Initial Stage of Novice Word Learning by Vocal Imitation and Repetition: fMRI Study

吉田，晴世

(Degree)
博士（学術）

(Date of Degree)
2010-03-06

(Date of Publication)
2010-05-24

(Resource Type)
doctoral thesis

(Report Number)
乙3095

(URL)
https://hdl.handle.net/20.500.14094/D2003095

A Doctoral Dissertation

Initial Stage of Novice Word Learning by Vocal Imitation and Repetition: fMRI Study

September, 2009

Haruyo YOSHIDA

Table of Contents

Acknowledgements

I would like to thank all those who have supported me and were involved in one way or another in the preparation of this dissertation. My deepest appreciation goes to Prof. Norihiro Sadato whose enormous support and insightful comments were invaluable during the research project.

I would like to also convey my grateful thanks to Prof. Hirokazu Yokokawa for encouraging me to go ahead with my dissertation. I would also like to express my gratitude to my family for their moral support and warm encouragements.

Chapter 1

Introduction

1.1 We Acquire a Language by Imitating and Repeating

Imitation is the most common way of learning many different skills, including the use of language, during development (Iacoboni and Dapretto, 2006). By imitating and repeating what they hear, children can quickly acquire a language from their surroundings (Bohannon and Stanowicz, 1989). Under the assumption that second-language learning is similar to first-language learning, the audio-lingual method was established in English as a second language/English as a foreign language (ESL/EFL) education in the 1950s. This method addresses the need for people to learn foreign languages rapidly, and is best suited to beginners-level learning of English as a second language (Lado, 1964). In contemporary Japan, language teachers frequently use a “shadowing method” that requires learners to repeat a presented native utterance as quickly and as accurately as possible (Murphey, 2001). Shadowing has been shown to be effective for improving speaking and listening abilities, and is used for the training of simultaneous translators (Kurtz, 1992).

According to some theories, such as the motor theory of speech perception, imitation is the basis of speech perception (Liberman and Whalen, 2000). Recognition of the phonological aspects of an utterance may follow imitation once resonating circuits are activated (Fadiga et al., 2002; Watkins et al., 2003); imitation might be involved in spoken lexical representation through resonance, which describes the process by which a listener interacts closely, but unconsciously, with a visible and audible interlocutor(Bernardis and Gentilucci, 2006). This resonance can be observed behaviorally with the string-of-phonemes recognition and repetition tasks (Gentilucci and Bernardis, 2007). When repeating a string of phonemes presented by a visible and/or audible speaker, the features of both lip kinematics and vocal spectra tend to be imitated automatically (Gentilucci and Bernardis, 2007). Gentilucci and Bernardis postulated that these resonant circuits are activated by the mirror neuron system (MNS).

Mirror neurons make up a particular class of visuomotor neurons, originally discovered in area F5 of the monkey premotor cortex, which discharge both when a monkey performs a particular action and when it observes another individual (monkey or human) performing a similar action (see Rizzolatti and Graighero, 2004, for a review). Neuroimaging studies in humans have revealed similar activation patterns for both self- and other-executed actions in the inferior parietal and inferior frontal areas, regions thought to be part of the human MNS. It has been suggested that the MNS plays a role in action understanding and imitation, particularly in humans.

The MNS responds not only to visual but also to auditory stimuli. In research with non-human primates, Kohler et al. (2002) found that premotor cortex (F5) neurons are activated when an animal performs a specific action, when it hears the sound associated with the action, and also when it observes the action. In humans, using transcranial magnetic stimulation (TMS), Fadiga et al. (2002) provided evidence that listening to speech specifically modulates the excitability of neurons for the tongue muscles. When a subject listens to spoken words, an increase in motor evoked potentials can be recorded from the listener’s tongue muscles if the words involve strong tongue movements. Watkins et al. (2003) also used TMS to show that perceiving visual and auditory speech increased the excitability of the orofacial motor system during speech perception. By combining positron emission tomography (PET) with TMS, Watkins and Paus (2004) further showed that during auditory speech perception there is increased excitability of the motor system underlying speech production, and that the increased excitability is significantly correlated with activity in Broca’s area. They suggested that Broca’s area “primes” the motor system when speech is heard, and thus operates at the interface of perception and action. These data strongly suggest that Broca’s area acts as a node for vocal imitation.

Acquisition of the phonology or syntax of a first language by imitation and repetition is supported by procedural memory (Cohen and Eichenbaum, 1993), which is defined as the long-term learning of skills, often through repeated practice. The results or content of such learning or knowledge are typically unavailable to introspection or recollection, and this differentiates procedural and declarative memory (Gupta and Cohen, 2002). Procedural memory has two observable aspects: repetition priming and skill learning (Poldrack et al., 1999a; Gupta and Cohen, 2002). Repetition priming refers to facilitation (as indicated by greater accuracy or faster performance) in the processing of specific items during a task as a result of previous exposure to those items. Skill learning refers to improvement in task performance that is not restricted to specific items and extends to new items, and thus refers to the development of generalized abilities as a result of practice. Skill learning and repetition priming have three common characteristics: first, both appear to increase as a power function of the number of exposures; second, they are stimulus specific; and third, they depend on associations between stimuli and responses. Skill acquisition is accompanied by automaticity, such that certain activities can be performed quickly, effortlessly, with little awareness, and with little need for conscious thought (Logan, 1988). Thus, the phenomena of skill learning and repetition priming are highly relevant to understanding language learning. However, it is not clear how the MNS is related to the learning effects resulting from repeating words or phrases in audio-linguistic training.

In this paper, the author first describe the importance of imitation, with a potential role for mirror neurons, in children's language acquisition. Mirror neurons are involved in action recognition, understanding others' intentions, rather than understanding in general, and linguistic processes. Secondly, motor theory is described, which postulates that that there is a direct relation between speech perception and motor control. Imitation also involves regions of the brain responsible for motor control. Any description of language learning requires an understanding of the shared mechanisms underlying verbal short-term memory and vocabulary acquisition, and therefore insights from research related to working memory, including behavioral in nature, are described. Thirdly, from the perspective of language learning strategies, the Audio-Lingual method and the Shadowing method are introduced. Fourthly, three experiments are described, presenting data from investigations of (i) psychophysical testing of unfamiliar Uzbek phrase learning, (ii) regional fMRI signals associated with unfamiliar Uzbek word processing, and (iii) psychophysical testing of unknown English words. Finally, based on these three experiments, the theoretical vocabulary learning model is introduced, emphasizing the effectiveness of imitation and repetition in learning a foreign language.

1.2 Statement of the Problem

Generally, foreign language teaching has capitalized on the usefulness of imitation and repetition as effective language-learning tools, seen to be particularly crucial in early stages of foreign language learning. From the standpoint of a behaviorist view, Starch (1915) states, "Apparently imitation and repetition of correct expression are far more efficacious in forming correct habits than grammatical knowledge." In Skinner's "Verbal Behavior" (1957), he assumes that behavior can be fully described by conditioned and associated responses. Learning depends on the frequency with which the responses are repeated, on consistent reinforcement by suitably rewarding correct responses, and on careful sequencing of stimulus-response associations so as to minimize the chance of mistakes. Programming into easily assimilable and minimal steps allows control and conditioning of responses and effective construction of behavioral patterns. Behavior involving purposeful use of motor muscular activity (a skill) cannot be learned without practice. The more practice the more successfully it is learned. Based on this behavioristic theory, the Audio-Lingual method was developed in the 1940s and dominated FLT in the 1950s and 1960s.

For fluent communication, learners need to have abundant vocabularies. Effective methods for learning and retaining of vocabulary over the long term, as well as knowing how to use vocabulary items, are important. With this goal in mind, many teachers and researchers are looking for new ways to develop these skills. According to Kadota (2007), the shadowing method, the act or task of tracking the speech and, simultaneously, repeating it as precisely as possible, is useful for retention of vocabulary. Here also imitation and repetition are found at the center of the language learning.

However, the role that imitation and repetition play in the language acquisition process has not yet been fully resolved. Linguistic processes are closely related to the human mind. The human mind, however, is very difficult to study, as it cannot be observed directly. But it leaves its traces everywhere, particularly in language. Language has been a window of the mind. Many people have tried to discern the workings of the mind from children’s patterns of development. Psycholinguists are concerned with the mental processes that are involved in learning to speak, and are also interested in the underlying knowledge and abilities which children must have in order to use language and to learn to use language in childhood. Is there any biological foundation for language? Like the question of children’s language acquisition, the studies of ESL/EFL language learning methods described above have not clearly revealed how the human brain works while learners are imitating and repeating. Pedagogically there are, two key issues to be addressed; why imitation and repetition are useful, and what area or areas of the brain are working while processes of imitating and repeating are taking place.

Chapter 2

Review of Related Literature

2.1 Mirror Neurons: Imitation and Language Acquisition

2.1.1 What Are Mirror Neurons?

The discovery of mirror neurons has been well publicized and investigation of the function of these neurons, as well as the putative mirror neuron system in which they operate, provide the rationale for many research projects. Mirror neurons are visuomotor neurons characterized by their heretofore unique ability to respond both to the perception of an action, as well as to the motor planning and execution of the action. A number of studies (di Pellegrino, Fadiga, Fogassi, Gallese, and Rizzolatti, 1992; Gallese, Fadiga, Fogassi, and Rizzolatti, 1996) have shown that these neurons are activated both when the monkey is executing a movement, as would happen in a motor related area, and when the monkey is simply watching the same movement being performed by another.

Fadiga et al. (2002) provide evidence that listening to speech specifically modulates the excitability of tongue motor neurons. This suggests that the processing of speech is associated with activation of the listener’s motor regions responsible for generation of articulatory gestures, and provides support for Liberman’s psychophysical experiments in 1967 that is the listener understands the speaker when neural centres for his/her own articulatory gestures are activated. Fadiga et al. used transcranial magnetic stimulation (TMS) to demonstrate that, while a subject listens to spoken words, an increase in motor evoked potentials can be recorded from the listener’s tongue muscles when the words involve strong tongue movements. They initially showed that listening produces activation for specific phonemes in speech motor centers. The mirror system is thought to be the physiological, neural basis on which it is possible to understand of the actions of others. This mechanism for understanding others depends on shared motor representations between the agent and the observer. The phoneme recognition mechanism is thought to be involved in phonetically understanding others’ speech because the speaker (agent) and the listener (observer) share representations of articulatory motor actions.

In research with monkeys Kohler et al. (2002) found that many object-related actions were recognized by distinctive sounds associated with those actions. Multifunctional neurons in the premotor cortex are activated when the animal performs a specific action, when it hears the sound associated with the action, and when it observes the action. If humans also have these mirror neurons we can link understanding others actions with their intended actions (and why they make those actions), which they describe in language. The lexical presentantions contain not only schemas on how an action should be executed but also action ideas. It also can be a key to gestural communication and the evolution of spoken language.

It is not feasible to find individual mirror neurons in humans since electrophysiology in man can only occur in rare instances and at specific brain sites. However, activity consistent with mirror neuron function has been observed when humans execute or observe grasping actions, and these neural effects as measured by PET (positron emission tomography), fMRI (functional magnetic resonance imaging), and MEG (magnetoencephalography), have identified putative mirror regions or what is sometimes called a mirror system. Other regions of the brain may support mirror systems dedicated to different actions. An increasing number of human brain mapping studies refer to the existence of a mirror system. Collectively this data indicate that action observation activates certain regions involved in the execution of those actions. Unlike monkeys, intransitive actions have also been shown to activate motor regions in humans. Because of the overlapping neural substrates for action execution and observation in humans, many researchers have attributed high level cognitive functions to mirror neurons such as imitation (Carr, Iacoboni, Dubeau, Mazziotta, and Lenzi, 2003; Miall, 2003), action understanding (Umilta et al., 2001), and intention attribution (Iacoboni et al., 2005).The discovery of a mirror system for grasping in or near human Broca’s area has lead some to suggest a possible role for mirror neurons in the evolution of language (Rizzolatti and Arbib, 1998).

2.1.2 Mirror Neurons and Action Understanding

Mirror neurons, when initially discovered in macaques, were thought to be involved in action recognition (Fogassi et al., 1992; Gallese et al., 1996; Rizzolatti et al., 1996). However this created a basis for later work attributing a role in imitation to the human mirror system. Although the term of action understanding was often used, its meaning could range from the act being in agreement with what you see to inferring the intentions guiding to the observed action. It is important to refer to the fundamental neurophysiological observation that a mirror neuron activates both when the monkey executes a certain action and when he observes roughly corresponding actions. The question remains: how do mirror neurons mediate an understanding of actions executed by others? The proposed mechanism is rather simple. According to Rizzolatti et al. (2001), each time an individual sees an action performed by another individual, neurons representing that action are activated in the observer's premotor cortex.

Recently, the discovery of a special class of neurons in the primate premotor cortex has provided some clues to how mirror neurons provide a solution to the problem of understanding others' actions and intentions. In fact, even the sound of an action in the dark activates these neurons (Kohler et al., 2002; Keysers et al., 2003). In macaques, two major areas containing mirror neurons have been identified, area F5 in the inferior frontal cortex and inferior parietal area PF (Brodmann area (BA)7b) in the inferior parietal cortex (Rizzolatti et al. 2001). Inferior frontal and posterior parietal human areas with mirror system properties have also been described with different techniques in several labs (Nishitani et al., 2002; Binkofski et al., 1999; Buccino et al., 2001; Grafton et al., 1996; Iacoboni et al., 1999; Rizzolatti et al., 1996; Johnson-Frey et al., 2003; Heiser et al., 2003; Koski et al., 2003; Koski et al., 2002; Grezes et al., 2003; Grezes et al., 1998; Nishitani et al., 2000).

The automatically induced motor representation of the observed action corresponds to that which is spontaneously generated during active action generation, the outcome of which is known to the individual acting. Through personal experience, therefore, the mirror system transforms visual information into knowledge of the others' state. The activity of mirror neurons correlates with action understanding. The visual features of the observed actions are fundamental to trigger mirror neurons only as far as they provide one form of sensory information about the perceived action. If action comprehension is possible on another basis (e.g., action sound), mirror neurons signal the action, even in the absence of visual stimuli.

2.1.3 Mirror Neurons and Imitation

Until recently, the possible relationship between mirror neuron function and imitation has been neglected. It is now becoming a major field in its own right. One issue is the role that imitation plays in the origin of language, in the acquisition of language by children, in the historical development of language, and in the social uses of language. There are also several questions concerning imitation related to the individual's learning and use of the lexicon, phonology and syntax. There are many arguments about whether the origin, source, acquisition and functioning of language are innate and modular, or the product of more general human cognitive capacities.

It seems clear that imitation does not operate in any deliberate way through the training of children to acquire speech sounds, words, phrases and grammatical rules, as adults have to do in learning a foreign language. At best it seems to be a matter of unconscious imitation; however it is unclear, even given the operation of subconscious imitative processes, how young children extract the right words and grammatical rules from the stream of speech sound. If the capacity for imitation and use of language are innate, then examining imitation more closely may reveal a mechanism. There are further questions about the exact nature of the human capacity for imitation and how it developed over the course of human evolution. There are specific questions regarding how imitation can operate in the acquisition of distinct aspects of language, such as phonology, the lexicon, and syntax.

Learning by imitation is an important part of human motor behavior, which requires complex set of mechanisms (Schaal, Ijspeert, and Billard, 2003). Wolpert, Doya, and Kawato (2003) emphasize some of those mechanisms:

(i) The mapping of sensory variables into corresponding motor variables,
(ii) Compensation for physical differences between the imitator and the demonstrator,
(iii) Understanding the actor's intention in executing an observed movement.

Many cognitive neuroscientists view imitation as the product of mirror neuron function in humans. Since most mirror neurons are found in motor areas it is plausible to imagine a role in motor control for mirror neurons. One possibility is that these neurons execute an internal model for control. Some researchers suggest that the central nervous system uses internal models for movement planning, control, and learning (Kawato and Wolpert, 1998; Wolpert and Kawato, 1998). A forward model is one that predicts the sensory consequences of a motor command (Miall and Wolpert, 1996; Wolpert, Ghahramani, and Flanagan, 2001) while an inverse model transforms a desired sensory state into a motor command that can achieve it. The proposal of Arbib and Rizzolatti (1997) that mirror neurons may be involved in inverse modeling plays a central role in recent hypotheses about the neural mechanisms of imitation.

Iacoboni and Dapretto (2006) presented a schematic overview of the frontoparietal mirror neuron system (MNS) (Fig 1; shown in red) and its main visual input (shown in yellow) in the human brain. An anterior area with mirror neuron properties, situated in the inferior frontal cortex, encompasses the posterior inferior frontal gyrus (IFG) and adjacent ventral premotor cortex (PMC). A posterior area with mirror neuron properties is located in the rostral part of the inferior parietal lobule (IPL), and can be considered the human homologue of area PF/PFG in the macaque areas 55 and 109. The main visual input to the MNS originates from the posterior sector of the superior temporal sulcus (STS). Together, these three areas form a 'core circuit' for imitation. The visual input from the STS to the MNS is represented in Fig. 1 by an orange arrow. The red arrow represents the information flow from the parietal MNS, which is mostly concerned with the motor description of the action to the frontal MNS, which is more concerned with the goal of the action. The black arrows represent efferent copies of motor imitative commands that are sent back to the STS to allow matching between the sensory predictions of imitative motor plans and the visual description of the observed action.

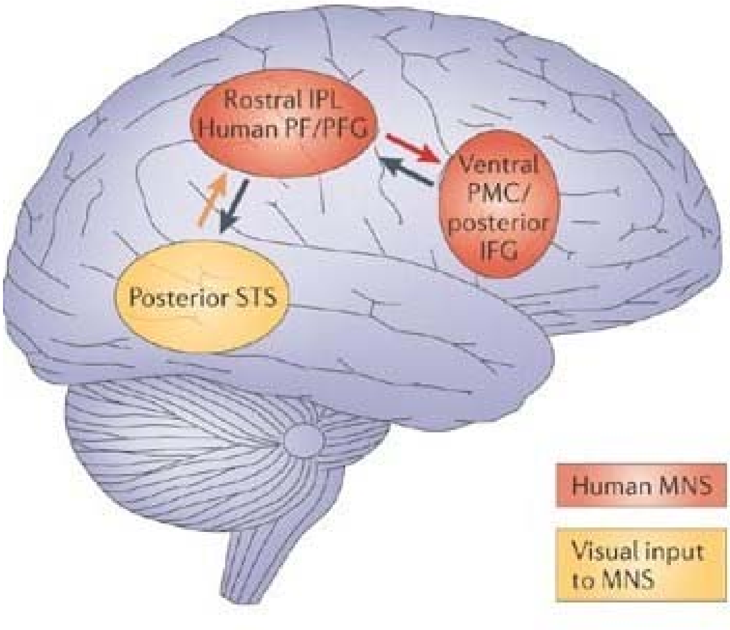


Fig. 1 Neural circuitry for imitation. (Adapted from Iacoboni and Dapretto, 2006)

2.1.4 Mirror Neurons and Language

Rizzolatti and Arbib (1998) have argued for a relationship between mirror neuron function and linguistic communication. They support this link with two primary lines of evidence: (i) macaque area F5, known to contain mirror neurons, correspond to cortical regions in human which include Broca’s area, a region often associated with speech production, and (ii) non-invasive brain imaging methods suggesting the existence of a mirror system in humans for grasping recognition which includes areas in or near Broca’s area. The proposal is that brain mechanisms supporting language in humans developed on top of a primitive mirror neuron system similar to that found in monkeys. According to their mirror system hypothesis, the endowment of Broca’s area with mirror system properties provides humans with the evolutionary basis for language parity, the approximate meaning-equivalence of an utterance for both speaker and listener. Arbib (2002, 2005) has expanded the hypothesis into seven evolutionary stages:

Stage 1: Simple grasping,
Stage 2: A mirror system for grasping,
Stage 3: A simple imitation system for grasping,
Stage 4: A complex imitation system for grasping,
Stage 5: Protosign, a hand-based communication system,
Stage 6: Protospeech, a vocalization-based communication system,
Stage 7: Language, which required little or no biological evolution beyond stage 6, but which resulted from the cultural evolution in Homo sapiens.

In stage 2, the evolutionary introduction of mirror neurons may appear to occur abruptly; however mirror neurons are proposed to be originally involved in motor control. These neurons are located in the premotor cortex, interleaved with other motor neurons, where such things as hand movements are controlled. The importance of goal-directed grasping in primates may provide a basis for development of neural systems which use visual feedback to comprehend the actions and states of others, and also provide a mechanism for imitation.

2.2 Motor Theory of Language

Motor theory states that there is a direct relation between the functioning of speech and motor control in general, with language depending on pre-existing motor primitives combined in the operation of motor equivalence (Allott 1991). As described by Liberman & Mattingly (1985), according to the motor theory of speech perception a listener decodes auditory speech input by encoding it in terms of the articulatory gestures from which the speaker presumably constructed it, rather than more directly from the acoustic patterns of the speech. According to this influential view, the “generative” listener activates neural structures that counter-intuitively do not correspond to particular patterns of perceived sounds, but to the vocal gestures that generated them. Listening, one covertly imitates the speaker’s articulatory gestures and monitors one’s covert action in order to decode its meaning.

The motor system internally represents actions in terms of motor ideas, as well as executes them. Brain imaging methods such as PET and fMRI have demonstrated that bodily action is preceded by a mental picturing of the proposed action. A perceptually organized pattern is transduced into a motor program and executed by way of the changes in position and changes in limb positions which constitute human action. The concept of the motor program has evolved from earlier understanding of motor action in terms of schemas and motor patterns. At the highest level, the motor program can be well described as an action program incorporating all individual components of a single goal-directed action. At the lowest level there are motor sub-routines which are motor elements that can be executed in chains and in combination to create actions.

In newborn infants, many aspects of motor behavior and communicative motor behavior in particular are found. The neural connections to support such behavior must have been established before birth. The speech and skeletomuscular systems share common neural control modes. Kinematic patterns for movements of the tongue dorsum are similar to those of voluntary flexion extension movements at the elbow. The motor control of the tongue is a section of the total system of motor control for all bodily movement. Tongue movements can be explained in terms of activation of a small number of independent muscle groups, each corresponding to an elementary or primitive movement, making use of a relatively small number of motor actions.

Bernstein (1967) described the concept of motor equivalence, in which the motor image or motor idea is independent of the particular limb or particular muscle assembly usually employed to execute the motor program. Motor equivalence is demonstrated in the relation between speech and gesture, and can operate from speech to gesture or from gesture to speech. It also seems likely that it can operate between other modalities and speech or more precisely between speech and perception and between perception and motor action. Words formed from primitive speech sound elements are motor programs. Movements of the arm are the motor equivalents of the speech primitives. Before we produce a sentence there is a motor image of the sentence. A sentence is a high-level motor program or action plan.

2.2.1 Imitation in Shadowing Words

Shadowing has been used as a tool not only to explore how quickly an individual can repeat an utterance made by another, but also to explore the conditions under which repetitions are imitations of the model speaker. Goldinger (1998) asked participants to produce utterances which, in one condition, could not be imitations of a model's speech and, in another condition, were repetitions of a model's speech that could be imitative. He found that listeners judged the repetitions to be better imitations of the model than were the baseline productions. Moreover, this research discovered variables, such as word frequency and the number of times the participant heard a model's word before repeating it, which affected imitative accuracy. On the basis of these findings he proposed a theory of speech perception in which listeners store traces of perceived words that retain information about the contexts in which they occur, such as the voice quality of the speaker and other peculiarities. The stored traces of a word are activated when the word is spoken, and they combine with the spoken word to guide shadowing responses. For that reason, the shadowed words are imitative.

When an utterance is perceived, traces in memory are activated to a degree dependent on their similarity to the input utterance. Activated traces from prior encounters with speech input, combined with the trace of the new input into an echo, influence a speaker's shadowing response. Repetition of utterances in the familiarization phase of the experiment provides a way to shape the characteristics of the echo. The more repetitions occur, the more repetitions increase the number of identical traces in memory, which should make the echo for that utterance type increasingly similar to the repeated utterance in such properties as voice, speaking rate, and so forth. According to Goldinger's research, if the echo guides the shadowing response, imitative fidelity should increase with repetitions during familiarization.

There are two theories of accounting for the process of imitation in shadowing. One account is provided by the motor theory of speech perception (Liberman & Mattingly, 1985), described above. In that theory, listeners perceive the linguistic vocal tract gestures of the speaker, and they do so in a way that involves their own speech motor system in the perceptual process. According to this theory, others' speech first activates one's own speech motor system, and thus speech can serve as a prime for an imitative response.

The other account is provided by direct realist theory (Fowler, 1994), also a form of gesture theory. It states that, in effect, listening to speech includes receiving instructions for the shadowing response. That is, for example, if a listener hears a certain positioning of the tongue with lips shaped, that can serve as instructions to produce a shadowing response, and if it does, the response is likely to be an imitation. The direct realist account (Porter and Hogue, 1998) is based on the assumption that speech production and speech perception involve nonlinear, dynamic, self-organizing systems, reducing the potentially large number of variables which need to be controlled for or detected during linguistic processing to a relatively small number of parameters. In both speech production and speech perception, when combined in different ways, it is assumed that the parameters to be controlled or detected correspond to those psychologically specified segments which convey meaning. This theory suggests that the parameters themselves may correspond to vocal tract gestures.

2.2.2 Repeating and Remembering Foreign Language Words

Unlike children, adults find learning non-native phonemes difficult. According to Werker and Logan (1985), non native speakers of a language demonstrate poorer performance on listening discrimination tasks compared with native speakers of the language from which the phonemes were chosen, when they are presented with speech stimuli from phonetic categories not used in their own language.

Based on the idea that the optimal learning situation provides the listener with a range of identification cues, specifically with a view of the face of the speaker, so called “visible speech”, several experiments have been conducted (Munhall and Vatikiotis-Basteson, 1998; Reisberg, McLean and Goldfield 1987).

One experiment (Davis and Kim, 2001) examined the proposal that visible speech can help with processing a difficult speech signal such as when listening to a foreign language. They investigated this issue by examining whether presenting the face of the speaker improved the accuracy of repetitions of short phrases of a language participants had not heard before, in this case Korean, as well as examining whether this manipulation facilitated performance on a subsequent old and new recognition task. The results showed that both repetition accuracy and the subsequent memory of foreign language phrases were improved by showing the speaker’s face. The perception and repetition of unknown speech sounds was improved by visual information of the tongue, teeth, lips, jaw and cheeks of the speaker. The implication of this finding is that during foreign language learning there are benefits of using a presentation method that includes orofacial information simultaneous with, and corresponding to, speech.

2.3 Verbal Short-term Memory and Procedural Memory

2.3.1 Vocabulary Acquisition and Verbal Short-term Memory

The learning of new words, or vocabulary acquisition, is one of the most essential processes during human development. Without a system for vocabulary learning we could never acquire language, and without language, human culture could not be developed or sustained. A second crucial and characteristic human skill is the ability to retain sequences of words in short-term memory. Without this additional mnemonic ability, it would be impossible to understand anything but the simplest of sentences.

Many scholars suggest that human vocabulary acquisition processes and aspects of human working memory may be closely related (Gathercole & Baddeley, 1989, 1990, 1993; Papagno, Valentine, & Baddeley, 1991). This work has been conducted within the framework of the working memory model (Baddeley, 1986), which offers a valuable theoretical perspective for investigation of vocabulary acquisition. What makes it an especially useful framework is the further fact that a good deal is known about the neural substrates of verbal short-term memory, from neuropsychological investigation (Baddeley, Papagno, & Vallar, 1988; Shallice & Vallar, 1990; Waters, Rochon, & Caplan, 1992).

The search for shared mechanisms underlying verbal short-term memory and vocabulary acquisition is important for several reasons. First, it offers a new processing-oriented approach to examining vocabulary acquisition. Second, investigation of word learning can enlighten the relations between short- and long-term memory systems. Third, examination of this linkage of two critical human faculties would provide us with a valuable way of understanding the details of language learning. Fourth, to the extent that the neural substrates of such processing can be specified, this relationship has relevance to the investigation of the brain and language.

2.3.1.1 The Working Memory Perspective

The working memory model has been one of the most widely influential theories of short-term memory for more than 20 years (Baddeley & Hitch, 1974; Baddeley, 1986). The model has three major components:

visuospatial short-term memory, verbal short-term memory, and a central executive system, the latter of which controls the flow of information to and from the other components.

The verbal short-term memory system has been termed the ‘‘articulatory loop’’ and consists of two parts (Baddeley, 1986). One subcomponent of this processing system consists of mechanisms that enable rehearsal such as repeating a phone number to oneself until it can be dialed; the second subcomponent consists of a system for phonological representation of individual spoken words and their temporary storage (Baddeley, 1990). The rehearsal process ‘‘refreshes’’ the decaying traces in the memory store; it is believed to rely on articulatory mechanisms and has therefore been called articulatory rehearsal (Baddeley, 1986). The memory buffer stores verbal material, but its memory traces decay within 1–2 sec, which is why a refresh mechanism is needed.

2.3.1.2 Behavioral Evidence for a Relationship between Verbal Short-term Memory and Vocabulary Acquisition

One easy way of measuring verbal short-term memory is to ask subjects to recall lists of words. In the immediate serial recall (ISR) task, the subject is presented with sequences of unrelated verbal items such as digits or words and is required to recall the sequence in correct order, immediately following its presentation. Presentation of the list may be auditory or visual. The subject may be required to respond in speech, in writing, or in some other fashion. The subject’s digit span is measured as the length of the longest list that the subject can recall at some criterion of performance. This task has played a central role in development of the working memory model.

Digit span as measured by ISR is widely accepted as the standard measure of verbal short-term memory. Gathercole and Baddeley (1989) have described a group of children with specific language impairments (SLI), whose digit span was found to be highly correlated with their poor nonword repetition performance. This finding suggests that common abilities are involved in immediate serial recall and nonword repetition, and that nonword repetition can be used as an alternative to ISR to gauge verbal short term memory abilities. To further examine their verbal short-term memory abilities, the SLI group was compared with a control population of normal children matched in terms of vocabulary and reading skills. The SLI children’s performance was significantly impaired, relative to that of the controls, on nonword repetition as well as digit span. The SLI group, whose mean age was 8 years, had nonword repetition ability equivalent to that of normal 4-year-olds, while their vocabulary abilities were equivalent to those of normal 6-year-olds. These findings are important because they suggest that the same abilities are involved in immediate serial recall, nonword repetition, and vocabulary acquisition, as the SLI group’s performance was impaired in all three of these areas.

Gathercole and Baddeley (1990) have also described a longitudinal study of 4- and 5-year-old normal children, in which nonword repetition scores (Gathercole, Willis, Baddeley, & Emslie, 1994) were highly correlated (a) with receptive vocabulary scores and (b) with the time taken to learn unfamiliar names in a simulated vocabulary acquisition task. Thus again, this suggests that the mechanisms underlying verbal short-term memory play an important role in vocabulary acquisition.

In a study of 9-year-old Finnish children learning English as a foreign language, Service (1992) found a close association between nonword repetition ability and English grades 2 years later. Repetition scores were not, however, correlated with arithmetic scores. This suggests that the correlation between verbal short-term memory as measured by nonword repetition and foreign-language learning reflects reliance on some common processing component and not merely common reliance on general intellectual abilities. These studies by Gathercole, Baddeley, and Service have focused on the long-term correlations between verbal short-term memory abilities and vocabulary acquisition.

Another approach to the study of the linkage between these skills is the study of adult language within in a controlled experimental context. Papagno et al. (1991) have examined normal adult subjects' paired associate learning, defined as their ability to learn associations between pairs of words, so as to be able to recall one member of the pair when prompted with the other. The experiment tested learning of English–English paired associates as well as learning of English–Finnish paired associates. The subjects' task was to learn to associate known English words with an unrelated phonological form, which was either an English word or a Finnish word. Subjects were asked to perform these tasks simultaneously with one of two other tasks. In a simultaneous articulation condition, subjects had to repeatedly utter the word 'bla' while simultaneously performing the primary paired-associate tasks. In a simultaneous finger tapping condition, subjects had to tap their finger repeatedly while simultaneously performing the primary tasks. That is, subjects had to perform paired-associate learning of either English–English pairings or English–Finnish pairings, while at the same time performing a simultaneous interference task involving either articulation or finger tapping.

Extrapolating from previous studies using simultaneous articulation interference tasks (Baddeley, Thomson, & Buchanan, 1975; Salame & Baddeley, 1982; Baddeley, Lewis, & Vallar, 1984; Hanley &

Broadbent, 1987), Papagno et al. (1991) predicted that learning of English–Finnish pairings would be affected more by simultaneous articulation than by simultaneous finger-tapping. This prediction was confirmed. There was a significant difference between the impact of the two interference tasks on English–Finnish learning, but no significant difference in the impact of the two tasks on English–English learning. These findings are important because they provide further evidence that articulatory rehearsal plays a role in vocabulary acquisition and that simultaneous articulation interferes with this rehearsal.

To summarize, there is now considerable evidence suggesting that aspects of rehearsal and the storage of phonological representations underlying verbal short-term memory also underlie vocabulary acquisition. The articulatory loop model suggests that there is a partnership between the storage system for phonological representations and speech output planning mechanisms, such that the latter can serve to generate and/or refresh representations within the former. The partnership of these two components provides a very simple model of the system underlying performance in both verbal short-term memory and vocabulary acquisition. In immediate serial recall, this system sets up a loop of rehearsal activations which can maintain phonological representations of the recall stimuli in an active state. In vocabulary acquisition, likewise, a similar rehearsal loop aids the formation of a phonological representation for a new vocabulary item, by allowing the learner repeated access to the new word, and thereby aiding consolidation of the new lexical item. (Bishop, 1992)

2.3.2 Skill Acquisition and Procedural Memory

Cohen and Squire (1980) introduced the term declarative memory, which is memory for facts or things which can be explicitly discussed or declared, to refer to procedural memory, and it is now well accepted that such memory relies on the hippocampus and medial temporal lobe structures, and that it is profoundly impaired in patients with amnesia (Cohen & Eichenbaum, 1993). Impairments in this memory system can be revealed by direct tests of memory, which require explicit retrieval of the contents of specific experiences; a typical example would be memory for arbitrary pairings of words. Indication of a second memory system came in the form of a striking pattern of preservation of certain abilities in patients with amnesia (Cohen & Squire, 1980). The forms of memory and learning that are spared by hippocampal damage include the acquisition of skills that are acquired gradually over several sessions of practice, and facilitation of priming in processing of a stimulus, following prior exposure to that stimulus. Patients with amnesia exhibit normal patterns of skill acquisition and priming in such tasks, provided there is no requirement for direct recollection of previous exposure or practice. This second kind of learning and memory has been termed implicit memory (Schacter, 1987), nondeclarative memory (Squire, 1992), or procedural memory (Cohen & Eichenbaum, 1993), and it is examined by indirect tests, which do not require explicit recollection of previous experiences.

Early findings from patients with amnesia were followed by the discovery that dissociations between declarative and procedural memory could be obtained in normal populations (Graf & Schacter, 1985). These results generated much interest in the research community, as evidenced by the large body of research on “explicit” and “implicit” memory or on declarative and procedural memory (Roediger & McDermott, 1993). Much of the interest in procedural (implicit) memory arises from the fact that a great deal of everyday human learning appears to have the character of procedural memory: it occurs gradually, as a result of practice over many exposures; and the results or contents of such memory, learning, or knowledge are typically unavailable to introspection or recollection. For example, learning to ride a bicycle appears to have these characteristics, as does acquisition of the phonology or syntax of a first language. Thus the study of procedural memory touches on traditions of psychological inquiry, such as the investigation of skill acquisition and investigation of language learning. Further specification of the nature and mechanisms of procedural memory therefore holds the promise of providing fresh insight into fundamental and pervasive human learning processes.

2.4 Language-learning Strategy Perspectives

2.4.1 The Audio-Lingual Method

The Audio-Lingual method (e.g., Fries, 1945; Lado, 1964) of teaching English as a second language has its origins in the time during World War II, when it became known as the Army method. It was developed as a reaction to the grammar-translation method of teaching foreign languages. Grammar-translation had been used to teach for many years, but the method was perceived as ineffective because students' progress was seen to be slow. The Audio-Lingual method set out to achieve quick communicative competence through innovative methods. From about 1947-1967 the Audio-Lingual method was the dominant foreign language teaching method in the United States.

This approach is based on the principles of behavioral psychology. To the behaviorist, the human being is an organism capable of a wide repertoire of behaviors. The occurrence of these behaviors is dependent on three crucial elements in learning: a stimulus, which serves to elicit behavior; a response triggered by a stimulus; and reinforcement, which serves to mark the response as being appropriate and encourage the repetition of the response in the future (Skinner 1956; Brown 1980). A representation of this can be seen in Fig. 2.

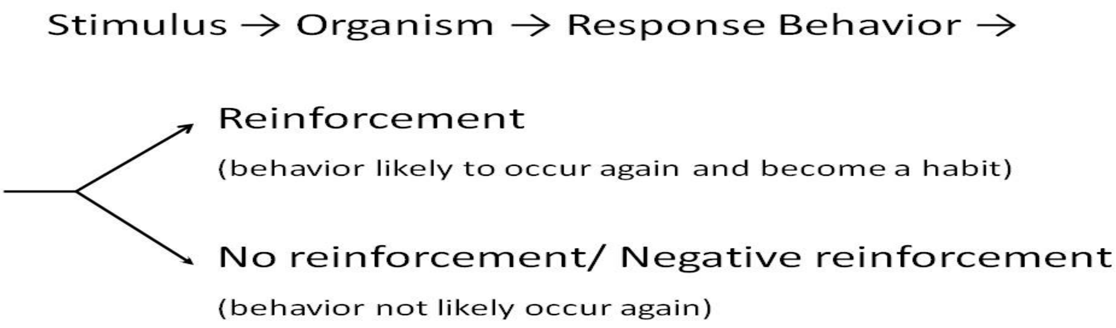


Fig.2 Stimulus, response, reinforcement model (Adopted from Richards, 1986).

Reinforcement is a vital element in the learning process, because it increased the likelihood that the behavior will occur again and eventually become a habit. To apply this theory to language learning is to identify the organism as the foreign language learner, the behavior as verbal behavior, the stimulus as what is taught or presented of the foreign language, the response as the learner’s reaction to the stimulus, and the reinforcement as the extrinsic approval and praise of the teacher or fellow students, resulting in the acquisition of a set of appropriate language stimulus-response chains.

New material is presented in the form of a dialogue. Based on the principle that language learning is habit formation, the approach fosters dependence on mimicry, memorization of set phrases and over-learning. Structures are sequenced and taught one at a time. Structural patterns are taught using repetitive drills. Little or no grammatical explanations are provided; grammar is taught inductively. The use of drills and pattern practice is a distinctive feature of the Audio-Lingual method. Brooks (1964) introduced the use of repetition in drills. The student repeats an utterance aloud as soon as he has heard it. He does this without looking at printed text. The utterance must be brief enough to be retained in phonological memory. Sound is, according to principles of this teaching method, as important as form and order.

One of the key principles of the Audio-Lingual method is that the language teacher should provide students with a native-speaker-like model. By listening, students are expected to be able to mimic the model. Based upon contrastive analyses, students are drilled in pronunciation of words that are most dissimilar between the target language and the first language. Grammar is not taught directly by rule memorization, but by example. The method presumes that second language learning is very much like first language learning (Celce-Murcia,1991; Larsen-Freeman, 1986).

In Lado's influential book 'Language Teaching. A Scientific Approach '(1964), he sums up the principles of the Audio-Lingual method (Lado, 1964). Seven principles are selected for the language learning effectiveness, with particular application in Japanese EFL classes:

Principle 1. "Speech before Writing" Teach listening and speaking first, reading and writing next. This principle is the basis for the audio-lingual approach. Corollary Speech cannot be invented by the student; it has to be imitated. Poor models produce poor imitations. Good models do not guarantee good imitations, but they are necessary to permit good responses. In the light of this corollary, native speakers who possess an acceptable variety of speech are desirable models, as are those who, though not native speakers, have achieved a high level of mastery.

Principle 2. “Basic Sentences” Have the students memorize basic conversational sentences as accurately as possible. This practice, advocated by linguists, has a strong psychological justification not dealt with in published experiments but tested repeatedly otherwise. Linguists support the use of conversations because they present words in sentence structures and in context. Conversational dialogues are preferable to poetry or formal prose because conversations show a greater range of the basic constructions of the language in matter-of-fact context. Poetry uses more of the unusual constructions and the less typical variants of common constructions. Prose makes little use of questions, requests, and answers; it is characterized by longer statement patterns.

Principle 3. “Patterns as Habits” Establish the patterns as habits through pattern practice. Knowing words, individual sentences, and/ or rules of grammar does not constitute knowing the language. Talking about the language is not knowing it. The linguist, the grammarian, and the critic talk and write about the language; the student must learn to use it. To know the language is to use its patterns of construction with appropriate vocabulary at normal speed for communication. Understanding or even verbalizing a pattern may help a student to learn it but will never take the place practicing the patterns through analogy, variation, and transformation to establish them as habits. This is pattern practice.

Principle 4. “Sound System for Use” Teach the sound system structurally for use by demonstration, imitation, props, contrast, and practice. Observation repeatedly shows that merely listening to good models does not produce good pronunciation after childhood. Partial attempts, props in the form of articulatory clues, and minimal contrasts to focus sharply on the phonemic differences eventually result in satisfactory responses, but to increase facility and fluency, practice becomes indispensable.

Principle 5. “Practice” The student must be engaged in practice most of the learning time. This principle has a psychological justification, since, other things being equal, the quantity and permanence of learning are in direct proportion to the amount of practice. Linguists have demonstrated the importance of practice through mimicry-memorization and pattern practice. The strongest support for this principle is in the success of the intensive courses in English developed under Fries’s direction and the success of the intensive language programs under linguistic auspices during the Second World War. Although these did not include controlled experiments, the evidence was convincing.

Principle 6. “Speed and Style” Linguistically, a distorted rendition is not justified as the end product of practice. Psychologically, partial experiences and props are necessary as intermediate steps to a full experience. The principle makes sure that the practice ends in a

> linguistically acceptable and psychologically full experience.

Principle 7. "Immediate Reinforcement" let the student know immediately when his response has been successful.

Audiolingualism holds that language learning is based on the principles of behavior psychology. Since language is a formal, rule-governed system, it can be formally organized to maximize teaching and learning efficiency. Audiolingualism thus stresses the mechanistic aspects of language learning by imitating and repeating a model speaker's utterance and language use.

2.4.2 The Shadowing Method

By the end of the twentieth century, mainstream language teaching no longer regarded the specific method employed as being the key factor in accounting for success or failure in language teaching. However, in Japan, the shadowing method has been widely used in classrooms from the end of twentieth century. Shadowing can generate effective listening if it is used correctly. It is also a difficult practice for novices because they have to listen to English and repeat it simultaneously. Takizawa (1999) explained that shadowing is a practice where students repeat a model sound, listening to the model without provision of a script. Tamai (1998) defined shadowing as the act or task of listening in which the learner tracks the speech and repeats it as exactly as possible while listening attentively to incoming information. Nema (1999) stated "shadowing is to repeat the target utterance or repeat it soon after hearing several words as accurate as possible". In short, shadowing is to repeat what has been heard as accurately as possible. Someone might think that shadowing is similar to a simple repetition process, but there are key differences. Takizawa (1999) pointed out that during ordinary repetition, learners voice the sentences during pauses in the model sound, and will often read the dialogue with an auxiliary script. In contrast, shadowing requires repetition of the model sound without the presence of pauses or aid of a script.

Tamai (2005) stated, “the purpose of listening is to understand the meaning from the sound which the listener can listen to, and the other is to comprehend the necessary information”. Listening competence can be categorized into two types. One relates to knowledge — vocabulary, grammar, and phonemes. This is like a word processor, which can shift information from one language to another. The other relates to cognition and processing. This works as if it were a central processing unit (CPU) for analyses or processing. Metaphorically, improving listening ability is to reinforce one’s word processor software by increasing the linguistic knowledge and capability of the CPU. Exercises involving repeating and listening comprehension might be useful for reinforcing the word processor software, but shadowing may be the effective activity focusing on the latter category.

2.4.2.1 Working Memory and Processing of Speech Inputs

The working memory model defined by Baddeley argues in support of shadowing as an effective method of language teaching. Baddeley (1986) defined a system, termed working memory, for the temporary holding and manipulation of information during the performance of a range of cognitive tasks such as comprehension, listening, and reasoning.

In both working and short-term memory, information from various inputs is preserved temporarily. However, working memory is thought to perform operations on the maintained information which short-term memory does not. In other words, working memory concerns not only the holding of information, but also the processing of information for the purposes of comprehension, learning, and reasoning. According to Gathercale and Baddeley (1993), there are three components in working memory: a central executive, a visuo-spatial sketch pad, and a phonological loop. The central executive component is the most important. They stated its importance as follows:

> Some of its primary functions are regulatory in nature: It coordinates activity within working memory and controls the transmission of information between other parts of the cognitive system. In addition, the executive allocates inputs to the phonological loop and sketchpad slave systems, also retrieves information from long-term memory. (Gathercale & Baddeley, 1993, p. 5)

The central executive fulfills many different functions, supplemented by the phonological loop and the visuo-spatial sketch pad. The putative function of the phonological loop is to maintain verbally coded information, whereas the visuo-spatial sketch pad is involved in the short-term processing and maintenance of material which has a strong visual or spatial component. Activation of working memory might relate to listening processes. If so, this may influence the reinforcement of cognition and processing mentioned above. It can be suggested that shadowing is effective for activation of working memory.

Tamai (1998) observed, “Shadowing is not an automatic activity, but a highly cognitive activity as basing on a working memory” (p. 56). He also modified the working memory model that Futaya proposed and provided in 1994 (see Fig.3). In the model, the sound information selected by attention is recognized, and then the part of the information selected by attention is processed in the phonological loop. Next, within the phonological loop, subvocalized input remains here. If input to the phonological loop becomes close to maximum capacity, understanding will be compromised. In the case of shadowing, Tamai (2002) stated that the phonological loop is recruited for this task. Kuramoto (2004) proposed that shadowing could develop listening competence by recruiting the series of operations required for working memory.

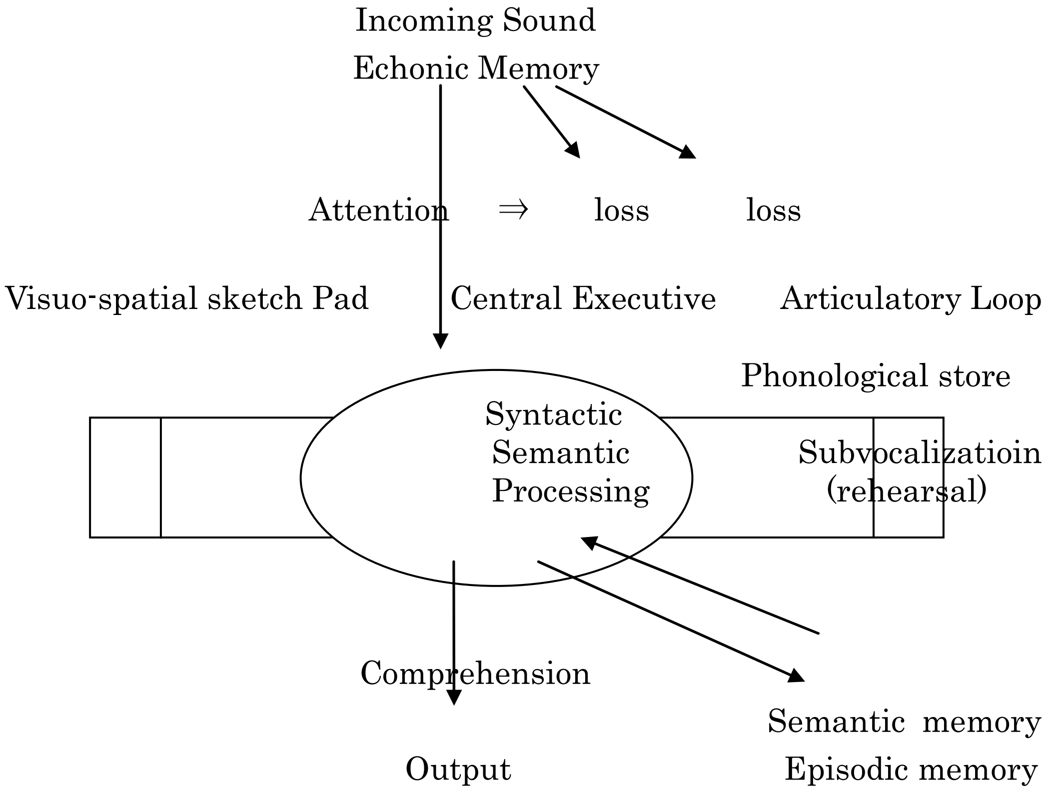


Fig.3 Baddeley’s working memory revised by Futaya (1994) and then by Tamai (2002)

2.3.2.2 Automaticity and Internalization

Shadowing has two effects; one is the phonological processing and articulation of speech and the other is the internalization of new words. These are both lower-level processes. Shadowing can train the learner's speech perception through listening, and make this process more automatic. Automation describes when a process or task can be completed with less effort, and be more resistant to interference. Shadowing is, for novice learners, an act that has a high cognitive load. The more that learners practice, the earlier linguistic processed attain automaticity. This is called habituation. If a student can habituate to shadowing, they can shadow speech effectively but also understand the sentence meaning and grammatical structure. Therefore, shadowing can reinforce the student's listening ability and contribute to their speech-knowledge database — which may comprise of prosodic features, pauses, and stress, which can be confined to long-term memory.

Internalized new items, such as new vocabulary, and grammatical rules, are stored in long-term memory when the learners' subvocal rehearsal process, which corresponds to activity of the phonological loop of working memory, can be streamlined and actualized in the process of articulation during shadowing. In other words, simultaneously processing and repeating phonological input reinforces linguistic storage processes during shadowing. The temporal storage of vocalized phonological information relies on one component of working memory, the phonological loop, which consists of two subsystems: a phonological short-term store and subvocal rehearsal. The phonological short-term store retains phonological information. This can be a passive function. The information in the phonological short-term store disappears within 2 seconds without rehearsal. In order to retain the information, subvocal rehearsal is needed, which enables prolongation of the storage duration for phonological information. This processing sequence is the same that learners use unconsciously when they acquire new knowledge. Shadowing is the act that vocalizes subvocal rehearsal. Therefore, shadowing can aid learning of new vocabulary and grammatical items. In other words, learners can ,more rapidly understand phonological information, and also reinforce already-established lexical stores, if they train in shadowing as a listening activity.

There is a report concerning shadowing and word retention (Funayama, 1998). In the report, Funayama insists that shadowing contributes to retention of vocabulary. That is, shadowing can retain the word throughout sound material: (i) where is the sound input maintained, and (ii) how long is it maintained – throughout what unit of speech (word, sentence, and conversation). In a laboratory setting, subjects had the opportunity to encounter words during a shadowing exercise which they had previously been asked to try to memorize, and following this memory retention for trained words was evaluated. He concluded that phonetic processes such as like shadowing enable learners to better retain vocabulary.

Chapter 3

Experiments: Nonword Imitation and Repetition Tests

This experiment uses Uzbek words, which constitute non-words to native Japanese participants, as experimental stimuli. Participants were asked to listen to unfamiliar Uzbek words of varying length, spoken one at the time, and attempt to repeat them correctly immediately afterwards. Non-word tests, as well as digit span tests, provide measures of short-term phonological memory, which has been shown in several studies to correlate with and sometimes to predict language development, as measured by vocabulary acquisition (see for example Adams & Gathercole, 2000; Alloway, Gathercole, Willis & Adams, 2004; Gray, 2003; Marton & Schwartz, 2003; Masoura & Gathercole, 1999; Baddeley, Gathercole & Papagno, 1998). Almost without exception, non-word repetition has been found to be more strongly associated with vocabulary knowledge than digit span.

The theoretical starting point for much of the recent research in this area has been the working memory model proposed originally by Baddeley and Hitch (1974) and developed further by Baddeley (1992, 2000). According to this model, working memory is a system providing temporary storage and information processing in support of complex cognitive tasks such as language learning, linguistic comprehension, reading, problem-solving and reasoning. Information processing is putatively performed by an attention-controlling central executive of limited capacity, whereas temporary storage is considered to involve use of two specialized “slave systems”, the visuospatial sketch pad and the phonological loop, as well as an episodic buffer capable of storing information in terms of a multidimensional code. According to the version of the model described by Baddeley (2000),, the episodic buffer has an integrating function, combining information from different sources. The phonological loop is regarded as being important for early vocabulary learning. Indeed, Baddeley, Gathercole and Papagno (1998) have proposed that the phonological loop serves primarily as a language learning device by providing temporary storage of unfamiliar sound patterns during the construction of more permanent memory representations.

Hence, the Uzbek word repetition test described here was designed, in which Japanese EFL learners were instructed to repeat each Uzbek word within 2 seconds of presentation, and thereby theoretically activating the phonological loop. According to Baddeley’s (1992, 2000) working memory model, verbal material is temporarily stored in the phonological loop, a system composed of two parts: the phonological store, capable of holding verbal information for about 2 seconds, and a subvocal rehearsal process that refreshes the decaying representation contained in the phonological store. The total time for hearing a non-word in the test situation and commencing to repeat it tends to be no more than 2 seconds. Baddeley et al. (1998) claim that non-word repetition measures the capacity of the store, rather than measuring subvocal rehearsal. They also maintain that the store plays an essential role in learning the phonological forms of new words.

In this experiment, we were interested in the number of repetitions needed to remember target words. Previous findings are encouraging in that humans can remember novel words after relatively few repetitions. Crothers and Suppes (1967) discovered that almost all of their participants remembered all 108 Russian-English word pairs after 7 repetitions, and about 80% of 216 word pairs were learned by most participants after 6 repetitions. Similarly, Lado, Baldwin and Lobo (1967) presented their intermediate level college students of Spanish with a list of 100 words, and found that only one exposure sufficed for an average of 95% recognition and 65% recall. In general, results on this issue show that, if remembering word pairs is the aim, a surprising amount can be learned within a relatively short time, and not many repetitions are needed before word pairs can be remembered. In the current study, the number of the Uzbek word repetitions was therefore limited to 8 times, a number which, based on previous investigations, should be sufficient for learning of novel words.

3.1 Psychophysical Testing of Uzbek Phrases

In order to ensure the effectiveness of imitating and repeating the foreign language, psychophysical testing of Uzbek phrases was conducted.

3.1.1 Methods

3.1.1.1 Subjects

In total, 13 native Japanese speakers participated in this study. The subjects were recruited from among the college students majoring in English education at Osaka Kyoiku University (Osaka, Japan). All participants had normal or corrected-to-normal vision, normal hearing and were naive to the purpose of the experiment. None of them knew the

Uzbek language.

In total, two groups of Uzbek phrases were used as stimuli, termed groups A and B. Both groups consisted of fifteen phrases. Participants were divided into two groups: Group Ⅰ consisted of four third year students and Group Ⅱ consisted of nine graduate students.

3.1.1.2 Recording equipment and stimulus presentation

During task completion, participants' repetition of the target words was recorded using a PC voice recorder. The stimulus phrases were spoken by a male native Uzbek speaker as the model voice, and auditory stimuli were presented as well as corresponding facial expressions in the center of a 17-in. monitor.

3.1.1.3 Procedure

Both groups were exposed to fifteen different phrases, presented seven times each, in a random order. Group Ⅰ were presented with phrase group B, and were instructed to watch and listen to the model voice without repeating the phrases as they were presented. Group Ⅱ were presented with phrase group A, and were instructed to repeat the phrases as they were presented, imitating the model voice as much as possible. After completing the task, all participants receive a 16th presentation, and repeated the word each time after the model speaker. Their voices were recorded and collected for data analyses. A week later, participants returned and the experiment was repeated, but this time each participant was allocated to the group (Ⅰ or Ⅱ) which they had not previously experienced. We therefore used a cross-over design.

Participants' performance at repeating the Uzbek phrases was evaluated by the native speaker of Uzbek. The standards of evaluation were accuracy, fluency, and prosody, from A to D (A: excellent and meaningful, B: good and meaningful, C: average and meaningful, D:

meaningless).

3.1.2 Results

Each participant’s performance was listed below.

Participant: S1～S13
Voice with repeated imitation: I
Voice without repeated imitation : NI
S1: I=A NI=C
S2: I=A NI=B
S3: I=A NI=C
S4: I=A NI=B
S5: I=B NI=C
S6: I=A NI=B
S7: I=A NI=B
S8: I=A NI=B
S9: I=C NI=D
S10: I=C NI=C
S11: I=A NI=C
S12: I=A NI=B
S13: I=A NI=C

Ten As (excellent and meaningful), one B (good and meaningful), two Cs (average and meaningful) and no Ds (meaningless) comprised the scores for performance after repeated imitation. In comparison, no As, six Bs, six Cs, and one D comprised the scores when participants completed the task without reaped imitation. Except S1, all participants performed better after repeating imitation of the model voice than without repeating imitation. The results revealed a significant effect of repeating imitation when learning novel phrases of the target language of Uzbek.

3.2 fMRI Testing of unfamiliar (Uzbek) word learning

To test the hypothesis that vocal imitation learning of a foreign language is mediated by the MNS, we conducted a functional magnetic resonance imaging (fMRI) study in which native subjects practiced the imitation of unknown foreign words. We utilized the adaptation effect (Grill-Spector et al., 2006) to depict practice-related neural substrates. The practice-related decline in neural activity has been shown in both motor learning (Matsumura et al., 2004) and non-motor learning (Raichle et al., 1994). The adaptation or priming method is based on the observation that stimulus repetition induces a decrease in brain activity particularly in regions associated with processing of that stimulus. Although the neural mechanism of the adaptation effect is not completely understood, the reduction has been linked to performance improvements due to repetition, and thus represents some aspect of the learning process, related in particular to procedural memory (Grill-Spector et al., 2006). The same decrease in activity is also observed when the repeated stimuli are not identical but share a common property, allowing the neural correlates of that property or related process to be extracted. Thus, by varying the property that is repeated, it is possible to identify specific neural representations. Priming designs have been widely utilized to depict the neural substrates of specific psychological processes such as visual object perception (Grill-Spector and Malach, 2001), reading (Dehaene et al., 2004), number representation (Naccache and Dehaene, 2001; Piazza et al., 2004), sentence comprehension (Dehaene-Lambertz et al., 2006), and motor processes (Dinstein et al., 2007).

During the present fMRI study, audio-visual movie clips of an individual speaking an Uzbek word were repeatedly shown to the participants, who were naïve to the Uzbek language, followed by instructions to imitate or not to imitate the speaker. In this setting, the perception of action (in this case, articulatory action) should activate the neural structures involved in the motor planning of the actions to be imitated (Chaminade and Decety, 2001). If vocal imitation learning of foreign words presented audio-visually was mediated by the MNS, regions of the motor system would show adaptation effects during both imitation and observation. In particular, we hypothesized that effects of imitation learning would be observable in Broca’s area, based on previous work suggesting this region “primes” the primary motor cortices in response to heard speech (Watkins and Paus, 2004). In contrast, constant activation would be observed in the areas involved in auditory perception and speech vocalization.

3.2.1 Methods

3.2.1.1 Subjects

In total, 21 healthy volunteers (10 male and 11 female; mean age = 19.38 years; age range = 18–22 years) participated in this experiment. Three subjects’ data were excluded due to technical difficulties and poor performance during the fMRI session. Hence, the data from 18 subjects (9 men; mean age = 19.6 years; age range = 18–22 years) were included in the analyses. All subjects were right-handed according to the Edinburgh handedness inventory (Oldfield, 1971), and were native Japanese speakers. No participant had ever been exposed to the Uzbek language, or had a history of neurological or psychiatric illness. The protocol was approved by the ethical committee of the National Institute for Physiological Sciences, Japan, and all subjects gave their written informed consent to participate.

3.2.1.2 Stimulus Preparation

Digital recordings were made using a video-recorder (Sony, Tokyo, Japan), and were edited using Adobe Premiere software (Adobe, San Jose CA, USA). The recordings were of a male speaker pronouncing Uzbek words. We generated 30 video clips (23–34 frames at 33.3 ms/frame), each of which included 1 Uzbek word. The 30 movie clips were divided into 2 groups (stimulus sets A and B). Each group consisted of 15 words and their average frame rate was 25.6 (0.853 s). Additionally we made a video clip (40 frames at 33.3 ms/frame) of the still face of the speaker without any voice or sound.

3.2.1.3 Experimental Design

The subjects learned Uzbek words by vocal imitation during an event-related fMRI session that consisted of 4 runs. Each run included 3 types of event condition: imitation, observation, and baseline. In the imitation condition, each trial was 5 s in duration and was synchronized to the echo-planar imaging (EPI) scanning: functional images were obtained during the first 2 s followed by a 3 s silent period (Fig. 1). During the 2 s active-scanning period, a black screen without any cue signal was presented. In the 3-s silent period, a video clip of a stimulus (e.g., A) was presented for 1.55 s, and then replaced by a blue crosshair cue that prompted vocal repetition for 1 s. Subjects were required to complete the vocalization of the Uzbek word before the crosshair disappeared, so that the collected voice recording did not overlap with the scanner noise. This was followed by a black screen for 0.45 s. The inter-trial interval (ITI) was fixed at 5 s. Therefore, both stimulus presentation and repetition were performed during the 3-s silent period. Fifteen words were presented twice each; hence, 30 trials were included in each condition. The observation condition was identical to the repetition condition except that a video clip of the stimulus was shown followed by a green crosshair to prompt subjects to stay still. The baseline control condition was identical to the non-repetition condition, except that a video clip of a still face without vocalization was presented, and only 15 trials were completed. Each run consisted of 2 repetition conditions of 15 words (stimulus A), 2 observation conditions of 15 words (stimulus B), and 1 baseline condition of 15 still faces. Each run took a total of 6.25 min and included 75 trials. Each subject completed 4 runs, and within each run the order of stimuli presentation was pseudo-randomized. Thus, each Uzbek word was presented 8 times per subject. We adopted a rapid event-related design, in which the efficiency of the design was highly dependent on the temporal pattern of the stimulus presentations. The detailed method used to obtain a high-efficiency design has been described in previous studies (Sadato et al., 2005; Saito et al., 2005; Morita et al., 2008). To counterbalance for possible effects of crosshair color, half of the subjects (5 men and 4 women) were instructed to repeat the presented words in the presence of the blue crosshair and to stay still in the presence of the green crosshair, as mentioned above, and the opposite instructions were given to the other half of the subjects (4 men and 5 women).

3.2.1.4 Stimulus Presentation and Recording

All stimuli were presented using Presentation software (Neurobehavioral Systems, Albany, CA, USA) on a microcomputer (Dimension 8200; Dell Computer Co., TX, USA). Using a liquid crystal display (LCD) projector (DLA-M200L; Victor, Yokohama, Japan). Visual stimuli were projected onto a half-transparent viewing screen located behind the head coil, and the subjects viewed the stimuli through a mirror. The auditory stimuli were presented through MRI-compatible headphones (Hitachi, Yokohama, Japan). For each subject, the volume of the sound was adjusted to an appropriate level (approximately 50 dB) for task execution. Subjects' vocalizations were digitally recorded via a MR-compatible microphone-recording system (FORMI II; Optoacoustics Ltd., Or-Yehuda, Israel).

3.2.1.5 MRI Scanning

A time-course series of 80 volumes was acquired using T2*-weighted gradient-echo EPI sequences with an Allegra 3 Tesla MR imager (Siemens, Erlangen, Germany). Each volume consisted of 34 slices (slice thickness = 4.0 mm) with a 0.5-mm gap to cover the entire cerebral and cerebellar cortices. Oblique scanning was used to exclude the eyeballs from the images. To minimize the effects of the noise of the image acquisition during the presentation of the task stimuli and the verbal repetition, we adopted a "sparse sampling" technique. The time interval between 2 successive acquisitions of the same slice was 5000 ms with a flip angle (FA) of 85° and an echo time (TE) of 30 ms. The cluster volume acquisition time was 2000 ms, leaving a 3000-ms silent period. The field of view (FOV) was 192 mm and the in-plane matrix size was 64 × 64 pixels. For anatomical reference, T1-weighted magnetization-prepared rapid-acquisition gradient-echo (MPRAGE) images (repetition time [TR] = 2500 ms; TE = 4.38 ms; FA = 8°; FOV = 230 mm; matrix size = 256 × 256 pixels; slice thickness = 1 mm; a total of 192 transaxial images) were collected at the same positions as the echo-planar images.

3.2.1.6 Post-MRI Test: Fluency Evaluation

After the fMRI experiment, all of the video clips (except for the still face) that had appeared in the fMRI session were presented to the participants, and the subjects were asked to imitate them. The procedure was similar to the tasks during the fMRI experiment, with the exception that the color of the cross hair changed to red, requiring subjects to vocalize the word. Thus, subjects vocalized both observed and imitated Uzbek words that were presented during fMRI scanning. Subjects' vocalizations were digitally recorded using a microphone system (FORMI II).

3.2.1.7 Performance Analysis

Measurements of the Reaction Time

The reaction time (RT) and execution time (ET) of the imitated utterances, collected both during fMRI and during post-fMRI test, was analyzed with voice-analyzing software (SUGI Speech Analyzer, Animo Co., Yokohama, Japan) (Fig. 4). Statistical analysis with a repeated measures ANOVA, with repetition and items as within-subject factors, was performed by SPSS 10.0J software (SPSS Japan Inc., Tokyo, Japan).

3.2.1.8 Image Data Analysis

Preprocessing

The first 2 volumes from each fMRI session were discarded, to allow for stabilization of the magnetization, and so 78 volumes remained per session. In total, 312 volumes per subject were included in the analysis. The data were analyzed using Statistical Parametric Mapping software (SPM5; Wellcome Trust Centre for Neuroimaging, London, UK; Friston et al., 2007) implemented in MATLAB (Mathworks, Natick, MA, USA). The EPI images were realigned for motion correction, spatially normalized into Montreal Neurological Institute (MNI) stereotaxic space, and smoothed with an isotropic Gaussian kernel of 8 mm full width at half maximum in the x, y, and z axes.

3.2.1.9 Statistical Analysis

Statistical analysis was conducted at two levels. First, individual task-related activation was evaluated. Second, individual data were summarized and incorporated into a random-effects model to make inferences at a population level (Friston et al., 2007). In the individual analyses, the signal time course for each subject was modeled including each of the conditions (imitation, observation, and control) × repetition (2/run) using a delta function convolved with a hemodynamic response function, run effect, and high-pass filtering (128 s). Grand mean scaling was performed instead of proportional scaling. To test hypotheses about regionally-specific trial effects, the estimates for each model parameter were compared with the linear contrasts. The predefined contrasts for individual analysis are shown in Table 1.

In the group analysis using the random-effects model, the weighted sum of the parameter estimates in the individual analysis constituted the “contrast” images (Friston et al., 2007), which represented the normalized task-related increment of the MR signal at each repetition in each subject. Using the contrast images, a 2 (imitation vs. observation) × 8 (repetition) factorial design incorporating the subject effects was performed for every voxel within the brain, in order to obtain population inferences. To evaluate the constant activation of the execution-related activation versus the observation condition at the same points in the repetition order, conjunction analysis was performed with contrasts #17 to #24 (Table 2).

To evaluate the repetition effect, the contrasts shown in Table 2 (#25 to #27) were utilized. To depict the repetitive suppression effect common to both observed and imitated movement, we searched for a significant linear decrease across the repetitions, irrespective of the observe or imitate condition (#27), within the volume defined by the intersection of the areas showing a linear decrease during the execution condition (#25) and those showing a linear decrease during the observation condition (#26). Statistical significance was set at $p < 0.05$, corrected for multiple comparisons at the cluster level (Friston et al., 1996).

3.2.2. Results

3.2.2.1 Performance

The RTs of imitated words decreased gradually as the subjects repeated them (Fig. 5). A repeated-measures ANOVA showed a significant main effect of repetition (F(1397.118, 6.128) = 18.594, $p < 0.001$, Greenhouse-Geisser corrected) but no interaction of repetition × word (F(1397.118, 177.704) = 1.167, $p = 0.290$, Greenhouse-Geisser corrected). The ETs of imitated words increased gradually as the subjects repeated them (Fig. 5), such that they became closer to the ETs of the presenter (853 ms ± 94 ms, mean ± standard deviation [S.D.]). A repeated-measures ANOVA showed a significant main effect of repetition (F(1362.125, 6.027) = 10.855, $p < 0.001$, Greenhouse-Geisser corrected). The interaction effect of repetition and word on RT was not significant (F(1362.125, 174.786) = 0.949, $p = 0.665$, Greenhouse-Geisser corrected). In post-fMRI testing, ETs of repeated and observed words were equivalent (Fig. 5). A two-way ANOVA showed that the main effect of repetition versus observation on ET was not significant (F(536, 2) = 0.363, $p = 0.696$). The main effect of word (F(536, 29) = 7.428, $p <0.001$) and the interaction of repetition x word (F(536, 31) = 5.785, $p < 0.001$) were significant.

Neural Correlates of Imitation Learning

Both imitation and observation of the novel speech activated bilateral fronto-temporal areas (Fig. 6). In the first trial, the contrast of the imitation and observation conditions showed activation in the bilateral primary sensorimotor cortex (SM1), superior temporal gyrus (STG), insula, pre-supplementary motor area (pre-SMA), anterior cingulate gyrus (ACG), and left globus pallidus (Table 3). The extent of the activation decreased as the word was repeated (Fig. 6). The contrast of imitation with observation showed relatively constant activation across repetitions centered in the bilateral primary motor areas (Fig. 6).

The repetitive suppression effect common to both observed and imitated movement was analyzed using conjunction analysis, and showed effects extending from the left ventral premotor cortex (PMv; Brodmann's area (BA) 6) to Broca's area (BA 44) (Fig. 7, Table 4). The effect of imitation irrespective of repetition was evaluated by conjunction analysis with the contrast of the imitation condition and the observation condition at each repetition. Symmetrical activation in the bilateral primary motor cortex was found (Fig. 8, Table 5).

3.2.3 Discussion

3.2.3.1 Performance

Decreasing RTs indicated faster preparation of speech responses, and thus represented an effect of learning (Gupta and Cohen, 2002). The ETs of the subjects were initially shorter than the ETs of the presenter; however, with repetition they gradually increased towards the ETs of the presenter. This indicates that the subjects did not completely imitate all components of the presented stimuli, but rather the degree of the imitation gradually increased, again indicating a learning effect. Importantly, this learning effect was detected even when subjects observed the speech without making overt utterances. This might have been due to mental rehearsal during the observation condition, because the experimental setup required subjects to prepare or plan imitative actions in both the imitation and observation conditions. This finding implies that the learning effect is brought about by both observation and execution.

3.2.3.2 Neural Activation

In the first trial, in contrast to the control condition, the observation condition showed prominent activation in temporal-parietal-frontal regions (Fig. 6). This is consistent with previous studies showing the activation of motor speech areas during speech perception (Wilson and Iacoboni, 2006). Wilson and Iacoboni (2006) suggested that the co-activation of the superior temporal region and the motor speech area represents a sensorimotor transformation during speech perception. During observation, the activation in the fronto-parieto-temporal regions decreased as the trials were repeated, consistent with the observations of Rauschecker et al. (2008), who found decreasing activity in the left premotor cortex, SMA, inferior frontal gyrus (IFG), STG, and cerebellum during repetitive covert imitation of pseudo-words. They suggested that the changes reflect the development of a more efficient representation of the articulation pattern of these novel words in two connected systems: one for perception (left temporal cortex) and one for processing speech output (left frontal cortex).

In the first trial, the contrast of the imitation and the observation conditions showed prominent temporal-parietal-frontal activation. With the exception of the primary motor cortex, regional activation decreased as the trials were repeated. This might indicate that, while auditory verbal feedback of self-generated words is crucial to the acquisition of proper phonation and articulation, it is not necessary once the vocalization is acquired (Hirano et al., 1996). Supporting this, Hirano et al. (1997) found an absence of activity in the STG when normal adult subjects vocalized sentences that were used daily.

Within the areas activated by the first presentations of the words, there was a significant repetitive suppression effect during imitation and observation in Broca’s area, extending to the PMv. This finding suggests that the MNS is related to vocal imitation learning.

Broca's Area: BA 44/45

Cytoarchitectonic areas 44 and 45 are located in the IFG, and occupy the pars opercularis and triangularis (Amunts et al., 1999). It is widely accepted that both of these areas constitute the anatomical correlates of Broca's region (Aboitiz and Garcia, 1997), although macroscopic features are not reliable landmarks of their cytoarchitectonic borders (Amunts et al., 1999). Functionally, it has become clear that, in addition to BA 44 and 45, at least BA 47 (and probably also the ventral part of BA 6) should be included in the left frontal language network (Bookheimer, 2002; Hagoort, 2005). From a functional–anatomical perspective, it therefore makes sense to refer to the left IFG as the language-relevant part of the frontal cortex (Hagoort, 2005).

The left IFG might consist of three different modular regions related to phonology, semantics, and syntax (Bookheimer, 2002). Investigating the comprehension of garden-path sentences, Uchiyama et al. (2008) found that in the left IFG, verbal working memory was located more dorsally (BA 44/45), semantic processing was located more ventrally (BA 47), and syntactic processing was located in between the verbal working memory and semantic processing areas (BA 45). These findings indicate a close relationship between semantic and syntactic processes, and suggest that BA 45 might link verbal working memory and semantic processing via syntactic unification processes. Poldrack et al. (1999b) suggested that phonological processing is automatically engaged during performance of a semantic task, despite the identification of separate regions for phonology and semantics in the IFG.

Working Memory

In the present study, after the presentation of the Uzbek word, the subject had to retain the heard utterance until presentation of the go (imitation) or no-go (observation) signal. Therefore, the imitation condition involved delayed imitation. Delayed imitation is one of the most important capacities for social species, such as humans, that live in groups. Many species are capable of mimicry or immediate imitation of a particular act. However, delayed imitation may be a specifically human achievement (Donald, 1991; Donald, 1993). Imitation depends on three cognitive components (Barkley, 2004): first, the inhibition of pre-potent responses; second, an evolved mental mechanism for carrying past sensory perceptions of others' behaviors forward in time across a delay interval; and third, the capacity to reconstruct motor responses on the basis of remembered actions of others. The latter represent the retrospective and prospective aspects, respectively, of the working memory system. Thus, verbal working memory is an important component of the delayed vocal imitation task performed in the present study.

In Baddeley's (2003) model, verbal working memory involves a "phonological loop", which is considered to be analogous to covert speech. It has been shown that impaired non-word repetition is a reliable marker for children with persistent language impairment (Gathercole and Baddeley, 1990; Bishop et al., 1996). Thus, it is conceivable that verbal working memory, i.e., the ability to hold transient linguistic or non-linguistic mental representations (Johnston, 1994), is important for acquiring linguistic competence. Covert speech is a form of motor imagery of overt speech (Aziz-Zadeh et al., 2005), which is tightly linked to the motor system (McGuigan and Dollins, 1989). Neuroimaging data indicate that the premotor areas (Rizzolatti et al., 2002), and occasionally the primary motor areas, are active during motor imagery (Rueckert et al., 1994; Porro et al., 1996; Dechent et al., 2004). In particular, the motor system and the ventral premotor areas caudal to Broca's area are involved in covert speech, together with the posterior IFG (BA 44) (Paulesu et al., 1993; Nixon et al., 2004 using TMS). Using TMS and a syllable-counting task, Aziz-Zadeh et al. (2005) found that the left PMv and the left IFG are essential to language elaboration even when motor output is not required, because TMS of these areas arrested both covert and overt speech. Region F5 in non-human primates, which is known to be part of the MNS, is sensitive to visual, motor, and auditory components of the same action (Kohler et al., 2002; Keysers et al., 2003). This area is thought to be homologous to BA 44, which is a good candidate for processing the imagery involved in covert speech. The inferior frontal junction (BA 6/44) is related to task preparation (Brass and von Cramon, 2002). Thus, the PMv and Broca's area are related to the planning and preparation of speech movements, and are essential for short-term phonological memory via covert speech (Aziz-Zadeh et al., 2005).

Neural Adaptation and Learning

Neural adaptation and verbal learning have been studied with PET (Raichle et al., 1994) while subjects generated verbs from a visually-presented list of nouns. Just 15 min of practice of generating verbs caused a significant reduction in response times, and the occurrence of stereotyped responses across practice blocks. Correspondingly, task-related activation in the anterior cingulate, left prefrontal and left posterior temporal cortices, and right cerebellum rapidly declined, such that the cortical circuitry used for verbal response selection was indistinguishable from that involved in simple word repetition. As stimulus-response pairings through repeated practice resulted in automaticity, it can be considered the outcome of successful learning. After learning, once initiated, a response can be carried through to completion without further control. Practice allows tasks to be performed using processes that are less attention-demanding and enables a sequence of responses to be produced automatically following stimulus presentation. This learning process is represented by the practice-related reduction of the activity in the network including the left prefrontal cortices (Raichle et al., 1994). This is consistent with the results of the present study. Considering that the PMv and Broca's area are related to the planning and preparation of speech movements (Watkins et al., 2003), the learning process is likely to be related to procedural learning of the vocalizations. Furthermore, the present study has shown that even without overt pronunciation, the left BA 44 and PMv show learning-related effects, as indicated by the reduction in task-related activation in these regions. Thus, the phonological loop, which supports a type of working memory, also supports covert speech or silent shadowing (Murphey, 2001), and is represented by part of the MNS (i.e., the left BA 44 and PMv).

Execution Related Component

Execution versus observation revealed constant activation in the primary motor cortex. As primary motor cortex (area M1) did not show adaptation effects, this activity probably does not relate to learning, at least not during vocal imitation learning (but see Karni et al., 1995). By contrast, M1 showed constant activation during observation without imitation. This finding is consistent with a previous TMS study that showed an increase in M1 excitability when listening to sentences (Watkins et al., 2003). Considering that Broca's area "primes" the motor system in response to heard speech, and thus may operate at the interface of perception and action as part of the MNS (Watkins and Paus, 2004), the present data suggest that it also mediates the learning effect through interaction with the premotor cortices. The interaction of Broca’s area and the premotor regions represents the “motor vocabulary” of speech.

3.2.4 Conclusion for This Experiment

In summary, vocal imitation learning is mediated by the left PMv and Broca’s area, which are the part of the human MNS.

Table 1. Predefined contrasts for individual analysis

	Run	1					2					3					4				
	Condition	Imit		Obs		CTL	Imit		Obs		CTL	Imit		Obs		CTL	Imit		Obs		CTL
	Repetition	1	2	1	2		3	4	3	4		5	6	5	6		7	8	7	8	
1	Imitation x repetition 1	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Imitation x repetition 2	0	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Imitation x repetition 3	0	0	0	0	0	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
4	Imitation x repetition 4	0	0	0	0	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0	0
5	Imitation x repetition 5	0	0	0	0	0	0	0	0	1	0	1	0	0	0	-1	0	0	0	0	0
6	Imitation x repetition 6	0	0	0	0	0	0	0	0	0	1	0	1	0	0	-1	0	0	0	0	0
7	Imitation x repetition 7	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	-1
8	Imitation x repetition 8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	-1
9	Observation x repetition 1	0	0	1	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	Observation x repetition 2	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Observation x repetition 3	0	0	0	0	0	0	0	1	0	-1	0	0	0	0	0	0	0	0	0	0
12	Observation x repetition 4	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0
13	Observation x repetition 5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	-1	0	0	0	0	0
14	Observation x repetition 6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0
15	Observation x repetition 7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	-1
16	Observation x repetition 8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1

Imit: imitation; Obs, observation; CTL, control.

Table 2. Predefined contrasts for random effect model

	condition	Imitation								Observation							
number		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
1	Imitation x repetition 1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Imitation x repetition 2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Imitation x repetition 3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Imitation x repetition 4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
5	Imitation x repetition 5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
6	Imitation x repetition 6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
7	Imitation x repetition 7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
8	Imitation x repetition 8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9	Observation x repetition 1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
10	Observation x repetition 2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
11	Observation x repetition 3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
12	Observation x repetition 4	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13	Observation x repetition 5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
14	Observation x repetition 6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
15	Observation x repetition 7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	Observation x repetition 8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
17	Imitation - Observation at repetition 1	1	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
18	Imitation - Observation at repetition 2	0	1	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
19	Imitation - Observation at repetition 3	0	0	1	0	0	0	0	0	0	0	-1	0	0	0	0	0
20	Imitation - Observation at repetition 4	0	0	0	1	0	0	0	0	0	0	0	-1	0	0	0	0
21	Imitation - Observation at repetition 5	0	0	0	0	1	0	0	0	0	0	0	0	-1	0	0	0
22	Imitation - Observation at repetition 6	0	0	0	0	0	1	0	0	0	0	0	0	0	-1	0	0
23	Imitation - Observation at repetition 7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-1	0
24	Imitation - Observation at repetition 8	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-1
25	Repetitive suppression during Imitation	7	5	3	1	-1	-3	-5	-7	0	0	0	0	0	0	0	0
26	Repetitive suppresion during observation	0	0	0	0	0	0	0	0	7	5	3	1	-1	-3	-5	-7
27	Repetition suppression	7	5	3	1	-1	-3	-5	-7	7	5	3	1	-1	-3	-5	-7

Table 3. Imitation - Observation on the first trial

cluster P	cluster size	voxel (FDR)	voxel Z	x	y	z	location
<0.001	6210	<0.001	7.47	-44	-12	32	SM1
		<0.001	6.47	-62	-8	-2	STG
		<0.001	5.9	-44	4	-6	insula
<0.001	5933	<0.001	7.43	50	-8	34	SM1
		<0.001	6.55	42	22	-6	insula
		<0.001	6.42	54	-4	22	SM1
		<0.001	6.05	58	-12	4	STG
<0.001	1317	<0.001	4.83	0	16	46	pre-SMA
		<0.001	4.67	4	24	52	pre-SMA
		<0.001	4.53	4	20	30	ACG
0.005	351	<0.001	4.27	-12	6	-2	globus pallidus

Significant activation during the first imitation trial compared with the first observation trial. P values are corrected for multiple comparisons at cluster or voxel level (FDR corrected). ACG, anterior cingulate gyrus;SM1, the primary sensorimotor cortex; pre-SMA, pre-supplementary motor area; STG, superior temporal gyrus.

Table 4. Repetitive suppression during both Execution and Observation conditions over 8 repetitions

cluster P	cluster size	voxel (FDR)	voxel Z	x	y	z	location	BA	% prob
0.002	411	0.004	5.23	-56	-4	42	PMv	6	100
		0.006	4.67	-42	4	12	IFG	44	20
		0.007	4.52	-54	12	30	IFG	44	60

Significant linear decrease irrespective of conditions was searched within the volume defined by the linear decrease during execution condition and observation condition ($P<0.05$ corrected at cluster level). P values are corrected for multiple comparisons at cluster or voxel level (FDR corrected). IFG, inferior frontal gyrus;PMv, ventral premotor cortex. % Probability of the BA was calculated with Anatomical Toolbox (Eickhoff et al. 2005).

Table 5. Constant activation during Execution compared with Observation conditions at each repetition

cluster P	cluster size	voxel P	voxel Z	x	y	z	side	location	BA	% prob
0.001	433	<0.001	5.69	-46	-10	30	L	Precentral gyrus	3a	60
									4p	50
0.002	409	<0.001	5.73	48	-8	32	R	Precentral gyrus	4p	70
									3a	40

Conjunction analysis (with conjunction null hypothesis) of the contrasts 17 - 24 (Table 2)

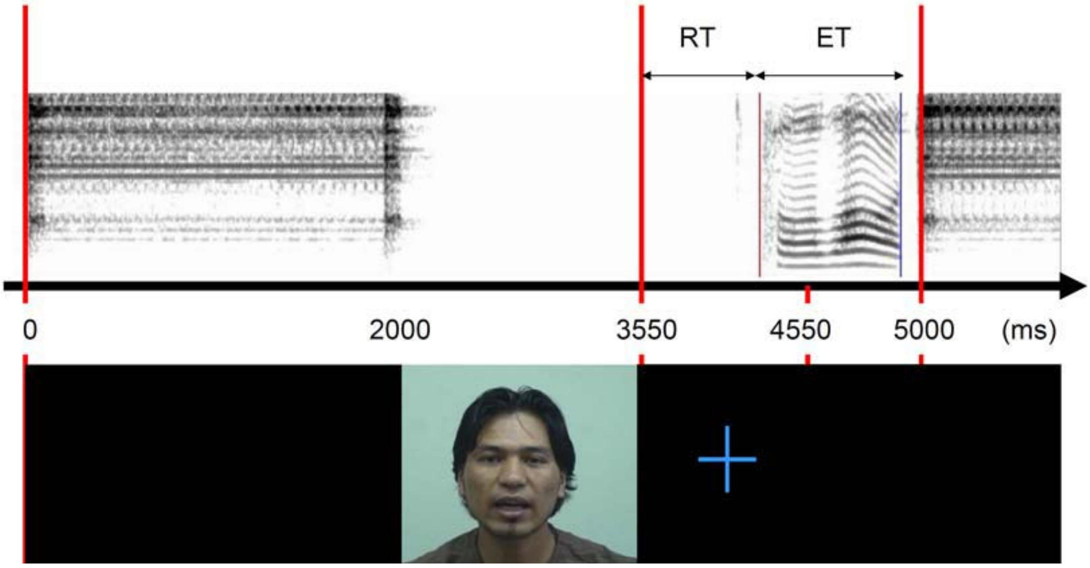
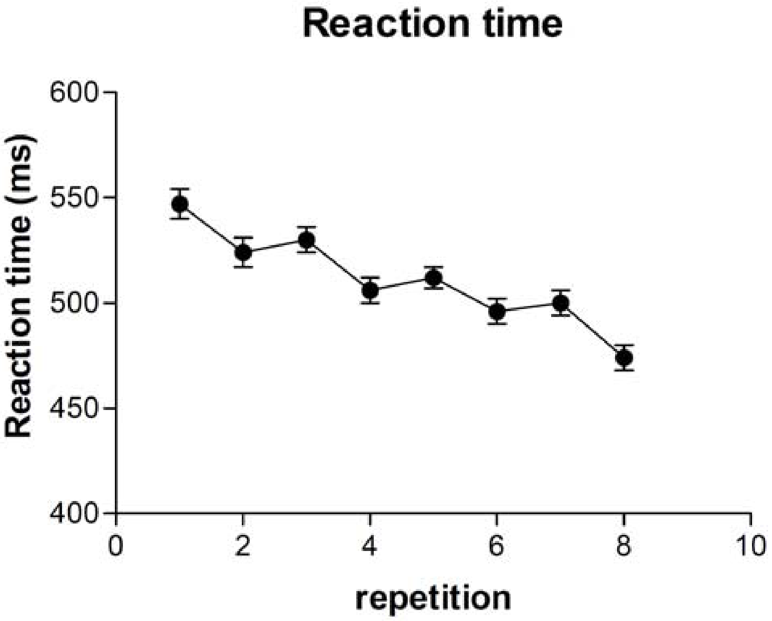


Fig. 4 Time line of a trial.

In the top raw is plotted the power spectrum of the recorded scanning noise (from 0 to 2000 ms) followed by the silent period (from 2000 to 5000 ms) when subjects had to repeat or observe the presented vocalization. Video clip of the Uzbek speech was presented for 1550 ms (from 2000 to 3550 ms), followed by blue cross for 1000 ms which prompted either imitation or observation (in this example, prompting imitation). RT, reaction time; ET, execution time.



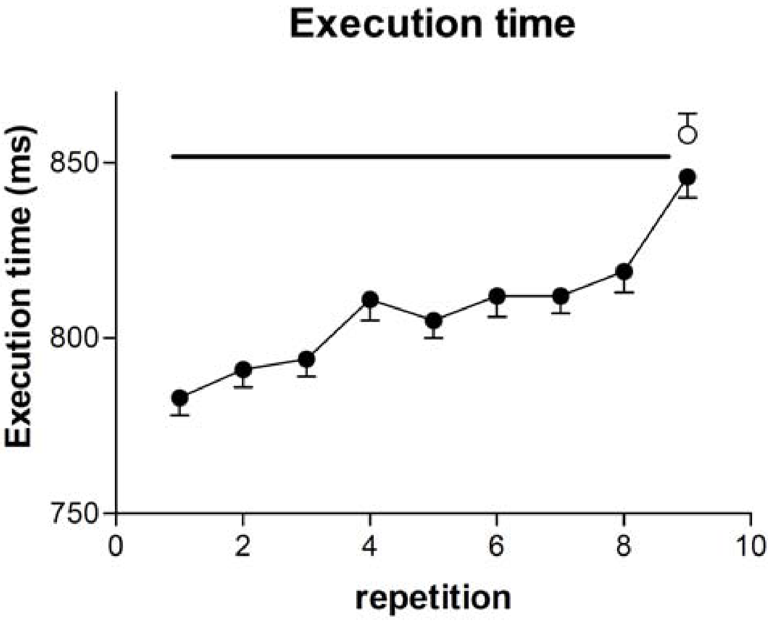


Fig. 5 (Top) Reaction time of the vocalization of the repeatedly presented Uzbek words during functional MRI. Error bars indicate 1 standard error of mean. (Bottom) Execution time of the vocalization of the repeatedly presented Uzbek words during functional MRI (open circle). Horizontal bar indicates the execution time of the presenter. The post-fMRI test (9th repetition) showed that the execution time of

the repeatedly vocalized words and those observed without vocalization (closed circle) show similar execution time. Error bar indicates standard error of mean.

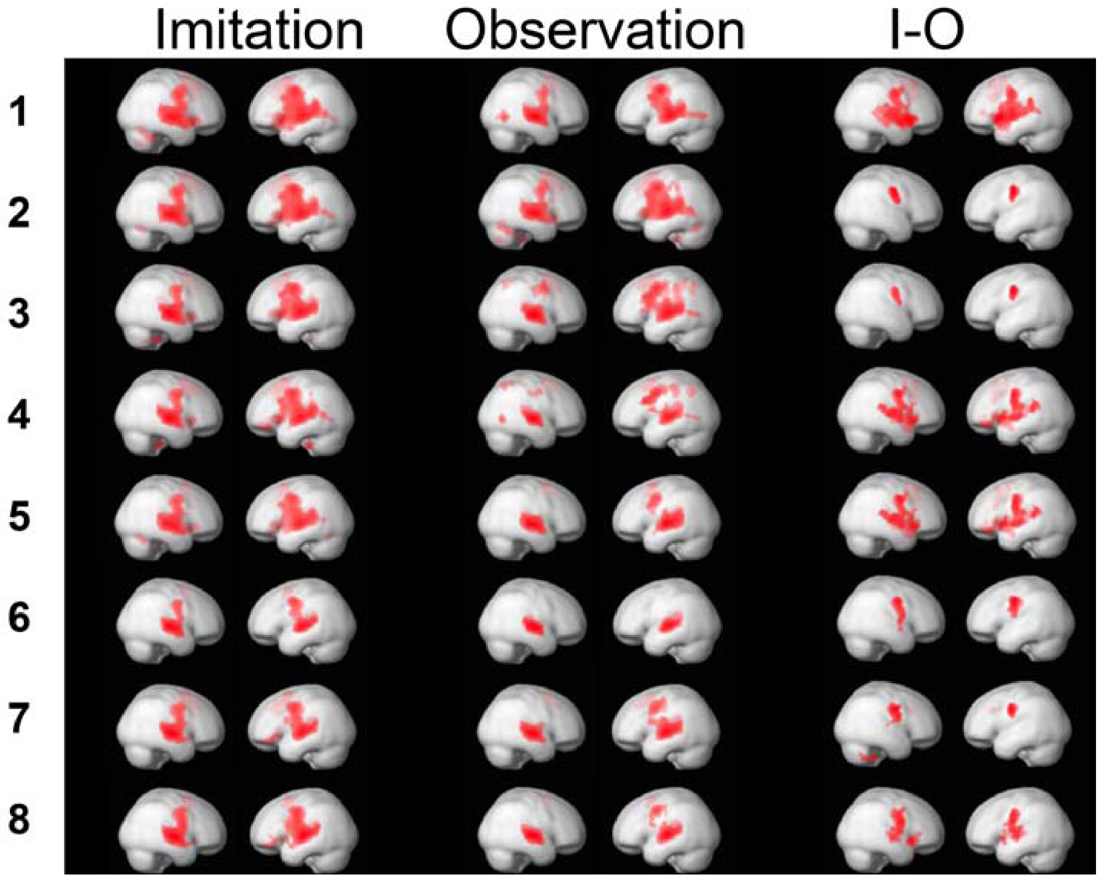


Fig. 6 Task related activation of each repetition. Imitation related activations (left column), observation-related activations (middle column), and their contrast (right column) were superimposed on the averaged surface rendered anatomical MRI. *P*<0.05 corrected for multiple comparisons at cluster level.

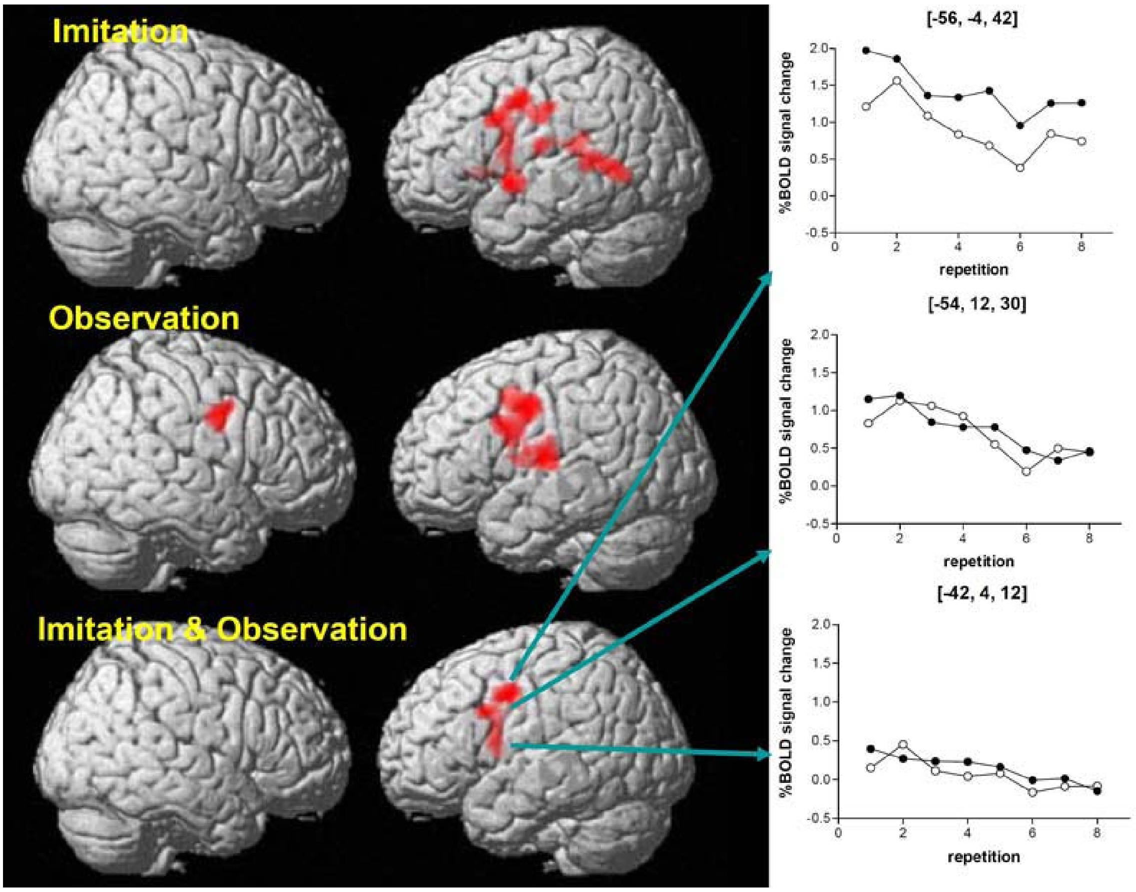


Fig. 7 The attenuated task-related activation during repeated overt pronunciation (middle column, top), observation without overt pronunciation (middle column, 2nd row), and their intersection.

The activated foci were superimposed on the surface-rendered 3D high resolution MRI implemented in SPM5. $P < 0.05$, corrected for multiple comparison at cluster level. Right column shows the % BOLD signal change across repetitions in the PMv (top) and BA44 (2nd and third row).

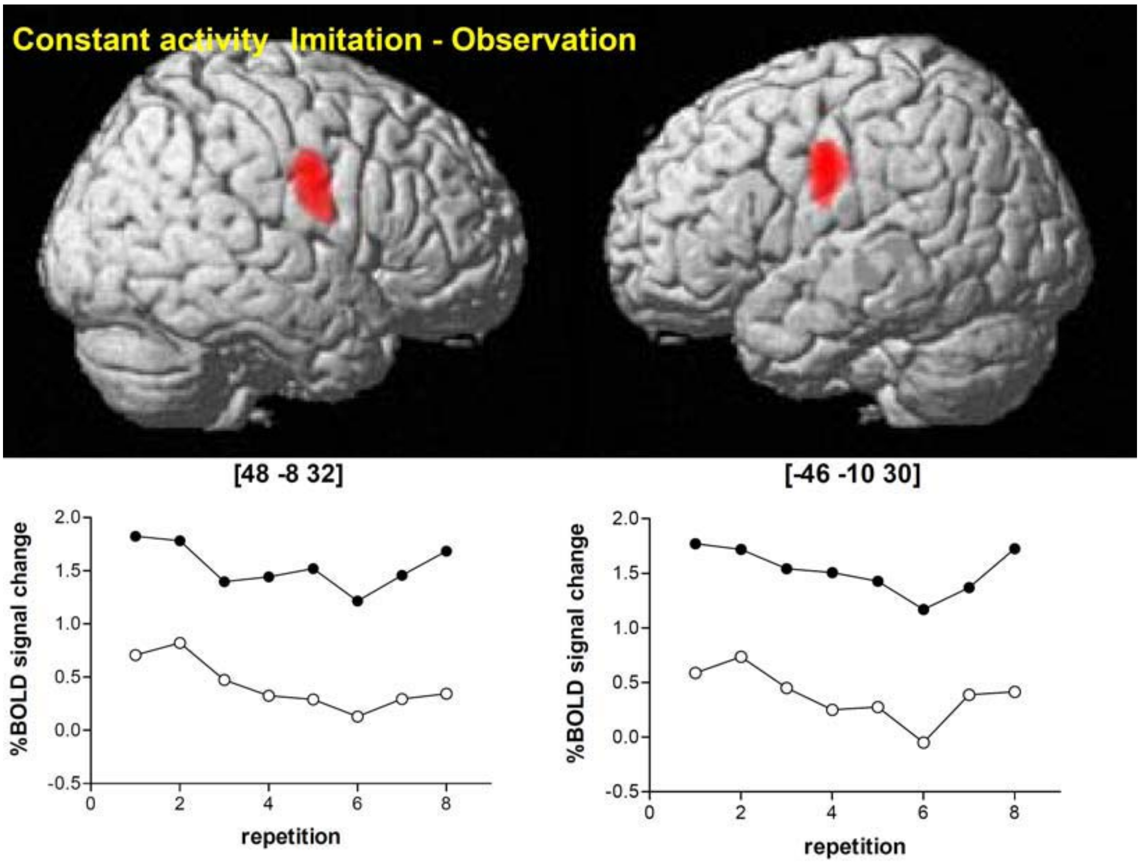


Fig. 8 Constant activity across repetitions by means of the contrast Execution – Observation.

The activated foci were superimposed on the surface-rendered 3D high resolution MRI implemented in SPM5. *P*< 0.05, corrected for multiple comparison at cluster level. Bottom row shows the % BOLD signal change across the repetition in the right SM1 (left) and left SM1 (right).

3.3 Psychophysical Testing of Unknown Medical Words

In order to investigate the effectiveness of overt/covert repeating when learning foreign language words, according to the fMRI results, we conducted psychophysical tests of learning of unknown technical medical words in healthy adults.

3.3.1 Subjects and Methods

In total, 52 native Japanese speakers, 34 male and 18 female, participated in this study. The subjects were in their first year college education majoring in medicine at the medical college (Osaka, Japan). None of them knew the target medical words.

A total of 12 medical words were employed, three used for each of four training days. Participants were divided into two groups. Each group (Group Ⅰ and Group Ⅱ) consisted of 17 male and 9 female students. The tasks were administered four times, on Oct. 14, Oct. 21, Nov. 7, and Nov. 14 in 2009. The stimulus words were presented to classes of participants using a video which produced a Japanese teacher’s model voice with his facial expressions. The video and corresponding auditory stimuli were presented at the front of the class. In the videos, the teacher presented three target words and their meanings in Japanese. Group Ⅰ were instructed to repeat the three target words five times after each of three presentations, imitating the model voice as much as possible. Group Ⅱ were instructed to watch and listen to the model voice without repeating the three target words. 80 minutes later after completion of this task, all participants were instructed to repeat after the teacher following one further presentation of the model sound, to write the meaning of the three target words in Japanese and identify them by their correspondence to different parts of the body.

On November 21, a week after the fourth training session, all participants were instructed to repeat each of the target words after a single video presentation of each word, indicate the meaning of the three target words, and identify them by their correspondence to different parts of the body.

On December 12 and January 9, all participants completed delayed recognition tests, where they were instructed to identify the twelve target words among 32 distracters to check if they learnt each word or not.

The schedule of the experiments is shown in Fig. 9.

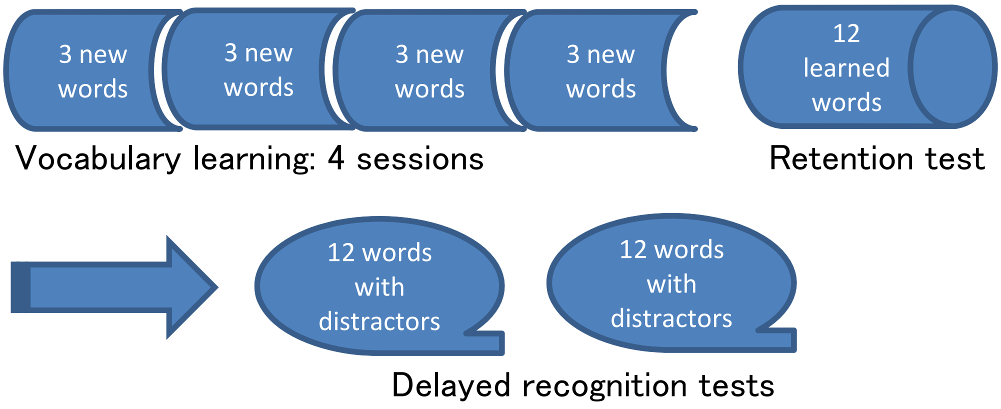


Fig.9 The schedule of the experiments.

3.3.2 Results

We first examined whether there were any group differences in the learning of novel medical words. Table 6 shows descriptive statistics of Group I and Group II by target word on the first training session. The scoring system was as follows:

Five points: both the meaning and the part of the body are accurate.

Four points: only the meaning is accurate.

Three points: only the part of a body is accurate.

Two points: the meaning is not exact but still acceptable.

One point: Incorrect both in meaning and part of body.

Table 6. Descriptive Statistics of the Vocabulary Test results for Group Ⅰ and Group Ⅱ by Target Word (cervical, sterna, cephalic)

	Group	N	Mean	SD
Cervical	G Ⅰ	21	4.5714	1.07571
	G Ⅱ	22	4.4545	1.18431
Sternal	G Ⅰ	21	4.0952	1.33809
	G Ⅱ	22	3.6364	1.61968
Cephalic	G Ⅰ	21	4.0000	1.44914
	G Ⅱ	22	4.1818	1.40192

An examination of the means and standard deviations (SDs) for the vocabulary test suggested that scores of Group Ⅰ and Group Ⅱ followed a normal distribution.

Independent samples t-tests were conducted in order to compare the performance on the vocabulary test in Group Ⅰ and Group Ⅱ for this first training session. (see Table 7).

Table 7. Results of t-tests comparing scores on the Vocabulary Test for Group Ⅰ and Group Ⅱ

Vocabulary Test	Levene's Test			Difference of Two Samplers		
	Group	*F*-value	Sig.	*t*-score	df	Sig.
Cervical	G Ⅰ G Ⅱ	.463	.500	.338	41	.737
Sternal	G Ⅰ G Ⅱ	4.597	.038	1.015	41	.316
Cephalic	G Ⅰ G Ⅱ	.612	.439	-.418	41	.678

Table 7 shows that there are no significant differences between Group Ⅰ and Group Ⅱ in the vocabulary test for "Cervical" (*t*=.338, df=41, *p*=.737), "Sternal" (*t*=1.015, df=41, *p*=.316 or "Cephalic" (*t*=-.418, df=41, *p*=.678).

Table 8 shows descriptive statistics of Group Ⅰ and Group Ⅱ performance in the second training session.

Table 8. Descriptive Statistics of the Vocabulary Test results for Group Ⅰ and Group Ⅱ by the Target Word (umbilical, axillary, brachial)

	Group	N	Mean	SD
Umbilical	G Ⅰ	23	4.6087	1.07615
	G Ⅱ	24	3.7917	1.64129
Axillary	G Ⅰ	23	4.4348	1.03687
	G Ⅱ	24	4.5417	1.10253
brachial	G Ⅰ	23	4.3478	1.30065
	G Ⅱ	24	3.4583	1.74404

An examination of the means and standard deviations (SDs) for vocabulary scores on the second session also demonstrates that this data follows a normal distribution for Group Ⅰ and Group Ⅱ.

Independent samples t-tests were conducted in order to compare the performance on the vocabulary test between Group Ⅰ and Group Ⅱ in the second session. (see Table 9).

Table 9. Results of t-tests comparing scores on te Vocabulary Test for Group Ⅰ and Group Ⅱ

Vocabulary Test	Levene’s Test			Difference of Two Samples		
	Group	*F*-value	Sig.	*t*-score	df	Sig.
Umbilical	G Ⅰ G Ⅱ	16.521	.000	2.026	45	.049
Axillary	G Ⅰ G Ⅱ	.126	.724	-.342	45	.734
Brachial	G Ⅰ G Ⅱ	13.701	.001	1.988	45	.053

Table 9 shows that there are no significant differences between Group Ⅰ and Group Ⅱ in the vocabulary test scores for “Umbilical” (*t*=2.026, df=45, *p*=.049), “Axillary” (*t*=-.342, df=45, *p*=.734), or “Brachial” (*t*=1.988, df=45, *p*=.053).

Table 10 shows descriptive statistics of Group Ⅰ and Group Ⅱ by the target word on the third training session.

Table 10. Descriptive Statistics of the Vocabulary Test results for Group Ⅰ and Group Ⅱ by the Target Word (pelvic, coxal, and carpal)

	Group	N	Mean	SD
Pelvic	G Ⅰ	12	4.7500	.86603
	G Ⅱ	22	4.7273	.88273
Coxal	G Ⅰ	12	4.7500	.86603
	G Ⅱ	22	4.2727	1.27920
Carpal	G Ⅰ	12	4.6667	.88763
	G Ⅱ	22	4.8636	.63960

An examination of the means and standard deviations (SDs) for the vocabulary tests shows that scores for Group Ⅰ and Group Ⅱ followed an approximately normal distribution.

Independent samples t-tests were conducted in order to compare the performance on the vocabulary test between Group Ⅰ and Group Ⅱ (see Table 11).

Table 11. Results of t-tests comparing scores on the Vocabulary Test for Group Ⅰ and Group Ⅱ

Vocabulary Test	Levene's Test			Difference of Two Samples		
	Group	*F*-value	Sig.	*t*-score	df	Sig.
Pelvic	G Ⅰ G Ⅱ	.002	.962	.072	32	.943
Coxal	G Ⅰ G Ⅱ	5.771	.002	1.290	32	.207
Carpal	G Ⅰ G Ⅱ	1.795	.190	-.747	32	.460

Table 11 shows that there are no significant differences between Group Ⅰ and Group Ⅱ in the vocabulary test for "Pelvic" (*t*=.072, df=32, *p*=.943), "Coxal" (*t*=1.290, df=32, *p*=.207), and "Cephalic" (*t*=-.747, df=32, *p*=.460).

Table 12 shows descriptive statistics of Group Ⅰ and Group Ⅱ by the target word on the fourth training session.

Table 12. Descriptive Statistics of the Vocabulary Test results for Group Ⅰ and Group Ⅱ by the Target Word (tarsal, pattelar, crural)

	Group	N	Mean	SD
Tarsal	G Ⅰ	21	3.9524	1.53219
	G Ⅱ	17	3.9412	1.71284
Pattelar	G Ⅰ	21	4.0000	1.44914
	G Ⅱ	17	4.8235	.72761
Crural	G Ⅰ	21	3.7619	1.48003
	G Ⅱ	17	3.8824	1.45269

An examination of the means and standard deviations (SDs) for the vocabulary test suggested that scores of Group Ⅰ and Group Ⅱ followed an approximately normal distribution .

Independent samples t-tests were conducted in order to compare the performance on the vocabulary tests between Group Ⅰ and Group Ⅱ .

(see Table 13).

Table 13. Results of t-tests comparing scores on the Vocabulary Test for Group Ⅰ and GroupⅡ

Vocabulary Test	Levene's Test			Difference of Two Samples		
	Group	*F*-value	Sig.	*t*-score	Df	Sig.
Tarsal	G Ⅰ	.217	.644	.021	36	.983
	G Ⅱ					
Patellar	G Ⅰ	30.102	.000	-2.274	36	.030
	G Ⅱ					
Crural	G Ⅰ	,456	.504	-.251	36	.803
	G Ⅱ					

Table 13 shows that there were no significant differences between Group Ⅰ and GroupⅡ in the vocabulary tests for "Tarsal" (*t*=.021, df=36, *p*=.983), "Coxal" (*t*=-2.274, df=36, *p*=.030), and "Cephalic" (*t*=-.251, df=36, *p*=.803).

We then examined whether there were any group differences in word retention a week after 4 training sessions. For this analysis fewer participants' data were included, due to their absences from class. 27 out of 52 participants' data were excluded in total. Therefore, analyses were based on data collected from 25 participants.

Table 14 shows descriptive statistics for performance of Group Ⅰ and GroupⅡ in the word retention test.

Table 14. Descriptive Statistics of the Vocabulary Test results for Group Ⅰ and Group Ⅱ by the Target Word (Patellar, Crural, Tarsal, Cephalic, Cervical, Sternal, Carpal, Pelvic, Coxal, Axillary, Umbilical, and Brachial)

Target Word	Group	N	Mean	SD
Patellar	G Ⅰ	10	3.9000	1.44914
	G Ⅱ	15	4.0667	1.48645
Crural	G Ⅰ	10	2.7000	1.25167
	G Ⅱ	15	3.1333	1.64172
Tarsal	G Ⅰ	10	3.2000	1.54919
	G Ⅱ	15	3.7333	1.62422
Cephalic	G Ⅰ	10	3.2000	1.54919
	G Ⅱ	15	3.2667	1.79151
Cervical	G Ⅰ	10	4.1000	1.44914
	G Ⅱ	15	3.9333	1.57963
Sternal	G Ⅰ	10	2.9000	1.44914
	G Ⅱ	15	2.7333	1.48645
Carpal	G Ⅰ	10	3.2000	1.75119
	G Ⅱ	15	3.8667	1.68466
Pelvic	G Ⅰ	10	4.4000	.96609
	G Ⅱ	15	4.6000	1.05560
Coxal	G Ⅰ	10	3.1000	1.66333
	G Ⅱ	15	2.4667	1.30201
Axillary	G Ⅰ	10	3.2000	1.61933
	G Ⅱ	15	4.1333	1.50555
Umbilical	G Ⅰ	10	3.3000	1.49443
	G Ⅱ	15	2.5333	1.40746
Brachial	G Ⅰ	10	3.2000	1.54919
	G Ⅱ	15	2.9333	1.53375

An examination of the means and standard deviations (SDs) for the vocabulary test suggested that these followed a normal distribution for both Group Ⅰ and Group Ⅱ.

Independent samples t-tests were conducted in order to compare the performance on the vocabulary retention test between Group Ⅰ and Group Ⅱ (see Table 15).

Table 15. Results of t-tests comparing scores on the Vocabulary Retention Test for Group Ⅰ and Group Ⅱ

Vocabulary Test	Levene's Test		Difference of Two Samplers		
	F-value	Sig.	*t*-score	df	Sig.
Patellar	.086	.772	-.277	23	.784
Crural	4.142	.054	-.707	23	.487
Tarsal	.281	.601	-.819	23	.421
Cephalic	.667	.422	-.096	23	.924
Cervical	.477	.497	.267	23	.792
Sternal	.003	.956	.277	23	.784
Carpal	.146	.706	-.954	23	.350
Pelvic	.070	.794	-.480	23	.636
Coxal	3.442	.076	1.067	23	.297
Axillary	.380	.544	-1.474	23	.154
Umbilical	.790	.384	1.302	23	.206
Brachial	.100	.754	.424	23	.675

Table 15 shows that there were no significant differences between Group Ⅰ and Group Ⅱ in the vocabulary retention test for the 12 target words.

Finally, we examined whether there were group differences on recognition of target words approximately one month and two months after the fourth word-training session. The test consisted of presentation of the 12 target words and 20 distracters presented in the same way as previously, and participants were required to indicate whether each word had been presented during their training sessions. For scoring purposes, 1 point was awarded for each correct response, up to a possible total score of 32 points.

In this analysis, the number of participants included was again reduced due to student absences from class. 29 out of 52 participants' data were excluded in total. Therefore, analyses were based on data collected from 23 participants who attended all the lessons.

Table 16 shows descriptive statistics for Group Ⅰ and Group Ⅱ for the first delayed recognition testing session.

Table 16. Descriptive Statistics for the Word Recognition Tests for Group Ⅰ and Group Ⅱ, first and second delayed retention testing sessions

	Group	Mean	SD
1st delayed Test	G Ⅰ	8.000	1.773
	G Ⅱ	8.267	2.120
2nd delayed Test	G Ⅰ	8.000	1.414
	G Ⅱ	7.667	2.257

An examination of the means and standard deviations (SDs) for the vocabulary test suggested that data follow a normal distribution for both Group Ⅰ and Group Ⅱ.

To investigate inter-group effects on the delayed word recognition tests, a repeated measures ANOVA was conducted on the results of the word recognition tests. This showed no significant main effect of word recognition test (F (1,21) = .551, p = .466), and the group(F (1,21) = .002, p = .966) . An interaction of word recognition x group was not significant (F (1,2) = .551, p = .466). These ANOVAs show that the task differences neither interfere with nor promote word retention one or two months after training.

Summary of participant comments

"I found myself imitating and repeating the target word after teacher's sound model even though I was told not to repeat.”

"I was not comfortable restraining myself from imitating and repeating the target word.”

”I recall more words using the visual imagery strategy.”

“When I try to recall the target word, the teacher’s face and voice popped up.”

“When I try to recall the target word, I could see mentally my own face and voice saying the test words..”

3.3.3 Discussion

The results of this psychophysical experiment testing the learning of unknown medical words support the results of the fMRI study testing learning of Uzbek words. During fMRI, post-fMRI test showed that the execution time of repeated and observed words were equivalent. Similarly, in this experiment, both overt and covert imitation groups showed no significant difference in retention tests given immediately after the class, nor a week after the semester session, nor in the delayed recognition test given after a further one month and two month period. This implies that the learning effect was produced even when subjects observed with instruction to not overtly imitate the teacher. The comment by a student that she was not comfortable restraining herself from imitating and repeating the target word demonstrates the positive influences of covert speech. This suggests that covert speech the form of a motor image of overt speech, tightly linked to the motor system.

This experiment also supports the utility of vocabulary leaning models which emphasize attention, imitation and repetition.

Chapter 4

Useful Vocabulary Learning Models

4.1 Implication for a Longitudinal Vocabulary Learning Model

4.1.1 AIR Model Emphasizing MN Importance at Initial Stages

Practically all second and foreign language learners and their teachers are well aware of the fact that learning a new language involves the learning of large numbers of words. It is not surprising that many learners are apprehensive when faced with such an enormous task and teachers as well as learners have always shown a keen interest in finding out how words can best be learned. For many years, instructional practice has been based solidly on the view supported by psychologists, that elaboration of features of new words promotes their retention (Anderson, 1995; Baddeley, 1997). This means that the more attention that is paid to the formal and semantic aspects of words and the richer the associations that are made with existing knowledge, the higher the chances are that the new information will be retained.

Based on this idea, for vocabulary learning in the initial stages, the AIR (Attention, Imitation and Repetition) with MN (Mirror Neuron) model has been formulated (See Fig. 10).

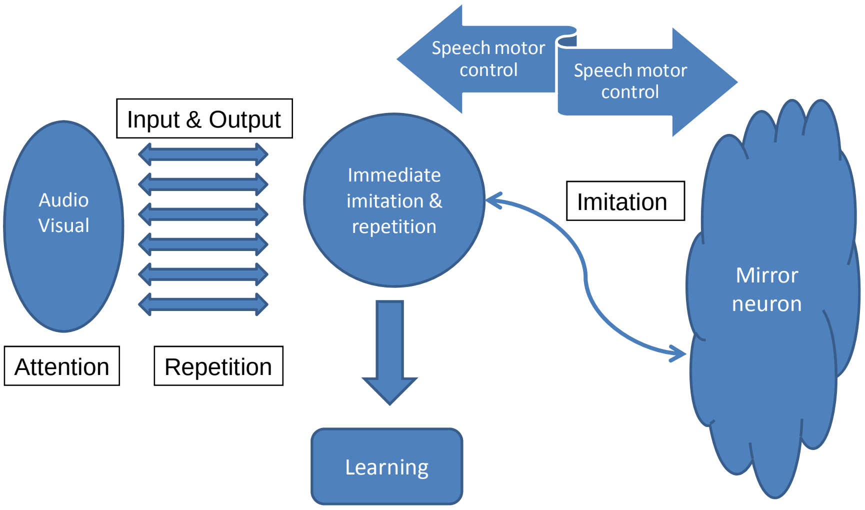


Fig. 10 AIR Model with MN for the early stages in language learning.

4.1.1.1 Attention and Noticing

Discussion within the second language acquisition (SLA) literature on explicit and implicit acquisition has more recently related to development of hypotheses concerning ‘noticing’ and ‘attention.’ Schmidt (2000) defines ‘noticing’ as the subjective correlate of what psychologist call ‘attention’ roughly equivalent to ‘clear perception’ and ‘detection within selective attention’ (Tomlin and Villa 1994). They encapsulate these views as follows: “Since many features of [second/foreign language] input are likely to be infrequent, non-salient, and communicatively redundant, intentionally focused attention may be a practical (though not theoretical) necessity for successful language learning.” To date, almost all the SLA research on noticing and attention has been concerned with acquisition of grammatical systems. However, there is recent interest in applying these notions to vocabulary learning (Laufer and Hulstijn, 2001).

According to Baddeley's influential model (1983), working memory is assumed to comprise a central executive which acts as an attentional controller assisted by 2 subsidiary systems (the visual spatial sketchpad, and the central executive which has limited capacity) and which directs attention and the flow of information and deals with cognitively demanding tasks. The phonological (articulatory) loop is responsible for retaining information in an auditory form, that is to say, manipulation of speech-based information. According to this model, paying attention triggers learners' motivation to process information about novel words.

4.1.1.2 Imitation and Repetition

Imitation is considered an indication that the acquisition process of a linguistic structure has started, that is, it is an initial step before comprehension and production of the said structure is possible (Fraser et al., 1963). The exemplars imitated may be represented and stored and, therefore, internally analyzed for statistical regularities (MacClelland & Plaut, 1999; Seidenberg, 1997). In addition, the imitated exemplars must have been individually perceived, that is, identified within the sound signal and extracted from it. But the system must also have focused on these exemplars and selected them.

Many kinds of research show that vocabulary acquisition needs and involves direct imitation. Roughly, between one in twenty and nine in twenty words produced by infants at around 24 months are mimicked (Bloom, Hood, & Lichtbown, 1974). These figures are likely to be underestimations since they concern only immediately overheard words. Many words that seem spontaneous are in fact delayed imitations overheard days or weeks previously (Miller, 1977). A major predictor of vocabulary increase in older children is their skill in repeating nonword phonetic sequences, which is a measure of mimicry and storage (Gathercole & Baddeley, 1989). Language acquisition problems are linked with impairments in vocal imitation (Bishop, North, & Donlan, 1996).

In addition to aiding vocabulary development, imitation aids the acquisition of speech in other ways. Imitation provides the basis for making longer sentences than children could otherwise spontaneously make on their own (Speidel & Herreshoff, 1989). Children analyze the internal linguistic rules and patterns of speech by repeating to themselves phrases and sentences overheard previously in so-called crib talk (Kuczaj, 1983). Many proto-conversations involve children and parents repeating what the other has said in order to keep alive social and linguistic interaction. Repetition enables immigrant monolingual children to learn a second language by taking part in conversations (Fillmore, 1979). The capacity to imitate overheard words therefore is closely linked with our capacity to learn language.

4.1.1.3 Mirror Neurons

In 1992, mirror neurons, premotor neurons that not only instruct motor actions but detect and observe them in others, were discovered in Broca’s area (Di Pellegrino, Fadiga, Fogassi, Gallase, & Rizzolatti, 1992). Rizzolatti, the discoverer of mirror neurons (1998) has already linked them to the evolution of speech. Here mirror neurons and imitation are linked to the nature and evolution of speech.

Empirical evidence exists for a special imitation pathway. Research on speech shadowing finds a privileged input/output speech loop that is separate to the rest of the speech system (McLeod & Posner, 1984). Research on brain injured and learning impaired individuals finds a direct link between phonological analysis input and motor programming output (McCarthy & Warrington, 1984).

How do we possibly build an information processing system that could imitate vocalizations? First, a clue exists in the fact that children can imitate adults: vocal imitation therefore cannot directly imitate physical movements. Vocalizations, however, are made with regard to motor targets (Shaffer, 1984): for example, if a lip is blocked while articulating a consonant, the vocal apparatus makes a target-related correction (Gracco & Löfqvist, 1994). This suggests that imitation can be based upon copying of motor goals. This is convenient in two ways since not only are these goals independent of the particular shape and size of a vocal tract but they might be directly detectable in spoken sound. The articulatory information of phonemes links to the above properties required for a target to be readily imitated by children since they are (i) contrastive (i.e., [b] vs. [d]; [t] vs. [d]), (ii) categorical (Harnad, 1987), and (iii) innate; infants can initially detect categorical distinctions in unfamiliar foreign languages, though this ability is usually lost by their first year (Eimas, Miller & Jusczyk, 1987).

The existence of language is necessarily linked with the existence of a vocabulary acquisition device. Therefore, presently unknown means must exist to do this. Mirror neurons look like a possible candidate but they need complimentary specialized information, motor targets to imitate that overcome the problem that vocal tracts vary in size and shape, are hidden, and yet their movements must be copied by children. Theoretically, we can predict what the vocal tract will look like in speech. Possibly, speech could contain both motor targets and phonemes as separate phenomena. The vocabulary acquisition device theory offered here is consistent with the claims of the motor theory of speech perception. For example, our theory agrees that 'the objects of speech perception are the intended phonetic gestures of the speaker'

(Liberman & Mattingly, 1985). The theory presented here also corroborates with the idea that speech perception triggers representations of the coarticulated gestures that produced it. These gestures are the primitives which allow speech production to be translated into actual articulator movements, and they are also the primitives from which the specialized mechanisms of speech perception can compute this motor signal (Liberman & Mattingly, 1989).

To summarize, the authoe propose that the vocabulary acquisition device evolved from mirror neurons in Broca’s area. Human language could not begin to develop until patterns of vocalizations of words were being passed down from generation to generation. Such vocal replication would put different kinds of motor goals into selective competition. Those that were most easily copied and therefore more easily transmitted would be selected, reproduced and so end up forming the phonetic/vocalization basis of speech. Due to the advantage bestowed by having a good ability to acquire vocabulary, mirror neuron circuits themselves would have also become selected for, thus expanding Broca's area.

4.1.2 ICon (Incremental Consolidation) Larning Model for the Longitudinal Schedule

In order to maintain vocabulary knowledge, the need for a longitudinal schedule, or a longer-term plan of language training and use, is inevitable. As Gathercole and Baddeley (1993) put it, a properly functioning phonological working memory is key to the long-term learning of languages. When learning new vocabulary, for example, the phonological information is first and temporarily held in phonological working memory and then transferred into some more permanent knowledge structure in the lexical-semantic memory system (Gathercole and Baddeley, 1993). If the temporary trace in phonological working memory is not distinct and durable enough, it is less likely to form a more permanent trace in long-term memory. If there are problems with phonological working memory, the phonological material can either be encoded wrongly or it can be lost quickly (Gathercole and Baddeley, 1993).

4.1.2.1 Incremental Learning

Vocabulary knowledge constitutes an integral part of learners' general proficiency in a second/foreign language (L2) and is essential for successful communication (Nation, 2001). Research on vocabulary acquisition demonstrates that systematic rehearsal is essential for effective vocabulary learning. Rehearsal is defined as an activity which facilitates encoding of new information into long-term memory through overt or silent articulation (Hulstijn, 2001). Unless they are rehearsed frequently, most new words will eventually be forgotten, no matter how deeply they are processed at the first encounter, due to the fragile nature of human memory (Ellis, 1995; Hulstijn, 2001; Nation, 2001).

It is widely known that people frequently do not learn a word from a single encounter and learners need multiple contacts with new words to acquire them. The real question is how many exposures which does not mean a repetition for a meeting are necessary to learn a word. Nation (1990) surveys the research and finds results ranging from 5 to 16 or more exposures. This variation is likely to be a result of the different types of exposure found in different studies. Evidently, the number of exposures required to learn a word depends on type of exposure and level of engagement.

One way of learning words which has been studied extensively is incidental learning from reading. In native language (L1) reading studies, conducted mainly with school children, research designs have tended to focus on the possibility of single-exposure word learning.

The range of probabilities of single-exposure word learning extends from about 5% to 14% (Nagy, 1997). Thus the chances of learning a word from a single exposure are small, but young readers in school typically read a relatively large amount of text, and so the amount of learning from single-exposure encoding could, overall, be sufficient.

In L2 reading studies, research designs typically expose learners to texts in which novel words occur with varying frequencies, and then test which words are learned. There is a great amount of variation in the results, but a consistent conclusion is that learning under these conditions can happen, but that the memory for learned words in those studies is generally not robust (Paribakht and Wesche, 1993; Horst, and Meara, 1999; Pigada and Schmitt, 2006). Given the low rate of absorption, any meaningful incidental learning requires a program being in place which maximizes the amount of reading which is done, such as an extensive reading program (Day and Bamford, 1998).

The other way of learning vocabulary is through an intentional method where words are explicitly taught and/or intentionally focused on by the learner through learning strategies. Not surprisingly, research has shown that when learners' attention is explicitly focused on learning vocabulary, the absorption is greater than during incidental learning. However, the efficiency depends on the level of engagement with the vocabulary learning task. With high-engagement techniques like the Keyword Method (Hulstijn, 1997), relatively few meetings may be enough to establish new items' form-meaning association. Other techniques, which involve less mental effort and engagement with the word, may require many more encounters with novel words.

Because mastery of several types of word knowledge is necessary to use a word well, and because word learning is incremental in nature, vocabulary programs need to build recycling into the curriculum. This can be conducted by selecting textbooks where vocabulary recycling is a design fundamental. Unfortunately, many teachers have to use a prescribed textbook, many of which do not recycle vocabulary in any principled way. In this situation, the teachers will have to insert supplementary activities into their classes. These could include vocabulary games, explicit review sessions, or something as simple as using previously learnt vocabulary in the example sentences which the teacher emphasizes on illustrating the highlighted language points of the day.

4.1.2.2 Consolidation

It is clear that the recycling of vocabulary is essential to learning. However, consolidation of knowledge about lexical items entails more than just recycling. It is also important *how* the learners review and revise their vocabulary, and the way human memory works plays a part in this, particularly in how the mind forgets information. It seems that when learning new information, most forgetting occurs soon after the end of the learning session. After that major loss, the rate of forgetting decreases (Fig. 11).

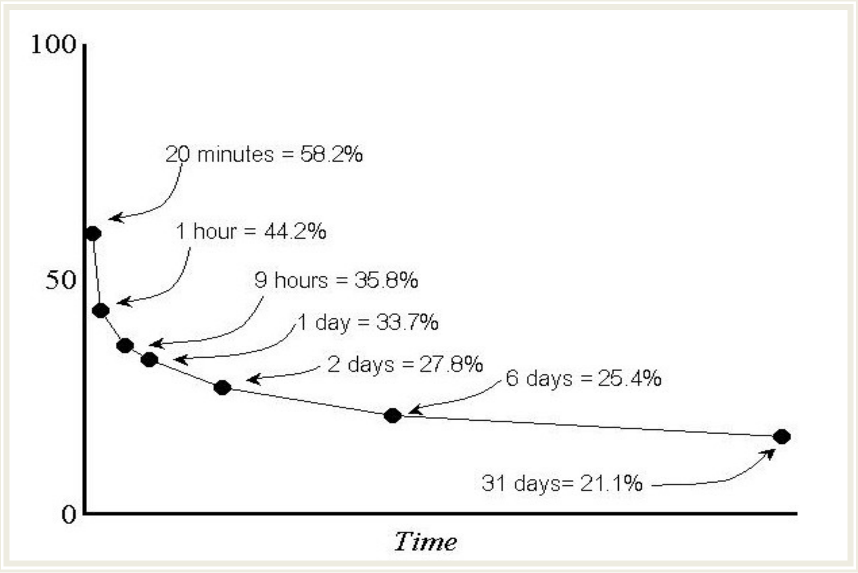


Fig. 11 Ebbinghaus forgetting Curve, adapted by Pimeleur (1976) and

modified by the author.

By understanding the nature of forgetting, we can better organize a recycling program which will prevent against loss of knowledge and therefore be more efficient for learning languages. The forgetting curve in Fig. 11 indicates that it is critical to have a review session soon after the learning session, but less essential as time goes on. The principle of expanding rehearsal is based on this insight, which suggests that learners review new material soon after the initial meeting and then at gradually increasing intervals (Pimsleur, 1967; Baddeley, 1990). One explicit memory schedule proposes reviews 5 to 10 minutes after the end of the study period, 24 hours later, one week later, one month later, and finally 6 months later (Russell, 1979: 149). In this way forgetting should be minimized (Fig. 12). Learners can use the principle of expanding rehearsal to individualize their learning. They should test themselves on new words they have studied. If they can remember them, they should increase the interval before the next review, but if they can't, they should shorten the interval.

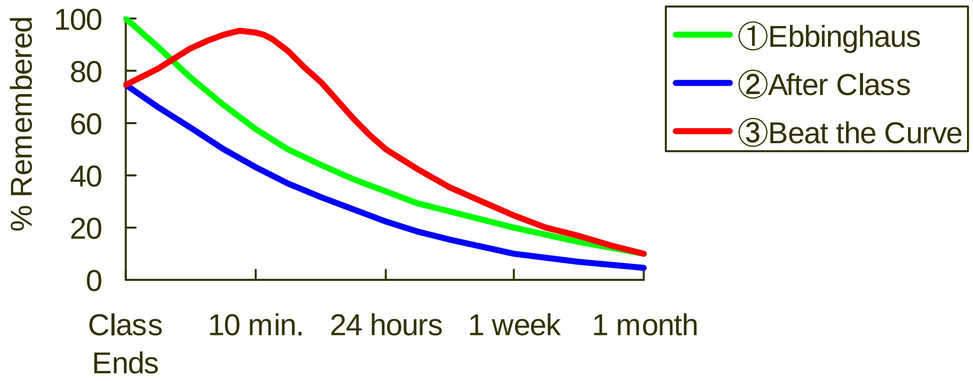


Fig. 12 Pattern of forgetting adapted by Russell (1979) and modified by the author.

① The green line indicates the forgetting curve if it were the case that all new items were recalled at the end of the class.

② The blue line indicates the more typical forgetting curve, as only about 75% of material is recalled at the end of the class.

③ The red line shows the forgetting curve for a typical learner who completed one review session immediately after the class. Less material is forgotten, and indeed material can be regained which was lost in the initial teaching.

Landauer and Bjork (1978) combined the principle of expanding practice with research results demonstrating that the greater the interval between presentations of a target item, the greater the chances it would be subsequently recalled. From this, they suggest that the ideal practice interval is the longest period that a learner can go without forgetting a word (Fig. 13).

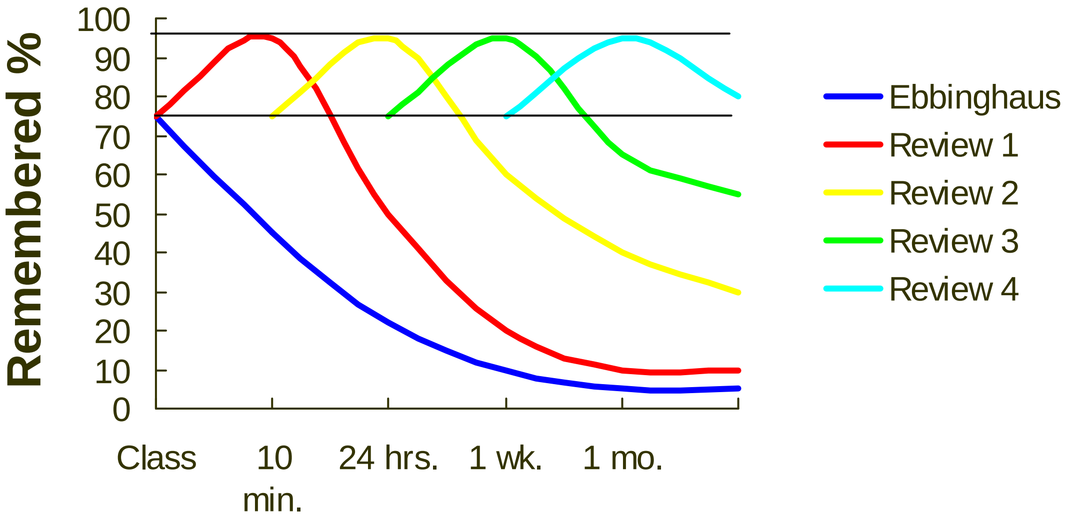


Fig. 13 Pattern of Forgetting with expanding rehearsal adapted by Landauer and Bjork (1978) and modified by the author.

Research by Schouten-van Parreren (1991) shows that some easier words may be overlearned in the sense that more time is devoted to them than necessary, while more difficult abstract words are often underlearned. A practice schedule based on the expanding rehearsal principle may help in avoiding this problem.

Chapter 5

Conclusion

The results of two studies, fMRI testing of Uzbek words and psychophysical testing of unknown medical words, support two conclusions: (1) the learning of vocal imitation is mediated by the left premotor region and BA 44, both of which showed a learning-related reduction in the task related activity, and (2) overt and covert imitation and repetition produce equal performance on word recognition tasks. Pedagogically there remain two issues; namely, why imitation and repetition are useful, and what area of brain is working while imitating and repeating.

Here, the roles of imitation and repetition are summarized. The importance of imitation has been discussed by many scholars not only in the field of language acquisition, but also in many scientific, biological, and psychological areas. How does imitation occur? How can the motor plans necessary for imitating an action derive from the observation of that action? Imitation has a central role in human development and learning of motor, communicative, and social skills

(Piaget, 1962). Models of imitation based on instrumental learning, associative learning, and more complex cognitive processes have been proposed. Because imitation is not a unitary phenomenon, it is possible that different imitative behaviors, subsumed under this name, arise are supported by different underlying mechanisms. The ability to copy elementary actions, however, should be based on simple neural mechanisms.

Concerning repetition, a useful way of assessing the capacity of the phonological loop is be using nonword repetition tasks, participants are asked to repeat unfamiliar spoken forms. Thus, no lexical support is available and one has to rely on the functioning of the phonological loop. In the present studies, Uzbek words for fMRI testing and unknown medical words for psychophysical testing were adapted as a substitute for nonwords. Baddeley et al. (1998) presents data from various studies indicating that scores on nonword repetition tasks are strongly associated with vocabulary knowledge in early and middle childhood.

Service (1992) studied Finnish children learning English at school at the age of nine or ten. She found that their scores on a nonword repetition test were a very strong predictor of English language learning when their proficiency was tested two years later. The long-term learning of the sound structures of novel, unfamiliar words seems to depend on adequate representations of the sound pattern in phonological short-term memory, and ready availability of these representations. Dealing with familiar lexical items, however, does not seem to rely on the phonological loop to the same extent. Thus, an individual with relatively limited phonological loop capacity may be able to deal with familiar words in the native language without difficulty but will encounter considerable problems in acquisition of vocabulary of a foreign language. Some individuals with dyslexia may have normal vocabulary in L1, whereas they may be a deficit in L2 vocabulary. Such individuals may have used intact general cognitive abilities to compensate for early language deficits.

Papagno and Vallar (1995) compared university students who were able to speak several languages with students who only spoke one native language. Although the two groups of students performed equally well on visual spatial tasks and were equivalent in general intellectual skills, the polyglots performed significantly better on auditory digit span and nonword repetition. This study gives additional evidence that the phonological loop is of critical importance in learning a second language. A natural talent for language learning is obviously based on excellent phonological short-term memory functioning.

There is also a possibility that the acquisition of syntax is related to the phonological loop. Syntactic rules are abstracted on the basis of language patterns consisting of strings of words. These word strings must first be held in phonological working memory. A low capacity will impede the construction of more permanent, long-term memory representations.

In conclusion, the present work provides new evidence about the relationship between immediate nonword imitation and repetition, and word-learning in Japanese EFL learners. It also highlights the need for more detailed specification of underlying mechanisms. Further empirical and computational work is needed to forward our understanding of these important relationships, and efforts in this direction are currently under way.

Bibliography

Aboitiz, F., Garcia, R. (1997). The anatomy of language revisited. *Biological Research, 30*, 171-183.

Adams, A.-M. & Gathercole, S.E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language Communication Disorders, 35*, 95–116.

Allott, R. (1991). The motor theory of language. In W. von Raffler-Engel, J. Wind and A. Jonker (Eds.). *Studies in Language Origins II*(pp.123-157). Amsterdam: Benjamins.

Alloway, T., Gathercole, S., Willis, C. & Adams, A. (2004). A structural analysis of working memory and related cognitiveskills in young children. *Journal of Experimental Child Psychology, 87*, 85–106.

Arbib, M. (2002). The mirror system, imitation, and the evolution of language. In C. Nehaniv, & K. Dautenhahn (Eds.). *Imitation in Animals and Artifacts* (pp. 229–280). Cambridge, MA: MIT Press.

Arbib, M. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences, 28*, 105–124.

Arbib, M. & Rizzolatti, G. (1997). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition*, 29, 393–424.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H. & Zilles, K. (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *The Journal of Comparative Neurology,* *412*, 319-341.

Anderson, J. (1995). *Cognitive psychology and its implications* (4th edn.), New York: Freeman.

Aziz-Zadeh, L., Cattaneo, L., Rochat, M. & Rizzolatti, G. (2005). Covert speech arrest induced by rTMS over both motor and nonmotor left hemisphere frontal sites. *The Journal of Cognitive Neuroscience, 17*, 928-938.

Baddeley, A. (1983). *Your memory: A user's guide*. London: Penguin.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (1990). *Human memory: Theory and practice*. MA: Allyn and Bacon.

Baddeley, A. (1992). Working memory. *Science, 255*, 556–559.

Baddeley, A. (1997). *Human memory: Theory and practice* (rev Edn.), Hove: Psychology Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–422.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829-839.

Baddeley, A., Gathercole, S. & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Baddeley, A. & Hitch, G. (1974). Working memory. *The Psychology of Learning and Motivation, 8*, 47–89.

Baddeley, A., Lewis, V. & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology A, 36*, 233–252.

Baddeley, A., Papagno, C. & Vallar, G. (1988). When long-term learning depends on shortterm storage. *Journal of Memory and Language, 27*, 586–595.

Baddeley, A., Thomson, N. & Buchanan, M. (1975). Word length and the structure of short term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575–589.

Barkley, R. (2004). Attention-deficity/hyperactivity disorder and self-regulation. In R. Baumeister & K. Vohs (Eds.). Handbook of self-regulation research, theory, and applications (pp. 301-323). New York: The Guilford Press.

Bernardis, P. & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia, 44,* 178-190.

Bernstein, N. (1967). *The Co-Ordination and Regulation of Movements*. Oxford : Pergamon Press.

Binkofski, F., Buccino, G., Stephan, K., Rizzolatti, G., Seitz, R., et al. (1999). A parieto-premotor network for object manipulation: Evidence from neuroimaging. *Experimental Brain Research, 128*, 210–213.

Bishop, D., North, T. & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *The Journal of Child Psychology and Psychiatry, 37*, 391-403.

Bloom, L., Hood, L. & Lichtbown, P. (1974). Imitation in language: If, when, and why. *Cognitive Psychology, 6*, 380-420.

Bohannon, J. & Stanowicz, L. (1989). Bidirectional effects of repetition in conversation: A synthesis within a cognitive model. In G. Speidel & K. Nelson (Eds.). *The many faces of imitation in language learning* (pp.121-150). Springer-Verlag: New York.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25*, 151-188.

Brass, M. & von Cramon, Y. (2002). The role of the frontal cortex in task preparation. *Cerebral Cortex, 12*, 908-914.

Brooks, N. (1964). *Language and language learning: Theory and practice* (2nd ed.). New York: Harcourt Brace.

Brown, H. (1980). *Principles of language learning and teaching*. N.J.: Prentice Hall.

Buccino, G., Binkofski, F., Fink, G., Fadiga, L., Fogassi, L., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13*, 400–404.

Carr, L., Iacoboni, M., Dubeau, M., Mazziotta, J. & Lenzi, G. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *PNAS, 100*, 5497–5502.

Celce-Murcia, M. (1991). Grammar pedagogy in second and foreign language teaching. *TESOL Quarterly, 25*, 459-480.

Chaminade, T. & Decety, J. (2001). A common framework for perception and action: Neurimaging evidence. *Behavioral and Brain Sciences, 24*, 879-882.

Cohen, N. & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge: MIT Press.

Cohen, N. & Squire, L. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science, 210*, 207–209.

Crothers, E. & Suppes, P. (1967). Experiments in second-language learning. New York: Academic Press.

Davis, C. & Kim, J. (2001). Repeating and remembering foreign language words: Implications for language teaching systems. *Artificial Intelligence Review, 16*, 37–47.

Day, R. & Bamford, J. (1998). *Extensive reading in a second language classroom*. Cambridge: Cambridge University Press.

Dechent, P., Merboldt, K. & Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Cognitive Brain*

Research, 19, 138-144.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J., Le Bihan, D. & Cohen, L. (2004). Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *Psychological Science, 15*, 307-313.

Dehaene-Lambertz, G., Dehaene, S., Anton, J., Campagne, A., Ciuciu, P., Dehaene, G., Denghien, I., Jobert, A., Lebihan, D., Sigman, M., Pallier, C. & Poline, J. (2006). Functional segregation of cortical language areas by sentence repetition. *Human Brain Mapping, 27*, 360-371.

Dinstein, I., Hasson, U., Rubin, N. & Heeger, D. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology, 98*, 1415-1427.

Di Pellegrino, G., Fadiga, L., Fogassim, L., Gallase, V. & Rizzolatti, G. (1992). Understanding motor events. *Experimental Brain Research, 81*, 176-180.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge: Harvard University Press.

Donald, M. (1993). Precis of origins of the modern mind: Three stages in the evolution of culture and cognition. *Behavioral and Brain Sciences, 16*, 737-791.

Eimas, P., Miller, J. & Jusczyk, P. (1987). On infant speech perception and the acquisition of language. In S. Harnad, (Ed.). *Categorical perception* (pp.161-195). New York: Cambridge University Press.

Ellis, N. (1995). The psychology of foreign language vocabulary acquisition: Implications for CALL. *Computer Assisted Language Learning, 8*, 103-128.

Fadiga, L., Craighero, L., Buccino, G. & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience, 15*, 399-402.

Fillmore, L. (1979). Individual differences in second language acquisition. In C. Fillmore, D. Kempler & W. Wang (Eds.). *Individual differences in language ability and language behavior* (pp.203-228). New York: Academic Press.

Fogassi, L., Gallese, V., Dipellegrino, G., Fadiga, L., Gentilucci, M., Luppino, G., et al. (1992). Space coding by premotor cortex. *Experimental Brain Research, 89*, 686–690.

Fowler, C. (1994). Speech perception: Direct-realist theory. In R. Asher & J. Simpson (Eds.). *The Encyclopedia of Language and Linguistics, 8* (pp. 4199-4203). Oxford: Pergamon.

Fraser, C., Bellugi, U. & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior, 2*, 121-135.

Fries, C. (1945). *Teaching and learning English as a foreign language*. MI: University of Michigan Press.

Friston, K., Ashburner, J., Kiebel, S., Nichols, T. & Penny, W. (2007). *Statistical parametric mapping: The analysis of functional brain images*. London: Academic Press.

Friston, K., Holmes, A., Poline, J., Price, C. & Frith, C. (1996). Detecting activations in PET and fMRI: Levels of inference and power. *Neuroimage, 4*, 223-235.

Funayama, C. (1998). Goi syutoku katei ni okeru shadowing no koukendo. *Shadouingu no Ouyoukenkyu*, 12-25.

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593–609.

Gathercole, S. & Baddeley, A. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children:

A longitudinal study. *Journal of Memory and Language, 28*, 200-213.

Gathercole, S. & Baddeley, A. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336–360.

Gathercole, S. & Baddley, A. (1993). *Working memory and language*. Hove: Lawrence Erlbaum Associates Publishers.

Gathercole, S., Willis, C., Baddeley, A. & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*, 103–127.

Gentilucci, M. & Bernardis, P. (2007). Imitation during phoneme production. *Neuropsychologia, 45*, 608-615.

Goldinger, S. (1998). Echoes of echoes? Shadowing words and nonwords in an episodic lexicon. *Psychological Review, 105*, 251-279.

Gracco, V. & Löfqvist, A. (1994). Speech motor coordination and control: Evidence form lip, jaw, and laryngeal movements. *The Journal of Neuroscience, 14*, 6585-6597.

Graf, P. & Schacter, D. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 501–518.

Grafton, S., Arbib, M., Fadiga, L. & Rizzolatti, G. (1996). Localization of grasp representation in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research, 112*, 103–111.

Gray, S. (2003). Diagnostic accuracy and test-retest reliability of nonword repetition and digit span tasks administered to preschool children with specific language impairment. *Journal of Communication Disorders, 36*, 129–151

Grezes, J., Armony, J., Rowe, J. & Passingham, R. (2003). Activations related to “mirror” and “canonical” neurons in the human brain: An fMRI study. *Neuroimage, 18*, 928–937.

Grezes, J., Costes, N. & Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A PET investigation. *Cognitive Neuropsychology, 15*, 553–582.

Grill-Spector, K. & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amsterdam), 107*, 293-321.

Grill-Spector, K., Henson, R. & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*, 14-23.

Gupta, P. & Cohen, N. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psycohological Review, 109*, 401-448.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, *9*, 416-423.

Hanley, R. & Broadbent, C. (1987). The effect of unattended speech on serial recall following auditory presentation. *British Journal of Psychology, 78*, 287–297.

Harnad, S. (1987). *Categorical perception*. New York: Cambridge University Press.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J. & Mazziotta, J. (2003). The essential role of Broca's area in imitation. *European Journal of Neuroscience, 17,* 1123–1128.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H. & Konishi, J. (1996). Cortical speech processing mechanisms while vocalizing visually presented languages. *Neuroreport, 8*, 363-367.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H. & Konishi, J. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *Neuroreport, 8*, 2379-2382.

Horst, M. & Meara, P. (1999). Test of a model for predicting second language lexical growth through reading. *Canadian Modern Language Review, 56*, 308-328.

Hulstijn, J. (1997). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.). *Second Language Vocabulary Acquisition* (pp. 203-224). Cambridge: Cambridge University Press.

Hulstijn, J. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal, and automaticity. In P. Robinson (Ed.). *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.

Iacoboni. M & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience, 7*, 42-951.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *Public Library of Science, 3*, 529-535.

Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J., et al. (1999). Cortical mechanisms of human imitation. *Science, 286*, 2526–2528.

Johnson-Frey, S., Maloof, F., Newman-Norlund, R., Farrer, C., Inati, S., et al. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron, 39*, 1053–1058.

Johnston., J. (1994). Cognitive abilities of children with language impairment. In R. Watkins & R. Rice (Eds.). *Specific language impairments in children* (pp 107-121). Baltimore: Paul H. Brookes.

Kadota, S. (2007). *Shadowing to ondoku no kagaku*. Tokyo: Cosmopia

Karni, A., Meyer, G., Jezzard, P., Adams, M., Turner, R. & Ungerleider, L. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377*, 155-158.

Kawato, M. & Wolpert, D. (1998). Internal models for motor control. *Sensory Guidance of Movement, 218*, 291–307.

Keysers, C., Kohler, E., Umilta, M., Nanetti, L., Fogassi, L. & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research, 153*, 628-636.

Kohler, E., Keysers, C., Umilta, M., Fogassi, L., Gallese, V. & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science, 297*, 846-848.

Koski, L., Iacoboni, M., Dubeau, M., Woods, R. & Mazziotta, J. (2003). Modulation of cortical activity during different imitative behaviors. *Journal of Neurophysiology, 89*, 460–471.

Koski, L., Wohlschlager, A., Bekkering, H., Woods, R., Dubeau, M., et al. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral Cortex, 12*, 847–855.

Kuczaj, S. (1983). *Crib speech and language practice*. New York: Springer-Verlag.

Kuramoto, A. (2004). Tekisuto hyouji o tomonau shadowing no yuuyousei – ninchi mekanizumu no kanten kara. *Kwansai review, 22*, 153-162.

Kurtz, I. (1992). ‘Shadowing' exercises in interpreter training. In C.

Dullerup & A. Loddegaard (Eds.). *Teaching translation and interpreting: training, talent and experiencee* (pp.245-250). Amsterdam/Philadelphia: John Benjamins Publishing.

Lado, R. (1964). *Language teaching: A scientific approach*. New York: McGraw-Hill.

Lado, R., Baldwin, B., & Lobo, F. (1967). Massive vocabulary expansion in a foreign language beyond the basic course: The effects of stimuli, timing and order of presentation. *Washington, D.C.: U.S. Department of Health, Education, and Welfare, Project No*. 5-1095.

Landauer, T. & Bjork, R. (1978). Optimum rehearsal patterns and name learning. In M. Gruneberg, P. Morris & R. Sykes (Eds.). *Practical aspects of memory* (pp. 625-632). London: Academic Press.

Larsen-Freeman, D. (1986). *Techniques and principles of language teaching*, Oxford: Oxford University Press

Laufer, B. & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*, 1-26.

Liberman, A., Cooper, F., Shankweiler, D. & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 117–123.

Liberman, A. & Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition, 21*, 1-36.

Liberman, A. & Mattingly, I. (1989). A specialization for speech perception. *Science, 243*, 489-494.

Liberman, A. & Whalen, D. (2000). On the relation of speech to language. *Trends in Cognitive Sciences, 4*, 187-196.

Logan, G. (1988). Toward and instance theory of automatization. *Psychological Review, 95*, 492-527.

Marton, K. & Schwartz, R. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *46*, 1138–1153.

Masoura, E. & Gathercole, S. (1999). Phonological short-term memory and foreign language learning. *International Journal of Psychology*, *34*, 383–388.

Matsumura, M., Sadato, N., Kochiyama, T., Nakamura, S., Naito, E., Matsunami, K., Kawashima, R., Fukuda, H. & Yonekura, Y. (2004). Role of the cerebellum in implicit motor skill learning: a PET study. *Brain Research Bulletin*, *63*, 471-483.

McCarthy, R. & Warrington, E. (1984). A two-route model of speech production: Evidence from aphasia, *Brain*, *107*, 463-485.

McClelland, J. & Plaut, D. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Science*, *3*, 166-168.

McLeod, P. & Posner, M. (1984). Privileged loops from percept to act. In H. Bouma & D. Bouwhuis (Eds.). *Attention and performance X: Control of language processes* (pp. 55-66). N.J.: Erlbaum

McGuigan, F. & Dollins, A. (1989). Patterns of covert speech behavior and phonetic coding. *Pavlovian Journal of Biological Science*, *24*, 19-26.

Miall, R. (2003). Connecting mirror neurons and forward models. *Neuroreport*, *14*, 2135–2137.

Miller, G. (1977). Spontaneous apprentices: Children and language. New York: Seabury Press.

Morita, T., Itakura, S., Saito, D., Nakashita, S., Harada, T., Kochiyama, T. & Sadato, N. (2008). The role of the right prefrontal cortex in self-evaluation of the face: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *20*, 1-14.

Murphey, T. (2001). Exploring conversational shadowing. *Language Teaching Research, 5*, 128-155.

Munhall, K. & Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In R. Cambell, B. Dodd & D. Burnham (Eds.). *Hearing by Eye, Part 2: The Psychology of Speechreading and Audiovisual Speech*(pp.123-139). London: Taylor and Francis, Psychology Press.

Naccache, L. & Dehaene, S. (2001). The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex, 11*, 966-974.

Nagy, W. (1997). The role of context in first- and second-language vocabulary learning. In N. Schmitt & M. McCarthy (Eds.). *Vocabulary: Description, Acquisition, and Pedagogy* (pp. 64-83). Cambridge: Cambridge University Press.

Nation, P. (1990). *Teaching and learning vocabulary*. New York: Heinle & Heinle.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nema, H. (1999). Shadowing to chikujituuyaku – sono jissen houkoku. *Senshu Journal of Foreign Language Education, 27*, 51-81.

Nishitani, N. & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 913–918.

Nishitani, N. & Hari, R. (2002). Viewing lip forms: Cortical dynamics. *Neuron, 36*, 1211–1220.

Nixon, P., Lazarova, J., Hodinott-Hill, I., Gough, P. & Passingham, R. (2004). The inferior frontal gyrus and phonological processing: An investigation using rTMS. *Journal of Cognitive Neuroscience, 16*, 289-300.

Oldfield, R. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*, 97-113.

Papagno, C., Valentine, T. & Baddeley, A. (1991). Phonological short-term memory and foreign-language learning. *Journal of Memory and Language, 30*, 331–347.

Papagno, C. & Vallar, G. (1995). Short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology* A, *48*, 98-107.

Paribakht, T. & Wesche, M. (1993). Reading comprehension and second language development in a comprehension-based ESL programme. *TESL Canada Journal, 11*, 9-27.

Paulesu, E., Frith, C. & Frackowiak, R. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*, 342-345.

Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York: Norton

Piazza, M., Izard, V., Pinel, P., Le Bihan, D. & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44*, 547-555.

Pigada, M. & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language, 18*, 1-28.

Pimsleur, P. (1967). A memory schedule. *Modern Language Journal, 51*, 73-75.

Poldrack, R., Wagner, A., Prull, M., Desmond, J., Glover, G. & Gabrieli, J. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage, 10*, 15-35.

Porro, C., Francescato, M., Cettolo, V., Diamond, M., Baraldi, P., Zuiani, C., Bazzocchi, M. & di Prampero, P. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study. *Journal of Neurophysiology, 16*, 7688-7698.

Porter, R. & Hogue, D. (1998). Nonlinear dynamical systems in speech perception and production. *Nonlinear Dynamics, Psychology, & Life Sciences*, *2*, 95-131.

Raichle, M., Fiez, J., Videen, T., MacLeod, A., Pardo, J., Fox, P. & Petersen, S. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, *4*, 8-26.

Reisberg, D., McLean, J. & Goldfield, A. (1987). Easy to hear to understand: A lip-reading advantage with intact auditory stimuli. In B. Dodd & R. Cambell (Eds.). *Hearing by eye: The psychology of lip-reading (*pp. 97–113). London: Lawrence Erlbaum.

Rauschecker, A., Pringle, A. & Watkins, K. (2008). Changes in neural activity associated with learning to articulate novel auditory pseudowords by covert repetition, *Human Brain Mapping*, *29*, 1231-1242.

Richards, J. & Rogers, T. (1986). Approaches and methods in language teaching: A description and analysis. NY. Cambridge University Press

Rueckert, L., Appollonio, I., Grafman, J., Jezzard, P., Johnson, R. & Turner, R. (1994). Magnetic resonance imaging functional activation of left frontal cortex during covert word production. *Journal of Neuroimaging*, *4*, 67-70.

Rizzolatti, G. & Arbib, M. (1998). Language within our grasp. *Trends in Neuroscience*, *21*, 188-194.

Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169-192.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., et al. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, *111*, 246–252.

Rizzolatti, G., Fogassi, L. & Gallese, V. (2001). Neurophysiological mechanisms underlying action understanding and imitation. *Nature Reviews Neuroscience*, *2*, 661–670.

Rizzolatti, G., Fogassi, L. & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex, *Current Opinion in Neurobiology, 12,* 149-154.

Roediger, H. & McDermott, K. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.). *Handbook of neuropsychology, 8,* (pp. 63–131). Amsterdam: Elsevier.

Rueckert, L., Appollonio, I., Grafman, J., Jezzard, P., Johnson, R. & Turner, R. (1994). Magnetic resonance imaging functional activation of left frontal cortex during covert word production. *Journal of Neuroimaging, 4*, 67-70.

Russell, P. (1979). *The brain book*. London: Routledge and Kegan Paul.

Sadato, N., Okada, T., Honda, M., Matsuki, K., Yoshida, M., Kashikura, K., Takei, W., Sato, T., Kochiyama, T. & Yonekura, Y. (2005). Cross-modal integration and plastic changes revealed by lip movement, random-dot motion and sign languages in the hearing and deaf. *Cerebral Cortex*, *15*, 1113-1122.

Saito, D., Yoshimura, K., Kochiyama, T., Okada, T., Honda, M. & Sadato, N. (2005). Cross-modal binding and activated attentional networks during audio-visual speech integration: a functional MRI study. *Cerebral Cortex, 15,* 1750-1760.

Salame, P. & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior,* 21, 150–164.

Schaal, S., Ijspeert, A. & Billard, A. (2003). Computational approaches to motor learning by imitation. *Philosophical Transactions of* *the Royal Society of London. Series B, Biological Sciences*, *358*, 537–547.

Schacter, D. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–518.

Schmidt, R. (2000). Attention. In P.Robinson (Ed.). *Cognition and Second Language Instruction (pp.3-32)*. Cambridge: Cambridge University Press.

Schouten-van Parreren, C. (1991). Psychological aspects of vocabulary learning in a foreign language. *Paper presented at the Vocabulary Acquisition in L2 Symposium*, Malaga.

Seidenberg, M. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, *275*, 1599.-1603.

Service, L. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology A*, *45*, 21–50.

Shaffer, L. (1984). Motor programming in language production. In H. Bouma & D. Bouwhuis (Eds.). *Attention and performance, X* (pp.17-41). London: Erlbaum.

Shallice, T. & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. In G. Vallar & T. Shallice (Eds.). *Neuropsychological impairments of short-term memory* (pp.11-53). Cambridge, England: Cambridge Univ. Press.

Skinner, F. (1957). *Verbal Behavior.* New York: Appleton-Century-Crofts.

Speidel, G. & Herreshoff, M. (1989). Imitation and the construction of long utterances. In G. Speidel & K. Nelson (Eds.). *The many faces of imitation in langauge learning* (pp.181-197), New York: Springer-Verlag.

Squire, L. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting memory and learning. *Journal of Cognitive Neuroscience,* 4, 232–243.

Starch, D. (1915). The measurement of efficiency in reading. *The Journal of Educational Psychology, 6*, 1-24.

Takizawa, M. (1999). Application of interpreter training techniques to English language learning(1) Emphasis on shadowing. *Bulletin of Hokuriku University,* 22, 217-232.

Tamai, K. (1998a). Assessment of the shadowing skill in the classroom. *Theoretical and applied linguistics at Kobe Shoin, 1*, 55-69.

Tamai, K. (1998b). Shadowing no rironteki haikei to hyoukahou. *Shadouingu no Ouyoukenkyu*, 1-16.

Tamai, K. (2002). Risuninguryoku koujyou niokeru shadowing no kouka ni tuite. *Interpretation Studies, 2*, 178-192.

Tamai, K. (2005). Shadowing wa bannouyaku nanoka. *The English Teachers' magazine, 53*, 28-30.

Tomlin, R. & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16,* 183-204.

Uchiyama, Y., Toyoda, H., Honda, M., Yoshida, H., Kochiyama, T., Ebe, K. & Sadato, N. (2008). Functional segregation of the inferior frontal gyrus for syntactic processes: A functional magnetic-resonance imaging study. *Neuroscience Research, 61*, 309-318.

Umilta. M., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., et al. (2001). I know what you are doing. A neurophysiological study. *Neuron, 31*, 155–165.

Waters, G., Rochon, E. & Caplan, D. (1992). The role of high-level speech planning in rehearsal: Evidence from patients with apraxia of speech. *Journal of Memory and Language, 31*, 54–73.

Watkins, K., Strafella, A. & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia, 41*, 989-994.

Watkins, K. & Paus, T. (2004). Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of Cognitive Neuroscience, 16*, 978-987.

Werker, J. & Logan, J. (1985). Phonemic and phonetic factors in adult cross-language speech perception. *Perception and Psychophysics, 37*, 35–44.

Wilson, S. & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. *Neuroimage, 33*, 316-325.

Wolpert, D., Doya, K. & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 358*, 593–602.

Wolpert, D., Ghahramani, Z. & Flanagan, J. (2001). Perspectives and problems in motor learning, *Trends in Cognitive Sciences, 5*, 487–494.

Wolpert, D. & Kawato, M. (1998). Multiple paired forward and inverse models for motor control, *Neural Networks, 11*, 1317–1329.