction	More complicated and co-constructed conversation. More negotiation of meaning, which facilitates second language acquisition.	Uncontrollable interlocutor effects such as personality and proficiency level, which may disturb other speaker(s).	c.f. Swain & Lapkin, 2001

Second, an interview is closer to real conversation as compared to a single speaker talking, and an interviewer or a teacher can control which sentence patterns or vocabulary should be used. However, interviewees tend to merely answer the questions posed by the interviewer; therefore, some researchers have criticized the interview's asymmetric power balance (c.f. Egyud & Glover, 2001; Taylor, 2001), which make it difficult for negotiation of meaning to occur.

Lastly, interactions between more than two speakers are considered to equalize the asymmetric power balance of the interview (c.f. Skehan, 2001; Taylor, 2001): Speakers become relatively relaxed in such interactions, as compared to interviews where they sometimes feel nervous or upset (Nakatsuhara, 2013). Luoma (2004) reported that when speakers are at equal positions, the atmosphere becomes different from that of the interview, and more realistic interaction occurs. Language learners can also display their interactional competence better in interactions than in interview scenarios

[1] Negotiation of meaning is considered to facilitate second language acquisition. Examples of negotiation of meaning include asking for clarification or confirmation, providing clarification, checking for comprehension, asking for or responding to help, correcting or suggesting words, and incorporating corrections or suggestions.

where they merely answer questions (Ducasse & Brown, 2009). Further, they can exercise their abilities while using a wide range of language functions (Skehan, 2001) and activate their conversation management ability (Galaczi, 2008). Taylor (2000), for example, analyzed speaking data obtained from interviews and paired oral interaction using the C2 Proficiency Speaking Test (formerly known as Cambridge English: Proficiency (CPE)) and reported that the paired oral tasks produced more language among the subjects than the interviews, although there was no statistically significant difference. In other words, interactions bring about more complicated and co-constructed conversation, which increases the negotiation of meaning and, accordingly, facilitates second language acquisition. Furthermore, students show positive responses to interactive activities (c.f. Egyud & Glover, 2001; Van Moere, 2006) because they can control the conversation themselves and use natural language (c.f. Fulcher, 1996). Nonetheless, there are disadvantages in such interactions, including uncontrollable interlocutor variables such as gender, age, intimacy, social status, introvert/extravert personalities, willingness to speak, and proficiency level, which are entwined with each other and may disturb other speakers (Brooks, 2009; Van Moere, 2006).

Thus, there are positive and negative aspects among these speaking activities; however, it is necessary to understand which speaking activities are suitable for which level of learners. Studies concerning different speaking types have mainly addressed the effects of interlocutors, assessment, and conversation analyses, and no studies have used NIRS to analyze speaking types from a brain science standpoint.

2.5 Study purpose

The purpose of this study was to answer the following three research questions by measuring the changes in brain blood flow using NIRS and objectively analyzing the collected data:

RQ1 What speaking activities activate brain blood flow in the following four tasks?

Task 1: Single speaker's picture description
Task 2: Interview (Q&A)
Task 3: Paired oral interaction on an easy topic
Task 4: Paired oral interaction on a difficult topic

RQ2 Among four groups of different English proficiency levels, which group's brain blood flow is activated while undertaking these tasks?

Group A: Native speakers of English or near native speaker proficiency level
Group B: Advanced proficiency level learners of English who have studied abroad
Group C: Intermediate proficiency level learners of English

Group D: Elementary proficiency level learners of English
RQ3 Which areas in the prefrontal lobe are activated during these speaking activities?

3 Research Method

3.1 Using the NIRS device to measure changes in brain blood flow

Changes in brain blood flow were measured by placing a 16-channel NIRS device (Spectratech OEG-16, Spectratech Inc.) on each participant's forehead. This device is able to measure the change of hemoglobin concentration in the prefrontal area while ensuring the participant is in close-to-natural conditions with little constraint; therefore, the author regarded NIRS as best suited for face-to-face speaking activities. Since the School of Dental Medicine at the author's university owned the device, the experiment was conducted in the room where the NIRS was located after the dental treatment hours. The measurement was carried out by research associate/clinical fellows of the School of Dental Medicine who had experience measuring their patients' NIRS data. They also processed the collected data from the OEG-16.

3.2 Participants

Table 2. Groups, Participants, and Tasks

Group	English Level	TOEIC®	Male	Female	Tasks 1,2,4
A	Native, Native-like	Native(-like)	5	2	Eiken Pre-1st
B	Advanced, Studied abroad	600–975	6	7	Eiken Pre-1st
C	Intermediate	400–595	8	6	Eiken 2nd
D	Elementary	< 395	8	8	Eiken 2nd

Fifty adult participants, who were divided into four groups according to their proficiency level, participated in the study. Table 2 shows the basic data of the participants. Seven participants in Group A were English teachers (including one former teacher), four were native English speakers, and three were near-native level Japanese speakers of English who had received higher education in English-speaking countries. In a pilot study, the changes in brain blood flow among the seven participants were rarely different; thus, they were classified as Group A. In Group B, there were 12 university students and one English teacher who had no higher education experience abroad. The TOEIC® scores of Group B ranged widely from 600 to 975. Almost everyone in Group B had participated in the university's study abroad program for more than eight months, and they spoke English fluently. Group C consisted of 14 university students with TOEIC® scores ranging from 400 to 595, who had no study abroad experience longer than three months. There were 16 university students in Group D whose proficiency levels were elementary. All participants except for the eight teachers were students of the same university and their English learning background was similar, with the exception of two advanced level students: one was born in the States and lived there for three years and the other student's second language was Cantonese. However, these two students' proficiency levels were similar to other students. The average age of the students was 21.3 years old. The ratio of male to female students was 27 to 23. Forty-eight participants were right-handed and two were left-handed, the left-handed students' data were compared with right-handed subjects; there was no difference between them, and they were rather average participants in terms of the data used for analysis. For this reason, they were included in the analysis. For participants without TOEIC® scores, their Eiken or IELTS® scores were converted into TOEIC® scores using a correspondence table.

The author explained the research project to each participant in advance, and they signed a written consent form to confirm their agreement to participate in the experiment. The project was approved by the Ethical Review Board at the authors' university (#2017-1).

3.3 Measurement procedure and tasks

The measurement was carried out in pairs. First, the OEG-16 (NIRS) was attached to participant A's forehead, who was then asked to keep their head still. The participants of Groups A and B took an Eiken Pre-1st level interview test, and those of Groups C and D took the same test type at the 2nd level. Because the participants' English speaking level varied widely, the author assigned easier tasks for Groups C and D, although Task 3 was identical for all participants. They did not practice the four tasks beforehand given that Sakai, Hashimoto, and Homae (2001) reported that more brain activation occurs in the case of a novel task, and the brain tends to be less activated when it has become accustomed to a specific task from repeated practice. The author used Eiken's old interview tests with Eiken's permission (#2019-0109), except for Task 3. These tasks are included in the Appendix.

The design of the measurement and tasks undertaken is shown in Figure 2.

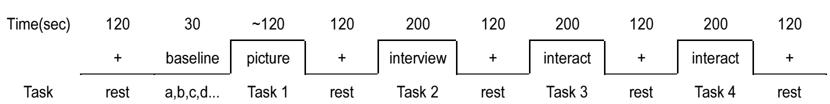


Figure 2. Block design of the speaking tasks

1) After attaching the NIRS device to participant A's forehead, the research associate/clinical fellows confirmed that all channels were ready for measurement. Participant A was placed at rest for 120 seconds to allow the hemoglobin concentration to sufficiently decrease.
* The participant was placed at rest for about 120 seconds after each of the four tasks.
2) Participant A repeated "a, b, c, d, e, f" for 30 seconds. The level of hemoglobin concentration during this repetition was used as a benchmark.
3) **Task 1 (picture description)**: Following Eiken's protocol, a four-frame cartoon (Eiken's Pre-1st level interview test for Groups A and B) or a three-frame cartoon (Eiken's 2nd level interview test for Groups C and D) was provided to Participant A, who was allowed 60 seconds to think about the story while viewing the cartoon and 120 seconds thereafter to narrate the story. The participants could complete the task as fast as they liked, so many of them stopped narrating the story before 120 seconds had passed.
4) **Task 2 (interview)**: The first author whose first language is Japanese assumed the role of the interviewer and asked two questions in English. The first concerned the participant's preference about a trip, and the second question asked about the content of the cartoon in Task 1. The length of Task 2 was approximately 200 seconds and was followed by a 120-second rest.
* Participant B (an interlocutor of Participant A) entered the measurement room, and A and B were asked to undertake paired oral interactions.
5) **Task 3 (paired oral interaction on an easy topic)**: The randomly assigned topics were either school, family, or English, on which two speakers could have a conversation for 200 seconds, followed by a 120-second rest.
6) **Task 4 (paired oral interaction on a difficult topic)**: The topics were taken from other Eiken interview questions and were the same difficulty level as the picture description. Participants were given 200 seconds to talk about the topic, followed by a 120-second rest (see Appendix).
* Only one NIRS device was available for the experiment, and Participant A and B exchanged the device when A finished the four tasks. In order to take a counterbalance, half of the members of each group participated in the measurement in a [Task 3 → Task 4] → [Task 1 → Task 2] cycle.

4 Results and Discussion

4.1 Change in brain blood flow caused by different groups and tasks

In this study, the amount of change in the brain blood flow was shown via oxygenated hemoglobin (Oxy-Hb; mM) levels. There were two reasons for not using deoxygenated hemoglobin (Deoxy-Hb), which indicates a non-activated brain state. The first was that little difference was observed between the groups when analyzing Deoxy-Hb for each task. The second was based on Hoshi's (2005, p. 61) description that "Oxy-Hb change always follows the same direction with blood flow change. When measuring the brain blood flow, Oxy-Hb is the best index for the blood flow change except for infants." Oxy-Hb data were obtained for each task individually on 16 channels every 0.655 second. Results were calculated first for each participant, then for each of the 16 channels, and subsequently averaged for each group. Finally, the averaged value of the 16-channel data was used as the measurement value, following Taura's (2016a) procedure.

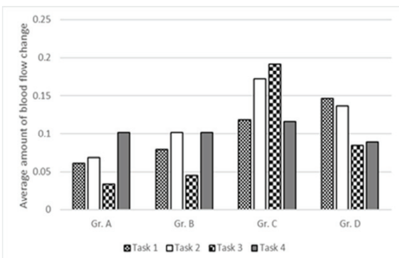


Figure 3. Averaged Oxy-Hb change displayed by task difference and sorted by group

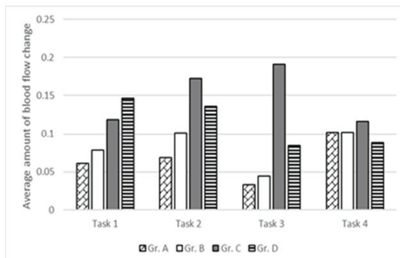


Figure 4. Averaged Oxy-Hb change displayed by group difference and sorted by task

Figure 3 shows the averaged Oxy-Hb change displayed by task difference and sorted by groups. Group A, consisting of the native and native-like proficiency level speakers, showed lower activated brain states in all tasks, with Task 3 the lowest and Task 4 the highest. Task 3 was an interaction on an easy topic, which means it was similar to a daily conversation for these participants, and, thus, their brains might have worked economically or conserved energy. By contrast, in Task 4, participants were required to describe their own thoughts on a given topic (e.g., "Should more be done to improve the healthcare system in Japan?"), which may have activated their brain blood flow. Group B, comprising advanced learners, displayed a similar tendency. In the intermediate Group C, the participants' blood flow was the most activated among the four groups, particularly in Task 3, indicating that participants undertook the task to their best ability. Task 4, on the other hand, resulted in the least activated blood flow for Group C, suggesting that the task was difficult for them. The author sometimes heard these participants muttering about the task's difficulty. As for elementary Group D, their activated brain states were seen in Tasks 1 and 2, since they were more accustomed to these kinds of tasks than interactions. Group D seemed to find Tasks 3 and 4 difficult, as did Group C, because they often stagnated in their own thoughts. This might be the closer state indicated by Oishi (2006) as the non-activated state.

Figure 4 shows the averaged Oxy-Hb change displayed by group difference and sorted by tasks. Task 1 clearly demonstrated the activation difference caused by the English proficiency level of each group; in other words, the lower the English proficiency, the more brain activation occurred. Task 1, the picture description task, is likely to be a good indicator of the speakers' English proficiency level. Although Task 2 appears similar to Task 1, the blood flow change of Group D did not increase proportionally, which may be because elementary-level learners regarded the interview as more difficult than other groups. In Task 3, the brain activation level of Groups A and B was much lower because the task may have been very easy for them, which resulted in them "saving energy," as observed by Sakai (2009). By contrast, activated blood flow was seen in Group C. When intermediate proficiency level learners undertake an interaction activity on an easy topic, the brain is most activated. As for Task 4, although the brain activation levels seemed to be similar among the four groups, we should not regard them as similar but rather should compare them with other tasks. For example, in Groups A and B, the activation levels in Task 4 were higher than Task 3, which means that the two groups' brains were working actively in the interaction activity on a difficult topic. As for Groups C and D, it is likely that they found Task 4 difficult for their blood flow to be activated.

4.2 Total amount of averaged Oxy-Hb change across the four tasks

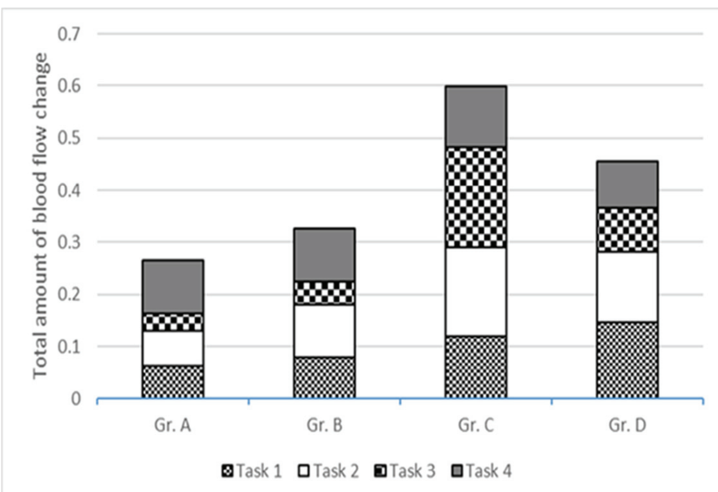


Figure 5. Total amount of averaged Oxy-Hb change across the four tasks

Figure 5 displays the total amount of averaged Oxy-Hb change across the four tasks. As per Oishi's (2006) report on reading, the reason for the lowest levels of Oxy-Hb found in Group A may be based on their brains working economically or saving energy since native-level speakers can deal with such tasks easily. Group B somewhat shared this phenomenon with Group A, although their brains seemed to work more actively in Tasks 1 and 2 than Group A. Group C demonstrated the highest Oxy-Hb change among the four groups, that is, the interview (Task 2) or the interaction on an easy topic (Task 3) were the speaking activities that might be most appropriate for intermediate level learners. The main difference between Groups C and D was shown between Tasks 3 and 4, while no significant difference was recorded between Tasks 1 and 2. As for Group D, an easy task such as Task 1 may lead to brain activation because they were likely not prepared to tackle the interaction activities.

4.3 Statistical analysis

Table 3. Results of Mixed Effects Modeling

	Coefficient	Std. error	*t*	95% confidence interval		*p*
				Lower bound	Upper bound	
Intercept	-1.536	0.021	-73.967	-1.577	-1.495	<0.001
Task-level explanatory variable						
Task 2/Task 1	0.142	0.016	8.684	0.110	0.174	<0.001
Task 3/Task 1	0.129	0.016	8.290	0.099	0.160	<0.001
Task 4/Task 1	0.061	0.015	3.978	0.031	0.092	<0.001
Participant-level explanatory variable						
Group B/Group A	0.066	0.017	3.839	0.032	0.100	<0.001
Group C/Group A	0.277	0.017	16.309	0.244	0.311	<0.001
Group D/Group A	0.106	0.017	6.257	0.073	0.139	<0.001
Time	0.001	<0.001	8.368	<0.001	0.001	<0.001

Note: 1) For task types, Task 1 data was used as a reference. For proficiency level, Group A data was used as a reference. 2) The results were obtained from the analysis of average brain blood flow by time.

The results of the multilevel modeling are summarized in Table 3. Since the measured blood flow was nested in the task and participants, mixed effects modeling was carried out to determine the effect of the participants' proficiency level and task types on brain activity, and the average amount of each participant's Oxy-Hb was used as the outcome variable. These analyses were carried out using IBM SPSS statistics v. 24.0 (IBM, Tokyo, Japan).

First, the explanatory variables were significantly associated with task types in a random intercept model. All coefficients of the task types were statistically significant ($p < 0.001$) when Task 1 (single speaker's talk) was used as a reference. The coefficient of Task 2 was the highest at 0.142, while that of Task 4 was the lowest at 0.061, that is, the task types were positively related to brain blood flow.

For participant level, the coefficients among the groups were 0.066 (Groups A and B; the lowest among the three), 0.277 (Groups A and C; the highest), and 0.106 (Groups A and D), respectively. The table indicates that the participants' proficiency levels significantly affected the brain blood flow when Group A was used as a reference ($p < 0.001$).

4.4 Relationship between the activated channels and total amount of averaged Oxy-Hb change

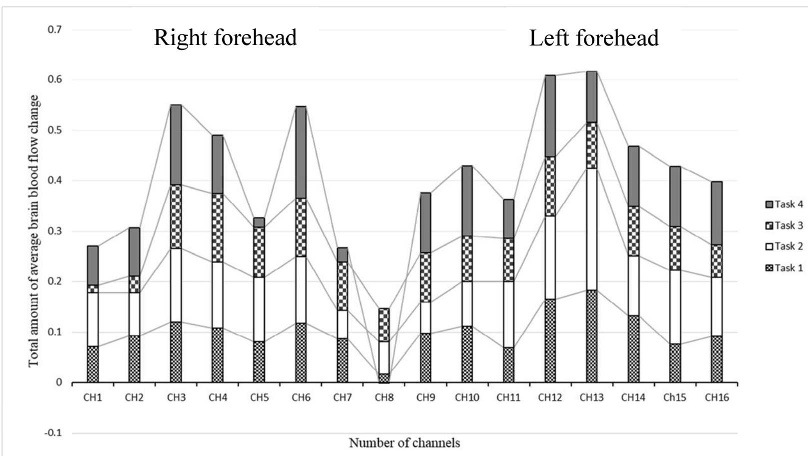
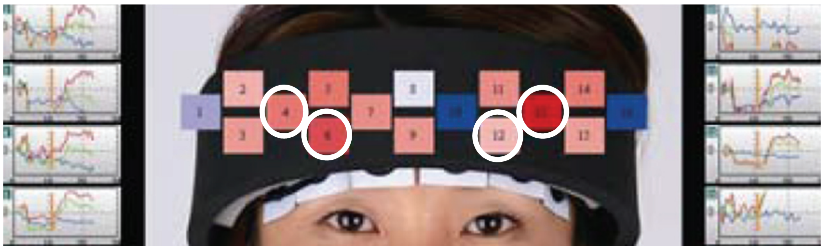


Figure 6. Relationship between the activated channels and total amount of averaged Oxy-Hb change

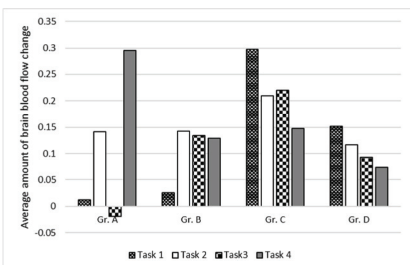
Figure 6 shows the relationship between the activated channels and total amount of average Oxy-Hb change. On the left side of the forehead, the most activated channels were 12 and 13, with Tasks 1 and 2 seeming to activate this area. On the right side, the Oxy-Hb change of channels 3, 4, and 6 was higher, and they were relatively activated by Task 4.



Right (4, 6) Left (12, 13)

Figure 7. NIRS channel layout on participants' foreheads

Figure 7 illustrates the channel layout when the NIRS device was attached to participants' foreheads. In the next two figures, channels 12 and 13 are used for the left side of the forehead, and channels 4 and 6 are used for the right.



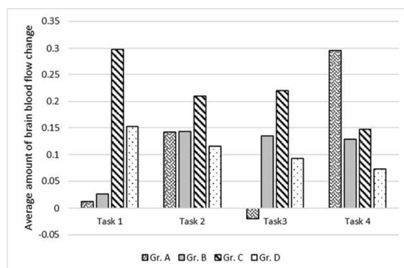


Figure 8. Averaged Oxy-Hb change sorted by group on channels 4, 6, 12, and 13

Figure 9. Averaged Oxy-Hb change sorted by task on channels 4, 6, 12, and 13

Figure 8 displays the averaged Oxy-Hb change of the activated channels 4, 6, 12, and 13. Every bar chart demonstrates a larger amount of average Oxy-Hb when compared with Figure 3. Task selection by difficulty level may be possible by checking the specific channels rather than looking at all channels; for instance, Tasks 1, 2, and 3 activated the learners' brain blood flow in Group C, whereas Tasks 3 and 4 did not increase Oxy-Hb in Group D. Consequently, learners such as those in Group D should tackle tasks such as Task 1 or 2.

Figure 9 demonstrates the averaged Oxy-Hb change of the activated channels 4, 6, 12, and 13. Likewise, the amount of Oxy-Hb change is larger when compared with Figure 4. In Tasks 1, 2, and 3, Group C demonstrated the highest blood flow, whereas Group A's brains were most activated in Task 4. The graphs indicate that it may be possible to choose appropriate tasks for each student by checking specific groups of channel. Recently, a 2-channel NIRS has become available at a lower price, which would make it easier to select more suitable tasks by attaching this simplified type of device.

5 Conclusion

In this study, 50 adult speakers of English of various proficiency levels participated in three types (four tasks) of English-speaking activities while their brain blood flow (and, thus, their brain activation level) was measured using NIRS. The tasks comprised a picture description task by a single speaker (Task 1), an interview (Task 2), and two paired oral interaction activities on an easy and a difficult topic (Tasks 3 and 4). By analyzing the obtained data, the authors observed which speaking activities facilitated increased brain activity, which speaking activity was suitable for which level of learner, and from which area of the brain the activation data was retrieved. The following presents the findings in the context of this study's three research questions.

5.1 RQ1: What speaking activities activate brain blood flow in the four tasks?

RQ1 is closely associated with RQ2, and the learners' proficiency level greatly affected the impact of the four tasks. In terms of Task 1 (a single-speaker story narration undertaken while the participant looked at a three/four-panel cartoon), the lower the participants' proficiency level, the greater their brain activation. In other words, English proficiency and the amount of Oxy-Hb change showed an inverse relationship. Task 1 could, therefore, be used to determine English learners' proficiency level. Task 2, where the author interviewed the participants, demonstrated similar results to Task 1; however, the elementary level Group D seemed to find the interview task somewhat difficult. Task 3, where the participants interacted on easy topics such as school, family, and English, activated the brains of the intermediate learners in Group C. By contrast, Task 3 did not cause an activated brain state in the proficient speakers of Groups A and B. For advanced level learners of Group B who still need to practice speaking English, Task 3 (the interaction on an easy topic or daily conversation) might not cause sufficient brain activation. By contrast, the brain blood flow increased less among the elementary level learners in Group D, showing that they were not ready to maintain interaction with each other by employing their listening and speaking skills simultaneously. Task 4, the paired oral interaction on a difficult topic, increased the amount of Oxy-Hb the most in Groups A and B. By contrast, the intermediate speakers of Group C displayed the least activated blood flow in Task 4 among the four tasks: For this proficiency level learners, Task 4 resulted in brain deactivation. Both Tasks 3 and 4 were likely to be too difficult for the elementary speakers in Group D.

5.2 RQ 2: Among four groups of different English proficiency levels, which group's brain blood flow is activated while undertaking the four tasks?

As described above, a close relationship was found between the tasks and English proficiency. For the advanced learners of Group B, the interview (Task 2) and the paired oral interaction task on a difficult topic (Task 4) seemed to be suitable as these tasks activated their brains. Task 3, by contrast, did not result in an increase of Oxy-Hb. An increase in brain blood flow was seen in Group C during all tasks. Various speaking activities seemed to be effective for the intermediate learners, especially the interviews and the paired oral interaction on an easy topic (Tasks 2 and 3). It would be most effective to incorporate various speaking activities for intermediate proficiency level students. In terms of the elementary level group, they were not ready to undertake flexible interactive conversation; thus, teachers should assign patterned speaking activities for introductory practice.

However, we do not suggest that teachers stop assigning interaction activities to elementary level learners because interactions using negotiation of meaning facilitate second language acquisition, it is also important for learners to become accustomed to interacting with each other in English.

5.3 RQ3: Which areas in the prefrontal lobe are activated during the speaking activities?

Kanzaki et al. (2019) reported that there is a difference in activation between the left and right frontal lobe during different tasks. The present study also revealed that certain channels were activated and differences were seen between the left and right forehead: The left forehead showed an activation tendency for Tasks 1 and 2, whereas the right forehead showed a similar tendency for Task 4. Although it is not practical to measure brain blood flow of the entire brain for learners on a normal basis, a simplified NIRS, which can measure activated forehead blood flow, might provides learners the opportunity to select speaking activity tasks.

A newspaper reported that a project was implemented at a private high school in conjunction with Tokyo University (Yomiuri Shimbun, 2019). The high school students wear simplified NIRS on their forehead while studying social science, mathematics, and English. The project intends to collect data for four years. The results of the project may provide more practical suggestions in the future.

5.4 Study limitations and scope for future research

One limitation of this study was that it was difficult to determine whether the blood flow activation was caused by task difficulty or cognitive load or both. Further research will be necessary to identify which cause applied and to what extent. A second issue was that there was a difference in each participant's measurement data. It is possible to observe tendencies by using the averaged data, but not every participant displayed the same results; therefore, individual changes and features should be analyzed in future research. Furthermore, the NIRS used in this study did not analyze the entire brain but only the forehead, which resulted in insufficient data because we could not measure the speech areas directly. In future studies, correlations between the forehead and other areas of the brain, and between the 16-channel and simplified 2-channel data, should be compared.

References

Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing, 26*(3), 341-366. doi: 10.1177/0265532209104666

Council of Europe. (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment.* Cambridge: Cambridge University Press.

Ducasse, A. M., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. *Language Testing, 26*(3), 423-443. doi: 10.1177/0265532209104669

Egyud, G., & Glover, P. (2001). Oral testing in pairs: A secondary school perspective. *ELT Journal, 55*(1), 70-76. doi: 10.1093/elt/55.1.70

Elllis, R. (1995). Interpretation tasks for grammar teaching. *TESOL Quarterly, 29*(1), 87-105.

Fulcher, G. (1996). Testing tasks: Issues in task design and the group oral. *Language Testing, 13*(1), 23-51. doi: 10.1177/026553229601300103

Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., & Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *Journal of Biomedical Optics, 16*(7), 077011-1-9. doi: 10.1117/1. 3602853

Galaczi, E. D. (2008). Peer-peer interaction in a speaking test: The case of the First Certificate in English examination. *Language Assessment Quarterly, 5*(2), 89-119. doi: 10.1080/15434300801934702

Hatta, T., Kanari, A., Mase, M., Nagano, Y., Shirataki, T., & Hibino, S. (2009). Strategy effects on word searching in Japanese letter fluency tests: Evidence from the NIRS findings. *Reading and Writing, 22*(9), 1041-1052. doi: 10.1007/s11145-008-9143-9

Hoshi, S. (2005). Kinoteki kinsekigai bunko-ho: Genkai to kanosei [Functional near-infrared spectroscopy: Limitation and potentiality]. *The Journal of the Japanese College of Angiology, 45*(2), 61-67.

Johnson, M., & Tyler, A. E. (1998). Re-analyzing the OPI: How much does it look like natural conversation? In R. Young, & A. W. He (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency* (pp. 27-51). Philadelphia, PA: John Benjamins.

Kanzaki, H., Wada, S., M. Kumazawa, Yamada, Y., Sudo, T., Ozawa E., Seko, T., Akaike, S., Murakami, M., Oikawa, T., Okumura, S., Nakamura Y., & Tomonari, H. (2019). Mandibular prognathism attenuates brain blood flow induced by chewing. *Scientific Reports, 9*(1), 1-8. Retrieved from https://www.nature.com/articles/s41598-019-55553-5.pdf

Kikuchi, Y., & Taura, H. (2011). Bairingaru sutorupu tesuto suikouji no no-fukatsu bui ni kansuru fNIRS no imejing kenkyu: daini gengo shutoku nenrei hikaku oudan kenkyu [Research on active area of a brain using fNIRS brain imaging over performance of bilingual stroop tests: Comparison and cross-sectional analysis of second language acquisition age]. *Studies in Language Science, 1*, 91-145.

Krashen, S. (1977). The monitor model for adult second language performance. In M. Burt, H. Dulay, & M. Finocchiaro (Eds.), *Viewpoints on English as a second language* (pp. 152-161). New York: Regents Publishing Company.

Kubota, Y., Toichi, M., Shimizu, M., Mason, A. R., Coconcea, C. M., Findling, R. L., Yamamoto, K., & Calabrese, J. R. (2005). Prefrontal activation during verbal fluency tests in schizophrenia: A near-infrared spectroscopy (NIRS) study. *Schizophrenia Research, 77*(1), 65-73. doi.org: 10.1016/j.schres.2005.01.007

Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.

MEXT (2003). "Eigo ga tsukaeru nihon jin" no ikusei no tameno kodo keikaku [The action plan to cultivate "Japanese people who can use English"]. Retrieved from http://www.mext.go.jp/b_menu/shingi/chousa/shotou/082/shiryo/attach/1301980.htm

MEXT (2017a). Heisei 29 nendo "Eigo kyoiku jisshi jyokyo chosa" Kotogakko no kekka [Heisei 29, "Implementation status survey of English education" results of high schools]. Retrieved from http://www.mext.go.jp/coponent/a_menu/education/detail/__icsFiles/afieldfile/2018/04/06/1403469_02.pdf

MEXT (2017b). Heisei 29 nendo "Eigo kyoiku jisshi jyokyo chosa" Chugakko no kekka [Heisei 29, "Implementation status survey of English education" results of junior high schools]. Retrieved from http://www.mext.go.jp/component/a_menu/education/detail/__icsFiles/afieldfile/2018/04/06/1403469_03.pdf

Nakatsuhara, F. (2013). *The co-construction of conversation in group oral tests*. Frankfurt am Main: Peter Lang.

Oishi, H. (2006). *Nokagaku karano daini gengo shutokuron [Second language acquisition theory from neuroscience]*. Kyoto: Showa-do.

Sakai, K. L. (2009). *No no gengo chizu [Language map of a brain]*. Tokyo: Tokyo Meiji Shoin.

Sakai, K. L., Hashimoto, R., & Homae, F. (2001). Sentence processing in the cerebral cortex. *Neuroscience Research, 39*(1), 1-10. doi: 10.1016/s0168-0102(00)00205-4

Schecklmann, M., Ehlis, A-C., Plichta, M. M., & Fallgatter, A. J. (2008). Functional near-infrared spectroscopy: A long-term reliable tool for measuring brain activity during verbal fluency. *NeuroImage, 43*(1), 147-155. Retrieved from https://pubmed.ncbi.nlm.nih.gov/18657624/

Shindo, M., Kitamura, H., Tachibana, A., & Someya, T. (2013, July). *Dual functional near-infrared spectroscopy (2): An application in face-to-face conversation.* Paper presented at the World Federation of the Societies of Biological Psychiatry, Kyoto, Japan.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (1st ed., pp. 167-185). London: Routledge.

Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (1st ed., pp. 99-118). London: Routledge.

Taura, H. (2016a). Bairingaru no imejingu kenkyu: Koremade no kenkyu seika [Brain imaging research on bilinguals: Current study outcomes]. *Studies in Language Science, 1*, 81-116.

Taura, H. (2016b). Ogata fNIRS ki (OMM–300) to keitaigata fNIRS (LIGHTNIRS) tono sokansei kenkyu [Correlative research between large fNIRS (OMM–300) and portable fNIRS (LIGHTNIRS)]. *Ritsumeikan Studies in Language and Culture, 27*, 133-143.

Taura H. (2016c). Ogata fNIRS ki (OMM–300) to kan-i fNIRS ki (Pocket NIRS) tono sokansei kenkyu [Correlative research between large fNIRS (OMM–300) and simplified fNIRS (Pocket NIRS)]. *Ritsumeikan Studies in Language and Culture, 27*, 145-148.

Taylor, L. (2000). Investigating the paired speaking test format. *UCLES Research Notes, Vol. 2.* Retrieved from http://www.cambridgeesol.org/ rs_notes/rs_nts2.pdf

Taylor, L. (2001). The paired speaking test format: Recent studies. *UCLES Research Notes, Vol. 6.* Retrieved from http://www.cambridgeesol.org/rs_notes/rs_nts6.pdf

Van Moere, A. (2006). Validity evidence in a university group oral test. *Language Testing, 23*(4), 411-440. doi: 10.1191/0265532206lt336oa

Yamashita, Y., Maki, A., Watanabe, E., & Koizumi, E. (1998). Development of Optical Topography for noninvasive measurement of human brain activity. *MEDIX*, *29*, 36-40. Retrieved from http://www.hitachi.co.jp/products/healthcare/products-support/contents/medix/pdf/vol29/29_06e.pdf

Yokoyama, S. (2010). *No karano gengo kenkyu nyumon — saishin no chiken kara kenkyu hoho made [Introduction of language research from brain — from current perception to research methods]*. Tokyo: Hitsuji-shobo.

Yomiuri Shimbun [Yomiuri Newspaper] (June 7, 2019) Jyugyo shuchu, no ketsuryu de kakunin [Brain blood flow confirms concentration in classes].

Young, R. (2000, March). *Interactional competence: Challenges for validity* [Paper presentation]. Annual Meeting of the American Association for Applied Linguistics, Vancouver, Canada. ERIC 444361. Retrieved from http://www.wisc.edu/english/rfyoung/IC_C4V.Paper.PDF

Zhang, H., Duan, L., Zhang, Y-J., Lu, C-M., Liu, H., & Zhu, C-Z. (2011). Test-retest assessment of independent component analysis – derived resting – state functional connectivity based on functional near-infrared spectroscopy. *NeuroImage, 55*(2), 607-615. Retrieved from https:// pubmed.ncbi.nlm.nih.gov/21146616/

Appendix

Task 1 for Groups A and B (Eiken Pre-1st level interview test)

Instruction: You have one minute to prepare. This is a story about a man who thought he should lose some weight. You have two minutes to narrate the story. Your story should begin with the following sentence: "One day, a man was coming home from the grocery store with his wife."

Task 1 for Groups C and D (Eiken 2nd level interview test)

Instruction: You have one minute to prepare. This is a story about Aiko, who moved into a new house. You have two minutes to narrate the story. Your story should begin with the following sentence: "One day, Aiko and her parents were moving into their new house."

Task 2 for Groups A and B (Interview)

1) Do you like to travel?
   Yes → Why? Where would you like to go?
   No → Why? When you have free time, what would you like to do?
2) Shall we talk about the picture?
   First, please look at the fourth picture. If you were the man, what would you be thinking?
   Do you do any specific exercise for your health? If so, please describe it.

Task 2 for Groups C and D (Interview)

1) Same as for Groups A and B
2) Shall we talk about the picture?
   First, please look at the third picture. If you were the girl, Aiko, what would you be thinking?
   Do you do any specific things to get along with your friends? If so, please describe them.

Task 3 for all groups (Interaction on an easy topic)
One of the following topics was randomly assigned:
School, English, Family

Task 4 for Groups A and B (Interaction on a difficult topic; Eiken Pre-1st level)
One of the following topics was randomly assigned:

1) Should the government introduce stricter food-safety laws for restaurants and supermarkets?
2) Do you think that advertising for alcohol should be banned?
3) Should more be done to improve the healthcare system in Japan?

Task 4 for Groups C and D (Interaction on a difficult topic; Eiken 2nd level)

One of the following topics was randomly assigned:

1) Some people say that schools should give students tablet computers to use in class. What do you think about that?
2) Today, many Japanese people work in foreign countries. Do you think the number of these people will increase in the future?
3) Some people say that junior high schools should provide school lunches. What do you think about that?

Junko Negishi, Associate Professor
Department of English and American Language and Literature,
School of Literature, Tsurumi University
2-1-3 Tsurumi, Tsurumi-ku, Yokohama, Kanagawa pref., 230-8501, Japan
Phone: +81-45-580-8125
E-mail: negishi-j@tsurumi-u.ac.jp

Hiroyuki Kanzaki, Associate Professor
Department of Orthodontics, School of Dental Medicine, Tsurumi University
2-1-3 Tsurumi, Tsurumi-ku, Yokohama, Kanagawa pref., 230-8501, Japan
Phone: +81-45-580-8507
E-mail: kanzaki-h@tsurumi-u.ac.jp

Yuko Yamada, Research associate
Department of Orthodontics, School of Dental Medicine, Tsurumi University
Phone: +81-45-580-8507
E-mail: b1931126yuko@gmail.com

Masumi Murakami, Clinical Fellow
Department of Orthodontics, School of Dental Medicine, Tsurumi University
Phone: +81-45-580-8507
E-mail: m9s1m7@me.com

Erika Ozawa, Clinical Fellow
Department of Orthodontics, School of Dental Medicine, Tsurumi University
Phone: +81-45-580-8507
E-mail: e.ozawa725@gmail.com

Yoshiki Nakamura, Professor Emeritus
Department of Orthodontics, School of Dental Medicine, Tsurumi University
Phone: +81-45-580-8507
E-mail: nakamura-ys@tsurumi-u.ac.jp

Received: April 30, 2020
Revised: June 27, 2020
Accepted: June 30, 2020