The Process of Internalization in Japanese Speech Production: An Activity Theory Approach to Script Construction for Skit Presentations

小村親英

要旨

　第二言語習得研究では、日本語の発話を促すためのクラス活動として「スキット・プレゼンテーション」があるが、その発表台詞の作成に関わる発表者達の精神機能をアクティビティー理論 (Activity Theory) の視座で考察する。本稿では特に、社会文化理論の枠組みの中で利用される、ヴィゴッツキーの物体－他者－自己制限、ミードの社会行動主義、バフチンの腹話的相互作用を用いて、精神機能の内化過程 (internalization) を論じるものである。発表で発話を促すためには、台詞作成時に発表者が個人的な目標を建てて会話に臨むことが好ましいと説く。

【キーワード】Activity Theory, Vygotsky's object-other-self regulations, Mead's social behaviorism, Bakhtin's ventriloquation

1. Introduction: classroom activities for speaking proficiency

It is often acknowledged in the study of second language acquisition (SLA) that there are certain pedagogical sequences in classroom activities for enhancing speaking proficiency. For example, one of the typical sequences in a Japanese as a second language (JSL) class consists of "grammar explanation," "mechanical drills," "contextualized drills," "pair work," and "role play" in order (Kawaguchi & Yokomizo, 2005, p. 98). (*Figure 1*)　In this configuration, the ultimate goal is located in the "role play," which exists to facilitate the native-like speaking proficiency.　From the viewpoint of Watsonian behaviorism, it is natural to presume that those "mechanical drills" and "contextualized drills" are both arranged in this sequence for the preparation of the later "pair work" and "role play."　Those drills are

conceptually framed by the behavioristic principles of stimulus-reinforcement designs which comprise various kinds of repetitious practices.

| Grammar Explanation | → | Mechanical Drills |

→ | Contextualized Drills | → | Pair Work | → | Role Play |
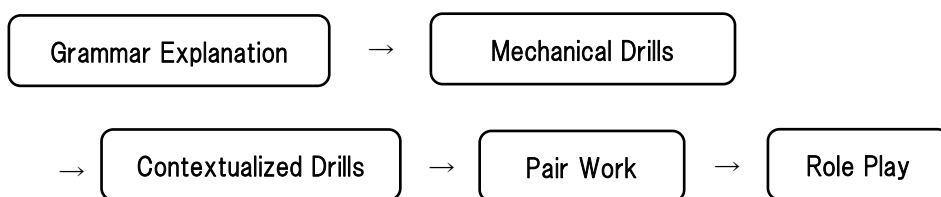
(*Figure 1*) (trans. Komura)

To ultimately attain native-like speaking proficiency, there are several types of drills in each component of this sequence:

a. Mechanical drills: designed to seek grammatically and phonetically correct utterances in pattern practices, in which specific grammatical items to be practiced are provided.   These include (1) repetition drills, (2) expansion drills, (3) development drills, (4) substitution drills, (5) conversion drills, (6) combination drills, and (7) talk-back drills

b. Contextualized drills: to seek grammatically and phonetically correct utterances while exchanging information in certain pre-determined contexts, (1) teacher (questioner) to student (answerer) interaction, (2) student (questioner) to teacher (answer) interaction, and (3) student to student peer interaction

c. Pair work: to facilitate native-like speech production when exchanging personal information using pre-determined discussion topics

d. Role play: to facilitate native-like speech production by exchanging information and playing out certain roles in certain contexts.   These can include, (1) a type of role play where students have to complete a partially-finished script, (2) a type of

role play built upon a situation, roles and vocabulary, (3) a type of role play built

upon a situation and roles, (4) a type of role play built upon only a speech topic

(Kawaguchi & Yokomizo, 2005, pp. 98-127)

One of the types of "role play," in JSL class there is a classroom activity called "skit

presentation." Skit presentation involves working outside of class to write a short sketch

utilizing grammatically and pragmatically appropriate Japanese sentences. The participating

students, as a pair or a group of three play out their sketches in class after memorizing their

parts. These short sketches should be original and constructed through their conjoint effort.

In this paper, the focus will be placed on the student's internalization process in

constructing original scripts for skit presentations. Part of the conceptual foundations of

exploring the process of internalization are drawn from Vygotsky's sociocultural study of

mind. Vygotsky (1978) postulates that learning is a fundamentally social process (Daniels,

2008; del Rio & Alvarez, 2007; Johnson, 2004). Focusing on the social process, Vygotsky

emphasized "tool mediation as necessary to carry out cognitive and material functions"

(Thorne, 2005, p. 393). In this Vygotskian sociocultural study of the human mind, the

Activity Theory of A. N. Leont'ev (1981b), in particular, will be drawn upon for the

theoretical framework. The purpose of this study is to demonstrate different facets of the

internalization process used in constructing original scripts for skit presentations. Using

students' negotiations from their collective activity of constructing scripts, this paper will

attempt to elucidate how the participating students internalize their voices (speech thoughts)

through the negotiations.

2. Theoretical framework: Activity Theory

Focusing on the socio-genetic notion of mental functioning, Vygotsky (1978) asserted that human action and thinking are mediated by socioculturally constructed tools and signs. To articulate the relationship between mental functioning and sociocultural context, Vygotsky emphasized language (a sign system) as the primary tool for mediating human action and thinking (Brooks & Donato, 1994). In this Vygotskian socio-genetic perspective, language used in mediated action is considered to be a vehicle to generate socially conformed individual consciousness. Lantolf and Thorne (2007) remarked:

> If we want to dig a hole in the ground in order to plant a tree, it is possible, following the behavior of other species, to simply use our hands. However, modern humans rarely engage in such nonmediated activity; instead we mediate the digging process through the use of a shovel, which allows us to make more efficient use of our physical energy and to dig a more precise hole. We can be even more efficient and expend less physical energy if we use a mechanical digging device such as a backhoe. Notice that the object of our activity remains the same whether we dig with our hands or with a tool, but the action of digging itself changes its appearance when we shift from hands to a shovel or a backhoe. … The material form of a tool as well as the habitual patterns of its use affect the purposes to which it is put and methods we use when we employ it. Thus, a shovel requires one type of motion and a backhoe another. Physical tools, which are culturally constructed objects, imbue humans with a great deal more ability than natural endowments alone. (p. 203)

Within the Vygotskian sociocultural perspective, human mental process are mediated by

cultural tools and these mediational tools transform human action. In light of the study of the relationship between a human mental process and the social contexts in which it occurs, Lattuca (2002) points out:

> We tend to study contexts and persons in contexts separately, often extracting the individual from his or her context in order to get a better look at the phenomenon of cognition. Many traditional psychological theories of learning and studies based on these theories manifest this assumption about the separate spheres of thinking and being. In contrast, sociocultural anthropological theories focus intently on contexts and culture: They are more apt to assume that analytic strategies should begin with an account of social phenomena and then, on the basis of these, develop analyses of individual mental functioning. Today theorists from various fields have begun to think about how to repair the mind-body duality, to argue that learning cannot be separated from the contexts in which it occurs, and to reconceptualize cognition and learning as activities that occur through social interaction. (pp.711-712)

Using the Vygotskian concept of socio-genetic human mental functioning, A. N. Leont'ev (1981b) constructed Activity Theory which proclaims that the notion of activity should be considered a goal-oriented process. He remarked that "the emergence in activity of goal-oriented processes or actions was historically the consequence of the transition of human to life in society" (1981b, p. 39). There are fundamental principles of Activity Theory:

(1) The human mind is formed and functions as a consequence of human interaction with the culturally constructed environment

(2) The cultural environment is as objective as any physical, chemical, or biological property.

(3) Activities are oriented to objects (concrete or ideal) and impelled by motives or needs (physical, social, and psychological). Actions are directed at specific goals and are socioculturally designed means of fulfilling motives. Operations are the specific processes through which actions are carried out.

(4) Mental processes are derived from external actions through the course of appropriation of the artifacts made available by a particular culture, both physical and semiotic (signs, words, metaphors, narratives).

(5) Mediation through the use of culturally constructed tools and others' voices (or discourses) shape the way people act and think as a result of internalization.

(6) To understand human activity, including mental activity, means to know how it developed into its existing form. (Lantolf & Pavlenko, 2001, pp. 144-145)

Drawing on these principles, Activity Theory "posits three levels of analysis of individuals at work on a particular task: *activity*, *action* and *operation*" (Darhower, 2004, p. 325). To explicate the components and internal relations of an activity system (activity-action-operation), it is appropriate to formulate analytical representations for each unit of analysis in an activity system. Lantolf and Throne (2006) clarify the relationship based on Bødker's formulation (1997). (*Table 1*) It is important to distinguish between an action and an activity. Here is a classic quotation from Leont'ev pointing out the difference between an action and an activity:

A beater, for example, taking part in a primeval collective hunt, was stimulated by a
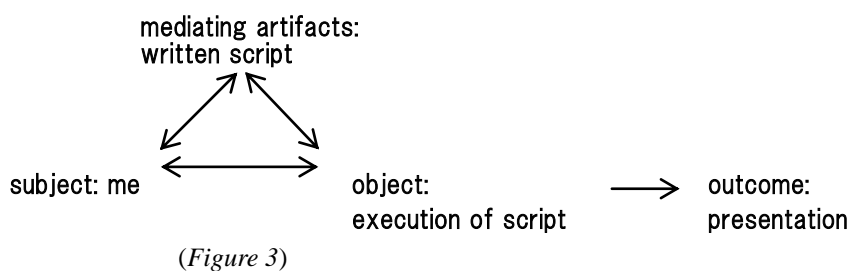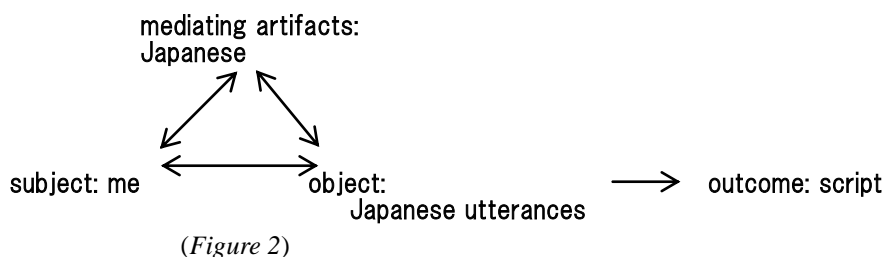
Unit of analysis	Everyday description	Oriented toward	Carried out by
Activity	Why something takes place	Motive, transformation of object	Community and/or society
Action	What is being done	goal	Individual or group
Operation	The actual doing	condition	individual

(*Table 1*) (2006, p. 217)

need for food or, perhaps, a need for clothing, which the skin of the dead animal would meet for him. At what, however, was his activity directly aimed? It may have been directed, for example, at frightening a herd of animals and sending them toward other hunters, hiding in ambush. That, properly speaking, is what should be the result of the activity of this man. And the activity of this individual member of the hunt ends with that. The rest is completed by the other members. This result, i.e., the frightening of game, etc., understandably does not in itself, and may not, lead to satisfaction of the beater's need for food, or the skin of the animal. What the processes of his activity were directed to did not, consequently, coincide with what stimulated them, i.e., did not coincide with the motive of his activity; the two were divided from one another in this instance. Processes, the object and motive of which do not coincide with one another, we shall call "actions." We can say, for example, that the beater's activity is the hunt, and the frightening of the game his action. (Leont'ev, 1981a, p. 210, cited from Engeström and Miettinen, 1999, p. 4)

Placing the analysis within the context of the role of discourse in classroom activities, the school and the classroom are an activity system (action-activity-operation) (Smit, Fritz & Mabalane, 2010). Within an activity system, for example, constructing a dialogue or "sketch" with the purpose of utilizing expressions is considered to be action done by an individual

student. It is safe to assert that this activity system is made up of the student's individualized

action mediated by Japanese language. That is to say, the object of the action is to make

Japanese utterances to fulfill the goal of creating the conversation in the sketch. On the

other hand, after the completion of the script, the object of the action is shifted when

students produce Japanese utterances by means of the written script. The outcome is to

present a sketch in class. As proposed by Y. Engeström (1991, 1999), it is useful to present

two diagrams (*Figures 2 & 3*) for clarifying the two faces of mediation in preparing a skit

presentation (Hiruma, Wells & Ball, 2007; Wells, 2007). It is obvious that different focuses

in object will generate the different outcomes.

```
          mediating artifacts:
          Japanese
              ↗ ↖
           ↙      ↘
subject: me  ⟷  object:            ⟶  outcome: script
                 Japanese utterances
          (Figure 2)
```

```
          mediating artifacts:
          written script
              ↗ ↖
           ↙      ↘
subject: me  ⟷  object:            ⟶  outcome:
                 execution of script     presentation
          (Figure 3)
```
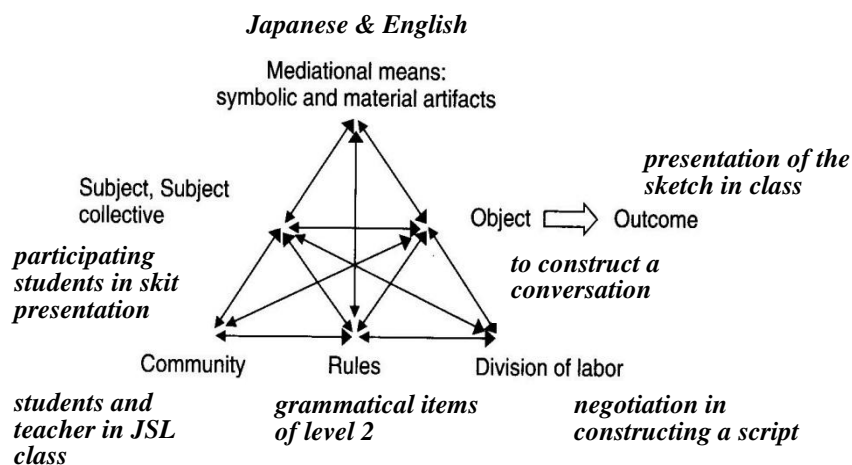
However, it becomes necessary to treat the mediated activity of constructing a script as

a collective activity since the action is conjointly formed by the participating students to

perform their sketch in class. It is considered important to distinguish between collective

activity and individual action (Leont'ev, 1981b). For the discussion of collective activity, I

use an example of an activity system in which the students and their native Japanese teacher discuss a draft of a sketch. The central idea of the co-construction of the draft is to elucidate the process of interaction between the students and their native Japanese teacher in the consultation sessions. To review and revise the scripts, the students and the teacher, in these sessions, discuss the drafts in Japanese and English. Seen in the discourse of the consultation sessions, the interrelated aspects of the collective activity serve to illustrate a shift from an individual activity to collective activity. (*Table 2*) The point is that the object of each action is either to construct the individual Japanese sentences which make up the entire conversation or to first construct an outline of the conversation which is then filled in with individual sentences. Based on the different focuses on the objects, it is natural to claim that the collective activity embraces the purpose of playing out the sketch conjointly, not of playing out individual roles in the conversation.

Activity	Subject	Object of the action	Outcome
Individual	A student	Construction of her/his specific utterances	Playing out her/his role in the sketch
Collective	Participating students and their teacher	Construction of their conversation	Playing out the sketch

(*Table 2*)

In addition to the core of the three-level model in individual mediated activity, other important components, "community," "rules" and "division of labor," constitute the expansive model of collective mediated activity. (Nelson & Kim, 2001) (*Figure 4*) New interrelated components are presented to clarify the phases of interrelations in a mediated activity.

Japanese & English

Mediational means:
symbolic and material artifacts

Subject, Subject
collective

*presentation of the
sketch in class*

Object ⟹ Outcome

*participating
students in skit
presentation*

*to construct a
conversation*

Community   Rules   Division of labor

*students and
teacher in JSL
class*

*grammatical items
of level 2*

*negotiation in
constructing a script*

(*Figure 4*)

3. Internalization

3. 1 Vygotsky's concept of object-other-self regulations

It seems necessary to clarify how an individual speaker internalizes her or his voices in speech exchanges playing out in skit presentations. There are three theoretical orientations which attempt to explain how each speaker's utterances are shaped and presented in the chains of conversation. Internalization is not done simply to attain new information as an input to one's own speech repertory, but rather is to acquire qualitatively different intake to one's own speech production.

Vygotsky (1978) claimed language (a sign system) is the primary tool for mediating human action and thinking. Mediated by socio-culturally constructed signs, human mental functioning emerges in social interaction with others (Ellis & Barkhuizen, 2005; Holland & Lachicotte, 2007; Kozulin, et al., 2003; Nassaji & Swain, 2000; Wells, 1999). In this Vygotskian socio-genetic perspective, language used in mediated action is considered a vehicle to generate socially conformed individual consciousness (Komura, 2008).

Vygotskian psycholinguistic theory holds that:

> Every function in the learner's cultural development appears twice, on two levels.
> Some first, on the social, and later, on the psychological levels. First, between people
> as an interpsychological category, and then inside…as an intrapsychological category.
> (Vygotsky, 1978, p. 86)

Citing this connection, SLA researchers started investigating L2 acquisition focusing on the
course of language socialization that happened in a particular sociocultural setting (Anton,
1999; Donato, 2000; Mondada & Doehler, 2004; Nassaji & Wells, 2000; Ohta, 2001; Storch,
2002; Swain, 2000; Yoshida, 2009). They tried to demonstrate how individuals acquire an
L2 in the Vygotskian scheme of transformation from interpsychological (between
individuals) to intrapsychological (inside individual) planes. L2 acquisition occurs in the
internalization process in which interpsychological function occurs first and
intrapsychological function follows (Kozulin, 2003; Lantolf & Thorne, 2006).

For example, Vygotsky (1978) discussed the process of speech development of children.
Yoshida (2009) introduced Vygotsky's theory of regulation, pointing out that "any function
in a child's cultural development emerges first in the interpsychological plane and next in the
intrapsychological plane" (p. 8). As demonstrated in Wertsch's experiment, a model of
mother-child (female) interaction in putting together a puzzle using a model puzzle as a
guide (1991b), the child cannot figure out how to proceed with the problem solving task.
The child's attention is placed only on the task of finding the correct pieces (object) at first.
Yoshida (2009) points out that "the child's attention is attracted exclusively to its physical
environment" (p. 8). At this stage, the child is object-regulated. Moving onto the next stage,

however, through interaction with her mother, the child comes to realize how to carry out the task. The child participates in the interaction and carries out the task under the guidance of her mother. The child's cognitive function to solve the problem is facilitated and controlled by her mother's metacognitive advances in the interaction. At this stage, the child is other-regulated. Finally, the child gradually starts controlling her own metacognitive strategies to solve the problem without relying on her mother. It becomes possible for the child to solve the problem by herself with her independent strategic function provoked in the mother-child interaction. At this stage, the child is self-regulated.

Central to this account of internalization from inter- to intra- psychological function are studies investigating the role of interactions in dyads and small group work in JSL class (Komura, 2005; Ohta, 2001; Mori, 2002; Yoshida, 2009). In association with the relationship between interactional competence in language use and the social interaction in which it occurs, it seems feasible to integrate socio-genetic orientation into the analysis of L2 script construction in skit presentations. For example, on the level of object-regulation, each student seems incapable of gaining control over the construction of a script by herself or himself. In the next stage, however, each students co-construct the script through interaction – an exchange of ideas. Finally, each student takes responsibility for playing out her or his role in the presentation, internalizing the content of the co-constructed script. The main assumption here is that a student's potential language proficiency in the construction of a script emerges in the process of internalization in the regulatory shifts (from object-, other, to self-regulation). In other words, the internalization process involves a transformation shift from interpsychological to intrapsychological planes when the students conjointly

engage in their script construction. "There is thus a gradual process of internalization whereby a fully externalized social practice becomes a substantially internalized cognitive practice" (Atkinson, 2002, p. 537).

3. 2 Mead's social behaviorism

One of the most influential contributions to the sociocultural approach to the human mind is Mead's (1934) extensive investigations of the human mental mechanism (mind) and the ways of self-constitution as a member of society (self) in relation to its social environment. Focusing on the psychological mechanism in each individual acting in social activities, he proposed that the nature of human mental functioning is depicted in the process of interactions with the social environment without indulging in the socio-genetic nor introspective origins of human cognition. He investigated how the human mind and self emerge in the process of everyday activities in a social group.

The term, "social behaviorism," is different from Watsonian behaviorism in which human behavior could be explained by stimulus-response connections, where there is no mentalist view of the experience. It is behavioristic in terms of analyzing human experiences starting with observable human behavior, but it is not behavioristic in that it does not ignore the inner experience of the individual (p. 7). Mead said that, "The behavioristic psychology has tried to get rid of the more or less metaphysical complications involved in the setting-up of the psychical over against the world, mind over against body, consciousness over against matter" (p. 105). His point is that metaphysical complications cannot be discounted in the study of human conducts, emphasizing the importance of

studying "consciousness" or "mind."   On the other hand, an exclusive study of inner psychological phenomena in each individual (introspection) does not lead to a proper theoretical orientation because Mead claimed that human conducts are social products (p. 6). Social psychology, as Mead called it, deals with human conducts from the inside (introspection) as well as from the outside (sociocultural influences).

In addition, Mead claimed that the contents of the introspective study of mind, "consciousness," and physiological phenomena in the social act, "behavior," do not exist as independent entities to be analyzed.   Those factors are not static or time-free entities to be described in isolation, but rather they are in a constant dynamic process of social acts interpenetrating one another moment by moment.   Both entities do not preexist as embryotic forms to be later recognized before the process of interactions, but rather, "consciousness" emerges in the very process of interactions.   At the same time, with the emergence of "consciousness," the nature of a social act is constantly on a move to be modified.   Mead remarked that "We are rather forced to conclude that consciousness is an emergent from such behavior; that so far from being a precondition of the social act, the social act is the precondition of it" (p. 18).

In relation to the concept of dynamic, time-bound and reflective interactions between mind and the social act to which it belongs, Watsuji, a Japanese philosopher, in the same 1930s when Mead advanced his social psychology, constructed his unique theory of human action emphasizing interactions with its indigenous ecological surroundings.   To clarify Mead's social psychology, it seems meaningful to present Watsuji's discussion as an example.

We feel cold. We feel cold as one of our individual experiences. There is no doubt that everyone living outside tropical areas in the world has had the experience of feeling cold. But what is cold? Everyone knows what cold means in its physical sense as one of the meteorological characteristics. For example, the temperature goes down and cold wind blows. It may start snowing. That is to say, the cold exists as an object which acts upon our sensory system, and then we feel the cold. But, as Mead was dissatisfied with this Watsonian behavioristic perspective to explain the relationship between individual experience and objects based on stimulus-response connections.

On the other hand, it seems impossible to come to know the objective cold that exists outside our consciousness prior to our perception of the cold. We feel the cold in our consciousness first and then realize the cold. In other words, we experience the cold depending on our perceptional act either upon the degrees of temperature or the conditions of the weather. The objective cold exists only in the relation to our individual consciousness of the cold. Indeed, we act upon the environment and then claim that we have an experience of the cold. The cold is, then, not an objective existence but a subjective existence. Mead disagreed with this introspective orientation to illustrate human experience. His claim is that we do not experience the cold as a result of introspective composition, simply by imagining the cold in mind. The cold does not preexist in our consciousness waiting for introspective analysis to feel the cold, but rather the cold actually exists outside of the body.

Given such a parallel subjective and objective scheme, it is interesting to see that both Watsuji and Mead constructed the notion of an interaction between a subject and generalized others in a social group. Watsuji claimed that we (subjects) start seeing our objectified

selves outside in the cold where we thrust ourselves outside (*ex-sistere* [Ger.])(pp. 10-14).

Our objectified selves exist in the cold and feel the cold simultaneously. This overlaps with Mead's theoretical construct in his social psychology. I am a principal agent of my action, "I," and at the same time, I recognize myself as a member of the group, "me," by hypothesizing that other members would act and respond in a similar way. I am an agent of my own action, "I," and at the same time, I am a social product, "me," in a social group. The relationship between "I" and "me" is established on its reflective process. Neither one emerges without the other in its social act.

Watsuji's (1979) conclusion is that the process of recognizing "me" in its reflective relationship between "I" and "me" is embedded in its indigenous ecological environment (p. 15). Watsuji's description of ecological environment is not limited to its geographical characteristics but rather includes people's interactions with the environment. For example, people in rural communities in the north of Japan wear thick winter clothes and sit under the *kotatsu* (a Japanese foot warmer with a quilt over it). A Japanese person feels the cold when the person objectifies herself or himself by transplanting herself or himself to the group of other people who feel the cold in the same manner as other group members do. The person may not feel the same cold in New York City, even when she or he experiences the same temperature and weather conditions of winter in Japan. The person feels the cold only assuming other members of her or his community do not exist to share the experience in New York City.

Mead, on the other hand, emphasized the importance of using language in the reflective process between "I" and "me." A principal agent of action in a social group, "I," responds to

the external stimulus, "coldness" by way of internalizing the meaning of the word, "cold," which is socially inherited in her or his language repertory. And by uttering a sentence, "It is cold," the "I" changes into "me" because the language itself carries the social meaning already. In short, socially constructed speech becomes a shared symbolic tool to transform "I" to "me." By presenting the shared symbolic tool to others in the community, "I" becomes capable of predicting that others would understand the meaning of the message. In addition, this transformation is not a one-way direction from "I" to "me." In the process of internalization, "I" already interprets the external stimulus on the basis of the ongoing prediction of others' interpretation. The way of interpretation that "I" internalizes the stimulus is already influenced by "I's" prediction. The meaning of the language conveyed by the shared symbol emerges in the process (p. 6). My individual utterance is already "half-others," inherently providing the mechanism of role-taking in society (p. 161).

In the same vein, Mead elaborated the idea of a shared symbol to characterize the reflective process between "I" and "me," into the study of attitude and gesture. He said:

> If there is any truth in the old axiom that the bully is always the coward, it will be found to rest on the fact that one arouses in himself that attitude of fear which his bullying attitude arouses in another, so that when put into a particular situation which calls his bluff, his own attitude is found to be that of the others. If one's own attitude of giving way to the bullying attitude of others is one that arouses the bullying attitude, he has in that degree aroused the attitude of bullying himself. (1934, p. 66)

It is obvious that an individual member of a society is a principal actor "I," and at the same time, the individual is a role-taking constituent of society, "me." By using socially

constructed symbols, such as language and gesture, "I" starts seeing "me" in relation to others.

In relation to his discussion of a shared symbol, Mead pointed out that human consciousness appears in the process of how "I" sees "me" in relation to others. Consciousness is described as an emergent accessibility to decide how to use the socially constructed symbols to see "I" as "me" (p. 30). Consciousness emerges when "I" accesses certain context wishing to represent the meaning in relation to others. As the meaning arises in the reflective relationship between "I" and "me," consciousness appears as a driving force to materialize the process of seeing "I" as "me" in a social group by way of the shared symbols. Mead wrote:

> The individual experiences himself as such, not directly, but only indirectly, from the particular standpoints of other individual members of the same social group, or from the generalized standpoint of the social group as a whole to which he belongs. For he enters his own experience as a self or individual, not directly or immediately, not by becoming a subject to himself, but only in so far as he first becomes an object to himself just as other individuals are objects to him or in his experience; and he becomes an object to himself only by taking the attitudes of other individuals toward himself within a social environment or context of experience and behavior in which both he and they are involved. (1934, p. 138)

Guided by Mead's orientation in his social behaviorism, it is feasible to recognize the exchanges of speech communication as a reflective relationship in skit presentations. The internalization of each member's speech production playing out in her or his skit

presentations occurs in its reflective process between "I" and "me." Each participant's consciousness emerges when she or he is engaged in her or his presentations.

3. 3 Bakhtin's concept of ventriloquation

Bakhtin (1981, 1986, 1994), a Russian literary critic and semiotician, discussed semiotic (sign-based) mediation to explain the complexities of human mental function, attempting to delineate the relationship between human mental functioning for language communication and the sociocultural context in which it occurs. Based on his studies of semiotic mediation, Bakhtin investigated the nature of "the real unit of speech communication: the utterance" (1986, p. 71). Disregarding the analysis of sentences as an impersonalized abstract form of language, he examined actually produced utterances in real-life conversation. For example, the sentence, "*Okane ga arimasuka* (Do you have money?)," is analyzed by formal linguists from morphosyntactic and lexical points of views. Formal linguists tend to analyze sentences as abstracted representations of linguistic information. Bakhtin, on the other hand, tried to look at sentences as personalized utterances situated in real-life situations. According to his perspective, the sentence, "*Okane ga arimasuka*" is actually translated into "Do you have money (I would like to borrow money, if you have some)?" Therefore an utterance always includes the speaker's "voice" (inner speech or speech thought). Bakhtin (1986) pointed out:

> Speech can exist in reality only in the form of concrete utterances of individual speaking people, speech subjects. Speech is always cast in the form of an utterance belonging to a particular speaking subject, and outside this form it cannot exist. (p. 71)

By rendering the actual mode of existence of linguistic phenomena, Bakhtin investigated how individual speakers construct voices through the exchanges of utterances.

Bakhtin continued to write that "Utterances are not indifferent to one another, and are not self-sufficient; they are aware of and mutually reflect one another" (1986, p. 91). Central to his investigation of utterance was that an individual utterance inherently interacts with others, and because of their mutually reflective relations, an utterance begins to carry social nature in dialogue. "The utterance is filled with *dialogic overtones*" (1986, p. 92. italics original), which indicate that one's utterances are born and shaped in terms of dialogues with others.

It is essential to emphasize a categorization of speech events with which Bakhtin distinguished the nature of systematic speech communication in his analytical apparatus. Wertsch (1991b) pointed out that "By switching from dealing with utterances to dealing with languages, Bakhtin was moving from unique speech events (individual utterances produced by unique voices) to categories or types of speech events (types of utterances produced by types of voices)" (p. 56).

The first category of speech event is "social languages." A social language is a specific way of speaking operated in a particular social group in a particular sociocultural setting (Bakhtin, 1986; Wertsch, 1991a, 1998). For example, there are social languages of profession, such as those used by doctors and lawyers. There are also social languages of different generations and genders. There are certain ways of using language in a particular social group in the same language community. Bakhtin claimed that dialogic overtones of individual utterances are born and shaped by the nature of social language situated in the

environment.

The second category of language is "speech genre." In differentiating speech genre from social language, Bakhtin (1986) remarked:

A speech genre is not a form of language, but a typical form [a type] of utterance; as such the genre also includes a certain typical kind of expression that inheres in it. In the genre the word acquires particular typical expressions. (p. 87)

The social stratum of the speakers (i.e., doctors and lawyers) distinguishes social languages, while speech genres are characterized by "typical situations of speech communication." Speech genres are, for example, formulaic expressions in everyday greetings, congratulations at a wedding, conversations around a mahjongg table, etc. In the situation of social interaction between teacher and students in an L2 class, specific discourse between them is therefore characterized as a specific speech genre constructed in the class. Unlike social language, speech genres take identifiable forms in the particular situations, and teachers in class introduce those identifiable forms to students as L2 formulaic expressions, greetings, or model sentences.

The dialogic interrelation invoked by social language is termed by Bakhtin as "ventriloquation" (1994). Wertsch (1991a) pointed out that "with any type of social language, speakers ventriloquize through speech genres and are thereby shaped in what they can say" (p. 96). Presupposing an addressee's voice, one voice of a speaker is constructed through ventriloquation situated in a specific social language. Presupposing a hearer's voice, a speaker's voice already includes the hearer's voice when the speaker produces speech. The speaker's voice constructed through ventriloquation is embodied by speech genre that

both speaker and hearer share.   In the process of dialogic interaction, social language functions as a socio-ideological framework in which a speaker's ventriloquation takes shape, and speech genre operates as identifiable linguistic signs with which speakers manifest their voices in speech communication.

To examine an actual speech development in real-life situations, Bakhtin (1986) asserted that "an utterance is a link in the chain of speech communication" (p. 94).   The nature of an utterance is always explained in relation to the preceding and subsequent utterances in the chain of speech communication. A speaker's voice in her or his utterance is constructed in the reference of her or his addressee's voice presented in the preceding utterance. At the same time, the speaker's voice which emerges from the previous utterance is now constructed in a new utterance, presupposing possible reactions from the addressee in the subsequent utterance. The role of others (with respect to a speaker) thus becomes an important element to construct the entire process of dialogic interaction. If someone asks, "Who is doing the talking, then?" Bakhtin will answer, "Both speaker and hearer are talking simultaneously."

In the process of constructing original scripts for skit presentations, it is feasible to think that each student goes through the internalization that Bakhtin claimed. In order to carry on a coherent conversation; each student experiences the process of ventriloquation to present her or his partner's voice in their utterances. At the same time, she or he has to construct her or his own utterance for the continuation of the conversation.   This internalization occurs within the speech genre that all the participants in their skit presentations share in the JSL class.

4. Skit presentation: goal-oriented activity

To bring my argument to the fore, this is an attempt to analyze the discourse from several consultation sessions between the participating students and their teacher with the purpose of reviewing and revising drafts of the students'script. The goal of the consultation sessions is to complete the final draft of the presentation. The following items are what the students were instructed to do: (1) coordinate and collaborate on their 8-to-10 minute presentations; (2) include grammatical items that they have studied in the level 2 JSL class; (3) set up the situations of the skit, for example, "who you are," "when it is," "where it is," "what you want to achieve" and "what properties are to be used." The repertory of sentence structures and expressions for their script is given to the students prior to the onset of their sketch construction. The evaluation categories are: (1) appropriate expressions; (2) collaboration; (3) memorization and (4) pronunciation and accuracy. [1]

In the following excerpt, Laura (pseudonym), a female student from Columbia, and Gloria (pseudonym), another female student from Italy, both enrolled in their second semester of JSL class, are constructing their script for presentation. After several script construction meetings, they have completed the first draft of the script. The following discourse occurred between the two students and their Japanese teacher in a consultation session. The sessions were recorded and their discourse was transcribed. The purpose of this session (*excerpt 1*) is to review and revise the draft of the script at which the students and the teacher are looking while revising it. The collective activity of reviewing and revising was done by having those students' read out the draft while the teacher offered instructive interventions during the reading. The session unfolds like this:

(*excerpt 1*)

(1) Laura (L)   : G*onenkan Chūgoku ni sunde imashita. Appuru de hatarakimashita. Kyonen nihon ni kaerimashita. Chūgoku ni itta koto ga arimashita ka.*

(I lived in China for five years. I worked at Apple. I came back to Japan last year. Had you ever been to China?)

(2) Teacher (T) : *ari…? ari…?*   (Have …have …?)

(3) Gloria (G)   : *Arimashita ka.*   (Had you…?)

(4) L   : *Arimashita ka*   (Had you…?)

(5) T   : Experience?

(6) G   : *Arimasu ka.*   (Have you…?)

(7) L   : *Arimasu ka…itta koto ga arimasu ka.*

(Have you been…have you ever been to?.)

(8) L   : *Hai…?*   (Yes…?)

(9) G   : *Itta koto ga arimasu ka.*   (Have you ever been to?)

(10) L   : *Itta koto ga arimasu ka.*   (Have you ever been to?)

(11) G   : *Iie, arimasen…nna…demo, ikitai desu. Watashi wa fuyuyasumi ni iku tsumori *desho*.

(No, I have not…but, I want to go. I *probably go during the winter break…)

(12) T   : *Desu…desu…?*   (*Desu…, desu…*?)

(13) G   : *Desho…desho?*   (*Desho…, desho…*)

(14) T   : *Desu yo?*   ([suggesting student means the emphatic form])

(15) G   : Maybe, I plan to go?

(16) T   : *Iku tsumori desu yo…?*   (I probably will …? [emphatic])

(17) L   : I don't know…depends on the weather…? I don't know.

(18) T   : But, you wanna go…? Say statement?

(19) G   : I don't know…it's OK? (laughter)

(20) L   : (laughter)

(21) T   : *Iku tsumori desu…desu…yo.*   (I probably will [emphatic])

(22) L   : *Yo.* Ah, I see. Oh, I get it.

(23) G : Ah! this *yo*. *Iku tsumori desu yo.*   (I will probably go

[emphatically])

In this excerpt, Laura wants to tell about her trip to China, and Gloria wants to say that she wants to visit China during the winter break. In line (1), Laura makes a mistake using the wrong tense of the structure, "*arimashita ka.*" With the teacher's intervention, Gloria quickly responded and corrected Laura's utterance for her, saying "*arimasu ka.*" It seems reasonable to assume that Laura now internalizes the correction in the process of interaction and succeeds in producing the correct utterance, saying "Have you ever been to China?" in line (10). This interactive discourse serves to demonstrate how Laura internalized her speech voice in the chain of the conversation. It is feasible to presume that Laura's internalization emerged in social interaction mediated by other speakers' utterances. Focusing on this social interaction, three essential theoretical orientations which have been discussed, the Vygotskian object-other-self regulation, Mead's social behaviorism and

Bakhtin's ventriloquation, seem to provide important insights into the discussion of the process of internalization. To summarize:

(1) Vygotsky's object-other-self regulation:

Laura's potential language proficiency in the construction of the utterance emerged in the process of internalization in the regulatory shifts (from object-, other to self-regulation). The internalization process involves a transformative shift from interpsychological to intrapsychological planes when Laura engaged in social interaction with other participants in the session.

(2) Mead's social behaviorism:

By objectifying herself in the reflective process between "I" and "me," Laura recognized her appropriate utterance in the discourse. Laura seems to have succeeded in the production of the utterance in the way her other partners would say the utterance.

(3) Bakhtin's ventriloquation:

Presupposing her hearers' voices in the chain of speech communication, Laura produced her utterance as if she had ventriloquized other speakers' voices. In other words, Laura's utterance was already embedded in other participants' voices.

As another example of interactive utterance, from the viewpoints of the three orientations, it is possible to account for Gloria's internalization process in producing her utterances of "I probably will go," which is manifested from line (11) to (23). That is to say, Gloria's internalization seems to have occurred in a shift from interpsychological to intrapsychological plane, by means of her objectifying her "I" and "me." It is also possible

to assume that Gloria ventriloquized other speakers' speech voices simultaneously.

However, the discourse in the consultation sessions provides an argument for the importance of goal-oriented activity in the preparation of the script. With a careful monitoring of the development of the utterances between Laura and Gloria, the object in the activity system (activity-action-operation) engenders a different direction to the conversation. For example, in *excerpt 2*, as a continuation of talking about her experience in China, Laura speaks:

(*excerpt 2*)

(24) L   : *Hai, watashi wa takusan fuku o kaimashita. (?) Demo,okane o takusan*

           haraimasen deshita. Tokorode, anata wa doko ni ikimashita ka.

           (Yes, I bought a lot of clothes. But, (?) I did not pay much money. By the

           way, where did you go?)

(25) G   : *Itaria ni sunde imashita..ima, *nihon ni sundeimasu kara, nihon de watashi*

           wa Itaria-go no sensei desu. Itaria ni itta koto ga arimasu ka

           (I lived in Italy. Now, *since I live in Japan, I am an Italian teacher. Have
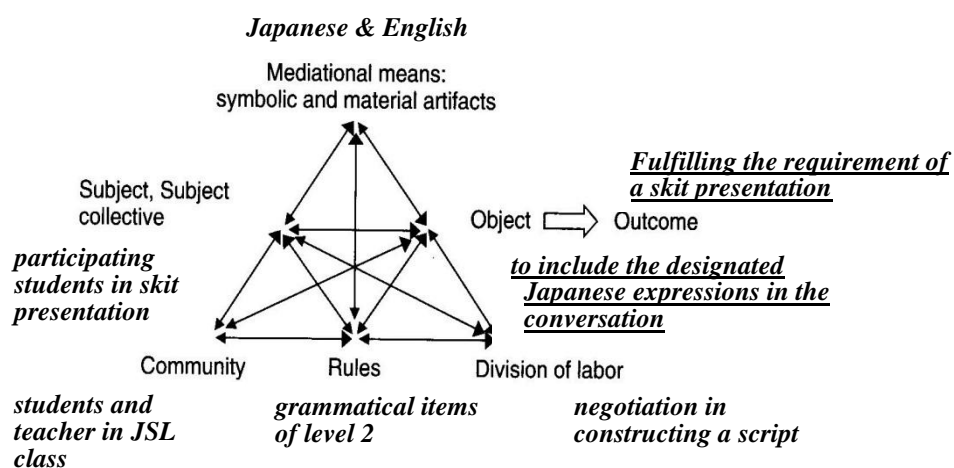
           you been to Italy?)

(26) L   : *Hai. Arimasu. *Kyonen, Itaria ni itta koto ga arimasu. ...*

           (Yes, I have. *Last year I have been to Italy….)

In line (24), Laura changes the subject, asking "By the way, where did you go?" The chain of the development of the conversation comes to an end abruptly and swings in a different direction. It is obvious to assume that Laura switched to ask the question, "Where did you go?" for the purpose of prompting the next utterance, "I lived in Italy, and now, since I live

in Japan, I am an Italian teacher," in which one of the Japanese expressions, the conjunction,

"since," is included. The effort of including the expression in the script seems to be central

to this account of the development. It is possible to proclaim that Laura's internalization

remains on the object-regulation stage because she seems to be too busy to include Japanese

expressions in her utterance as the ultimate goal of the activity. Her internalization process

does not seem to shift from object-regulation, other-regulation, and to self-regulation.

Rather, it seems to skip the other-regulation stage, jumping to her self-regulation directly.

(*Figure 5*)

Japanese & English

Mediational means:
symbolic and material artifacts

Subject, Subject
collective

*participating
students in skit
presentation*

*Fulfilling the requirement of
a skit presentation*

Object ⟹ Outcome

*to include the designated
Japanese expressions in the
conversation*

Community   Rules   Division of labor

*students and
teacher in JSL
class*

*grammatical items
of level 2*

*negotiation in
constructing a script*

(*Figure 5*)

On the other hand, after being reminded of the importance of setting up a goal in the

chain of speech development, Laura and Gloria revised their draft focusing on their

underlying motive of carrying on the conversation. It is critical to acknowledge the

existence of goals within the mediated activity to facilitate students' interactive utterances.

That is to say, Laura wants to recommend China to Gloria, and Gloria wants to visit China.

Excerpt 3 illustrates the speech development which embraces their goals in the revised draft:

(*excerpt 3*)

(27) L   : *Hai, watashi wa takusan fuku o kaimashita. (?) Demo,okane o takusan*

haraimasen deshita. Guroria-san, donna tokoro ni ikitai desuka.

(Yes, I bought a lot of clothes. But, (?) I did not pay much money . Gloria,

what kind of place do you want to?)

(28) G   : *Shiranai. Mada, kimete imasen. *Anata wa doko e itta hou ga ii desuka?*

(I don't know. I haven't decided yet. *Where do you think I should go?)

(29) L   : *Chugoku no machi no naka de, watashi wa Beijin to Shanhai ga suki desu.*

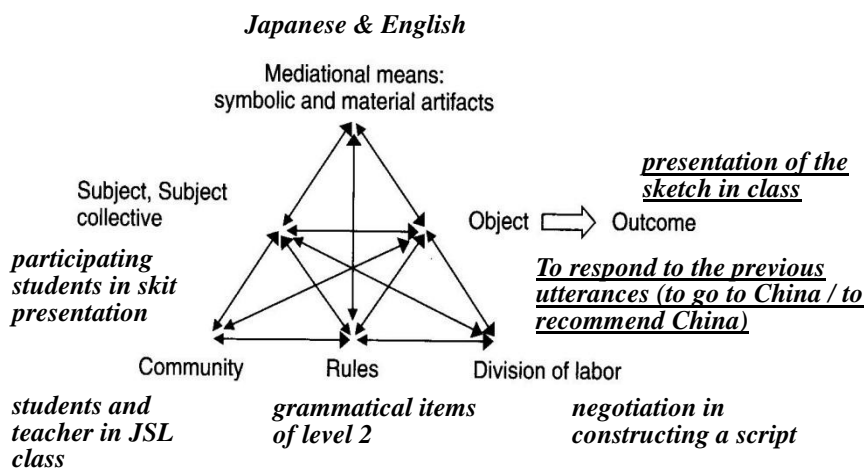*Anata wa Beijin to Shanhai ni itta hou ga ii desu*

(Among the cities in China, I like Beijing and Shanghai the best. You should

go to Beijing and Shanghai.)

(30) G   : *So desu ne. Kono machi de nani ga dekimasu ka?*

(All right. What can I do in this city?)

The heart of this dyadic process is the information exchanges between Laura and Gloria

seeking to achieve their goals in the conversation. Laura's goal is to recommend China for

Gloria to visit, while Gloria's goal is to visit China. It is important to note that the

collaborative process of interaction embraces the characteristics of a goal-oriented activity.

Within a framework shaped by goals, speakers' utterances are facilitated to respond to their

previous utterances, guiding them to their goals. The effort of responding to the previous

utterances provides an implication that these speakers internalize their speech voices from

object-, other- and self-regulations individually. Drawing on the concept of goal-oriented

activity, it seems clear that each speaker's speech voice was enhanced by their conscious

awareness of achieving their goals in the conversation. (*Figure 6*)

Japanese & English

Mediational means:
symbolic and material artifacts

presentation of the sketch in class

Subject, Subject collective

Object ⟹ Outcome

participating students in skit presentation

To respond to the previous utterances (to go to China / to recommend China)

Community   Rules   Division of labor

students and teacher in JSL class

grammatical items of level 2

negotiation in constructing a script

(*Figure 6*)

5. Conclusion

There are three interconnected concepts; (1) object-other-self regulation in the socio-genetic human function, (2) Mead's social behaviorism and (3) the Bakhtinian notion of ventriloquation. The process of internalization in constructing scripts for skit presentation was discussed in terms of the integration of these three concepts. It is essential to claim that three strata of cognitive shift (from object-, other- and to self-regulation) serve to elucidate the process of internalization. The focal point of the internalization process seems unable to manifest itself clearly, based on any of those three theoretical orientations independently. It is, rather, being suggested that there is an integrated model that accounts for the process of internalization which occurrs in the transformative shifts from interpsycological to intrapsychological planes.

In addition, the importance of explaining the internalization process resides in the fact that goal-oriented activities serve to facilitate speech production in the process of

constructing scripts of skit presentation. It is important to assert that speakers advance their speech voices for the purpose of achieving their goals in the conversation. For example, the question of "What do you want to achieve?" becomes a critical element in the conversation. Skit presentations should be a classroom activity in which participating students engage in interactive activities conjointly to accomplish their goals in the conversation. The role of social interaction in skit presentations is to construct speech voices interactively for the purpose of achieving each individual participant's goals.

Note

(1) The task requirements for skit presentation were created by a teacher who taught level-2 of spoken Japanese at Kansai Gaidai University in 2011.

References

Anton, M. (1999) The discourse of a learner-centered classroom: Sociocultural perspectives on teacher-learner interaction in the second-language classroom. *The Modern language Journal 83(3)*, pp. 303-318.

Atkinson, D. (2002) Toward a sociocognitive approach to second language acquisition. *The Modern Language Journal, 86(4)*, pp. 525-545.

Bakhtin, M. (1981) *The dialogic imagination* (C. Emerson & M. Holquist, Trans.). Austin, TX: University of Texas Press.

Bakhtin, M. (1986) *Speech genres and other late essays* (C. Emerson & M. Holquist, Eds.; V. W. McGee, Trans.). Austin, TX: University of Texas Press.

Bakhtin, M. (1994) The dialogic imagination. In P. Morris (ed.), *The Bakhtin reader* (pp. 74-80). New York: Routledge.

Bødker, S. (1997) Computers in mediated human activity. *Mind, Culture and Activity: An International Journal 4/3,* pp. 149-158.

Brooks, F. B., & Donato, R. (1994) Vygotskian approaches to understanding foreign language learner discourse during communicative tasks. *Hispania, 77(2)*, pp. 262-274.

Daniels, H. (2008) *Vygotsky and Research*. New York: Routledge.

Darhower, M. (2004) Dialogue journals as mediators of L2 learning: a sociocultural account. *Hispania, 87(2)*, pp. 324-335.

del Rio, P., & Alvarez, A. (2007) Inside and outside the zone of proximal development: An ecofunctional reading of Vygotsky. In H. Daniels, M. Cole., & J. Wertsch (eds.), *The Cambridge Companion to Vygotsky* (pp. 276-303). New York: Cambridge University Press.

Donato, R. (2000) Sociocultural contributions to understanding the foreign and second language classroom. In J. Lantolf (ed.), *Sociocultural Theory and Second Language Learning* (pp. 27-50). Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005) *Analysing Learner Language*. Oxford: Oxford University Press.

Engeström, Y. (1991) Activity theory and individual and social transformation. *Activity Theory, 7/8.* pp. 6-17.

Engeström, Y. (1999) Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R. Punamäki (eds.), *Perspectives on Activity Theory* (pp. 19-38). New York: Cambridge University Press.

Engeström, Y., & Miettinen, R. (1999) Introduction. In Y. Engeström, R. Miettinen, & R. Punamäki (eds.), *Perspectives on Activity Theory* (pp. 1-16). New York: Cambridge University Press.

Hiruma, F., Wells, G., & Ball, T. (2007) The problem of discoursing in activity. *An International Journal of Human Activity Theory, 1*, pp. 93-114.

Holland, D., & Lachicotte, W. (2007) Vygotsky, Mead, and the new sociocultural studies of identity. In H. Daniels, M. Cole., & J. Wertsch (eds.), *The Cambridge Companion to Vygotsky* (pp. 101-135). New York: Cambridge University Press.

Johnson, M. (2004) *A Philosophy of Second Language Acquisition*. New Haven: Yale University Press.

Kozulin, A. (2003) Psychological tools and mediated learning. In A. Kozulin, B. Gindis, V. Ageyev, & S. Miller (eds.), *Vygotsky's Educational Theory in Cultural Context* (pp. 15-38). New York: Cambridge University Press.

Lantolf, J. P., & Pavlenko, A. (2001) Second Language Activity theory: Understanding second language learners as people. In M. P. Breen (ed.), *Learner contributions to language learning: New directions in research* (pp. 141-158). Harlow, England: Pearson Education Limited.

Lantolf, J. P., & Thorne, S. L. (2006) *Sociocultural Theory and the Genesis of Second Language Development.* Oxford: Oxford University Press.

Lantolf, J. P., & Thorne. S. L. (2007) Sociocultural Theory and Second Language Learning. In B. VanPatten, & J. Williams (eds.), *Theories in Second Language Acquisition* (pp. 201-224). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Lattuca, L. R. (2002) Learning interdisciplinarity: Sociocultural perspectives on academic work. *The Journal of Higher Education, 73(6)*, pp. 711-739.

Leont'ev, A. N. (1981) *Problems of the development of the mind*. Moscow: Progress.

Leont'ev, A. N. (1981) The problem of activity in psychology. In J. V. Wertsch (ed.), *The concept of activity in Soviet psychology* (pp. 37-71). Armonk, New York: M. E. Sharpe.

Mead, G. H. (1934) *Mind, self, and society: from the standpoint of social behaviorist.* Chicago: University of Chicago Press.

Mondada. L., & Doehler, S. P. (2004) Second language acquisition as situated practice: Task accomplishment in the French second language classroom. *The Modern Language Journal, 88(4)*, pp. 501-518.

Mori, J. (2002) Task design, plan, and development of talk-in-interaction: An analysis of a small group activity in a Japanese language classroom. *Applied Linguistics 23(3)*, pp. 323-347.

Nassaji, H., & Swain, M. (2000) A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness 9(1)*, pp. 34-51.

Nassaji, H., & Wells, G. (2000) What's the use of 'Triadic Dialogue'?: An investigation of teacher-student interaction. *Applied Linguistics 21(3)*, pp. 376-406.

Nelson, C. P., & Kim, M. K. (2001) Contradictions, appropriation, and transformation: An Activity Theory approach to L2 writing and classroom practices. *Texas Papers in Foreign Language Education, 6(1)*, pp. 37-62.

Ohta, A. S. (2001) *Second language acquisition processes in the classroom learning Japanese*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Smit, B., Fritz, E., & Mabalane, V. (2010) A conversation of teachers: In search of professional identity. *The Australian Educational Researcher, 37(2)*, pp. 93-106.

Storch, N. (2002) Patterns of Interaction in ESL pair work. *Language Learning 52(1)*, pp. 119-158.

Swain, M. (2000) The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. Lantolf (ed.), *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford: Oxford University Press.

Thorne, S. (2005) Epistemology, politics, and ethics in sociocultural theory. *The Modern Language Journal 89 (3)*, pp. 393-409.

Vygotsky, L. S. (1978) *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, eds.). Cambridge, MA: Harvard University Press.

Wells, G. (1999) *Dialogic Inquiry: Towards a Sociocultural Practice and Theory of Education*. Cambridge: Cambridge University Press.

Wells, G. (2007) The mediating role of discoursing in activity. *Mind, Culture, and Activity 14(3)*, pp. 160-177.

Wertsch, J. V. (1991a) A sociocultural approach to socially shared cognition. In L. Resnick, J. M. Levine & S. D. Teasley (eds.), *Perspectives on socially shared cognition*. (pp. 85-100). Washington, DC: American Psychological Association.

Wertsch, J. V. (1991b) *Voices of the mind: A sociocultural research*. Cambridge, MA: Harvard University Press.

Wertsch, J. V. (1998) *Mind as Action*. New York: Oxford University Press.

Yoshida, R. (2009) *Learners in Japanese language classrooms; overt and covert participation.* London: Continuum.

川口義一＆横溝紳一郎 (2005) 『成長する教師のための日本語教育ガイドブック（上）』 東京：ひつじ書房

小村親英 (2005) A Sociocultural Study of Speech Development in Three Interactions: Adult-child, teacher-adult L2 learner, and adult L2 learner collaborative interaction. 『関西外国語大学留学生別科　日本語教育論集』第15号 pp. 65-85.

小村親英 (2008) A Sociocultural Explanation of the Overuse of Japanese Honorific Utterances in Speech Communication. 『関西外国語大学留学生別科　日本語教育論集』第18号 pp. 83-97.

和辻哲郎 (1979) 『風土－人間的考察』東京：岩波書店

(ckomura9@kansaigaidai.ac.jp.)