Tetsugaku

International Journal
of the Philosophical
Association of Japan

Volume 4, 2020

Special Theme
Analytic Asian Philosophy

ISSN 2432-8995

ISSN 2432-8995

Tetsugaku

International Journal of the Philosophical Association of Japan

Volume 4

2020

Special Theme

Analytic Asian Philosophy

The Philosophical Association of Japan

Tetsugaku
International Journal of the Philosophical Association of Japan
Volume 4, Special Theme: Analytic Asian Philosophy
May 2020

Edited and published by the Philosophical Association of Japan
Department of Philosphy, College of Humanities and Sciences, Nihon University
3-25-40, Sakurajosui, Setagaya-ku, Tokyo, 156-8550, Japan

ISSN 2432-8995

Table of Contents

Tetsugaku
International Journal of
the Philosophical Association of Japan
Volume 4, 2020

Philosophy in East Asia

How Could Ōmori Shōzō Use Wittgenstein to Fight against Wittgenstein?

Xu Yingjin[1]
Professor, Fudan University

Abstract: *Ōmori* Shōzō's philosophy *can be generally described a hybrid system composed of both a Wittgensteinian skin and a Husserlian core, in the sense that he systematically uses a Wittgensteinian philosophical methodology to fight against Wittgenstein's own publicity-oriented philosophical tendency. His first recipe for doing so, according to my reconstruction, is to appeal to the notion of tachiaraware (namely, "phenomena standing for themselves"), via which the gap between synthesizing activity and sense-data to be synthesized can be filled. Therefore, the first-personal character of tachiaraware could be easily transmitted to the formal features of "my language", without which no public language can be formed. Ōmori's second recipe for refuting Wittgenstein is to appeal to his Kasane-egaki (namely, "recoloring")-narrative, according to which the ordinary language (L2) is nothing but the "recoloring" of the phenomenal language (L1), while the scientific language (L3) is nothing but the "recoloring" of the ordinary language. Given that the L1-L2-L3-hierarchy has to be elaborated without implementing double standards, a Wittgensteinian emphasis on the putative primacy of public languages cannot be recommended due to its patent violation of the so-called "Double-Standard-Abominating Principle" (DSAP). Hence, since both the respect of the "tachiaraware" and DASP are required by a thorough implementation of the phenomenological principle itself, Ōmori's stance simply appears to be a natural result of radicalizing Wittgenstein's stance alongside the phenomenological route.*

1. Introduction

Due to the widely known philosophical affinity between continental philosophy and the Kyoto School, Japanese philosophy has long been viewed as a Japanese counterpart of European continental philosophy, rather than that of Anglophone

[1] e-mail: yjxu@fudan.edu.cn
His research interests include philosophy of AI, Wittgenstein, and Japanese philosophy.

analytic philosophy. Nonetheless, this view fails to do justice to the postwar philosophical development in Japan as a reborn country both politically and culturally connected to United States. Especially, the view is unfair to the achievements made by Ōmori Shōzō (大森荘蔵, 1921–1997),[2] who systematically introduced analytic philosophy into Japan.

Ōmori' s first major at the University of Tokyo was physics rather than philosophy. His interests in philosophy were triggered by his wartime service at the Institute of Technology of Imperial Navy, where he became fascinated with some philosophical problems related to optical issues. Hence, he re-registered as a philosophy student at his Alma Mater just after the war, and then he got a chance to study analytic philosophy in U.S. (at Stanford and Harvard), where he became interested in Wittgenstein. Notably, although English is widely assumed to be the primary linguistic tool for doing analytic philosophy, Ōmori had long been using Japanese as his working language ever since he began his teaching career at the Komaba campus of the University of Tokyo in 1953. Unfortunately, both Ōmori's own adherence to his mother tongue and the rarity of English translations of his writings prevented him from being widely recognized in the west.[3] However, his influence within Japan is still nonnegligible. One may appreciate such influence through the works of Ōmori's philosophical followers, such as Nagai Hitoshi (永井均), Noe Keiichi (野家啓一), Fujimoto Takashi (藤本隆志), Noya Shigeki (野矢茂樹), Tanji Nobuharu (丹治信春), Nakajima Yoshimiji (中島義道), Iida Takashi (飯田隆), etc.. In my view, his preference to his mother tongue may be still in accordance with the style of Wittgenstein himself, who also preferred to use German (which is his mother tongue) to do philosophy even at Cambridge, probably due to his consideration that the emergence of right types of philosophical intuitions do go hand in hand with a stubborn adherence to one's native language. But Wittgenstein is still luckier than Ōmori in the sense that his mother tongue, namely, German, is not as mysterious as Japanese to the English-speaking world. Put another way, in contrast to the "Anglicization" of Wittgenstein, more efforts need to be spent to make Ōmori "Anglicized".

This article attempts to make Ōmori "Anglicized". However, it might be natural for anyone ignorant of Ōmori (but still familiar with Wittgenstein) to ask the following question first: Why do I need to care about Ōmori, if his philosophy is

[2] Throughout this paper all Japanese names will be spelled in this way: Surnames first, then first names.

[3] As far as I know, Kobayashi (2019) is the only English literature that includes a brief introduction to Ōmori. There is no English translation of any of Ōmori's books yet.

nothing but a Japanese counterpart of Wittgenstein's philosophy? To respond to this query, my discussion will begin with a seemingly mysterious contrast between Ōmori and Wittgenstein: Ōmori seems to have used a Wittgensteinian methodology to fight against Wittgenstein himself.

2. How Could Ōmori Use Wittgenstein to Fight against Wittgenstein?

First of all, note that insofar as their writing styles are concerned, the link between Ōmori and later Wittgenstein is fairly visible. For instance, Ōmori's general view of philosophy and language, formulated in the preface of his *Language, Perception and World* (Ōmori 1971), looks simply like a Japanese re-writing of later Wittgenstein's corresponding views in his *Philosophical Investigation* (Hereafter *PI*. Wittgenstein 1958, cf. table-1)

Ōmori's expressions	**Wittgenstein's expressions**
Wittgenstein did metaphorically view philosophy as a fly catcher which catches a fly. If this metaphor makes sense, then one can use nothing—more appropriate than "spinning" to metaphorically describe the development of philosophy (Ōmori 1971, iii).	What is your aim in philosophy?—To shew the fly the way out of the fly-bottle (Wittgenstein 1958, §309).
Unlike science, philosophy is not intended to discover new facts or elaborate new theories. If there is "new facts" in philosophy, then that cannot be anything else than a case of seeing through the surface of a picture to perceive a hidden picture…. Philosophy is nothing but to see through what has been seen (Ōmori 1971, iv).	When we look into ourselves as we do philosophy, we often get to see just such a picture. A full-blown pictorial representation of our grammar. Not facts; but as it were illustrated turns of speech (Wittgenstein 1958, §295).
If there were no language, then human beings would not be able to	Here the term "language-game" is meant to bring into prominence the

exist qua human beings. The diversity of the functions of speech acts is nothing but the diversity of human life. Humans may shout loudly, give orders, make accusations, make threats, cheat others, sing songs, make inquiries, be silent…there is simply no way to make a complete list of infinitely many modes of speech acts (Ōmori 1971, 3).	fact that the speaking of language is part of an activity, or of a form of life. Review the multiplicity of language-games in the following examples, and in others: Giving orders, and obeying them—Describing the appearance of an object, or giving its measurements—Constructing an object from a description (a drawing)—Reporting an event—Speculating about an event—(Wittgenstein 1958, §23)

Table-1. The metaphilosophical similarity between Ōmori and Wittgenstein

However, it is noteworthy that in the same book, Ōmori also articulates his quasi-solipsist position, which looks patently conflicting with later Wittgenstein's hostility towards the possibility of a "private language":

> For confirming whether the "red impressions" in his tongue is the same as my impression, it is necessary to compare his impressions with mine. In order to do this comparison, I have to acquire his impression; but it is impossible to do this, because I simply have no access to what another person could perceive. In order to experience his perception, I have to be himself; but this is not what can be realized due to the constraints imposed on myself. There is simply no way to work this out. Even though I were one of the Siamese twins, I still could not perceive what my brother perceives, given that I am nobody else but myself, and I cannot be my brother (Ōmori 1971, 13–14).

Ōmori's mentioning of the case of "Siamese twins" in the preceding citation definitely refers to the same case used by Wittgenstein in *PI*:

> In so far as it makes sense to say that my pain is the same as his, it is also possible for us both to have the same pain. (And it would also be imaginable for two people to feel pain in the same—not just the corresponding—place.

> That might be the case with Siamese twins, for instance.) (Wittgenstein 1958, §253)

Ōmori looks quite aware of the fact that the same case of "Siamese twins" is used by himself and later Wittgenstein in entirely opposite directions. In Ōmori's own case, it is used in a solipsism-oriented direction, while in Wittgenstein's case, it is used in a publicity-oriented direction. Ōmori also formulates his puzzles about the soundness of the arguments underpinning Wittgenstein's relevant position as follows:

> The point mentioned in the main text (citer's note: it refers to the point that the attribution of "bellyache" to a person is based on the observation of his relevant behaviors, which cannot be transferred into explicit propositions. Cf. Ōmori 1971, 15), if I am not wrong, can be ascribed to Wittgenstein. But I cannot not accept his argument for the publicity of the mental experience, namely, an argument implied by the preceding point. (Ōmori 1971, 17 note 1)

Now a sharp question arises: How could Ōmori use a Wittgensteinian methodology to fight against later Wittgenstein's own position? Here is my answer: His use of the Wittgensteinian methodology, which can be described as a derived version of "phenomenological method" (or "linguistic phenomenology"), is more thorough than Wittgenstein himself, and such thoroughness in turn makes Wittgenstein's own emphasis on the primacy of public language fade away in Ōmori's narrative.

However, in what sense could Wittgenstein's methodology be categorized as "linguistic phenomenology"? My relevant observation is based on Spiegelberg's (1981) general account of the relationship between analytic philosophy and phenomenology. As he (Spiegelberg 1981, 83–93) insightfully points out, J. Austin's treatment of ordinary language, which is undoubtedly within the tradition of later Wittgenstein, is not only labeled by Austin himself as "linguistic phenomenology" (cf. J. Austin 1957) but substantially parallels Husserlian phenomenology in the sense that both philosophers intend to preclude transcendent entities which are beyond the scope of "the Given". (Though for Austin, "the Given" means the linguistic phenomena whereas for Husserl "the Given" means "the phenomena within consciousness", this distinction can be neglected from a high-level perspective. Cf. Spiegelberg 1981, 85). Hence, there is no reason not to

apply the label of "linguistic phenomenology" to later Wittgenstein, whose indifference to natural science in philosophical discourses is a perfect counterpart of Husserl's rejection of the so-called "naturalistic attitude" (cf. Wittgenstein 1980, §218, where he claims that a purely phenomenological color theory does not need to appeal to scientifically identifiable entities like "cones", "rods", "waves", etc.). The same phenomenological tendency could be also found in citations used by table-1, *PI* §295 (Wittgenstein 1958*,* §295), according to which the nature of philosophy is phenomenologically construed as an activity of "seeing just such a picture", a formulation fairly similar to the Husserlian notion of "eidetic intuition" (cf. Husserl 1913/1982*,* sec. 2).

Spiegelberg is very likely ignorant of Ōmori's philosophy, which is definitely precluded from Spiegelberg's own historical account of the "context of phenomenology". But if he could read Ōmori's philosophy, he, as I believe, would quickly identify it as a new variant of "linguistic phenomenology", which is featured both by a Wittgensteinian skin and a Husserlian core. More precisely, a more formal name of Ōmori's position is "*tachiaraware*-based monism". As I will explain immediately, it is via the notion of "tachiaraware", that a more thorough execution of the phenomenological method could be possible. This method further helps Ōmori to overcome the phenomenologically ungraspable dichotomy between sense-data and perceptual structures. By contrast, Wittgenstein's assimilation of perceptual structures to grammatical structures, which are supposed to be a part of a public language, may lead a Wittgensteinian to be committed to such a dichotomy, resulting in at most a lukewarm version of linguistic phenomenology.

3. Ōmori's *Tachiaraware*-based Monism

As aforementioned, a more formal expression of Ōmori's position is "*tachiaraware*-based monism". Here, the Japanese term "*tachiaraware*"(立ち現われ) literally means "phenomena manifesting/standing for themselves", which can be more succinctly but less precisely translated as "emergence" or "appearance". It looks somehow similar to the Russellian term "sense-data", but without the atomist implications of the Russellian logic atomism (Russell 1918*;* 1919). Hence, it looks more similar to Wittgenstein's conception of "phenomena", which is deeply interwoven into a priori grammars like that of color-space, and such a space can be

unpacked as a cluster of necessarily true propositions like "Red is warmer than green." (cf. Wittgenstein1929/1994, 213).

Ōmori is probably unfamiliar with Wittgenstein's comments on color-space. That said, as a reader of Husserl, he is aware of the relevance between his conception of *tachiaraware* and the Husserlian conception of "*Abschattung*". As he says:

> *Tachiaraware* is something equivalent to the Husserlian conception of "*Abschattung*". However, my notion is still different from Husserl's in the sense that in my case, in the mode of *tachiaraware*/*Abschattung*, "what is pointed at" is the *tachiaraware* "directly" standing for themselves. (Ōmori 1976/2015, 201)

A brief commentary is needed here. "*Abschattung*" means "profile", "adumbration" or "aspect" in English. It is used to highlight the "link between transcendence and time", or the point that "human perception always overruns itself with its anticipations and protentions on the one side as well as its retentions on the other" (Moran 2004, 161). Hence, this term helps to explicate "the idea of an action of a shadow that gradually presents defined contours" (Veríssimo 2016, 522). Accordingly, in Husserl's context, the functioning of *Abschattung* presupposes some form of mental activity which gradually makes the contours of perceived objects visible. In contrast, according to Ōmori's previous citation, contours of perceived objects will directly stand for themselves without being the results of some further mental activities synthesizing material which is supposed to be more primary. Here we can easily perceive the metaphilosophical similarity between Ōmori and Wittgenstein. It is obvious that the Husserlian dichotomy between synthesizing activity and perceived objects or projected meanings, or the so-called *Noesis-Noema* dichotomy, presupposes a form of the Aristotelian form-matter dichotomy. The application of this dichotomy requires some form of reflection, but such reflection may go beyond the scope of what phenomenological subjects can actually perceive. In contrast, Ōmori's strategy is just to appeal to *tachiaraware* as what can be directly perceived by subjects in a reflection-free manner. Assuming that the Husserlian intuition-reflection contrast could be metaphorically construed in terms of the dichotomy between "savages" and "civilized people", Ōmori's preceding strategy simply echoes Wittgenstein's following comment: "When we do philosophy we are like savages, primitive people, who hear the expressions of civilized men, put

a false interpretation on them, and then draw the queerest conclusions from it" (Wittgenstein 1958, §194).

However, the preclusion of the "civilized" reflection from *tachiaraware* by no means implies that perceptions are free from any form of thoughts. Rather, a rejection of the Aristotelian form-matter dichotomy, which is the result of the philosophical reflection of the preceding type, precisely reveals the thought-imbued nature of any perception from a *tachiaraware*-based point of view. This is also highlighted by Ōmori's following comment:

> Perceptions, or the *tachiaraware* of perceptions, definitely cannot exist without thoughts. In other words, the multiplicity of intuitions is not possible to exist, if all elements of understanding are precluded from intuitions. However, no matter how "intensive" thoughts are "injected into" the perceptions, it is still a cake work for a small kid to quickly grasp the differences among cases like "seeing a table", "touching a table" and "imaging a table in his mind's eye without seeing or touching it". Hence, it is unforgettable that pure perceptual *tachiaraware* cannot be regarded as something like sense-data, since all perceptions are thought-imbued. (Ōmori 1976/2015, 305)

Prima facie, the spirit of Ōmori's preceding comment bears some affinity with Wilfrid Sellars' (1997) criticism of the "myth of the given", since both Ōmori and Sellars' criticisms lead to the denial of the existence of sense-data *tout court,* if they are supposed to be immune to any reasoning on a higher level. But unlike Sellars' position, the Omori's version of the "rejection of the myth of the given" is substantially supplemented by his solipsism-oriented idea that the thoughts which are supposed to pervade perceptions are fundamentally "my" thoughts:

> However, no matter how much information language could acquire from external environments, and no matter how many times language was adjusted, all this is merely for one's purpose for learning or adjusting the language. For me, the meaning of language could be construed only from my perspective. Even the language of others is nothing but the language that I can understand. For instance, when somebody else says "red sedan", no matter how he understands the term "red", and whatever sensations that he has, in my own case, my understanding of "red" is always based on my

> understanding of my term, and the attribution of the meaning of the term is always up to me. Though language is sharable by many, and come into being due to its being shared, understanding a language is still one's own business. (Ōmori 1971, 21)

Accordingly, for Ōmori, it is always problematic to accept the Wittgensteinian assumption that the only type of grammatical structure that is permitted to exist is what belongs to a public language. The implicit argument underlying Ōmori's criticism can be reconstructed as the follows:

1. It is part of common sense that mutual misunderstanding between individuals does exist, even though individuals may speak the same public language.
2. There are only two possible explanations of this phenomenon: firstly, misunderstandings arise from the different sense-data privileged by different individuals, whereas the similar divergence does not appear on the grammatical level; secondly, divergences appear both on the levels of sense-data and grammar.
3. The first option looks hopeless, given that it entails a *tachiaraware*-theory-prohibiting dichotomy between sense-data and grammar.
4. The second option is permitted since it does not entail a dichotomy between sense-data and grammar, a dichotomy that is prohibited by the whole *tachiaraware*-narrative.
5. Hence, only the second option is left on the table. Hence, according to the duality-undermining feature of the whole *tachiaraware*-narrative, it will be very probable that the inter-subjective divergence of sense-data do go hand in hand with the inter-subjective divergence on the grammatical level. Hence, it is fairly legitimate to say that everyone speaks his own language. Accordingly, A's attempts to understand B should be construed as the reconstruction of B's language from the lens of A's language.
6. Accordingly, the sameness of A's language and B's language on the level of public language has to be reconstructed as something built on the overlapping place between A's personal language and B's personal language.
7. Therefore, the primacy of a public language cannot be taken as granted.

Nonetheless, it is noteworthy that the preceding conclusion does not imply the possibility of a private language. According to Wittgenstein, "The words of this

[private] language are to refer to what only the speaker can know—to his immediate private sensations. So another person cannot understand the language" (Wittgenstein 1958, § 243). In fact, this is not what Ōmori actually intends to mean. In particular, he definitely allows another person to understand the speaker's own language *in some degree.* What he intends to deny is just the possibility of understanding another person entirely and thoroughly. Hence, it seems that Ōmori stands precisely in the middle between a radical solipsist endorsement of private language (according to which only my language is understandable) and a Wittgensteinian endorsement of the public language (according to which understanding is possible only by resorting to a public language). This moderate position surely facilitates Ōmori's explanations of both the inter-subjective misunderstanding and understanding, whereas Wittgenstein's position can only explain mutual understanding alone.

However, there is a further question that needs be answered here, otherwise Ōmori would still be recognized as a "*radical* solipsist": how could the publicity of scientific language, rather than ordinary language alone, be accounted for with more details in Ōmori's framework? The key word included in the requisite response is "*Kasane-egaki*"(重ね描き).

4. Ōmori's Reconstruction of Publicity in Terms of "*Kasane-egaki*"

The literal meaning of *Kasane-egaki* is recoloring, or a procedure of applying a new layer of color on a previously painted surface to produce an effect of a mixed color. Thereby the initial layer of the color is not completely covered; rather, it can still reveal its own shades under the cover of the second layer. Ōmori uses *Kasane-egaki* to analogically explicate the relationship between the first-layer language and the second-layer language. For instance, if the first-layer language is construed as a *tachiaraware*-oriented language, then an operation of *Kasane-egaki*/recoloring will result in a second language through which the "shades" of the first layer could still be seen. Conceivably, the similar *Kasane-egaki*-relationship could hold between the second layer and the third layer, etc..

If Ōmori's whole philosophical career is taken into account, there are three layers of languages involved in the preceding Kasane-egaki-relationships. Not surprisingly, the most fundamental level is *tachiaraware*-oriented, and a more formal name of this language is "*chikakuzōgo*" (知覚像語), or " the language of the perpetual images" in English. The second-layer language is "*monogengo*" (物言語)

or "*nichijōgengo*"(日常言語), which respectively mean "the language of objects" and "the ordinary language" in English. The third layer is "*kagakugengo*" (科学言語) or "scientific language" in English (cf. Ōmori 1971:94–95). For the sake of brevity, hereafter I will call the preceding three languages as Language I (L1), Language II (L2) and Language III (L3), respectively. [4]

Insofar as the *Kasane-egaki*-relationship between L1 and L2 is concerned, it can be more specifically defined as a relationship between elements of a set and a set itself. For instance, the former could be illustrated as linguistic entities encoding specific perpetual images of, say, a cup, whereas the latter could be illustrated as a symbol representing "the cup itself"—a symbol which is nothing but an idealized set composed of infinitely many elements of images of the very cup. The "cup itself" is definitely the "*Kasane-egaki*" of the images of the cup in the sense that it covers the individual images in some degree on the one hand, and reveals the information of the individual images in some degree on the other. And the set itself does not come from nowhere. Rather, it is produced in accordance with a mental algorithm guiding the direction of *Kasane-egaki* (Ōmori 1971*,* 91). More specifically, which images have to be included as elements of the requisite set is "not merely defined in accordance with one's own habit or intellectual interests, but in accordance with one's way of betting on his own life and hence close to one's own life" (Ōmori 1976/2015*,* 219). Hence, Ōmori seems to indicate here that arbitrary and hence irresponsible decisions concerning the direction of *Kasane-egaki* have to be precluded for the sake of the security of life. Accordingly, *Kasane-egaki* definitely requires some prudence in producing the requisite sets to achieve minimal successfulness in intersubjective communications concerning the same set. This position is definitely less solipsism-oriented than his *tachiaraware*-narrative.

[4] The introduction of "*Kasane-egaki*" may be regarded by some researchers as a feature that has to be attributed to later Ōmori, but it is noteworthy that it occupies a salient position even in Ōmori (1971, 283–284), which is often regarded as the representative work of early Ōmori. Hence, as to the division of Ōmori's philosophical career, I am quite sympathetic to Noya's (2015, 33–34) point that the transition of Ōmori's philosophy is merely a transition of his methodology for treating experiences rather than that of his philosophical position. To be more specific, according to Noya (2015, 82), early Ōmori is more interested in treating the *Kasane-egaki*-relationship between L1 and L2, while later Ōmori is more interested in treating the *Kasane-egaki*-relationship between L2 and L3. And difference of this type is definitely less significant than the difference between early Wittgenstein and later Wittgenstein.

Now let's move on to the *Kasane-egaki*-relationship between L2 and L3, namely, that between ordinary language and scientific language. Ōmori's relevant comment goes as the follows:

> It is necessary to abandon the routinely conceived idea that "atoms do exist first of all, and we just use a language to express their existence afterwards". The truth is precisely the opposite: First of all, through the operation of *Kasane-egaki*, which is applied to the ordinary language, what is invented is a new way of talking, namely, the scientific language as a new language. Afterwards the meaning of the atoms could be developed within this language. Accordingly, I expect that people can think in the following way: objectivity is revealed just as the result of its being discussed, and the objectivity itself keeps on developing itself within its existence. Hence, a new meaning of existence is produced in a new linguistic narrative which is called as the "scientific language". It is the narrative of scientific language that creates the meaning of the existence of atoms. (Ōmori 1992, 142)

It is obvious that in this citation, Ōmori, being a *tachiaraware*-based monist notwithstanding, still shows his minimal respect to the ontological commitments made by scientists. These commitments are made merely within the sphere of a L3, which is produced via the *Kasane-egaki*-guiding operations of L2, and L2 is in turn the result of the *Kasane-egaki*-guiding operations of L1.

But what about the details concerning the creation of L3 on the basis of L2? Here, Ōmori (1994, 75–134) appeals to the dichotomy between "*ryakuga*"(略画) and "*mitsuga*" (密画), which mean "sketchy painting" and "meticulous painting" respectively. An example of the former is a sketchy representation of, say, the landscape, while an instance of the latter is a precise map produced with the aid of more complicated tools. According to Ōmori, the evolution from the sketchy paintings to meticulous ones looks unavoidable due to practical interests related to the mobilization of military troops or economical activities (Ōmori 1994, 97). Another motivation for the evolution is to eliminate unexplainable contradictions involved in sketchy paintings, such as the contradiction between, say, Kepler's new cosmological observation of the movement of Mars and the pre-Keplerrian sketchy cosmological law that the orbits of planets are round rather than elliptical (Ōmori 1994, 98–99).

A more contemporary-philosophy-of-mind-friendly label of the preceding *Kasane-egaki*-relationship may be "non-reductionism", according to which a description on a high level can be supervenient on a description on a lower level without being reducible to the latter. This stance can be further cashed out in terms of non-reductive physicalism, e.g., Donald Davidson's (1980) anomalous monism and John Searle's (2004) biological naturalism. According to non-reductive physicalism, mental properties are metaphysically physical events, but they are linguistically/epistemologically irreducible to physical events. Ōmori's *Kasane-egaki*-based narrative can be also viewed as an illustration of non-reductionism, but in an opposite direction: the most fundamental layer in Ōmori's sandwich-like *Kasane-egaki*-structure is L1, which corresponds to phenomena standing for themselves, while languages on higher levels, namely, L2&L3, are constructed out of L1 and are irreducible to L1. Hence, analogous to his non-reductive physicalist counterpart, Ōmori's position is also intended to achieve a compromise between the scientific narrative and folk theories (folk psychology in particular), though his starting point is Husserlian rather than naturalistic. Therefore, his position can be fairly labeled as "non-reductive phenomenology" due to the phenomenological nature of L1 and the irreducibility of L2&L3.

However, what Wittgenstein would say about Ōmori's *Kasane-egaki*-based account of L1-L2-L3-hierachy? He, according to my understanding of Wittgenstein's philosophy, would have to rearrange this hierarchy in a way that his starting point is neither Husserlian nor naturalistic but inter-subjective. Accordingly, the most fundamental layer in a Wittgensteinian narrative is L2, on which L3 is built, whereas L1 does not deserve an independent niche since it is supposed to play merely a marginal role in L2, a role that is ontologically ambiguous between "existence" and "non-existence" (cf. Wittgenstein 1956, §304). It is not hard to perceive that this rearrangement itself is based on two theoretical resources: firstly, insofar as the priority of L2 to L3 is concerned, this rearrangement is based on a general sense of "linguistic phenomenology", given Wittgenstein's observation that L3 is not something directly revealed in ordinary linguistic phenomena (cf. Wittgenstein 1980, §218); secondly, insofar as the marginalization of L1 within the sphere of L2 is concerned, it is simply based on Wittgenstein's famous private language argument, according to which sensation-relevant language-games can be played even without actually possessing the sensations in question.

But from Ōmori's perspective (surely from the lens of my understanding of his perspective), Wittgenstein's position is untenable since it cannot be compatible

with the following principle, namely, the Double-Standard-Abominating Principle (hereafter DSAP), which is self-evidently true:

> **DSAP**: If both objects A and B bear a desirable feature F, and A bears F more than B does (ceteris paribus), then there is no way to prefer B to A, otherwise double standards will be applied to A/B.

However, the following argument can easily show that Wittgenstein has violated DSAP:

1. L2 is preferable to L1, since L1 is peripheral to L2. (Wittgenstein's position)
2. L2 is preferable to L3, since L2 is more fundamental than L3. (Wittgenstein's position)
3. Wittgenstein's reason for preferring L2 to L3 is related to the affinity between L2 and linguistic phenomena.
4. Hence, insofar as the L2-L3 relationship is concerned, "bearing affinity with linguistic phenomena" (hereafter F_1) is the highlighted feature of any candidate language which is supposed to be properly located in a Wittensteinian *Kasane-egaki*-hierarchy.
5. Locally speaking, Wittgenstein's preference for L2 to L3 does not violate DSAP, since L2 does bear F_1 more than L3 does.
6. Wittgenstein's reason for preferring L2 to L1 is related to the affinity between L2 and openness to publicity.
7. Hence, insofar as the L1-L2 relationship is concerned, "openness to publicity" is the highlighted feature (hereafter F_2) of any candidate language which is supposed to be properly located in a Wittensteinian *Kasane-egaki*-hierarchy.
8. Locally speaking, Wittgenstein's preference for L2 to L1 does not violate DSAP, since L2 does bear F_2 more than L3 does.
9. Nonetheless, if L1&L2&L3 are all taken into account and F_1 is the highlighted feature, then L1 simply bears more F_1 than L2/L3 do, hence, L1 has to be preferred. And Wittgenstein's rejection of preferring L1 has violated DSAP.
10. Symmetrically, if L1&L2&L3 are all taken into account and F_2 is the highlighted feature, then L3 simply bears more F_2 than L1 or L2 does. Hence, L3 has to be preferred. And Wittgenstein's rejection of preferring L3 has violated DSAP again.
11. Therefore, for Wittgenstein, there is no way *not* to violate DSAP, no matter

whether F_1 or F_2 is highlighted.

Facing the preceding criticism, a Wittgensteinian may contend that Wittgenstein himself does not violate DSAP if the highlighted feature is neither F_1 alone nor F_2 alone but F_1 &F_2. Accordingly, the preference of L2 is completely compatible with DSAP. However, my rejoinder is this: this strategy can work only when having F_1 is not going to be conflicting with having F_2. But a further problem is: how to check the compatibility between F_1 and F_2, given that some seemingly mutually compatible feature-pair may turn out to be mutually incompatible on a deeper level? The best way to do the desired check is to appeal to the "extremity test" (ET), which can be used to pick out a considerable number, albeit perhaps not all, of feature-pairs which are potentially mutually incompatible:

> **ET**: If the intensities of both F_1 and F_2 are enhanced to extremities to result in, say, F_{1E} and F_{2E}, and if F_{1E} and F_{2E} are patently mutually incompatible, then F_1 and F_2 are at least potentially mutually incompatible as well. Otherwise F_1 and F_2 are very likely mutually compatible.

ET can be illustrated via the following cases. "Being red" is completely compatible with "being round" in the light of ET, since no matter how red or how round a patch is, a round-and-red patch is always conceivable. In contrast, "being loyal to a king" and "being prudent" are not mutually compatible due to the following consideration: Prima facie, surely there are many ministers who are both loyal to their kings and are judged as prudent by historians. But if the intensities of both "loyalty to the king" and "prudence" are enhanced to extremities to result in "*blind* loyalty to the king" and "*extreme* prudence", then it will not be hard to perceive the impossibility of being an extremely prudent minister who is simultaneously blindly loyal to the king, given that the intellectual autonomy required by "prudence" is not compatible with the preceding type of intellectual blindness.

For similar reasons, F_1 &F_2, namely, "bearing-affinity-with-phenomena-and-open-to-publicity", cannot pass ET either. Surely public ordinary language can carry this compound feature, but to pass ET requires more than this. Specifically, if the intensities of both "bearing affinity with phenomena" and "being open to publicity" are enhanced to extremities, then what we can get is nothing but: (1) *tachiaraware*, or phenomena standing for themselves *in a science-precluding manner*; (2) *scientific* accounts in a subjectivity-precluding

manner. The mutual incompatibility between the two is no less than that between "being blindly loyal to the king" and "being prudent". Therefore, Wittgenstein still violates DSAP.

Conceivably, a quick solution to Wittgenstein's preceding predicament is simply to acknowledge the primary status of either L1 or L3. To choose L3 will naturally lead to physicalist monism with a scientific flavor, which can be viewed as the result of carrying on later Wittgenstein's publicity-oriented proposal to the extreme. By contrast, choosing L1 will quickly result in Ōmori's *tachiaraware*-based monism with a Husserlian flavor, which can be viewed as a result of carrying on later Wittgenstein's phenomena-oriented proposal to the extreme. Between the two preceding remedies, Ōmori's route appears to be a bit more Wittgensteinian in the sense that his theory, especially through its *Kasane-egaki*-narrative, shows more respect to commonsensical language than what L3-adherents do, whereas to have this minimal respect is the bottom-line that any Wittgensteinian has to hold. Therefore, compared with physicalism, Ōmori's solution looks more like a remedy, rather than a replacement, of Wittgenstein's position.

5. Conclusion

Hitherto I have explicated how Ōmori uses a variant of Wittgensteinian linguistic phenomenology to fight against Wittgenstein's own publicity-oriented philosophical tendency. Ōmori's first recipe for doing so, according to my reconstruction, is to appeal to the notion of *tachiaraware*, by which the gap between synthesizing activity and sense-data to be synthesized can be filled. Therefore, the first-personal character of *tachiaraware* could be easily transmitted to the formal features of "my language", without which there is no public language can be formed. Ōmori's second recipe is to elaborate the L1-L2-L3-hierarchy within his *Kasane-egaki*-narrative in a non-DASP-violating manner, while there is no way for a Wittgensteinian emphasis of the putative primacy of public languages *not* to violate DASP. Hence, since both the respect of *tachiaraware* and DASP are required by the thorough implementation of any phenomenological principle, Ōmori's stance simply appears to be a natural result of radicalizing Wittgenstein's stance alongside the phenomenological route. Thus, as I have mentioned, Ōmori's philosophy could be described as a hybrid system composed of both a Wittgensteinian skin and a

Husserlian core. In this sense, Ōmori's philosophy has a special value for reviving the phenomenological tradition via a Wittgensteinian lens.

Acknowledgement:
Thanks to Haicheng Zhao and Ning Zong for their comments of different drafts of this paper. This research is sponsored by the National Foundation of Social Sciences of China (No. 15ZDB020) and the Foundation of Philosophy and Social Science of the Ministry of Education of China (No. 19JZD010).

References

Austin, John (1957) A Plea for Excuses. *Proceedings of the Aristotelian Society*, LVII, 1–30.

Davidson, Donald (1980) Mental Events reprinted in *Essays on Actions and Events,* (Oxford: Clarendon Press, 1980), 207–224.

Husserl, Edmund (1913/1982) *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*. Trans. F. Kersten. (The Hague: Nijhoff).

Kobayashi, Yasuo (2019) The Komaba Quartet: A Landscape of Japanese Philosophy in 1970s, in Bret W. Davis (ed.), *The Oxford Handbook of Japanese Philosophy*, (Oxford: Oxford University Press), 649–664.

Moran, Dermott (2004) Heidegger's Critique of Husserl's and Brentano's accounts of Intentionality, in Dermot Moran and Lester E. Embree (eds.), *Phenomenology: Critical Concepts in Philosophy*, (London and New York: Routledge), 157–183.

Noya, Shigeki (2015) *Ōmori Shōzō: Philosophical Sample* (Japanese tile:『大森荘蔵――哲学の見本』). (Tokyo: Kodansha).

Ōmori, Shōzō (1971). *Language, Perception and World* (Japanese tile:『言語・知覚・世界』). (Tokyo: Iwanami Shoten).

Ōmori, Shōzō (1976/2015) *Objects and Minds* (Japanese tile:『物と心』). (Tokyo: Chikuma shobō).

Ōmori, Shōzō (1992) *Time and Ego* (Japanese tile:『時間と自我』). (Tokyo: Seidosha).

Ōmori, Shōzō (1994) *The Construction of Knowledge and Constraining Spells* (Japanese tile:『知の構築とその呪縛』). (Tokyo: Chikuma shobō).

Russell, Bertrand (1918, 1919). The Philosophy of Logical Atomism. *Monist*, 28, 495–527; 29, 32–63, 190–222, 345–380.

Searle, John R. (2004) *Mind: A Brief Introduction.* (New York: Oxford University Press).

Sellars, Wilfrid (1997) *Empiricism and the Philosophy of Mind*, Robert Brandom (ed.), (Cambridge, MA: Harvard University Press).

Spiegelberg, Herbert (1981) *The Context of the Phenomenological Movement*. (The Hague: Martinus Nijhoff Publishers).

Veríssimo, Danilo Saretta (2016) Husserl's theory of perceptive donation according to profiles, *Psicologia USP. set-dez*, 27, 3, 521–530.

Wittgenstein, Ludwig (1929/1994). *Philosophische Betrachtungen; Philosophische Bemerkungen* (Wiener Ausgabe Band 2), Herausgegeben von Michael Nedo, (Wien, New York: Springer-Verlag).

Wittgenstein, Ludwig (1958). *Philosophical Investigations*, 2th edition, Trans. G. E. M. Anscombe, (Oxford: Wiley-Blackwell).

Wittgenstein, Ludwig (1980). *Philosophical Remarks*, Rush Rhees and Raymond Hargreaves (eds.), Trans. Maximilian A. E. Aue, (Chicago: University of Chicago Press).

Philosophy for Emperors: Some Lessons to Learn from the Imperial New Year's Lectures

Cheung Ching-yuen
Associate Professor, University of Tokyo

***Abstract**: One of the rituals of the Japanese imperial family is Kōsho Hajime no Gi. Nowadays, it is a ceremony in which the Emperor listens to lecture on human, social and natural sciences by distinguished scholars. Since early Meiji, there have been some lectures on philosophy or delivered by philosophers. The purpose of this paper is not only to report what the Emperors might have learnt from the philosophers, but also to see what we can learn from these lectures. To be precise, I would like to transform "Philosophy for the Emperors" into "Philosophy for Everyone".*

1. Introduction

According to Sakaguchi Ango, the emperor system is "extremely Japanese or even politically original".[1] For example, Toyotomi Hideyoshi took advantage of inviting Emperor Goyōzei (後陽成) to Jurakudai (聚楽第) to gain respect from other *Daimyō*. Yet we can say that Chinese elements are also essential in the making of the so-called *Tennō* system. Until the 7th Century, the emperor in Japan were called *Daiō* (大王), but the name was changed to *Tennō* (天皇) in order to emphasize *Ten* (heaven), a status only shared by *Tensi* (天子) in China. From 1868 to 1945, emperors in Japan were not only regarded as the "Head of the State", but also worshipped as a *Shintō* Deity. After the war, or precisely, on the January 1st 1946, Emperor Hirohito announced that he was and should only be regarded as a human being. Under the new Constitution, effective from 3rd May 1947, the Emperor is nothing more than "the symbol of the State and of the unity of the people" (Article 1).[2]

However, in November 2019 Japan was witness *to Daijōsai*, the "Great Thanksgiving Ceremony", which involves the Emperor eating and sleeping with Gods. When it comes to rituals of the Japanese imperial family, there is an annual

[1] Sakaguchi 2008, 217
[2] http://japan.kantei.go.jp/constitution_and_government_of_japan/constitution_e.html

ceremony of edification for the Emperor and Empress called *Kōsho Hajime no Gi*. The Imperial Household Agency*, or Kunaichō* explains the ceremony as follows:

> The Ceremony of the *Kousho Hajime* (Imperial New Year's Lectures) takes place every January at the Imperial Palace in the presence of Their Majesties the Emperor and Empress. This is a ceremony in which Their Majesties listen to experts' explanations in the fields of human, social and natural sciences in their respective field.[3]

The purpose of this paper is not only to report what the Emperors might have learnt from the philosophers, but also to see what we can learn from these lectures. To be precise, I would like to transform "Philosophy for the Emperors" into "Philosophy for Everyone" (also known as "P4E").

2. The Three Books

On the official website of *Kunaichō*, there is list of recent lectures to the Emperor, but the period is only limited to Heisei and Reiwa periods. I managed to request a full list from *Kunaichō*, and discovered that many lectures are related to philosophy. The lectures used to be categorized into three types: Chinese classics or *Kanjo* (漢書), Japanese classics or *Kokusho* (国書) and Foreign works or *Yōsho* (洋書). Here follows a report on philosophical lectures under these categories.

2.1 *Kanjo*

On the very first *Kōsho Hajime no Gi*, dated 23 January 1869, there were four lectures, of which two were on Chinese texts. These two *Kanjo* lectures are on *the Analects* delivered by two prominent scholars: Higashibōjō Tadanaga (東坊城任長) and Nakanuma Ryōzō (中沼了三). The *Analects* have always been one of the most important classical texts in Japan. Confucianism could be regarded as a moral philosophy, but it also discusses the political relationship between the ruler and the ruled. Idealistically, the ruler should be virtuous, and loyalty is regarded as a virtue for those who are being ruled, but what happens if the ruler was not virtuous? This is a topic that has never been covered in the *Kōsho Hajime no Gi*. Among the *Four*

[3] http://www.kunaicho.go.jp/e-culture/kosho.html

Books and Five Classics (四書五経), we cannot find any lectures on *Mencius* (孟子) and *Chunqiu* (春秋 Also known as *Spring and Autumn*). Mencius was clearly a taboo for the Japanese emperor system, because he suggested that the people are the most important, with the state coming next and the ruler being the least important. As for *Chunqiu*, it is a book on history of China during the Spring and Autumn Period, which was mainly about how one dynasty was usurped by another, something which has ostensibly never occurred in Japan.

We should notice a lecture on the study of the *Analects* in Japan by Takeuchi Yoshio (武内義雄) in 1941, the same year when Nishida Kitarō (西田幾多郎) also delivered his lecture to the emperor. Even in 1945 when Japan was preparing for a *Gyokusai* (scortched earth) -styled war of defense, there was still a lecture entitled "On the characteristics of Chinese culture that respects *li*" by Yano Jinichi (矢野仁一). Since 1951, there has been no more *Analects* lecture for the Emperor. However, it does not mean a complete farewell to Chinese culture. In the post-war era, there have in fact been lectures on Lu Xun and Chinese Literature. For the recently abdicated Emperor Akihito, his first taste of *Kōsho Hajime no Gi* was in 1991, when he received a lecture on "Technological thought in Ancient China" by Yoshida Mitsukuni (吉田光邦).

2.2 *Kokusho*

As we all know, Japan has just entered the new Era of Reiwa. This *Nengō was derived* from *Manyōshū* (万葉集), a collection of ancient Japanese poems. However, *Reiwa* is actually from a text written in ancient Chinese (*Kanbun*). The first two *Kokusho* lectures in 1869 were delivered by Tamamatsu Mahiro (玉松員弘) and Hirata Kanetane (平田鎖胤 the son-in-law of the infamous *Kokugaku* master, Hirata Atsutane). Both lectures were on *Nihon Shoki* (日本書紀), which is not only written in ancient Chinese, but also in the style of *Shiki* (史記): from ancient mythology to the author's contemporary regime. We should also note that Haga Yaichi (芳賀矢一), the author of *Kokuminsei Jūron* (国民性十論 1911), gave a lecture to the Emperor on *Manyōshu*, while Ueda Kazutoshi (上田萬年) delivered a lecture on the "Spirit of Japanese language and Native Studies" in 1924.

Shintoism can be seen as the "state religion" for modern Japan, but interestingly, Buddhist thought was once categorized under *Kokusho*. For instance, there was a lecture on K*ū*kai's Calligraphy-theory and Art by Taki Seiichi (瀧精一) in 1945, the final year of WWII. After the war, Buddhism continues to be a rare

topic for the emperor's lectures. In this category we can only identify *The Middle Way Thought* by Miyamoto Shōson (宮本正尊) in 1963, *Indian Philosophy in Japan* by Kanakura Enshō (金倉圓照) in 1972 and Nakamura Hajime (中村元)'s *Founding of Early Buddhism* in 1975.

2.3 *Yōsho*

The first lecture on *Yōsho* was in 1872 on National Law by Katō Hiroyuki (加藤弘之). Nishimura Shigeki (西村茂樹) also delivered talks on topics such as renaissance. Katō and Nishimura, as well as Nishi Amane (西周), were three important contributors to the *Meiroku Magazine,* an important journal for the promotion of new ideas set up in 1873. The first issue featured contributions by Nishi Amane and Nishimura Shigeki, and in the second issue, we can find Katō Hiroyuki's response to Fukuzawa Yukichi.

Fukuzawa was an advocate for practical knowledge or *jitsugaku*, but he suggested the *Meiroku Magazine* should be banned.[4] While he has never delivered a lecture to the Emperor, liberal thinkers were given the opportunity to deliver lectures. Many lectures were related to Western philosophy, for example Hozumi Yatsuka (穂積八束)'s 1912 Lecture on Aristotle's *Politics*, Tomii Masaakira (富井政章)'s 1918 lecture on Montesquieu's *L'esprit des Lois*, Hozumi Nobushige (穂積陳重)'s 1922 lecture on Kant and Bentham's theories on Perpetual Peace and the origin of the League of Nations, and Tajiima Kinji (田島錦治)'s 1925 Lecture on Adam Smith's *Wealth of Nations.*

3. Lectures by Japanese Philosophers

Up to now, we have yet to see any lectures related to Japanese philosophy. The first ever lecture by a Japanese philosopher was held in 1941 by Nishida Kitarō, followed by Watsuji Tetsurō (和辻哲郎) two years later. Other Japanese philosophers who delivered lectures to the emperor are Abe Yoshishige (安倍能成), Takahashi Satomi (高橋里美), Nishitani Keiji (西谷啓治), Shimomura Toratarō (下村寅太郎), Noda Matao (野田又夫) and Ueyama Shunpei (上山春平). Due to the word limit, I would like to focus on the lectures by Nishida, Watsuji and Takahashi.

[4] Fukuzawa 2009, 439–445.

3.1 Nishida's Lecture on Philosophy of History

Nishida's lecture is relatively well known to the Western academic world, thanks to an English translation by Yusa Michiko.[5] Nishida's lecture can be divided in four sections. At the beginning of the lecture, Nishida focuses on the topic "philosophy as a unifying discipline". According to Nishida, philosophy can be defined as the discipline that unites specialized fields and connects them to our daily lives. This is Nishida's basic position: knowledge is for life.

Nishida continues to see this idea in Western and Eastern philosophies: He suggests that Greek philosophy can be seen as a philosophy of *the polis,* centering in the city life of the Greeks, medieval philosophy can be seen as a religious philosophy, centering in the European Christian life, while modern Western philosophy is a scientific philosophy, centering in the recent scientific culture. On Eastern philosophies, Nishida thinks tradition thoughts such as Confucianism and Buddhism can be regarded as philosophy, and that these philosophical traditions have greatly influenced Japanese thought.

How did Nishida make sure that the emperor understood his "philosophy in a nutshell"? Nishida's strategy is to explain philosophy from a biological perspective, as he knows the emperor had some knowledge of biology. Nishida argues that biological life is always in connection the environment. Similarly, human life is also in constant relation to the environment. We can also create things to change the environment. Things created in the past are continuously having an impact on people in the present and future. In Nishida's words, "This is why we always possess a commonly shared tradition, centered in which we continue to develop our human life. Human life is different from biological life in that it is historical".[6] This is the very position of Nishida's later philosophy of historical life. Nishida's own position is clearly influenced by the biology of J. S. Haldane. Indeed, Nishida admitted that Haldane's position is closest to his philosophical position.[7]

In this talk, he mentioned two aspects: namely: the path of history, and the historical mission of Japan. Here, one of Nishida's key concepts is "globalization". In other words, Japan has to understand its role in the "globalized world". This is indeed the role of nationalism in globalization. In Nishida's own words,

[5] Yusa 2002, 314–318.
[6] NKZ 12: 269; Yusa 2002, 316.
[7] NKZ 11: 289.

> Today, however, because of the development of a global transportation network, the whole earth has become one world. Consequently, today's nationalism *(kokkashugi)* has to take into account what it means to be a nation in the global world. What I mean by "nationalism" is not that every country should retreat to itself [to the isolated idea of nation]; rather, each nation should have a place of its own in this [global] world. In other words, by "nationalism" I mean that each country ought to develop its global perspective within itself.[8]

Nishida's view of nationalism has a strong political message: it should not be a kind of "Japan First" nationalism, but a more globalized view of seeing the mission of Japan in the globalized world. Nishida continues,

> At the time when various ethnic groups enter into a global interaction, I suppose it is in the natural course of events that severe struggles among countries take place. I think, however, that the people who possess the most globally developed historical orientation will play the key role and lining stability to the epoch. What I mean by a nation-state that has a globally developed historical quality is a nation-state that, although subscribing to totalitarianism (*zentaishugi*), does not negate the [rights of] individuals, and whose collective life is mediated by the creative activities of individual persons.[9]

It is clear that Nishida is against the kind of totalitarianism that negates the creativity of individuals. Unlike propaganda emphasizing collective thinking during the war, Nishida emphasizes the importance of individuality. Nishida tried to use an analogical argument to link up biology and history. He writes, "Individuals are born of the historical society, to be sure, but as long as the historical society has the individual's creative activities as its medium [of development], that historical society has an eternal life in terms of its globally historical nature. It can be likened to how biological life continues to live on, being mediated by cellular activities".[10]

We can see a kind of coherence from Nishida's philosophical biology and philosophy of history. Nishida was trying his best to criticize a nation that does not

[8] NKZ 12: 270–271; Yusa 2002, 317.
[9] NKZ 12: 271; Yusa 2002, 317–318.
[10] NKZ 12: 271; Yusa 2002, 318.

respect to individual creativity. However, this is not the end of this talk. Nishida concludes by emphasizing the role of the imperial family:

> In the history of our country, the whole and the individual usually did not stand in opposition. Rather, [history] has unfolded with the imperial family (*kōshitsu*) as its center, while the individual and the whole mutually self-negated. Certainly, there were times when the power of the "whole" overshadowed that of the individual, but each time we returned to the founding spirit of Japan, and by maintaining the central presence of the imperial family, we took a step forward into the new era and created a new epoch.[11]

Nishida tried to justify his position in the notion of *Fukko Ishin*: "the restoration of the old ways" (*Fukko*) and "thoroughgoing renewal" (*Ishin*). Japan's role in the new era is not to negate the old completely, as is the case for a radical political revolution in the case of Modern China. We can say that Nishida was trying to push the envelope to try to educate the emperor on the need to avoid the isolation of Japan, but Nishida was still playing safe, or even politically correct, to justify that the imperial family should play a central role in leading Japan into a new era. As a philosopher who received "Order of Culture" in 1940, Nishida had no clue that Japan was preparing to attack Pearl Harbour in the next year, but he should know that modern Japan had already expanded its territories to Taiwan, Manchuria and Korean Peninsula. The colonizer had to face the others, the colonized.

3.2 Watsuji's Lecture on Shinkei's *Renga* Poetics

Two years after Nishida's lecture (under the *Yōsho* series), Watsuji was invited as the speaker to the Emperor (under the *Kokusho* series). At that time, Japan had already announced the idea of Greater East Asia Co-Prosperity Sphere, but militarily speaking, Japan was losing ground after the Battle of Midway. Some Kyoto School philosophers, such as Tanabe Hajime, Kōyama Iwao and Nishitani Keiji, began to imagine the possibility of Japan losing the war.[12]

Unlike Nishida who was already retired when he gave the lecture to the emperor, Watsuji was a full professor at Tokyo Imperial University. He had just

[11] NKZ 12: 271; Yusa 2002, 318.
[12] Ōhashi 2001.

published the second volume of *Rinrigaku* (倫理学 1937, 1942, 1949), in which he tried to justify the State as the highest social-organization. However, Watsuji in his younger age did criticize the State. For example, in *Koji Junrei* (古寺巡礼 1919), he showed his courage to publish a book on his pilgrimages to ancient temples in Nara, where a Cultural-Revolution-styled destruction of Buddhist temples was taking place due to the national policy of State Shintoism.[13]

Would Watsuji deliver a "political correct" lecture to the emperor? My reading is that Watsuji indeed tried to deliver a very subtle political message to the Emperor. In his lecture titled "On Shinkei's *Renga* Poetics (心敬の連歌論に就いて)", Watsuji gives us some details about Shinkei (心敬, 1406–1475), a monk who lived in Muromachi Period. This monk was familiar with Confucian thought, and was a famous *Renga* poet. *Renga* has two character: *Ren* (連) and *Ga* (歌), which means "linked-poems". In *Renga*, we can see two poets writing poems, with the second poet tries to link his poem to the one written by the first poet. In other words, *Renga* is a "genre of Japanese linked-verse poetry in which two or more poets supplied alternating sections of a poem. The *Renga* form began as the composition of a single *Tanka* (Japanese poem of thirty-one syllables) by two people and was a popular pastime from ancient times, even in remote rural areas".[14]

Here, Watsuji focuses on the essence of *Renga*. First of all, it is clear that *Renga* is a collective creation. Unlike Western art, *Renga* is unique in the sense that it is not to be reduced into one artist. Watsuji's examples are architecture, sculpture, painting and music. In these cases, all individual efforts were ultimately concluded under the name of one leader or a designer, who a person takes up the sole responsibility of the artistic creation. In Watsuji's words, these creations are by "one artist" who creates something with the hands of many others; they are not collection creation done by "many artists".[15]

In the case of Western literature, Watsuji mentioned the case of Ancient Greek epic poems (E.g. Homer's Odyssey). These poems were from various poets, but finally receive a single author. However, it is not the case in Japanese *Renga*. In *Renga*, it is crucial for the second poet to fully respect the poem written by the first poet. The second poet, for instance, cannot rewrite the poem by first. For example,

Frost falling. Dew vanishes from sight. Ninzei

[13] See my article, "In the Wake of 3.11 Earthquake: Philosophy of Disaster and Pilgrimage" in Yusa 2017, 133–149.
[14] https://www.britannica.com/art/renga
[15] WTZ 23: 258.

Cold sea breeze blows and the moon appears. Shinkei[16]

Here, Ninzei's poem is called the "previous poem" (前句), and Shinkei's poem is "my poem" (我句). From Ninzei's poem, we have the hint of a cold weather from "frost", and Shinkei suggests the location would be somewhere along the coast. Now we can see Shinkei tried to understand the scenario of the first poem, and to develop the second poem based on the first. Watsuji further argues that while two poems are "linked" to each other, they are not dissolved into one. This is to link two persons together, without losing their individualities. Besides, when the two poems are linked, it is not merely a unity of two poems, but also the unity of two persons. It opens the possibility of understanding the experience and feeling of the other. In other words, a poet will respect, understand and sympathize with the other. It is not merely a poetic creation, but a realization of human relationship.

For Shinkei, the most important thing of *Renga* writing is to "dissolve your heart in the first poem" (前句に心をくだくべきこと). Watsuji would see this as "the primary position when facing the other" (これまさに他の人格との交わりの第一の心得なり).[17] Without this deep understanding of the other, a *Renga* has no value even it shows good technique in writing. On the contrary, a *Renga* without sophisticated techniques can still be an expression of ultimate beauty, as long as it tried to reach the heart of the first poem. Here, to dissolve, literally, is to break your heart or break your bone. In other words, you will have to empty yourself in order to face the other. To be a *Renga* poet, it is necessary to suppress the ego or self. Similarly, it is necessary to be selfless as a person. It requires a training, which is not on technique but on the personality or spirituality.

What does it mean to be selfless? Is it merely a self-negation and a total acceptance of the other? Watsuji notices one should not merely focus on his or her own poem, but it does not mean to just follow the first poem. "To throw away oneself is not to the negation of one's character, but the way to become unique" (己れを捨てて前句にのみ心を砕く者かえってよく独自の句を制作し得という。己れを捨つるは個性の没却にあらずしてかえって個性の円成なり).[18] In other words, one should not "blindly follow the other" (付和雷同). An example would be that if one associates flowers to plum or cherry blossom, it would be agreed by all

[16] My translation. Original poems in Japanese:
霜のふるまがひに露や消えぬらん 忍誓
はま風さむしすみの江の月 心敬

[17] WTZ 23: 260.

[18] WTZ 23: 261–262.

(満座同心) but the meaning would be shallow (浅薄). In the case of Shinkei's *Renga*, we can see he did his best to understand the first poem, but did not just follow it without his own creativity.

Shinkei further develops his poetics with the theory of "close poems" and "distant poems" (親句疎句). Watsuji explains,

> Close poems mean the first poem and the second poem are closely related. Their link to each other is apparent. In contrast to this, distant poems mean the first poem and the second poem are seemingly unrelated. They are isolated to each other, but their hearts are together and linked subtlety. Shinkei calls distant poems are "connected by nothing (あらぬさまに継ぎたるもの);" they are more profound than the close poems. In *Sasamegoto* (1463), the two kinds of poems are compared as follows: "The close poems are [Confucian] teaching, the distant poems are *Zen*. The close poems are *sanimitta* (有相), the distant poems are *animitta* (無相)". "The poetics of the *sanimitta* should apply the poetics of the *animitta*-darma-body distant poems".[19]

What Watsuji actually trying to argue here, again, is to apply this poetics to how we deal with the other in a community. Shinkei's advice is that, we should not obsess with our own self and forget the unity with the other, but we cannot only agree with other and lose our own characters and creativity. To sum up, the philosophy of *renga* is to embrace the other. Indeed, Shinkei quotes the *Analects* in *Sasamegoto*: "The noble man is all-embracing and not partial. The inferior man is partial and not all-embracing (君子周而不比、小人比而不周)".[20]

Now it is clear that Shinkei's way of *Renga* is nothing but the principle of human relationship in a community. The way of *Renga* is, in other words, the way of human relationship. This is a typical way of a unity of art and morality. This opens up new ways in three directions. First, it demonstrates a unique aspect of Japanese art. Second, it can develop new theory of art. Third, it can provide strong reasoning for research on human relationships. It may be said that to talk about poetics during the Pacific War is rather ridiculous. However, Watsuji has this idea when he delivered his lecture to the Emperor: Japan should not forget the way they used to deal with others.

---

[19] WTZ 23: 262.
[20] WTZ 23: 263.

3.3 Takahashi's Lecture on Forms of Love

Takahashi Satomi is one of the pioneers of phenomenological research in Japan. "In 1921 he assumed a post in the science faculty of Tōhoku [Imperial] University in Sendai. He subsequently spent two years studying abroad in Germany with Rickert and Husserl".[21] He is the author of *Husserl's Phenomenology*, which was published in 1931. Takahashi is also well known as a critic of Nishida's philosophy. Back in 1912, he wrote a paper titled "The Fact of Consciousness-phenomenon and its Meaning" to review Nishida's *An Inquiry into the Good*, published a year earlier. It was one of the earliest philosophical debates in the history of modern Japanese philosophy.

Takahashi's basic position can be summarized as below: "The totality of enveloping, both in terms of content and in terms of experience, must be regulated as a love that is a single unity embracing will and action along with knowledge. In this way, the ultimate consists of absolute love as empirically regulated absolute nothingness. Hence, all things, at bottom, can be wrapped together in an absolute love in which at once all is one and one is nothingness".[22] According to Takahashi, Nishida's notion of love can be understood as a "dialectical love" (弁証法的愛). "Dialectical love is to see the self in self with the absolutely contradictory other, and to see the other in the other with the absolutely contradictory self".[23] Takahashi criticizes Nishida's dialectical approach, and develops his own philosophy of "one-being-love" (一在愛).

In an article titled "A System Which Includes Dialectic" (written in English), Takahashi examines different types of dialectic: "dialectic of process", "dialectic of field", "dialectic with two poles", "dialectic of pure negation or pure movement", "dialectic of the middle", "dialectic with three poles", "dialectic with an infinite number of poles" and "dialectic of the whole and parts", etc. Takahashi develops his own dialectic, which is a "wholeness which includes all the dialectic". Takahashi argues that "Hegel insisted that his absolute idea contained as negative-and-preserved (*aufgehobene*) moments all the dialectic processes which have occurred before becoming itself... Hegel's '*aufheben*' (sublation) implied in the idea of the Absolute means nothing more than the result, and so he did not succeed

[21] Heisig 2011, 822.
[22] Heisig 2011, 827–828.
[23] TSZ 5:226.

in attaining the full idea of '*aufheben*' which he had intended to realize. The realization of this includes all dialectics or 'wholeness which includes and transcends' all processes existing along the course of dialectic development". Takahashi further develops his own dialectics in which love is the unifying principle. He writes,

> Love unifies intellect, feeling and volition by including as well as transcending them, while enabling them to continue to exist. The author believes, moreover, that this all-inclusive whole itself should be included in "Absolute Nothingness" of which the ethical or religious counterpart is "Absolute Love". That is why in the last analysis he maintains that all is included in absolute love.[24]

In 1956, Takahashi was invited to deliver a lecture to the emperor. His topic is "Forms of love as basic motivation of culture (文化の根本動機としての愛の諸形態). In the beginning of this lecture, Takahashi suggests that "Love is the fundamental feeling of human being. Without love, it is inconceivable to have nation, state or their co-existence. It is the principle that connects human beings, and unites them into a community. Therefore, love should be understood as the basic motivation of world culture".[25]

For Takahashi, *eros* and *agape* are two classical philosophical concepts of love: the former is an upward movement and the latter a downward movement. To borrow Max Scheler's words, "All ancient philosophers, poets, and moralists agree that love is a striving, an aspiration of the "lower" toward the "higher"... The Christian view boldly denies the Greek axiom that love is an aspiration of the lower towards the higher. On the contrary, now the criterion of love is that the nobler stoops to the vulgar, the healthy to the sick, the rich to the poor, the handsome to the ugly, the good and saintly to the bad and common, the Messiah to the sinners and publicans".[26]

To overcome this "contradiction", we can expect a kind of love that unites *eros* and *agape*. According to Takahashi, Nishida's notion of love can be understood as a "dialectical love" (弁証法的愛). "Dialectical love is to see the self in self with the absolutely contradictory other, and to see the other in the other with the

[24] TSZ 5: 316–317.
[25] TSZ 5:202
[26] Scheler 1961, 85.

absolutely contradictory self" (TSZ 5:226). Takahashi criticizes Nishida's dialectical approach, and develops his own philosophy of "one-being-love" (一在愛).

It is noteworthy that Takahashi mentions the Schelerian notion of *Einsfühlung*: "What I called one-being-love (一在愛) is similar to what Scheler calls *Einsfühlung* (一体感). However, this love is not about the foundation of different forms of sympathy, as in the case of Scheler; rather, it includes all other things". (TSZ 5:231-232). Elsewhere, Takahashi mentions Scheler's *Nature of Sympathy* as "the most remarkable work on *Einsfühlung*. (TSZ 5: 197) Takahashi agrees with Scheler that "The ultimate love is consciousness-identification (*Einsbewusst*) and emotional-identification (*Einsfühlung*)" (TSZ 5:269). I believe Takahashi can be regarded as one of the earliest Schelerians in Japan, and his project is to interpret love as *Einsfühlung*. It is an important event in the history of modern Japanese philosophy.

Takahashi continues to argue that love is the very essence of Japanese culture, i.e. "harmony" (和). In postwar Japan, Japanese are facing "the suffering of the neighbours, nature disasters and the massive killing by nuclear bombs" (TSZ 5:42). As Scheler would emphasize the role of philosophical anthropology for providing a "unified" idea of man in the age of crisis, Takahashi suggests reflecting on a "unified" notion of love in a difficult postwar era. To borrow Takahashi's own words,

> One may be proud of the advantages of the Japanese notion of one-love, but she or he should also realize the drawbacks of this notion. In order to beware and avoid these shortcomings, sometimes we will have to emphasize on *eros*, while in other occasions it is necessary to emphasize on *agape*, *philia*, or even dialectic love. But eventually, we need to try to develop love as one-being-love, which encompasses all the other notions of love. (TSZ 5: 247).

In fact, Takahashi did mention *philia* as the fifth definition of love, followed by *eros*, *agape*, dialectical love and one-being-love. It comes to another difficult question: what is friendship? As discussed in Plato's *Lysis*, what does it mean to be friends? Do friends have all things in common, or have nothing in common? Philosophy, or the love of wisdom, is about *philia*. Like the cases of man and love, one will have to search for a "unified" idea of friendship, in which all notions of friendships can be included in one.

Takahashi is unique for being a Japanese philosopher who uses *Einsfühlung* to explain love. This attempt is similar to Nishida's early philosophy which tried to "unite" the differences between knowledge and faith, philosophy and religion, religion and culture, etc. But Takahashi and Nishida are still different in many ways. While Nishida argues *agape* is the foundation of *eros*, Takahashi does not agree with this standpoint. Besides, Takahashi noticed a different in the sense that Nishida is more influenced by *Zen* (禅), but Takahashi himself is rather influenced by *Jōdo Shinshū* (浄土真宗, the True Pure Land Sect of Buddhism).[27] Here, it is impossible to go into details of the two Buddhist sects, but it is clear that Nishida and Takahashi are not simply "Zen Buddhist" or "Pure Land Monk". They philosophize on various topics, including the problem of sympathy and love. Both philosophers tried to avoid a one-sided "nationalistic" approach to philosophical problems. For Takahashi, he understands Japanese philosophy as a "global Japanese philosophy" (世界的日本哲学).[28]

Takahashi's lecture could be understood as a development of the lectures by Nishida and Watsuji. These lectures provide some answers to the questions: What is the essence of Japanese culture? How can we (the self) deal with the other?

4. Concluding Remarks

We have just mentioned three lectures by Japanese philosophers, but what can we learn from them? In my opinion, it is not only to understand what they said, but also what they did not, or could not, say. In other words, these talks are indeed examples of an entangled philosophy, between philosophy and politics, between philosophy and literature, and between philosophy and love. Precisely speaking, it is about the impossibility of having a philosophy without politics, a philosophy without literature, or a philosophy without love.

Another thing we should learn from these lectures, is about their failure to embrace "the other". We can notice the absence of influential (but perhaps politically incorrect) philosophers such as Tanabe Hajime and Miki Kiyoshi. Although there were speakers from Keijō and Taiwan Imperial Universities (京城帝国大学 / 台湾帝国大学), there was an absence of non-Japanese speakers. Also, there was an absence of women philosophers.

[27] TSZ 5: 8.
[28] TSZ 5: 260.

Top scholars in Japan deliver lectures to the emperor, as well as other family members of the imperial family, but there has been no philosophical lecture for many years. It is time for a philosopher to deliver a talk including one of the above-mentioned neglected topics. We are also expecting a better gender balance in future lectures. Eventually, I hope philosophers can deliver lectures not only to the privileged, but also to everyone (including children). In this sense, I propose we should do not only *Tetsugaku* (the study of wisdom without love), but also *Kitetsu* (希哲 the love of wisdom),[29] for the very meaning of *kitetsu* is to provide a platform for more philosophical dialogue in our society.[30]

References

Fukuzawa, Yukichi (2009) A Bill to Ban the Publication of *Meiroku Magazine*, in *Meiroku Zasshi*, (Tokyo: Iwanami Bunko, 3), 439–445.

Heisig, James (et. al), (eds.) (2011) *Japanese Philosophy: A Sourcebook*. (Honolulu: University of Hawai'i Press).

Nishida, Kitarō (1979) *Nishida Kitarō Zenshū*. (Tokyo: Iwanami shoten). (Abbreviated as NKZ)

Ohashi, Ryōsuke (2001) *Kyōto Gakuha to Nihon Kaigun. [Kyōto School and Navy]*. (Kyoto: PHP Shinsho).

Sakaguchi, Ango (2008) *Sakaguchi Ango, Chikuma Nihon Bungaku 9*. (Tokyo: Chikuma Shobō).

Takahashi, Satomi (1973) *Takahashi Satomi Zenshū*. (Tokyo: Fukumura Shuppan). (Abbreviated as TSZ)

Scheler, Max (1961) *Ressentiment*. (New York: The Free Press).

Watsuji, Tetsurō (1991) *Watsuji Tetsurō Zenshū*. (Tokyo: Iwanami). (Abbreviated as WTZ)

Yusa, Michiko (2002) *Zen and Philosophy: An Intellectual Biography of Nishida Kitarō*. (Honolulu: University of Hawai'i Press).

Yusa, Michiko (2017) *The Bloomsbury Research Handbook of Contemporary Japanese Philosophy*. (New York: Bloomsbury Academic).

[29] I am the editor of a philosophical magazine called *Kitetsu*. Details can be found here: https://www.facebook.com/希哲雑誌-1396694247267021/

[30] Similar efforts have been done at Center for Philosophy University of Tokyo (UTCP): https://utcp.c.u-tokyo.ac.jp/projects/uehiro/2012/l3/index_en.php

Tetsugaku
International Journal of
the Philosophical Association of Japan
Volume 4, 2020

Special Theme
Analytic Asian Philosophy

On the Ontological Status of the Past: Ōmori Shōzō and Paul Ricœur

YAMANO Hiroki
The University of Tokyo

Abstract: *The aim of this paper is to examine the theory of the past by Ōmori Shōzō (大森荘蔵, 1921‐1997) through examining the meaning of "all is vanity (色即是空)" in Time and Being (1994) in order to consider the relationship between the past and historiography (the narrative of the past). In particular, this paper attempts to consistently interpret the theory of the past in Ōmori's philosophy and his pragmatic realism(実用的実在論) by reinterpreting the argument about "all is vanity" which Ōmori advocated in his later years.*

This paper is structured as follows. The first section of this paper examines the feature of Ōmori's philosophy in the later period. In this section, the characteristics of monism of "Tachiaraware (立ち現われ)" in Ōmori's philosophy are clarified. The second section presents the fundamental problem in the monism of Tachiaraware. In this section, it becomes clear that he makes a seemingly incongruous claim that the past is described negatively in terms of "all is vanity", while describing the present positively in terms of pragmatic realism. The third section reinterprets the argument about "all is vanity" in Ōmori's philosophy by confirming that he regards "kūbaku (空漠)" and infinite as the matrix of meaningful world. In this section, it is shown that he developed the argument about "all is vanity" as an argument compatible with his pragmatic realism. The fourth section highlights the features of the theory of the past in Ōmori's philosophy through examining the argument of the past in Paul Ricœur's work, Time and Narrative (1983–85) and Memory, History, Forgetting (2000). In the last section, the ontological status of the past in Ōmori's philosophy is clarified by comparing Ricœur's theory of the past.

Introduction

Ōmori Shōzō is a representative philosopher of postwar Japan. He first studied phenomenology at the University of Tokyo. Later, however, he studied analytic philosophy in the United States and actively imported the analytic philosophy into Japan. He taught philosophy to many of his students at the University of Tokyo, who

would lead the later philosophical society of Japan. His disciples include Iida Takashi (飯田隆, 1948-), Noya Shigeki (野矢茂樹, 1954-), Noe Keiichi (野家啓一, 1949-) and Nakajima Yoshimichi (中島義道, 1946-), who are still active on the front lines. In other countries, for example, Pierre Bonneels[1] and Michel Dalissier[2] have published some papers on Ōmori's philosophy and analyzed its features from their point of view. Thanks to their work, Ōmori's philosophy can be discussed in English and French.

In this paper, I would like to discuss a question concerning the concept of "the past (過去)" in Ōmori's philosophy, which has not been sufficiently argued in the past studies. Understanding Ōmori's concept of "the past" is an issue that cannot be avoided in order to understand Ōmori's philosophy because, in the monism of "Tachiaraware", which is a characteristic of Ōmori's philosophy, a paradox arises that "the past" appears *now* to us in the mode of "recall (想起)". Furthermore, to the nature of "the past", Ōmori gave a complicated characteristic of "all is vanity". What kind of philosophical stance did he try to establish by bringing up his worldview of the past with "all is vanity"? A few studies of Ōmori's philosophy have mentioned this point. It is surprising that even in "round-table discussion" where four Ōmori's disciples gathered, "all is vanity" was never argued.[3] Therefore, this paper aims to clarify the ontological status of "the past" in the philosophy of Ōmori Shōzō by examining the meaning of "all is vanity" in *Time and Being*. This paper also attempts to clarify the characteristics of Ōmori's theory of the past by comparing them with those of Paul Ricœur (1913–2005).

1. The Feature of Ōmori's Philosophy in the Latter Period

We begin our discussion by pointing out two major features of later Ōmori's philosophy (1. Overcoming mind/matter dualism and 2. Proposal of the monism of "Tachiaraware"). Let us examine these discussions.

---

[1] See Pierre Bonneels, L'empirisme tremblant du langage chez Ōmori Shōzō, in *European Journal of Japanese Philosophy*, no. 3, 2018, pp. 193‑214.

[2] See Michel Dalissier, Le bon sens est-il la chose du monde la mieux partagée? Sens commun et vie ordinaire chez Ōmori Shōzō, in *European Journal of Japanese Philosophy*, no. 3, 2018, pp. 215‑243.

[3] The round-table discussion can be read at the end of *Ōmori Shōzō Selection* (Tokyo: Heibonsha, 2011). There are four members, Iida, Tanji, Noe, and Noya.

We encounter various things in our daily lives. For example, you can take a walk and look at a house or trees in the park. But are the houses and trees thus "the object" itself? Immanuel Kant (1724–1804) would argue that the object is a "phenomenon", not the "thing-in-itself". In other words, it is an epistemological argument that the things we recognize are only phenomena (or representations) in consciousness, and they are not the objects as they are. If we adopt this worldview, we will first recognize the representation and use it as a medium for imagining what might be beyond it (thing-in-itself). Ōmori puts up opposition to this dualistic theory of representation and object by the monism of "Tachiaraware".[4] How, then, did Ōmori shift his argument from the dualism of representation and object to a monistic understanding of the world? Let us examine that point.

First, Ōmori abolished the distinction between genuine objects and representations as the copies of genuine objects. For Ōmori, all things which we can perceive by our senses are equal and they are risings (Tachiaraware) of the object. For example, gorgeous buildings, beautiful trees, and insects flocking to flowers are all examples of rising. These things possess a solid reality for human senses and life (TB, 167). In other words, we live in a world that should be called realism of rising. Rising, of course, has several aspects. For example, when you see the Kamo River (賀茂川) flowing through Kyoto, the Kamo River appears in the form of "perception (知覚)", and when you recall the sight of the Kamo River, the Kamo River appears in the form of "recall (想起)". Ōmori's philosophy is unique in that it explored various aspects of rising and developed a monistic view of the world.[5]

Here, we summarize the characteristics of the monism of rising. The important point of the monism of rising is to deny a dualism of "object" and

[4] According to Pierre Bonneels, the concept of "Tachiaraware (立ち現われ)" contains variable meanings such as "to appear" or "rising". Dalissier translates this concept into *l'apparaître* (Michel Dalissier, Le bon sens est-il la chose du monde la mieux partagée? Sens commun et vie ordinaire chez Ōmori Shōzō, in *European Journal of Japanese Philosophy*, no. 3, 2018, p. 224). This concept means the dynamism of appearance of being ("*se dresse dans l'apparence* (立ち現われる)") (p. 225). In this paper, I adopt the term "rising" as the translation of "Tachiaraware".

[5] Many people have criticized "the monism of Tachiaraware" in Ōmori's philosophy. For example, Nakamura Hideyoshi criticized Ōmori's philosophy as follows. "Ōmori's philosophy is George Berkeley's philosophy without God. And it does not seem that the fundamental issue is different from Berkeley's philosophy". (Noe Keiichi, *A Labyrinth of Philosophy: Critique and Response to Ōmori Philosophy*, Tokyo: Sangyō Tosho, 1984, p. 6.) However, this affinity with Berkeley in the Ōmori's philosophy is, rather, designed by himself. In fact, referring to Berkeley's expression "To be is to be perceived", Ōmori says, "To be in the past is to be recalled." (TE, 129).

"representation" in the Kantian sense and to reduce everything in this world to the classification of a mode of rising. Through thinking that an object directly appears before us without the intermediation of representation, Ōmori establishes his monistic philosophy. In the monism of rising, it is unreasonable to delineate a clear border between an external object and an internal representation.

However, in this monism, the following problems would be proposed. Is it possible to claim that all things which appear to us are equal at the ontological level? For example, in Ōmori's discussion, the mode of rising includes not only aspects of perception but also of imagination and "fancy (虚想)".[6] Does that mean that all the dreams and visions which appear to us in the mode of imagination and fancy also exist equally? Ōmori's answer is "Yes". But, in the monism of rising, how should we consider the problem of being and authenticity of rising? Will this theory abolish any distinction between dream and reality? Such objections could be posed. Let us look at Ōmori's answer to this point.

According to Ōmori, authenticity is determined practically in everyday and social life. Then, what guarantees that rising is authentic? It depends on the needs of human life.[7] Here, he emphasizes the practical superiority, not the epistemological superiority. In other words, for Ōmori, the authenticity of rising is conventionally determined in terms of whether it can support our actual lives. "We do not believe things from the viewpoint of authenticity. A thing we believe at the risk of our life is an authentic thing".[8] Thus, for Ōmori, the network of rising results from the core of our life.

From this viewpoint, Ōmori's philosophy was given the character of pragmatism. In fact, Ōmori referred to his position as "a coherence theory heavily contaminated by pragmatism"[9] in his theory of *Kotodama* (言霊). Ōmori named his position "pragmatic realism (実用的実在論)" in his later main book *Time and Being* (TB, 189–194). Therefore, it can be concluded that such pragmatism was one of the decisive characters of the monism of rising.

However, the monism of rising has two fundamental problems because of its monistic nature. That is the problem of "the past" and "the other".

---

[6] About the concept of "fancy", see Ōmori Shōzō, Beyond Hume's "Fancy, in *Revue Internationale De Philosophie 28*, no. 107/108 (1/2), 1974, pp. 99–115.

[7] Ōmori Shōzō, the Theory of *Kotodama* (ことだま論), in Iida Takashi, Tanji Nobuharu, Noe Keiichi and Noya Shigeki (Eds), *Ōmori Shōzō Selection*, Tokyo: Heibonsha, 2011, p. 289.

[8] Ibid., p. 290. (My translation, the same hereinafter)

[9] Ibid., p. 294.

2. The Fundamental Problem in the Monism of Rising

In the monism of rising, rising of all phenomena is described in terms of first-person ego experiences; however, as Noya and Nakajima have already stated, under these principles, we cannot explain "the past" and "the other" very well.[10] When and where does the "the past" appear to us? And, how can we compose "the other" in the experience of rising? Both are fundamental issues, but this paper addresses the former in accordance with its purpose.

We begin our discussion by examining the contradiction that arise when discussing the dimension of "the past" in the monism of rising. For Ōmori, the mode of rising that creates "the present" is perception. In the mode of perception, the present Kamo River in Kyoto and the present Nihonbashi (日本橋) in Tokyo appear to us. In contrast, the mode of rising that creates "the past" is recall. "The past" appears to us in the mode of recall. For example, through the experience of recall, the Kamo River, which was seen three days ago, appears to us. But the experience of recall is conducted now. Then, in the monism of rising, the Kamo River that was seen three days ago *now* appears to us. In other words, approving Ōmori's argument means that the "the past" *now* appears and exists; however, "the past" is a concept that means that an event no longer exists (that is, the pastness of the past). The explanation that the past now appears to us, therefore, seems to contain a serious contradiction for us.

Moreover, what is the object that we recall? Usually, it would be reminiscent of objects or events that have passed away. Then, the recalled past is drawn from the real past (実際の過去); however, this explanation of the past is dualistic and inconsistent with the monism of rising. In this way, the problem of the past becomes a big enigma in the monism of rising. How does Ōmori answer these questions?

To answer the problem of the past, Ōmori does not modify the monism of rising but rather strengthen its system. Surprisingly, he positively agrees that the past will appear to us *now and here* through the mode of rising.[11] This judgment is inevitable if Ōmori's philosophy has a monistic nature. For him, the only difference

[10] Noya Shigeki, *Ōmori Shōzō: An Example of Philosophy*, Tokyo: Kōdansha, 2007, p. 178; Nakajima Yoshimichi, *A living past: the theory of time of Ōmori Shōzō and its critical reading*, Tokyo: Kawade shobō, 2014, p. 174.

[11] Ōmori Shōzō, the Theory of Kotodama, in Iida Takashi, Tanji Nobuharu, Noe Keiichi and Noya Shigeki (Eds), *Ōmori Shōzō Selection*, Tokyo: Heibonsha, 2011, p. 248.

exists in the mode of rising. Depending on the type of rising, such as perception or recall, a phenomenon is constructed as the present or the past. Further, he insists that such recalls are verbal. According to Ōmori, we produce the past linguistically through recall. Of course, there is no dualistic distinction here between the real past and the linguistically produced past. In the monism of rising, the linguistically produced past is nothing but the past. That is why he reformulated the definition of "the past" as follows.

> I just mean that the past is existent, not independently from a conscience of recall, but in a linguistic meaning of a recalled proposition. (TE, 114-115)

This argument in *Time and Ego* (1992) was further advanced in *Time and Being* two years later.

> Then, when I recall something; besides when I recall something with a conviction of the reality, is it a kind of delusion that corresponds to no reality? That's right. We experience it as a recall what we call a dream in that case. In other words, as long as we do not obtain the meaning of the real past, all recalls are dreams. . . because there is no reality that corresponds to them. (TB, 200)

In this quotation, he calls the object of recall "delusion" or the kind of "dream". The reason is that "there is no reality that corresponds to them". For him, "A recall is not a perceptual reconstruction or reproduction" (TE, 45). And at this very point, he gives the past the essence of "all is vanity".[12] Moreover, Ōmori went as far as to say;

> Connecting the past to "all is vanity" instills "all is vanity" in pragmatic realism. (TB, 202)

[12] Ōmori himself did not give a clear explanation about the term "all is vanity" (TB, 12); however, it should be pointed out that Ōmori's philosophical and daily intuition might have been strengthened by Buddhist thought that completely denies the essence of things. At the Buddhist viewpoint, see Izutsu Toshihiko, *Consciousness and Essence: in search of the spiritual East*, Tokyo: Iwanami Shoten, 1991, pp. 19‐24.

> Pragmatic realism, including the realism of the past, is, of course, consistently the realism of "all is vanity". (色即是空の実在論) (TB, 204)

By introducing the element of "all is vanity" into his monism (TB, 185, 202–204), he reduced even his pragmatic realism to realism of "all is vanity"; however, should Ōmori's philosophy be understood as the philosophy which finally reached realism of "all is vanity" through the development of the theory of the past? We must consider that point.

Again, let us reconfirm the characteristics of Ōmori's philosophy. It does not admit the rising in recall or perception to be understood as a transcendental object independent of consciousness (TE, 51–52, 104, 108–111). Recall and perception were just the kinds of rising. Thus, based on this view, we can find at least the following commonalities in the rising of perception and recall.

1. In the mode of perception, the present appears internally in our consciousness; however, it is often replaced by the transcendental existence of the present.

2. In the mode of recall, the past appears internally in our consciousness; however, it is often replaced by the transcendental existence of the past.

A clear structural similarity can be found between the two propositions; however, we have just confirmed that Ōmori claims that the recalled past is like a dream. The reason for this is that there is no being corresponding to the past. If this is the case, the perceived present should be reduced to the sort of dream because there is no corresponding to the present (cf. TE, 110). At the same time, however, this position clearly contradicts the pragmatic realism that characterizes the later Ōmori's philosophy. This is because judging that even a perceived object is a "dream" is clearly against our daily beliefs and undermines the foundations of our daily lives (cf. TB, 189–194). Ōmori's assertion about the past seems at first sight to betray his own position that he naively accepted "there are mountains, rivers and plants".

This consideration indicates that there are difficulties in comprehensively interpreting the monism of rising. In other words, Ōmori makes a seemingly incongruous claim that while describing the present positively in terms of pragmatic realism (TB, 133–134, 166–169), the past is described negatively in terms of "all is vanity" (TE, 131–132; TB, 200–202).[13] However, for Ōmori, the present and the

[13] Despite that Sato regards the argument about "all is vanity" as the claim that abolishes all

past must have had the same ontological status in the respective modes of rising. How can we interpret this contradiction in the monism of rising?

3. "All is Vanity" as the Matrix of Meaningful World

It is true that Ōmori stated in his article "Realism of All is Vanity (色即是空の実在論)" (October 1993) that "the past" means "a sort of delusion that do not correspond to any reality" and that "all recalls are dreams" (TB, 200); however, in fact, at the stage of "The Past and Dream as Linguistic Product (言語的制作としての過去と夢)" (August 1991), the production of the past should have been more carefully positioned. On this point, let us confirm Ōmori's next words.

> Then, is the past literally just a dream like all recalls are the recalls of dreams? Of course not. Recalls are not founded but regulated and bound, because there is the past to believe in and dreams are sparsely embedded in it. (TE, 117)

It is true that Ōmori's explanations for recall and the past are not straightforward; however, in the theory of linguistic production, he redefined the past as follows. "The past" means "history shared by society", in other words, "a socially collaborated linguistic product" (TE, 119). In this sense, he understood "the past" as "narrative of the past".

Here, he pointed out "works by historians" and "disputes in court" (TE, 111) as examples of the narrative of the past. In that sense, the past is not created in a totally arbitrary way. Noya skillfully expressed the theory of the past in Ōmori's philosophy in the following way. "Social language practice gives measuring to the past. And the narrative of the past that are institutionally accepted as true are understood as the real world of the past".[14] So, what is the relationship between these social language practices and "all is vanity"? If Ōmori had understood "all is vanity" in a completely negative way, there would have been no discussion about the disciplined creation of the past. Perhaps we need to focus on the ambiguity in the

---

naïve realism as "delusion", he omits this interpretative difficulty. See Sato Masae, *Live in a Naive Way: The Philosophy of Ōmori Shōzō and the Path of Human Being*, Kanagawa: Seibunsha, 2009, pp.144‐146, 244.

[14] Noya Shigeki, *Ōmori Shōzō: An Example of Philosophy*, Tokyo: Kōdansha, 2007, p. 196.

term “all is vanity”. Therefore, we try to deal with the concept of “kūbaku (空漠)” as a clue to examine the ambiguity of “all is vanity”.

In “Semantics of Cubism (キュビズムの意味論)” (May 1992) Ōmori talks about the concept of “kūbaku”, which is an infinite space generated from the sound in hearing (TB, 120). The point of the argument here is that the auditory experience does not have a clear boundary between subject and object (TB, 118). First, in the case of visual and tactile experiences, there is a clear boundary in contact with the object (TB, 117–118). Ōmori asserts that this boundary creates a side of an object and a side of the subject. For example, when I touch a tree, there is a distinction between the tree being touched by my hand (a side of an object) and my hand touching the tree (a side of the subject). In contrast, auditory experiences do not have this boundary, making the distinction between self and object ambiguous (TB, 118). Pointing out the nature of this “frailty of sound”, he states that the sound is “transience (無常)” (ibid.). He insists that the meaning of “space” is generated from the experience of such sounds of “transience”. In other words, for Ōmori, “kūbaku” and “kū (空)” is not just nihility. Let us look at Ōmori’s words.

> This space, generated from soundscapes, is the foundation and the framework of our concept of “space”, say, kūbaku (infinite expansion). This kūbaku is not a vacuum, but a fertile matrix into which we draw infinitely variable figures. (TB, 120)

For him, kūbaku is the infinite basis of space for objects of all nature (森羅万象). According to Ōmori, in order to perceive such an infinite space, it is necessary to hear rather than see. The object in sight is always a three-dimensional finite object, but there is no boundary between any object in the sound itself. When we close our eyes and listen to the sound, we can feel the endless wave of sound. And “when the infinite space generated from the hearing is eventually overlapped with a visual or tactile scene, it becomes a space of things in which objects derived from visual or tactile sense are located” (TB, 119). Ōmori’s philosophical intuition is that there are two dimensions in the world. One is a world filled with three-dimensional objects, and we can perceive it through sight and touch. And the other is an empty space without boundaries of anything, and we can perceive it through hearing (TB, 118). In other words, we always live in this duality of the world. Therefore, the world we live in is both rich and empty. The two facts are not contradictory but compatible. To put it more simply, given the fabric of infinite space, we can weave finite

meaning onto it.[15] For Ōmori, the emptiness of being, in other words, the theory of the existence of "all is vanity" is not a negative situation, but rather a positive interpretation of the world.

We put meaning in a meaningless world. "All is vanity" is not the last word in Ōmori's philosophy, but rather the philosophical starting point where we can begin to explore the world of human meaning. Now, by discussing it so far, we have obtained a powerful clue to consider the relationship between pragmatic realism and realism of "all is vanity". First, pragmatic realism is a position to express aspects of the world of human meaning. According to Ōmori, human beings created practical meanings and values in the empty world to support their own lives. And the realism of "all is vanity" is the position to express the world before such human meaning was inserted. Here, let us recall the discussion in Ōmori's paper "Taste and Feeling (風情と感情)" (July 1990). There, he combined the impression of music with the infinite space (TB, 246). In other words, he argues that people are trapped by powerful emotions when they encounter the infinity of the world.[16] For him, infinity is nothing but a positive source of the world. In this sense, the realism of "all is vanity" in Ōmori's philosophy is compatible with pragmatic realism. Nor does it claim that "all is nothing". The realism of "all is vanity" is a theory that discusses a possibility of us creating a meaningful world in infinite space. In this sense, it is shown that, in the monism of rising, the realism of "all is vanity" is linked with pragmatic realism. Thus, the interpretation by Nakajima, who interprets the theory of the past in Ōmori's philosophy as a mere ideology, should be rejected.[17]

In this respect, we can comprehensively interpret three points of Ōmori's philosophy: poesis of the past, social language practice, and the realism of "all is vanity". For him, "all is vanity" is not a negative aspect of the absence of the world, but a fertile matrix to which we can project infinite meaning. And by being

[15] Of course, linguistic thinking activities will be required to create meaningful objects in infinite space. For example, Ōmori discusses this point in his article "The Past and Dream as Linguistic Product (言語的制作としての過去と夢)" (August 1991) and "The Meaning of Being (存在の意味)" (October 1992).

[16] At this point, see Sato Masae, *Live in a Naive Way: The Philosophy of Ōmori Shōzō and the Path of Human Being*, Kanagawa: Seibunsha, 2009, pp.126–127.

[17] Nakajima Yoshimichi, *A living past: the theory of time of Ōmori Shōzō and its critical reading*, Tokyo: Kawade shobō, 2014, pp. 206–207, 221. Similarly, Sato views Ōmori's argument of "all is vanity" as negative; however, as we have discussed in this paper, Ōmori's argument of "all is vanity" develops as an insight into the ontological basis on which humans create meaning. In other words, Ōmori' s argument does not recommend "withdrawal from social life" (p. 244). See Sato, pp. 244‐245.

supported by the infinite matrix and the meaning, we can do social language practice. In other words, we can create meaning for being in the world and recognize an object as an "object" (TB, 132–134). To create new meaning is to invent the possibilities of a new world. And the world we live in is a temporal horizon that includes the past and the future. As Ōmori says, the present we live in is the historical present, and it has a depth in time. "Being is already time, and time is already in being" (TB, 20). The time a man lives in is the historical time made by them in order to build a stable view of the world. In this very sense, pragmatic realism and the realism of "all is vanity" are combined. This is because "time is what we create individually for the needs of our lives" (TB, 31). For Ōmori, time is the totality of human history, and it is the stage of daily life where the past and the future can be included. And in this historical time, three aspects of time appear to us: the past, the present, and the future. Thus, according to Ōmori, "time does not flow".

In this sense, we can interpret realism of "all is vanity" as a position that forms a complementary relationship with pragmatic realism in the monism of rising.

4. Poiesis and Mimesis of the Past

In our previous discussions, we have examined the theory of the past in the monism of rising by reinterpreting Ōmori's discussion of "all is vanity". As a result, it became clear that Ōmori was developing the ontology of the past, which was based on the positive infinite. For him, the discussion of poiesis of the past is identical to the idea of human poetic creativity. Therefore, he regarded the act of describing the past as "poetry" (TE, 115).

Here, in order to clarify the characteristics of Ōmori's theory of the past, I try to compare Paul Ricœur's theory of the past with his position. There are three noteworthy similarities between Ōmori and Ricœur. First, both are influenced by Aristotle's concept of poiesis (cf. TE, 115; TN1, 66). Second, both are influenced by Husserl's phenomenological analysis concerning perception and time (cf. TB, 130; TN3, 23–44). The third point is that both sides have developed the argument of the narrative of the past (cf. TE, 53–56; TN1, 155–174). Despite these similarities, the theories of both are moving in opposite directions. To state it in advance, while Ōmori develops the monistic argument of poiesis (production) of the past, Ricœur develops the dualistic argument of mimesis (imitation) of the past. Let us start by examining Ricœur's theory of the past.

In his later year's work *Memory, History, and Forgetting* (2000), Ricœur begins his discussion by criticizing Plato's theory of memory. Plato argues about knowledge and truth without distinguishing memory from imagination in his *Theaetetus* and *Sophist*. But according to Ricœur, memory and imagination must be clearly distinguished, as Aristotle's analysis shows, by the presence of traces. Memory is clearly separated from the general function of the imagination by the feature of aiming at the anteriority of the "mark" (MHF, 12). Furthermore, he characterizes a recall of the past as a pair of "evocation" and "search" (MHF, 26). A recall is the experience of remembering the past. And this experience leads us to the quest to seek the absent past that has already passed. In fact, we ask what has happened in the past. It is an "effort to recall" (MHF, 28), in other words, "laborious recollection" (Ibid.).[18] Recollections of the past acquired in this way, of course, need to be preserved publicly in the form of narratives, not merely images. The vague knowledge sketched by the traces must now be described as the narrative of the past. So how does the narrative of the past (historiography) relate to the past?

To examine this point, we look at Ricœur's *Time and Narrative* (1983–85). There, Ricœur discusses the dialectic of historiography. That is, dialectic between *the same* and *the other* (TN3, 144–156). The former is the position to regard historiography and the past as the same ontologically. It is understood that historian's thought is psychologically identical to the person's thought of the past, and this leads to the oblivion of the otherness in history (TN3, 147). In contrast, the latter views historiography and the past as being different ontologically. It is argued that there will remain a critical gap between historiography and the past that cannot be bridged, thus unilaterally emphasizing the otherness in history. To overcome this dichotomy, Ricœur proposes the third path, *the analogous* (TN3, 151–156), which is created by combining the same and the other positions.[19] In other words, historians describe the past events *as* they were.

Ricœur's theory of the past is clearly dualistic. Historians try to recall the past, which has already been lost, in the indirect way of historiography. Moreover, Ricœur's criticism of Plato is also true of Ōmori. In the monism of rising, depending on the type of rising, everything appears to us. In fact, he makes little distinction between what is absent and what is past.

---

[18] In this point, see Jean Greisch, *Paul Ricœur. L'itinérance du sens*, Grenoble: éd. J. Million, coll « Krisis », 2001, pp. 288–292.

[19] See Johann Michel, *Paul Ricœur. Une philosophie de l'agir humain* (Passages), Paris: Cerf, 2006, pp. 192–199.

> If imagination is, in a broad sense, to appear things and events which are not presently perceived, to think perceptually(知覚的に思う) is nothing but imagination.[20]

As is clear from this quote, the expression "things and events which are not presently perceived" includes not only the Eiffel Tower a year ago (the object of memory) but also Pegasus (the object of imagination) that has not the anteriority of the past. These characteristics clearly imply the fact that Ōmori holds the analogous position as Plato, who does not explicitly separate imagination from memory. What is important to Ōmori is not the dualistic position that the narrative of the past imitates the past (mimesis) but the monistic position that the narrative of the past produces the past (poiesis). Here we can see the ontological difference between the two theories of the past.

Furthermore, from another perspective, we can highlight the difference between the two theories of the past. It is a difference in ethical dimensions. On the one hand, for Ricœur, historiography is the act of recovering the figure of the dead who have become victims of history (cf. TN3, 100, 118). It is the practice of ethical responsibility not to repeat similar events, while opposing the oblivion of past victims and fearful events. In this sense, Ricœur's theory of the past presents the ethics of a community that aims at the future and the past (TN3, 216, 227). On the other hand, for Ōmori, historiography is the process of constant creation of the past that enriches the meaning of the world. It is true that these tasks are constrained by the coherence of material evidence and testimony from others (TB, 201); however, the important point here is whether the past connects with the present or not. In other words, it is the connection between the past and the present that is the criterion for the selection of the past. Therefore, Ōmori's theory of the past shows the ethics of a community emphasizes the present (cf. TE, 48–49). The question is not which of the two is correct. What is important here is that Ōmori's view of the world gives us a possible model for thinking about the enigma of the past.

Conclusion

[20] Ōmori Shōzō, For the approval of fancy (虚想の公認を求めて), in Iida Takashi, Tanji Nobuharu, Noe Keiichi and Noya Shigeki (Eds), *Ōmori Shōzō Selection*, Tokyo: Heibonsha, 2011, p. 334.

We have discussed the characteristics of Ōmori's theory of the past by reinterpreting his argument about "all is vanity". For Ōmori, the ontological status of the past is the result of linguistic production of the world that makes everyday life possible, and it is always created and changed by human poetic imagination. This feature of Ōmori's theory of the past is a necessary result of his attempt to consistently develop the ontological argument of poiesis of the past within the framework of the monism of rising. This discussion became possible through a consistent interpretation of his theory of "all is vanity", which had not been scrutinized before. We also compared Ōmori's theory of the past with Ricœur. As a result, this paper approaches not only Ōmori's theory of the past but also the general problem of the ontological status of the past. In this sense, this paper would have contributed not just to the research for Ōmori's philosophy but to the problem of the ontological status of the past.

References

Bonneels, Pierre (2018) L'empirisme tremblant du langage chez Ōmori Shōzō. *European Journal of Japanese Philosophy*, 3, 193–214.

Dalissier, Michel (2018) Le bon sens est-il la chose du monde la mieux partagée? Sens commun et vie ordinaire chez Ōmori Shōzō. *European Journal of Japanese Philosophy*, no. 3, pp. 215–243.

Greisch, Jean (2001) *Paul Ricœur. L'itinérance du sens*, Grenoble: éd. J. Million, coll « Krisis ».

Iida Takashi, Tanji Nobuharu, Noe Keiichi and Noya Shigeki (Eds.) (2011) *Ōmori Shōzō Selection* 『大森荘蔵セレクション』. (Tokyo: Heibonsha).

Izutsu, Toshihiko (1991) *Consciousness and Essence: in search of the spiritual East* 『意識と本質：精神的東洋を索めて』. (Tokyo: Iwanami Shoten).

Michel, Johann (2006) *Paul Ricœur. Une philosophie de l'agir humain* (Passages). (Paris: Cerf).

Nakajima, Yoshimichi (2014) *A living past: the theory of time of Ōmori Shōzō and its critical reading* 『生き生きとした過去：大森荘蔵の時間論、その批判的読解』. (Tokyo: Kawade shobō).

Noe, Keiichi (1984) A *Labyrinth of Philosophy: Critique and Response to Ōmori Philosophy*『哲学の迷宮：大森哲学批判と応答』. (Tokyo: Sangyō Tosho).

Noya, Shigeki (2007) *Ōmori Shōzō: An Example of Philosophy* 『大森荘蔵：哲学の見本』. (Tokyo : Kōdansha).

Ōmori, Shōzō (1974) Beyond Hume's "Fancy. *Revue Internationale De Philosophie 28*, 107/108 (1/2), 99–115.

Ōmori, Shōzō (1994) *A construction of knowledge and its spell* 『知の構築とその呪縛』. (Tokyo: Chikuma shobō).

Ōmori, Shōzō (1992) *Time and Ego* 『時間と自我』. (Tokyo: Sēdosha). [TE]

Ōmori, Shōzō (1994) *Time and Being* 『時間と存在』. (Tokyo: Sēdosha). [TB]

Ricœur, Paul (1983) *Temps et Récit, t.1: L'intrigue et le recit historique*. (Paris: Seuil).

Ricœur, Paul (1984) *Time and Narrative Vol 1*, Trans. Kathleen McLaughlin and David Pellauer. (Chicago: The University of Chicago Press). [TN1]

Ricœur, Paul (1985) *Temps et Récit, t.2: La configuration du temps dans le récit de fiction*. (Paris: Seuil).

Ricœur, Paul (1985) *Time and Narrative Vol. 2*, Trans. Kathleen McLaughlin and David Pellauer. (Chicago: The University of Chicago Press).

Ricœur, Paul (1985) *Temps et Récit, t.3: Le temps raconté*. (Paris: Seuil).

Ricœur, Paul (1988) *Time and Narrative Vol. 3*, Trans. Kathleen McLaughlin and David Pellauer. (Chicago: The University of Chicago Press). [TN3]

Ricœur, Paul (2000) *La Mémoire, L'histoire, L'oubli*. (Paris: Seuil).

Ricœur, Paul (2004) *Memory, History, Forgetting*, Trans. Kathleen Blamey and David Pellauer. (Chicago: The University of Chicago Press). [MHF]

Satō, Masae (2009) *Live in a Naive Way: The Philosophy of Ōmori Shōzō and the Path of Human Being* 『素朴に生きる：大森荘蔵の哲学と人類の道』. (Kanagawa: Seibunsha).

A Wittgensteinian Approach to Reconsidering Nishida's Basho of True Nothing

Zhang Ligeng
Chinese Academy of Social Sciences

Abstract: *In this paper, I aim to reconsider Nishida Kitarō's concept of the Basho of True Nothing from the viewpoint of Ludwig Wittgenstein's theory of language games. First, I illustrate the necessity of introducing Wittgenstein to approach Nishida's theory, as well as the similarity between the two philosophers. On this basis, I argue that there is a crucial dilemma in Nishida's use of the Basho of True Nothing, which inevitably generates paradoxical formulations in his writings. The thrust of my argument is twofold: on the one hand, I advocate that the reason for such a dilemma lies in Nishida's potential confusion of the role of some essential words when he tries to describe something transcending language; on the other hand, in the contrast between Nishida's Basho of True Nothing and Wittgenstein's Form of Life, I argue that the special implication of the Basho of True Nothing reveals a fundamental discrepancy between the culture of East Asia and the so-called West. That said, this paper is a Wittgensteinian analysis rather than a comparative study, so Wittgenstein's methods and conceptions are used as a "microscope" with which to scrutinize Nishida's ideas. I make use of both Nishida's and Wittgenstein's ideas as building materials rather than simply seeing them as a maze in need of exploring. In summary, this paper is an introduction to a conceivable analytical reconstruction of Nishida's theory. Hopefully, this trial, the conclusion of which is still open, will contribute to the improvement of analytical philosophy in East Asia.*

1. Introduction

Nishida Kitarō (西田幾多郎) is said to be the most representative Japanese philosopher. In his philosophy, the "Basho of True Nothing" (真の無の場所) is one of the most significant concepts, and it is not only valuable in the history of thought but also of unique significance in contemporary philosophy. However, there are still some unsolved problems in clarifying the exact meaning of this terminology.

Generally, there seems to be three unavoidable and interrelated difficulties in the studies on Nishida: first, Nishida's writing is extremely obscure, and sometimes

it is even too hard to grasp the literal meanings of his expressions;[1] second, the source of his thoughts is so complicated that readers have to refer to many other philosophers, such as Neo-Kantianist Emil Lask and Heinrich Rickert, to appreciate his ideas; and third, his arguments are rarely expressed straightforwardly, making it even more difficult to evaluate the plausibility of his viewpoints. There is no doubt that Nishida provides numerous insights, but these difficulties thwart further exploration of his thoughts.

What are the roots of these difficulties? Apparently, Nishida's own obscure style of writing is responsible for them. However, as interpreters, we have responsibilities as well. There are already many interpretations of the Basho of True Nothing, but some blind spots still exist.

Among other things, this concept has seldom been considered in an analytical way, as it is an alien concept in the world of analytical philosophy. Nishida was greatly influenced by continental philosophy, so studies on him are naturally relevant to Kant, Hegel and Neo-Kantianists. Nevertheless, it should be acknowledged that continental philosophy is only "half" of Western philosophy, and sometimes we have to solve philosophical problems from the other "half" (i.e., analytic philosophy). Analytic philosophy emphasizes argumentative clarity and precision, often making use of conceptual or linguistic analysis. Such characteristics are conducive to reading Nishida to make his arguments clearer and easier to understand.

More specifically, in this paper I aim to reconsider the concept of the Basho of True Nothing from the perspective of Wittgenstein's theory of language games, identifying the reasons for Nishida's obscure and paradoxical articulations. Before moving on to further discussion, it is necessary to explain my approach. This approach may play an innovative role in clarifying Nishida's ideas.

2. Definition of a Wittgensteinian Approach

I am willing to call my approach Wittgensteinian, alluding to a method of scrutinizing the potential problems in other philosophers' thoughts from the perspective of language games. In other words, Wittgenstein's principles about how

[1] Many have noted that his articulations often seem quite paradoxical, with an idea often being accepted and denied at the same time. Not surprisingly, Botz-Bornstein criticized Nishida's Basho as being "closed and open at the same time" (Botz-Bornstein 2003, 53).

a word makes sense are used as a "hinge"[2] or criterion. It should be noted that, on the one hand, such an approach is not equivalent to a comparative study between two philosophers;[3] on the other hand, the significant similarities between Nishida and Wittgenstein will facilitate this discussion.

There is a commonality in the fundamental positions of the two philosophers. In brief, first, the starting points of both Nishida and Wittgenstein are in opposition to psychologism and in favour of a logical position. Second, both of their theories are established on a philosophical analysis of language, such as the structure of predicate or the role of linguistic expressions in daily life. Third, both philosophers try to explore the role of the elements beyond language, which can be seen as the prerequisite or background of our use of language.

The first aspect is significant, and in it lies their basic shared tenet.[4] A common enemy to both Nishida and Wittgenstein is psychologism, as neither philosopher is willing to base their theories on something psychological. For Nishida, in the period of Basho, even his earlier theory about "pure experience" was too psychological; for Wittgenstein, whether in the period of *Tractatus* or *Philosophical Investigations*, he never regarded psychology as a plausible starting point. Therefore, both of them attempted to start from logic instead of psychology, aiming to overcome stereotypes, such as the subject-object dichotomy in philosophy.

Even with their common ground, it is still not easy to find the point of penetration to read Nishida through Wittgenstein. As I see it, the point of penetration lies in the attainment of a transparent understanding of the Basho of True Nothing. This is not only because of the essential role of this concept in Nishida's theory but also because of its relevance to the theory of meaning, which is one of the themes of Wittgenstein's theory of language games.

Nishida does not provide a clear theory about the meaning of his theory of Basho, but it is obviously improper to construe the Basho of True Nothing as a lexical term denoting something that exists in the visible world.[5] It follows that an

---

[2] This terminology is used by Wittgenstein in *On Certainty*, see OC 341–343. In Japanese, it is translated as "蝶番"; see 冲永宜司 2009, 48.

[3] Of course, some illuminating comparative studies between Nishida and Western philosophers such as Wittgenstein have been made. See Botz-Bornstein 2003, 冲永宜司 2009 and Krummel 2017.

[4] The latter two aspects will be discussed in sections 3 and 4.

[5] The reason will be illustrated in the next section. It is, nonetheless, unfair to say that Nishida has no theory of meaning, e.g. see 朝倉友海 2018, 177–78. I only mean that he does not provide a distinct formulation of such a theory in the "Basho" monograph.

interpreter could not explain this term with a simple ostensive definition. Rather, the term is used by Nishida in a variety of contexts, which allows us to become acquainted with its meaning by learning how Nishida uses it. At this point, we encounter Wittgenstein's slogan "meaning is use". Wittgenstein advocates in his later philosophy (especially in *Philosophical Investigations*) that the meaning of a word consists in its uses.[6] In his view, language games should be construed as concrete examples of linguistic practice into which words are woven. Even though the Basho of True Nothing is not a castle in the air, the absence of its reference in the visible world requires us to consider its meaning with respect to its practical uses, which conforms to Wittgenstein's conception of language game.

In the following discussion, it should be borne in mind that Basho is by no means simply analogized to a language game. What I aim to deliver is a clarification of the Basho of True Nothing in terms of the language games in which it is involved. In other words, its meaning has to be interpreted within the network of concepts that contribute to its uses.

3. Nishida's Use of "Basho" and the "Basho of True Nothing"

First, we have to interpret the literal meaning of the word "Basho". Nishida explains this term in different ways, but there is something common across his statements. It is said that the original motivation for Nishida proposing this concept was as a response to the subject-object dualism. The introduction of Basho begins with a reflection on Aristotle's logic of the subject, initiating a reassessment of the structure of judgement. Nishida's approach is based on his unique understanding of predicate and judgement, focused on the predicate instead of the subject.

For Nishida, the predicate is the real foundation of knowledge claims or judgements. For example, in the proposition "red is a kind of colour", although the grammatical subject is "red", the real subject is "colour", because it is the universal "colour" mirroring itself as "red" (see NKZ3 428–429).[7] Similarly, when we say "this desk is made of oak", the true real subject is "reality" rather than "desk" (see NKZ3 431). The uniqueness of the theory of Basho stems from Nishida's reinterpretation of the role of subject and predicate. According to Nishida, in a

[6] This assertion will be illustrated in detail in sections 4 and 6, in which we will also be reminded that Wittgenstein's ideas are actually more complicated than this.

[7] When citing Nishida's own words, I use Krummel's translation in *Place and Dialectic: Two Essays by Nishida Kitarō*, but the page numbers still refer to the Japanese edition.

judgement, the predicate (which should be regarded as universal) subsumes the subject (which should be regarded as particular). Therefore, our knowledge claims are always expressed in such form (see NKZ3 390). In terms of this reinterpretation, the predicate is a place for us to make a judgement, and it is only in this place that a subject or object is allowed to emerge and play its grammatical role. Here, Basho appears to be something that can encompass almost everything. This is Nishida's response to the subject-object dualism.

Although the literal meaning of the word "Basho" is "place", it should not be simply translated as "place". This word refers to something epistemological, but Basho is much more than that. Furthermore, its meanings are varied,[8] and it does not denote any concrete "place" but rather alludes to the "placedness" or "implacement"[9] of our experience. Thus, it is better to describe Basho as something like a **mirror** that can reflect everything. Here is a concise summary of the comprehensive characteristics of Basho:

> (Basho) is the standpoint vis-à-vis reality, the most concrete entailing the non-distinction between experience and reality. . . At its most concrete level, presupposed by all other levels, basho envelops and encompasses all a priories, mental acts, categories, contexts, and perspectival horizons that constitute the world of objects. . . The physical field of forces, the field of consciousness, and the sociohistorical world (I and thou), then, all are understood in terms of basho. (Krummel 2015, 25)

A summary in Nishida's own words is as follows: "I want to conceive, at the root of all things, a seeing without a seer" (NKZ3 255).

As it has been said, Nishida's use of the predicate is quite different from our ordinary understanding. Nishida's predicate is inclusive and "means something more than the grammatical predicate or a conceptual universal, and he reminds us on occasion that both universals and particulars…are implaced in that final transcendent predicate-plane he equates with the Basho of true nothing" (Krummel 2012, 18–19). Not surprisingly, such a peculiar conception of the predicate may lead

[8] Nishida uses Basho to refer to all kinds of aspects, such as "place", "universe", "predicate", "nothing", and "self-determining act" (see Kummel 2012, 47). It should also be noted that Basho has two synonyms: one is "predicate-plane" (述語面), and the other is a term borrowed from Hegel, namely, "concrete universal" ("具体的一般者", see NKZ3 431).

[9] I also learned these two words from Krummel.

us to an abyss that devours our ordinary cognition or thought (see NKZ3 458). Standing face to face with this abyss, the natural law governing our cognitions will collapse, just as physical laws are invalidated in a black hole.

In fact, this peculiarity of Basho is reflected at the very beginning of the monograph "Basho":

> But, in order for objects to relate to one another, constituting a single system and maintaining themselves, we ought to consider not only what maintains that system but also what establishes the system within itself and wherein the system is implaced. That which is must be implaced in something. Otherwise, the distinction between is and is not cannot be made. . . there must be that which envelops the opposition between I and non-I within itself and makes the establishment of the so-called phenomena of consciousness possible within itself. (NKZ3 415)

Proceeding along Nishida's approach, we naturally reach the conclusion that there must be a predicate that cannot be a grammatical subject and thus inevitably leads to "Nothing" (無), even **"**absolute nothing**"** (絶対無, e.g., see NKZ3 432). According to Nishida, "The basho of true nothing must be that which transcends the opposition of being and nothing in every sense and enables them to be established within" (NKZ3 424). As a result, Basho, which can be seen as the concrete situation of our lived experience, has a hierarchy consisting of three levels or planes: "Basho of Being", "Basho of Oppositional Nothing" and "Basho of True Nothing".[10]

Nishida's argument leads us to "True Nothing", which entirely transcends language and can only be described in a paradoxical way, such as "seeing without a seer", "a circle without periphery" or "self-mirroring mirror". All of these articulations reveal the tension between "Being" (有) and "Nothing" (無) in the whole of his theory of Basho, which is more obviously presented in the concept of the Basho of True Nothing. Now, we arrive at the destination of Nishida's exploration: "That the universal predicate reaches its extremity means that the particular [grammatical] subject reaches its extremity and becomes itself" (NKZ3 477). We might be surprised to encounter such incomprehensible formulations, as all of the descriptions of the Basho of True Nothing seem totally paradoxical, the reasons for which must be determined.

[10] Please see the graphical representation in Krummel 2012, 27.

4. Nishida's Dilemma

Apparently, the paradoxical articulations are partly derived from the lack of exact definitions for the involved concepts. Nishida seldom provides such definitions, sometimes making use of concepts somewhat casually. For example, knowledge, volition and intuition are all concepts that contribute to the meaning of the Basho of True Nothing. Nishida defines them as:

> To go on subsuming the particular into the universal is knowledge, to subsume the universal into the particular is volition, and the unity of both directions is intuition. Although it would appear contrary to reason to say that the universal is subsumed into the particular, this sense must already be included when substance is conceived as that which becomes the [grammatical] subject but not the predicate. (NKZ3 453)

Unfortunately, such clear definitions rarely arise in his writings; worse still, his uses of these words in other paragraphs often do not completely conform to such definitions. In contrast, the second half of the quotation may represent his ideas more straightforwardly: he is fully aware that his statements are problematic ("contrary to reason"), but he does not seem willing to regard the problems as fatal.

However, this is only a small part of the reason for his paradoxical articulations. The deeper reasons remain to be discovered and might be related to his attitude towards contradictions. Needless to say, Nishida never seems to be worried about expressing his thoughts via apparently contradictory expressions, which are usually regarded as meaningless. For example, he repeatedly uses the mirroring as a metaphor containing contradictions: "If such reception or mirroring signifies in some sense an activity, this must be an activity without what is at work, a mirroring without *what* mirrors" (NKZ3 451). Another example is as follows: "I would instead like to start from the idea of self-awareness wherein the self mirrors itself within. I think that the fundamental meaning of cognition is that the self mirrors itself within itself" (NKZ 420).

Neither "mirroring without what mirrors" nor "self mirrors itself within itself" makes sense in ordinary language. From a logical point of view, contradictions are definitely meaningless. Nishida, however, advocates that such contradictions are actually the foundation or prerequisite for every meaningful expression. Such a

conception seems to be inherited from Hegel, who shares a parallel understanding of logic and contradiction. However, it is undeniable that even a seemingly paradoxical expression has to make sense, which means that we cannot seriously reconsider the conceivable meanings of such expressions in ordinary language.

Let us turn our attention to the "Basho of True Nothing", which is rife with paradoxical features. It is reasonable to construe such features as consisting of the following two aspects: first, Nishida makes use of this terminology in a quite different way than most Western philosophers; second, and more importantly, the role of the terminology in the framework of our language is very special.

Regarding the first aspect, True Nothing (or Absolute Nothing) does not simply mean "nothing" or "there is nothing". Actually, it is something that transcends both being and nothing: this is an entirely different way of thinking that stands in contrast to the thinking of most Western philosophers. Philosophers are inclined to express the transcendent in terms of "being", lacking a conception of "nothing" beyond being and not being. After all, "nothing" itself is derived from "thing", just as "infinite" is constructed from "finite". Therefore, the use of the Basho of True Nothing is entirely distinct.

Regarding the second aspect, the Basho of True Nothing does not take anything as its prerequisite; on the contrary, it is the precondition of every judgement. Considering his reference to Aristotle at the very beginning of his argument, Nishida seems to take for granted that there is an internal or intrinsic relation between language and reality. Nevertheless, when talking about True Nothing, such a relation seems to be neglected. It is said that the Basho of True Nothing plays an indispensable role in our language, but at the same time, it is also prevented from the framework of language due to lacking any reference.

These two aspects together create an apparent dilemma in Nishida's underlying thoughts, which can be seen as one of the deeper reasons for his paradoxical articulations: on the one hand, he is exploring the structure and nature of language, which means that he has to take a position outside of or beyond the language itself; on the other hand, he has to use words to articulate his ideas, which means the expressions of these ideas have to take root inside language so that all of the words involved make sense. It is no wonder that Nishida's status is similar to that of a physicist conducting research on black holes: both of them have to deal with something that transcends the limitations of their tools, but, of course, they can never abandon their tools. To manage this dilemma, resorting to a language game is a viable choice, for it provides a tool that is more functional.

5. How to Solve the Dilemma

In short, language games can be construed as concrete examples of linguistic uses into which the language is woven, and words have their meanings only in such games. For Wittgenstein, there are all kinds of language games (see PI 23). The games exist at different levels, because some games make sense only if other games are already given or accepted. Generally, however, non-linguistic elements become more essential in more fundamental games. Here, non-linguistic elements mainly refer to agreements on how to use words. In other words, before beginning to play a language game, we have already made some decisions that do not belong to the game itself, and to express or communicate successfully, we have to master the related rules in advance. Consequently, the propositions that are used to describe the acceptance of a rule and those that are used to describe something inside language games belong to different categories, which also means that we cannot construe the first propositions in an ordinary way.

Language games can be either very simple or very complex. For example, what Wittgenstein describes in the very beginning of *Philosophical Investigations* are "five red apples" and other primary games. Compared with these games, Nishida creates an extremely special language game for Basho and the Basho of True Nothing, which is much more complicated. It can be inferred that some potential problems in such a game have led to the aforesaid dilemma, and we have to identify them.

In general, when introducing a concept, we can either define it directly or describe its uses in certain contexts and explain its relation to other concepts that have been assigned exact definitions in advance. Nonetheless, if a concept completely alludes to something transcending language, lacking reference in the whole of our experience, the descriptions of its uses will become extremely difficult.

In fact, some of the concepts involved in Basho have actual references in our experience, while some do not. To be specific, the introduction of Basho starts from a reflection on Aristotle's logic of the "subject", apparently referring to the linguistic field. Some concepts involved in the process of Nishida's argument are only partly non-linguistic, such as "self-awareness" and "intuition". The end of the argument leads to True Nothing, which entirely transcends language. For the first and second kinds of concepts, it is possible to clarify their uses by means of a philosophical

analysis of language or their immediate definition. For example, we can analyse the structure of predicates or the roles of them in our lives. For the third kind of concepts, as they are "outside" (instead of "inside") our language (or better said, our language games), it is not proper to describe their meanings in an ordinary way. It seems that Nishida, however, is not fully aware of the differences here, so he frequently describes the "outside" concepts in the same way as the "inside" concepts. This might be the fundamental reason for the aforementioned dilemma, leading to many puzzling expressions.

This reason reminds us of a remark from Wittgenstein that concerns the role of philosophy. According to Wittgenstein's conception, philosophical problems are not empirical and have to be "solved through an insight into the workings of our language and in such a way that these workings are recognized despite an urge to misunderstand them" (PI 109). From the viewpoint of language games, most philosophical problems are caused by a variety of misunderstandings of the role of our language. Therefore, instead of finding something new to solve such a problem, we have to see how language actually works. Thus, Wittgenstein summarizes, "Philosophy is a struggle against the bewitchment of our understanding by the resources of our language" (PI 109).

Nishida struggles against the bewitchments of language as well. Unfortunately, sometimes he seems not to fully realize the situation and thus portrays something that cannot be portrayed. In summary, the problem does not lie in the introduction of a term as a prerequisite of everything but in the improper properties being attributed to it. This also means that Nishida has not realized that it is impossible to describe the Basho of True Nothing in a similar way as ordinary terms. Once seeing this clearly, we can solve the dilemma by not seeing the confusing or puzzling expressions as describing something but rather only as introducing special rules governing our logic and judgement. These rules stand outside our language and are not a part of it. Thus, their descriptions cannot be understood in an ordinary way. When we try to illustrate these rules, we assume that we are standing in a "superior" position in which we actually cannot stand: this is a paradox in and of itself. It is no wonder that so many paradoxical articulations arise. In this way, we may not completely resolve Nishida's dilemma, but we may at least attain a more positive perspective for reconsidering his way of expressions.

In contrast, Wittgenstein deals with the preconditions of our use of language more ingeniously, appealing to the field of practice and deeds instead of becoming entangled in linguistic expressions. According to him, to use language is to follow

some rules, which naturally are directed at something beyond language, such as customs, usages or institutions (see PI 199). Thus, following a rule is something practical rather than purely intellectual, and "to think one is following a rule is not to follow a rule" (PI 202). We can think, express and communicate with each other in terms of language, but in order to do all of these, we have to accept something outside or beyond language in advance. That is, we "follow the rule blindly" (PI 219), and such following does not require any further interpretation. In this way, Wittgenstein eliminates the articulation of something paradoxical.

6. One of Nishida's Insights: the "Basho of True Nothing" and "Form of Life"

I have thoroughly discussed the shortcomings of Nishida's writings, but these shortcomings do not fundamentally affect the illuminating force of his insights. Nishida is trying to explore a realm for which there seems to be no roads at all, so it is fair to say that he is very courageous. The impulse of his exploration might be partly owing to Buddhism, in which "nothing" is by no means outright nonsense, nor does it refer to nihilism. By Nishida's critical exposition, "nothing" even constitutes the background of being. For example, Nishida says,

> But if what becomes the substance of relations is simply something like a point, force would have to disappear. That which truly envelops the relationship of force within must be something like a field of forces. . . The nothing that opposes being by negating it is not true nothing. Rather true nothing must be that which forms the background of being. (NKZ3 422)

Nishida's background in Buddhism is certainly quite unfamiliar for most Western philosophers, including Wittgenstein.[11] From their viewpoints, it is odd or even unthinkable to derive "being" from "nothing". This can be identified as one of the essential divergences between the fundamental conceptions of Wittgenstein and Nishida.

In contrast to the Basho of True Nothing, the bedrock of Wittgenstein's system of language game is "form of life". For Wittgenstein, not all language games

[11] It has been noted that there are some potential connections and similarities between Buddhism and Wittgenstein's philosophy (see Gudmunsen 1977). Although it is illuminating to attempt to find out such connections, the essential difference between Buddhism and Western Philosophy should never be neglected.

are on the same logical level. Rather, they constitute a hierarchy: some language games might be more fundamental, and what lies at the bottom of the hierarchy is the "form of life".[12] It is a significant concept, even though it is only mentioned in *Philosophical Investigations* 3 times:

> . . . And to imagine a language means to imagine a form of life. (PI 19)
> The word "language-*game*" is used here to emphasize the fact that the *speaking* of language is part of an activity, or of a form of life. (PI 24)
> "So you are saying that human agreement decides what is true and what is false?" What is true or false is what human beings say; and it is in their language that human beings agree. This is agreement not in opinions, but rather in form of life. (PI 241)

Wittgenstein realizes that the unsayable experience plays an essential role in the foundation of our sayable behaviours, and form of life is the precondition of all kinds of language games and even of meaning itself. I would go further by saying that, without form of life, we cannot reach any agreement in our daily communications or activities. In this sense, it seems parallel to the Basho of True Nothing. This is why Botz-Bornstein asserts, "For Nishida, a form of life emerges within the basho. For Wittgenstein, a Lebensform[13] develops out of an 'unsayable *Erlebnis*'"(Botz-Bornstein 2003, 55). In general, form of life is always related to something cultural or historical, which should be the "riverbed" of our daily life. However, from the point of view of Basho, form of life is still something in need of further investigation, because even the "riverbed" has to be based on something more fundamental, such as the earth. It can be concluded that the end of such investigations inevitably leads to True Nothing.

As mentioned earlier, for Nishida, who was influenced by Buddhism, "nothing" is something (this expression sounds paradoxical, in Nishida's style) that can constitute a foundation of another thing. However, for Wittgenstein, and perhaps most Western philosophers, anything has to be placed on something, so it is unacceptable to regard "nothing" as a real foundation. Actually, the English word "nothing" itself is very interesting: literally, it alludes to a "thing" in the first place and then denies its existence, asserting that there is not anything, or "no thing". In contrast, Nishida is able to use the Japanese word "mu" (無) straightforwardly

[12] In Japanese it is translated as "生活形式".
[13] "Lebensform" is the German word for "form of life".

without admitting any "thing" in advance. This may be attributed to a difference in culture or way of thinking. Thus, the comparison between form of life and the Basho of True Nothing may reveal the limitation of Western thought as well.[14]

7. Summary and Supplementary Comments

It can be seen that both Wittgenstein and Nishida try to transcend certain inherent stereotypes in the traditional philosophy: Nishida wants to dispel the ingrained dichotomy of subject-object and propose a new style of logic, while Wittgenstein tries to reconstruct the framework of the theory of meaning. Subsequently, both of them provide something new to reassess our traditional way of thinking. It is not easy to assess whether they have gained an outright victory. However, from a positive perspective, both of their intellectual enlightenments stand out.

I prefer to see the discussions until now as an introduction or schema, leading to more in-depth research on Nishida's other ideas. Currently, there are at least two approaches to carrying out further studies. One is derived from sections 4 and 5. The appropriate use of a concept in general requires two prerequisites: an exact definition of the concept and tenable arguments to justify the definition. Regarding the Basho of True Nothing, neither of the prerequisites are fully articulated in Nishida's writings, but it is our duty to reconstruct his argument and clarify this concept. In this way, the theory of Basho will become more dynamic.

The other approach originates from section 6. Nishida's conception of "True Nothing" has a background in Buddhism. For example, the "self-differentiating undifferentiatedness" of the Basho of True Nothing shows the most conspicuous Buddhist aspect of Nishida's thinking (see Krummel 2012, 18). Nishida's theory, under the influence of Buddhism, provides a possibility beyond the traditional philosophical ways of thinking, which is quite unfamiliar for Wittgenstein and most other Western philosophers. As we are allowed to talk about something transcending contradictions, Nishida's idea can be used as a "mirror" to reflect the potential shortcomings or limitations in Western thought as a whole.

In fact, East Asian philosophers are in quite a similar situation, having to construct their own philosophy or system of thought inspired by Western philosophy.

[14] Thanks to Prof. Hamauzu Shinni for his suggestion concerning the difference between Form or Life and Basho as well as that regarding taking Husserl into consideration.

I have attempted to reveal the possibility of reconsidering Nishida from an analytic perspective; hopefully, this work will inspire more innovative investigations.

References

朝倉友海 (2018)「場所的論理の形成期における意味の問題」『神戸外大論叢』, 68. 2.

Botz-Bornstein, Thorsten (2003) Nishida and Wittgenstein: from 'pure experience' to Lebensform or new perspectives for a philosophy of intercultural communication. *Asian Philosophy: An International Journal of the Philosophical Traditions of the East*, 13. 1, 53–70.

藤田正勝 (2000)「場所―根底からの思惟」『日本の哲学』第 1 号『特集：西田哲学研究の現在』(京都: 昭和堂)

Gudmunsen, Chris (1977) *Wittgenstein and Buddhism.* (London: Macmillan).

Heisig, James W (2001) *Philosophers of Nothingness: An Essay on the Kyoto School.* (Honolulu: University of Hawai'i Press).

Krummel, John (2015) *Nishida Kitarō's Chiasmatic Chorology.* (Bloomington and Indianapolis: Indiana University Press).

Krummel, John (2017) On (the) nothing: Heidegger and Nishida. *Continental Philosophy Review*, 1, 1–30.

西田幾多郎 (2003)『西田幾多郎全集・第三巻』(abbreviated as NKZ 3) (東京：岩波書店).

Nishida Kitarō (2012) *Place and Dialectic: Two Essays by Nishida Kitarō*. Trans. John W. M. Krummel and Shigenori Nagatomo (New York: Oxford University Press).

冲永宜司 (2009)「規則と場所―ウィトゲンシュタインと西田における根拠なき根源についての考察」『比較思想研究』, 36.

Wilkinson, Robert (2009) *Nishida and Western Philosophy.* (Burlington, VT: Ashgate).

Wittgenstein, Ludwig (1969) *On Certainty* (abbreviated as OC). G. E. M. Anscombe and G. H. von Wright (eds.), trans. D. Paul and G. E. M. Anscombe (Oxford: Blackwell).

Wittgenstein, Ludwig (2009) *Philosophical Investigations* (abbreviated as PI). Trans. G. E. M. Anscombe, P. M. S. Hacker and Joachim Schulte, revised 4th edition (West Sussex: Blackwell).

The Problem of Meaning in Nishida's Early Writings

Miikael-Adam Lotman
JSPS Research Fellow, Kyoto University

***Abstract**: This paper revisits a traditional problem in Nishidian scholarship regarding the link between Nishida Kitarō's maiden work* An Inquiry into the Good *(1911) and his next major treatise* Intuition and Reflection in Self-Consciousness *(1917). In this paper I will show that the problem of fact vs. meaning, which is explicitly treated in Intuition and Reflection in Self-Consciousness, is influenced by the writings of Motora Yūjirō. I make the case that Nishida's adoption of Royce's thought experiment of "the perfect map of England" as a model for "self-consciousness" was instrumental in his attempt to explain the entirety of reality.*

In the first section, I show that Nishida's earliest ambition of explaining everything with the concept of pure experience accommodates his contemporaries' understanding of the scope and task of philosophy. However, Nishida developed his own explanation of philosophy's task. According to Nishida, philosophy arises in response to a sense of anguish one feels after contemplating on the meaning of the world.

In the second section, I explain why Nishida believed that a reflection on the world's meaning causes philosophical anguish. This is due to an incompatibility between Nishida's metaphysical views and a criterion of meaning that he inherited from Motora. This combination left Nishida with a dilemma: either the world has no meaning, or everything has meaning but the world does not exist as a single entity.

In the last section, I argue that Royce's "perfect map of England" offered resolved this dilemma, by showing how the world as a single entity can have meaning in relation to its proper parts. These proper parts are numerically distinct perfect images of the original image of the world. I also point to a discrepancy between Nishida's and Royce's models, which would become problematic in Intuition and Reflection in Self-Consciousness.

Introduction

The Problem of Meaning in Nishida's Early Writings

Nishida Kitarō is often hailed as Japan's foremost philosopher. Although it is uncontentious that Nishida's philosophy underwent a series of intellectual developments, it is not clear what kinds of problems motivated Nishida's thinking and which of the various concepts are essential for understanding his philosophy. According to one traditional view, the developments in Nishida's philosophy culminated with the advent of his "logic of place" as expounded in his seminal paper "Place" (1926). One of the major proponents of this view is Kosaka Kunitsugu, who claims that "even though the formative process of Nishida's philosophy can be categorized into several periods, any such attempt admits the standpoint of "place" as one of its turning-points" (Kosaka 1994, 78). Given that this view is correct, we must still tackle the problem of philosophical motivation. What kinds of problems left Nishida dissatisfied with his earlier standpoints (*e.g.* the theories of "pure experience" or "self-consciousness") and in what way did the "logic of place" serve as a viable solution?

It has been recently argued that the ontological and logical status of "meaning" is one of the underlying problems that lead Nishida to develop his "logic of place" (Asakura 2018, 161–179). According to Asakura, the philosophical weight of this problem can be recognized in *Intuition and Reflection in Self-Consciousness*, which he deems Nishida's "first genuinely philosophical treatise" (ibid, 163). It is well-known that in *Intuition and Reflection in Self-Consciousness* Nishida set out to solve the problem of fact *vs.* meaning. In this paper, I will supplement Asakura's interpretation by showing that the problem of meaning predates Nishida's 1917 opus and can be traced back to his earliest manuscripts that would later comprise his maiden work *An Inquiry into the Good*. However, according to my interpretation the problem of meaning is not exhaustively addressed by the question of its logical or ontological status. Rather, I will argue that the most fundamental problem for Nishida consisted in the task of giving meaning to the totality of life in the broadest possible sense.

In the first two sections, I will reveal the nature and background of the problem by clarifying Nishida's fundamental beliefs about the holistic character of true reality and his understanding of the nature meaning, which he inherited from the writings of Motora Yūjirō. In the last section, I will introduce Josiah Royce's thought experiment about the "perfect map of England" and explain why it served as a solution to Nishida's original problem.

1. To Explain Everything

In *An Inquiry into the Good* Nishida Kitarō attempted "to explain everything by deeming pure experience the sole reality" (Nishida 1990, xxx). In his next major work *Intuition and Reflection in Self-Consciousness* he re-addressed the problem of the "entirety of reality" with the concept of self-consciousness, narrowing his focus to the explication of "the connection of fact and meaning, existence and value" (NKZa 2, 3)[1]. But why did Nishida undertake the herculean task of "explaining everything", rather than confining himself to some specific problem in, say, metaphysics, ontology, ethics or aesthetics?

An indirect answer to the first question has to do with the way Meiji-era thinkers understood philosophy. In *Dictionary of Philosophy* (1905) — the first Japanese explanatory dictionary of philosophy — Tomonaga Sanjūrō defines the field in terms of its form, method and object. According to Tomonaga, philosophy deals with the form of cognition (as opposed to emotion), which is organized into a coherent system of knowledge. Thus, its form is cognitive and method systematic. Regarding its object, Tomonaga writes:

> **3 Object** Philosophy always appears in the form of systematic cognitions, *i.e.* as science. What then distinguishes it from other sciences? The difference lies in its object. Having said this, the difference is not necessarily of quality but [rather] of quantity. Other sciences deal with objects of local scope. In contradistinction, philosophy deals with objects of global and universal [全般的普汎] scope. . . . Other sciences take parts of everything [萬有] as their research objects. In contradistinction, the research object of philosophy is the totality of everything in the entire universe [萬有全般宇宙全体]. (Tomonaga 1905, 163)

The above understanding of philosophy as a field that ought to treat "the totality of everything in the entire universe" is echoed in Kihira Tadayoshi's "The division of labor in academia and the task of philosophy" (1905), published two years before Nishida's first article. In this article Kihira laments the decline of philosophy, which has fragmented into independent disciplines and become subservient to the progress of natural sciences. According to Kihira, philosophy is originally a holistic

[1] Nishida's writings in *Nishida Kitarō Zenshū* [Complete works of Nishida Kitarō] are cited as "NKZa", followed by volume and page numbers.

enterprise that is tasked with the pursuit of "ultimate knowledge" (終局的知識). As such, it is diametrically opposed to the direction of natural sciences, which pursue specific and partial knowledge of reality, thus providing us with a fragmented worldview at best. Therefore, Nishida's attempt at "explaining everything by deeming pure experience the sole reality" was in line with his contemporaries' understanding of the task and scope of philosophy.

This, however, does not mean that Nishida did not have his own reasons for believing in the holistic nature of philosophy. In the "Fragments related to pure experience" (NKZa 16, 276–572), which contains Nishida's personal research notes and paragraphs that he scrapped from the publication of *An Inquiry into the Good*, we find a passage where Nishida explains the emergence of philosophical thought as an essentially holistic enterprise:

> Philosophy emerges from the demand for new life rather than from the demands of pure knowledge. In this regard, philosophy shares its origins with religion. . . . [T]hose who know not of anguish, or those who do not seek for a deeper life even at the bounds of despair, have no need for philosophy. However, there isn't a person who — having been driven to the bounds of despair or otherwise thought about the totality of life — has not felt a particular anguish in their hearts. There are those who will thence try to discover a new meaning of the world and of life; these people feel the true demand of philosophy. (NKZa 16, 566–567)

As seen in the above, for Nishida "the true demand of philosophy" consists in an attempt at forging a new meaning of "the world and of life". That is, philosophy aims at making sense of everything. Although this paragraph did not make it to the 1907 article, he would remain convinced that true reality (*i.e.* everything related to the "the world and of life") is not "not simply an existence but something with meaning" (Nishida 1990, 49).

According to the scrapped paragraph, the holistic task of philosophy is perpetuated by a sense of anguish, which arises when one thinks about the meaning of true reality. Thus, the task of explaining everything condemns philosophers to a potentially infinite life of despair. Perhaps this overly gloomy upshot of philosophy is why Nishida thought it best to remove this paragraph from the final draft. In the published article, Nishida presented a more optimistic view of philosophy, promising the reader that the "clarification of the nature of the universe, human life,

and true reality" (ibid, 38) will be conducive to moral knowledge and peace of mind (安心).

In the first chapter of *An Inquiry into the Good* (published in 1909 as "Pure experience, thinking, will and intellectual intuition"), Nishida elaborated on the epistemological aspects of pure experience deeming it synonymous with true reality as the preconceptual foundation of all phenomena of consciousness. This state of experience is famously exemplified with the following:

> The moment of seeing a color or hearing a sound, for example, is prior not only to the thought that the color or sound is the activity of an external object or that one is sensing it, but also to the judgment of what the color or sound might be. (Nishida 1990, 3)

In brief, pure experience is the experience of *that* before the *what?* or the consciousness of a brute fact before the question of its meaning is allowed to arise. It is in this sense that Nishida characterizes pure experience as "simply the present consciousness of facts just as they are" with no meaning whatsoever (NKZa 1, 15). When read as a monograph this characterization is in contradiction with that of true reality as "something with meaning". But when read in the chronological order as a series of articles and with reference to the above scrapped paragraph, it can be interpreted as a final solution to the problem of philosophical anguish, albeit an escapist one. Given that philosophical anguish is caused by a reflection on the totality of life, and further that a clarification of the nature of true reality is conducive to peace of mind, then the solution to the problem of anguish can be seen in the vanishing of the question of meaning. That is, peace of mind would be achieved by accepting the totality of experience as it is, without any further inquiry into its meaning.

Even if the above interpretation is theoretically viable, Nishida himself was evidently dissatisfied with such a solution. In fact, in the 1911 preface to his maiden work Nishida admits to becoming gradually aware of the discrepancy between his initial goal and the direction he was heading in:

> At first I intended to develop my ideas in the section on reality and then publish what I had written. Hindered by illness and other circumstances, I failed to achieve this goal. In the following years, my thought changed somewhat, and I began to sense the difficulty of doing what I had initially

> intended. At that point I decided to publish this book just as it was. (Nishida 1990, xxix)

To summarize, Nishida's attempt at "explaining everything by deeming pure experience the sole reality" can be interpreted as an expression of Nishida belief in the intrinsically holistic nature of philosophy. For Nishida, the true demand of philosophy emerges as a response to the feeling of anguish that accompanies one's reflections on the totality of life. His further characterization of pure experience as something utterly meaningless can be interpreted as a provisional solution to the problem of philosophical anguish. This solution, however, threatens to undermine the whole project of philosophy.

2. Meaning and Totality

Why did Nishida believe that reflecting on the totality of life begets anguish in the first place? In this section I will argue that it follows from a conflict between Nishida's view of reality as a self-contained totality and his understanding of the nature of explanation. The latter, in turn, is related to a psychologistic criterion of meaning that Nishida adopted from Motora. Let us start from the latter.

By the time Nishida started drafting his earliest article in August 1906, Motora had already published a series of articles in *Tetsugaku-zasshi* about the definitions of experience and the relationship between fact and meaning. The articles published from May to July 1906 should be of particular interest for Nishida scholarship: "What is experience?", "What is experience? (cont.)", "The distinction between self-sufficient and incomplete experience", "The relation between fact and its meaning", "The relation between fact and its meaning (cont.)". For the purposes of this paper I will confine myself to "The relation between fact and its meaning".

One of Motora's main concerns in the essay is to offer a unified view of reality, which is neither biased towards traditional empiricism that grounds reality in matters of fact nor towards rationalism that grounds reality on the relation of ideas. According to Motora both views originate from the illegitimate dichotomy of mind and matter, and "have yet to evade the residual maladies of common-sense dualism" (Motora 1915, 890). His alternative draws from William James' notion of radical empiricism, which "must neither admit into its constructions any element that is not directly experienced, nor exclude from them any element that is directly experienced"

(James 1912, 41). What sets radical empiricism apart from its traditional counterpart is James' insistence that "the relations that connect experiences must themselves be experienced relations, and any kind of relation experienced must be accounted as 'real' as anything else in the system" (ibid). That is, even though reality is sought exclusively in experience, experience is to consist in both relations and their relata. Hence the prior are not viewed as any less real than the latter.

Motora analyzed the relation-relata structure by distinguishing between two stages in the development of experience: self-sufficient experience (自全経験) and incomplete experience (不全経験):

> Self-sufficient experience is something that does not relate to anything else but is self-subsistent; since it does not allow for any acts of inference nor expectation whatsoever it is a pure experience or a so-called "fact". In contradistinction, incomplete experience possesses a fringe[2] and points to something else. Moreover, since the means that constitute its necessary conditions for development are naturally distinct and varied depending on its purpose, incomplete experience can, at times, take the form of judgements, techniques or imagination. (Motora 1915, 884)

Motora's psychologistic criterion of meaning is based on the above characterization of experience: a given actual experience (*i.e.* something that is already revealed in experience; 既に経験となり) is meaningful if only if it relates to a discrete potential experience (*i.e.* something that is yet to appear as an experience; 経験に現れんとして未だに現れざる). If an actual experience fails to relate to a potential experience, then it is "self-sufficient" and hence barren of meaning. Motora refers to the latter state of experience with the term "fact", because it requires no further acts of inference or expectation but simply is there. For example, the "the clouds' evening glow" means "tomorrow's good weather", whereas the meaning of "ploughing fields" is "to harvest crops". The first is an example of inference, the

[2] The term originates from James' essay "Stream of Consciousness" (1892) referring to the "halo of relations" or "psychic overtones" that are said to accompany "all objects before the mind". James exemplifies this notion with instances of trying to recall a name that is on the tip of one's tongue but at the same time wholly absent from the mind. James considers these experiences counterexamples to the "ridiculous theory of Hume and Berkeley that we can have no images but of perfectly definite things" (James 1892, 254). For Motora, the fringe is not characteristic of all objects or mental images but only of underdeveloped stages of experience.

second exemplifies the purpose of an activity, but both kinds of meaning are supposed to signify a relation between an actual experience and a potential experience. Motora concludes from this analysis, that meaning is a state of "incomplete experience" and facts are states of "self-sufficient experience" (ibid, 892–893). Although Motora does not explicitly state this, his underlying intuition seems to be that nothing can have meaning in and of itself; meaning is essentially a relation between two discrete terms. This intuition is clearly reflected in Nishida's maiden work:

> The meanings of, or judgments about, an experience are simply expressions of its relation to other experiences. . . . [Meanings and judgments] indicate the relation between present consciousness and other consciousnesses, and therefore merely express the position of present consciousness within the network of consciousness. For example, when one interprets an auditory sensation to be the sound of a bell, one has merely established the sensation's position relative to past experiences. (Nishida 1990, 9)

In the above passage, Nishida substitutes "actual experience" with "present consciousness" and "potential experience" with "other consciousnesses" but the understanding of meaning as a relation between discrete terms remains the same. Furthermore, this criterion informs his understanding of the nature of explanation: "to explain is to be able to include other things into a single system" (ibid, 29).

While Motora's criterion could be used to "explain everything" for some metaphysical systems (*e.g.* James's pluralistic "mosaic philosophy"), it proves malignant for Nishida's theory of pure experience. To see that this is the case, we need to come to terms with two important features of Nishida's theory. Firstly, that it is a clear instance of immanence philosophy (内在哲学) and, secondly, that it reifies the totality of reality in terms of "the self-development of *a single entity*" (ibid, 57; emphasis mine). Immanence philosophy as defined by Tomonaga refers to:

> [t]he philosophy that is based on immediately given facts or pure experience. The "immanence" [in "immanence philosophy"] is opposed to "transcendence", signifying the scope of experience. One of the characteristics of immanence philosophy is that it attempts to explain the

> world by reducing the entirety of reality to the contents of consciousness, *i.e.* its immanent elements. (Tomonaga 1905, 285)

This definition is clearly reiterated in Chapter 2 of *An Inquiry into the Good*:

> From the perspective of pure experience, there are no independent, self-sufficient facts apart from our phenomena of consciousness . . . Our world consists of the facts called phenomena of consciousness, and all of the various philosophical and scientific systems are no more than explanations of these facts. (Nishida 1991, 44)

Now, let us assume that reality is wholly reducible to the facts of consciousness. Next, consider consciousness as a single entity Ω. Lastly, let us adopt Motora's criterion of meaning. Given these premises, we face an obvious dilemma: either Ω exists as a meaningless totality, for no fact can transcend Ω to make it meaningful, or everything has meaning but Ω does not exist as a self-contained reality. With this in mind, we can finally see why Nishida believed that reflecting on the totality of life is the cause of philosophical anguish. For, if life constitutes a self-contained totality in the state of pure experience, then all reflections on its meaning render our lives incomplete: "[t]he state of pure experience thus breaks apart and crumbles away. Such things as meanings and judgments are states of this disunity" (ibid, 9). To be sure, Nishida envisioned this tension between pure experience and its meaning as a dialectical process of self-development, whereby reality gains ever greater depth of meaning. However, even at that, this process can never terminate as a meaningful self-contained totality.

3. Roycean Solution

In the above sections I showed why Nishida's attempt at comprehending the meaning of everything led him to conclude that the experience of totality itself (*i.e.* pure experience) must be meaningless. After the publication of *An Inquiry into the Good*, however, Nishida started to speak of pure experience as "the world of understanding and meaning" (NKZa 1, 227) or "the world of value and meaning" (NKZa 1, 301). In this section, I will argue that Nishida's reading of Josiah Royce

allowed him to defend the idea of self-contained reality without having to give up its claim to meaning.

To escape from the meaninglessness of the totality of facts Nishida had to reject Motora's criterion of meaning as an external relation between two terms. In his response to Takahashi Satomi's criticism of *An Inquiry into the Good*'s conflicting accounts of meaning, Nishida writes:

> As Takahashi claims, consciousness as a fact ought to be able to have meaning. . . . For the fact [of consciousness] to have meaning, it should not refer to some transcendent thing that has no bearing whatsoever on its contents. Shouldn't we deem the fact of immediate experience as one furnished with meaning in itself? (NKZa 1, 311)

As revealed in a rather cryptic passage in "Fragments related to pure experience", Nishida had already grown aware of this problem while drafting the manuscripts of *An Inquiry into the Good*. In this passage he calls Motora's psychologistic account of meaning into question and suggests that an alternative account can be found in the form of Josiah Royce's absolute idealism:

> [Marginalia: Against Prof. Motora] Consciousness in general is not inference but fact. . . . Isn't it possible that within myself I feel the personal self as much as I feel the social self, or the cosmic self? . . . In what form does "the consciousness" exist? How does it function in the present? . . . How is everything (事々物々) absolute? Consult Royce. (NKZa 16, 544)

Here, Nishida criticizes Motora's view of consciousness as something that is characteristic of inference. To recall, for Motora all acts of inferring, expecting or imagining are meaningful states of consciousness or states of "incomplete experience" that seek to establish a link between an actual and a potential experience. While this view could account for partial psychological acts, Nishida contends that it cannot make sense of consciousness in its most general form. By "general consciousness" or "the consciousness" Nishida refers to consciousness of the broadest possible scope, *i.e.* of the "entire universe" (全宇宙; ibid, 543).

> When we are reading a book, the book is absolute; it is the entire universe. We do not seek for its external cause. . . . Should we accept this hypothesis

> of universal consciousness (which I call "the consciousness"), we will find that the universe . . . always exists in the present. It has no past or future. Thus, there is no need to demand for causal laws to govern it. (ibid, 543)

The above passages suggest that the most general form of consciousness, comprising the totality of facts, can be meaningful in relation to its parts: the cosmic self, the social self and the personal self. In other words, Motora's criterion of meaning can be revised without arriving at the precarious conclusion that facts can mean themselves. The key is to distinguish between discreteness and distinctness: the observation that some items are different does not entail that they are separate, for parts and wholes are examples of distinct yet non-discrete items.

The whole-part structure can be applied to facts of consciousness. For example, the fact that someone is ploughing a field means that there is a plough on the field. It is in this sense that Nishida can claim that facts have meaning *in* themselves without claiming that facts mean themselves. However, this insight by itself is insufficient for concluding that the fact of all facts can be meaningful. In order to claim this, we would need to arrive at a perspective where parts can mean wholes. Unfortunately, the fact that there is a plough on the field does not mean that someone is ploughing the field. In order to solve this problem, let us follow Nishida in consulting Royce's example of "the perfect map of England".

In his supplementary essay for *The World and the Individual* (1900), Royce offers a thought experiment devised in defense of the concept of actual infinity (Royce 1900, 475–476). He asks us to imagine the following scenario: "[u]pon and within the surface of England there exists somehow (no matter how or when made) an absolutely perfect map of the whole of England" (ibid, 506). The perfect map in question is in one-to-one correspondence with the mapped territory. Since the map itself lies on the surface of England it must be included as part of the territory. This will result in a "series of maps within maps such that no one of the maps was the last in the series" (ibid).

For our purposes, this experiment offers a way of vindicating the meaning of ultimate fact of consciousness. That is, it enables us to assign a meaning to the original self-contained totality Ω in relation to its proper part Ω', a meaning to Ω' in relation to Ω" and so on *ad infinitum*. The proper part Ω' can mean its whole because it is a numerically distinct perfect image of Ω. Ultimately, the totality will include an infinite series of subordinate totalities and is designed for the sole purpose of giving meaning to itself. This means that Nishida can, in theory, continue his quest for the

explanation of everything without the need to revise his fundamental belief about the nature of reality as a self-contained fact.

Alas, Nishida never explicated his motivations for adopting Royce's thought experiment as a model for "self-consciousness". Thus, we cannot be certain whether Royce's writings were instrumental in solving the problem of meaning. Given the historiographical evidence presented up to this point, it seems like a plausible explanation. There is yet another sense in which this explanation is charitable, owing to a fundamental difference between Royce's original thought experiment and Nishida's interpretation thereof.

As already mentioned, Royce's thought experiment served to vindicate the concept of actual infinity from traditional charges of self-contradiction. This, in turn, was meant as a defense of the reality of relations from Bradley's regress arguments in *Appearance and Reality* (1897). In brief, Royce's strategy was to grant Bradley the major premise that if one relation exists then an infinite number of relations should exist, but to deny the hasty conclusion that if something is infinite in number, then it cannot exist in actuality. It is for this reason that Royce adamantly insists that when it comes to infinite number series "the whole infinite series, possessing no last member, is asserted as an *existing fact*" (ibid, 506; emphasis mine). Again, in terms of the "perfect map": "there would be implied the assertion *not now of a process of trying to draw maps*, but of the contemporaneous presence, in England, of an infinite number of maps" (ibid, 507; emphasis mine).

In sharp contrast, Nishida interprets the "perfect map" as representative of a potentially infinite process of "reflecting the self within the self" (自己の中に自己を写す), noting that the only thing left is "to clarify its relation to the actual" (NKZa 16, 543). Nishida undertook this task in Chapter 10 of *Intuition and Reflection in Self-Consciousness* (NKZa 2, 67–72). *Prima facie*, this constitutes a gross misunderstanding of Royce's entire idea and it would have saved Nishida a great deal of trouble had he adopted Royce's "perfect map" as an example of actual infinity. However, when viewed in terms of our original problem, it would make sense for Nishida to adopt some of its structural features without the subscribing to the doctrine of the actually infinite. That is, Nishida's interpretation of Royce accounts for the possibility of making sense of the totality of life, while leaving room for the significance of our actual mortality. In other words, it shows us how an eternal Being could make potentially infinite sense of itself without the need for invoking another transcendent being. And yet, the meaning of life for us mortal beings is paradoxically grounded in the very negation of this possibility.

Conclusion

The broader aim of this paper is to contribute to Nishidian scholarship by supplementing a recent interpretation, according to which the problem of meaning was essential for Nishida's intellectual development. More specifically, I argued that Nishida adopted some aspects of Royce's "perfect map of England" for his concept of "self-consciousness" in order to give a meaningful explanation to reality as a self-contained totality. That is, I showed that a totality, which contains an infinite series of subtotalities, can satisfy a minimum requirement of meaning. The requirement in question dictates that nothing can mean itself. The totality, as described above, satisfies this requirement because its proper parts are numerically distinct perfect images of the original totality. For Nishida, this model satisfied two metaphysical features of his theory of pure experience: immanence philosophy and reification of the ultimate totality. The first requires that reality is wholly contained in consciousness and the second dictates that consciousness as the totality is itself an entity.

I also pointed to a major difference between Nishida and Royce. That is, Nishida's understanding of self-consciousness is potentially not actually infinite. I suggested that this might be due to Nishida's understanding of self-consciousness as an endless process of meaning-making, which is bounded by our actual mortality. However, the problem of actual *vs.* potential infinity became a central issue in *Intuition and Reflection in Self-Consciousness* and must await a separate treatment. For the purposes of this paper, it suffices to show that the problem of meaning links Nishida's maiden work to his second major treatise.

References

Asakura, Tomomi (2018) Basho-teki ronri no keiseiki ni okeru imi no mondai.*The kobe Gaidai ronso*, 68. 2, 161–180.

James, William (1892) *Principles of Psychology Vol 1*. (New York: Holt).

James, William (1912) *Essays in Radical Empiricism*. (New York: Longmans).

Kosaka, Kunitsugu (1994) *Tetsugaku to shukyo*. (Tokyo: Minerva-shobo).

Motora, Yūjirō (1915) *Shinrigaku gairon*. (Tokyo: Teimatsu-shuppansha).

Nishida, Kitarō (1990) *An Inquiry into the Good,* Trans. Abe, M. and Ives, C. (New Haven and London: Yale University Press).

Nishida, Kitarō (1965–1966) *Nishida Kitaro Zenshu*. 19 vols. (Tokyo: Iwanami Shoten, 2nd ed.)

Royce, Josiah (1900) *The World and the Individual, First Series, Supplementary Essay*. (New York: The Macmillan Company).

Tomonaga, Sanjūrō (1905) *Tetsugaku-jiten*. (Tokyo: Hobunkan).

Right Action: Development-Based Virtue Ethical Account

HONDA Masaya
Adjunct Professor, Doane University

Introduction

A virtue-ethical account of right action attempts to explain the rightness of an action by appealing to concepts such as goodness, excellence, and virtue. In this paper, I defend a version of the virtue-ethical account of right action that I call the "development-based account" (DBA): the rightness of an action is determined by the action's *conformity to the development*, wherein an agent manifests sympathy in ways relative to his/her nonmoral dispositions, stage of moral development, and final purpose. I develop this account based on Kitarō Nishida's account of the good[1] and, in so doing, I analyze the problem with virtue-ethical accounts as indicated by Robert Johnson. In "Virtue and Right", he argues that no virtue-ethical account of right action can sufficiently explain the idea that "we ought to become better people", [2] such that we (ordinary, nonvirtuous people) ought to pursue self-improvement, self-control, and advice from people in morally better positions. I demonstrate how the DBA resolves this problem.[3]

1. Overview: Development-Based Account of Right Action

In his book, *An Inquiry into the Good*, Nishida defends and develops an account of the good, saying, ". . . the good is the development and completion—the self-realization—of the self."[4] For Nishida, the good is not about merely achieving aspects of happiness, such as pleasure, well-being, and human flourishing, nor is it about fulfilling the will, through which decisions are wholly determined by moral demands. Instead, the good is the *whole process of development* through which

[1] Kitarō Nishida, *An Inquiry into the Good* (New Haven, Yale University Press, 1990), 100–145.
[2] Robert N. Johnson, "Virtue and Right", *Ethics* 113 (2003): 810–834, 810.
[3] I thank an anonymous reviewer who provided me with fruitful suggestions for developing this paper.
[4] Nishida, *An Inquiry into the Good*, 125.

agents manifest their dispositions in ways that allow them to become unique parts of a harmonious, well-coordinated whole.[5]

Nishida adopts this idea of the good from Plato, who holds that justice is a whole wherein each part works without impeding any other.[6] Like Plato, for Nishida, in a way, this is the final end of moral development, which is achievable only by manifesting virtues (though for Nishida, virtues are individualized, as seen below). However, Nishida's account of the good is quite different from Plato's (and, in fact, from that of Aristotle and his followers). Nishida initially characterizes the good as follows:

> Personality. . . which is the unifying power of consciousness. . . is first actualized in individuals. At the base of one's consciousness exists unanalyzable individuality. All activities of consciousness are an expression of this individuality: each person's knowledge, feeling, and volition possess qualities unique to the person.[7]

As this passage shows, for Nishida, reality itself does not have any structure. For him, the foundation of reality[8] is considered "pure experience:" the original experience based upon which we can come to be aware of, to make judgments about, and to know things at all.[9] Thus, we experience things without being aware of who we are and what is in front of us, that is, we simply experience things as they are—we sense, perceive, feel, will, understand, and know things, though, without being aware of any of those *distinctions*.[10] That is, at this level of experience, we do not distinguish ourselves from objects: in pure experience, there is no subjective–objective distinction.

Moreover, for Nishida, like Plato, moral development occurs along with our cognitive and emotional development. However, for Plato, moral development is primarily cognitive, wherein the mind comes to have an intelligible structure

---

[5] Masaya Honda, "Individualizing Virtues: Comparing Kitarō Nishida's Normative Naturalism with Neo-Aristotelian Naturalism", *International Philosophical Quarterly* 56, no. 1 (2016): 70–75.
[6] Plato, *Republic*, 433a–c.
[7] Nishida, *An Inquiry into the Good*, 136–7.
[8] Joel W. Krueger, "The Varieties of Pure Experience: William James and Kitaro Nishida on Consciousness and Embodiment", *William James Studies* 1, no. 1, 12
[9] Robert Edgar Carter, *The Kyoto School: An Introduction* (New York: State University of New York Press, 2013), 28.
[10] Ibid., 29–30.

(justice) through contemplation, in which reason redirects the goals of our desires (from "money-lover" to "honor-lover" to wisdom-lover).[11] In contrast, for Nishida, moral development occurs at the level of pure experience in which all mental activities are *individualized*—not given as distinctive "parts" (functions) but given *as a particular whole*, which we can analyze later. What consists of this "whole" using concepts such as "knowledge, feeling, and volition" is, for him, somehow artificial. Thus, moral development can be traced by observing how an agent develops his/her personality: how he/she develops (among other traits) his/her feelings, volitions, and knowledge, where ultimately, there is no subjective–objective distinction. Thus, Nishida's account of the good is initially characterized as follows:

(1) Reality (pure experience) has no intelligible structure (though, it is still intelligible).
(2) Moral development occurs at this level of reality.

Setting aside the truth of (1), it is a fact that pure experience is the terminal point of our experience, beyond which we have no experience at all. Moreover, it is plausible to say that our moral development occurs at this level of experience as a *whole-person* development rather than primarily occurring in one aspect of our lives, such as in cognitive development[12]—in how reason comes to redirect the goals of our desires. In the following sections, after I describe Robert Johnson's criticism of the virtue-ethical account of right action (section 2), I explore Nishida's theses, which I found to be interesting for the current purpose:

(3) Morality is characterized by a manifestation of sympathy (section 3).
(4) Sympathy is fully manifested if and only if this enables an agent to continuously be a unique part of a coordinate whole (section 3).[13]

---

[11] See C.D.C. Reeve, *Philosopher-Kings* (Indianapolis: Hackett Publishing Company, INC, 1988), 250–252.

[12] Plato's emphasis on cognitive development in moral education is criticized by, for instance, George Grote, *Plato, and the other companions of Sokrates* (London: John Murray, 1875), 399–400. Mark Jonas replies to this sort of objection in "Plato's Anti-Kohlbergian Program for Moral Education" (presentation, Philosophy of Education Society of Great Britain Annual Conference, New College, Oxford, March 26–29, 2015, 2–8).

[13] In this paper, I interpret Nishida as a virtue ethicist because he is sympathetic to Plato and Aristotle rather than to Mill and Kant, and sympathy as an expression of whole character plays a central role in his account of the good.

(5) The final end of morality is to fully manifest sympathy (section 3).
(6) Thus, moral development (the good) is characterized by an agent developing sympathy in ways that allow him/her to constantly remain a unique part of a coordinate whole (section 3).

Based on these, I shall argue:

(7) The rightness of an action can be explained by the action's conformity to this moral development (section 4).
(8) This solves the difficulty of the virtue-ethical account of right action, which is mentioned by Robert Johnson, who holds that moral obligation includes actions that pursue moral improvement (section 4).

(7) is the thesis I call the "development-based account of right action" (DBA), which can be initially formulized as follows:

> An agent's action is right for the agent at a given time (t), in given circumstances (C), if and only if it fits a segment of development (at t in C) wherein the agent realizes him- or herself.

In the DBA, each segment can be determined by analyzing an agent's pure experience in terms of *how* he/she manifests sympathy: whose and which needs he/she cares about and serves at a given time (t) and in given circumstances (C). Let us assume that developing sympathy requires *habituation*—constantly acting upon sympathy. Then, a moral development (for an agent per segment) can be explained as follows: (1) an agent fully and constantly manifests sympathy, though relative to his/her dispositions at t in C; and (2) manifesting sympathy thusly would enable him/her to achieve the final end of his/her moral development: to continuously be a unique "part" of the coordinated "whole" (though at the level of pure experience, the distinction between a "part" and a "whole" does not exist) (see section 3). I shall argue that the rightness of an action, for a given agent at t in C, can be explained based on whether the action could be performed by this agent's counterpart, who shares the same set of nonmoral dispositions, who is at the same level of development, yet who would successfully forward the moral development as (1) and (2) above describe (see section 4).

2. Johnson's Challenge

In "Virtue and Right", Robert Johnson argues that no virtue-ethical account of right action can sufficiently explain the idea that "we ought to become better people",[14] such that we (ordinary, nonvirtuous people) ought to pursue self-improvement, self-control, and advice from people in morally better positions.[15] As Johnson argues, there seems to be no significant and evident connection between being virtuous and pursuing these apparently morally permissible activities. Consider the following cases:

1. A person who habitually tells lies cannot help this behavior not because of vice but because of "insufficient appreciation of the value of truthfulness". He tells lies to please people but, based on a friend's advice, decides to change this habit. Based on a therapist's advice, he begins writing down his lies "to become more aware of his habits and to keep track of improvements".[16]
2. A person struggles to do what he should. His day-to-day life reveals a pattern of behavior characteristic of a person who is at war internally with his malicious and cowardly desires. This person takes measures to prevent the satisfaction of his vicious desires; after all, "in order to perform a just, brave, kind, or otherwise virtuous action, a nonvirtuous person will have to control himself in many ways".[17]
3. A person is morally insensitive in some area because of a moral blind spot; however, "he possesses enough self-awareness to know this, and when he has reason to doubt his perception, he asks for guidance from a friend who is in these respects more virtuous and whose vison is in these respects unhindered".[18]

In these circumstances, Johnson argues, we ought to do something a virtuous person would *not* do (or a virtuous person would not be motivated to do, or virtues would

[14] Johnson, "Virtue and Right", 810.
[15] Johnson later mentions that this criticism does not apply to virtue ethics that denies a theory of the right.
[16] Ibid., 816–817.
[17] Ibid., 820–821.
[18] Ibid., 822.

not aim to achieve[19]). Quoting Aristotle, Johnson argues that the not-yet-virtuous person must "develop the traits of taking the lesser evil, acting contrary to his natural tendencies, and avoiding what is pleasant", although virtue does not actually consist of any of these traits. Moreover, he argues, the aim of any virtue "does not include the acquisition of those self-same virtues, self-control, or the improvement in one's moral perception, nor could there be a special virtue of self-improvement".[20] Thus, even if we tend to believe that the activities involved in pursuing becoming a better person are somehow right or permissible, the virtue-ethical account cannot explain this.

Assuming that these difficulties are unavoidable, Johnson suggests three ways to modify the virtue-ethical account: (1) distinguishing two senses of "right"—as fully adequate and as morally excellent,[21] (2) adopting a version of the idealizing theory, wherein a right action is one that an ideal version of an agent would do (or would recommend or perceive to be right), and (3) focusing on "any completely virtuous person's history…about how she developed the virtues"—thus, right action is explained in relation to this development. While Johnson contends that none of these strategies work,[22] I nevertheless believe that to sufficiently explain the rightness of pursuing self-improvement, self-control, and advice from people in morally better positions, the concept of *moral development,* wherein one *becomes* virtuous (i.e., manifests sympathy), must be developed, and right action can be successfully explained based on this concept. For this, some sort of idealization must be made regarding how an agent manifests sympathy (see the following sections).

---

[19] Ibid., 830–834. Johnson briefly discusses the difficulties of other prominent virtue-ethical accounts of right action, such as Rosalind Hursthouse, *On Virtue Ethics* (New York, Oxford University Press, 2000), Michael Slote, "Agent-Based Virtue Ethics" *Midwest Studies in Philosophy* 20 (1995): 83–101, and Christine Swanton, "A Virtue Ethical Account of Right Action" *Ethics*, 112.1 (2001). In these accounts, right action is explained in relation to the *manifestation of virtues*, have already been developed. Johnson's criticism properly applies to all these accounts.

[20] Ibid., 833.

[21] Ibid., 825.

[22] Various scholars attempt to respond to Johnson's criticism of the virtue ethical account of right action. For this see, Varlerie Tiberius "How to Think about Virtue and Right", *Philosophical Papers* 35, no. 2 (2006): 247–265 and Sean McAleer, "Four Solutions to the Alleged Incompleteness of Virtue Ethics" *Journal of Ethics & Social Philosophy* 4, no. 3 (2010): 1–20.

3. Sympathy and the Coordinated Whole

Certainly, as Johnson indicates, a virtue-ethical account of right action that considers a concept of moral development needs further explanation. For this, let us discuss the moral status of pure experience. Apparently, this sort of experience does not characterize morality as a theory of right action must presuppose. Nishida says:

> Take, for example, a work of art. When does the true personality or originality of the painter appear? Insofar as the painter intends various things in his nature and the brush follows the will. The expression of personality in the moral realm is no different from this.[23]

Here, Nishida gives an example of one's actualizing personality. Previously, he agreed with Plato and Aristotle that "the satisfaction of reason [in the form of 'intellectual intuition'] is our highest good".[24] However, unlike Plato and Aristotle, Nishida claims this is a matter of love—a deep concern for the object's actualizing its own nature, regardless of whether it is a person or natural object.[25] He says, "each individual's true self is the system of independent, self-sufficient reality appearing before that person", and, in this way, "the sincerest demands of each and every person necessarily coincide at all times with the ideals of the objective world the person sees".[26]

Nishida's account of morality in this passage certainly sounds odd. For instance, the experience of an infant who indulges his/her appetite for milk or an artist who indulges him/herself in the ecstasy of creating an art piece, which might be counted as pure experience, does not appear to have any positive moral status. To clarify this, there are two key points: (1) the development in pure experience is characterized in terms of *sympathy*; and (2) sympathy is further characterized based on the final end of moral development: to constantly be a unique part of the coordinated whole. In this section, I shall discuss (1) and (2) to clarify Nishida's account of the good and the moral status of pure experience.

According to Nishida, an agent's moral purpose is to fully manifest his/her personality.[27] This is to manifest sympathy, wherein an agent has *immediate* access

[23] Ibid., 134.
[24] Ibid., 129.
[25] Ibid., 135.
[26] Ibid., 134.
[27] Nishida, *An Inquiry into the Good*, 125.

to the nature of an object (or objects) that he/she coexists with in the same environment. For Nishida, this means to manifest his/her mental traits, such as knowledge, feeling, and volition, in interacting with the objects of those traits such that all these traits and objects are unified in one "consciousness".[28] He states:

> Because our infinite spirit is never fundamentally satisfied by the unity constituted by an individual self, it inevitably seeks a larger unity, a great self that envelops both oneself and others. We come to express sympathy toward others and seek congruence and unity between oneself and others.[29]

Nishida seems to hold that an agent is capable of manifesting sympathy in pure experience because therein the distinction between "oneself" and "others" vanishes. That is, in manifesting sympathy, one is unsatisfied with being oneself and instead pursues a "larger unity" wherein one seeks "congruence and unity between oneself and others." Thus, an agent's moral purpose ("the good") is to undergo a development wherein his/she "is unified" with other objects and/or agents in fully manifested sympathy. For Nishida, sympathy is not just an awareness of others' needs as if they were one's own, as this understanding is too broad.

Rather, sympathy is an experience through which an agent, in *interacting* with an object (or objects), is purely absorbed into the object(s) (which Nishida calls "self-effacing action") and in which the distinction between the subject and the object disappears for *both parties*. At this level of experience, *not only* an agent (x) refers to what he/she experiences as one*,* a whole seamless reality prior to being either a subject or an object, *but also* other agents (y and z) *share* this experience with x, each from their own perspective (although x, y, and z have no recognition of the subject–object distinction).[30] Thus, at this level of experience, the needs of oneself (x) and others (y and z) should be experienced without any awareness of who, in particular, has those needs. For Nishida, this does not mean that no meaningful experience is possible at the level of pure experience (unlike, for instance, what William James holds[31]). Consider the following:

---

[28] Nishida assumes there is a psychic trait ("will") that unifies all particular psychic traits into one consciousness.
[29] Nishida, *An Inquiry into the Good*, 82–3.
[30] Ibid., 56.
[31] Nishida took the concept of "pure experience" from James, who claimed it is a chaotic whole. See Robert Edgar Carter, *The Kyoto School: An Introduction* (New York: State University of New York Press, 2013), 27–28.

> In the case of martial arts (e.g., Judo), the Judoka sometimes experiences a moment when he/she can flip down his/her opponent without all the might in conscious and doesn't have the awareness of flipping the opponent, nor the opponent has the awareness of having been flipped. What it seems that happens is that both bodies automatically move together. As they become nothing, their experiences are something they cannot express with words at that moment. But after having flipped the opponent decisively, they may try to articulate it or remember it.[32]

In such an experience, the needs of oneself and others are mutually accessible and therefore can be immediately and *mutually* satisfied (assuming that the agents involved need mutual flourishing).[33] Certainly, though, this does not make sense if we assume that the judoka and the opponent simply aim to win or to beat the other. Instead, as Jigoro Kano (1860–1938), the founder of judo, claims, judo should be practiced for the "mutual prosperity of the self and others".[34] In other words, the practice should be for promoting the practitioners' well-being (such as [though not limited to] promoting physical strength and mutual respect). Assuming this, at the very moment the judoka flipped the opponent, the needs of both were satisfied (assuming the opponent used a proper defense), that is, both practitioners fully exhibited their offensive and defensive skills spontaneously and simultaneously as the opportunity arose.

Nishida claims that a self-effacing action (occurring in pure experience) comes with an agent's "intellectual intuition" in which knowledge is evident in how the person immediately grasps how things should be developed and completed.[35] Thus, a skilled agent, such as an artist or a craftsman, can immediately see how to develop and complete their work. Notice that in the judoka example, the execution of skills presupposes the knowledge of judo—how to control one's body to suppress one's opponents' as well as the knowledge of the doctrines found in Confucianism,

---

[32] Koyo Fukasawa, "The Potentiality of Empathy with Others in Competitive Sport: A Suggestion from Nishida's 'Pure Experience' and 'I' and 'Thou'" *International Journal of Sport and Health Science* 12 (2014): 47–52.

[33] Involvement of awareness in "pure experience" seems to be controversial. See Robert Edgar Carter, *The Kyoto School: An Introduction* (New York: State University of New York Press, 2013), 29.

[34] Fukusawa, "The Potentiality of Empathy", 47.

[35] Ibid., 31–33.

Daoism, Buddhism, and Shintoism.[36] In particular, Zen Buddhism (through Bushido, "the way of warriors") has influenced judo such that judokas are required to focus on mental and spiritual development for the good or value of a community.[37] Thus, in judo training, judokas "know" (though without deliberation) what they are doing—how to mutually satisfy their needs as well as the community's. Let us say that Nishida's sympathy consists of a set of dispositions including (1) having immediate access to the needs of oneself and others, (2) recognizing which needs and in which way these needs should be cared for and served (for the good or value of a community), and (3) acting upon these needs such that this promotes the mutual flourishing of anyone involved in that action.

Furthermore, Nishida's account of sympathy presupposes the *final* end of moral development: a *certain* sort of mutual flourishing. As the previous examples of the infant and artist show, being a self-effacing action alone is not enough to characterize the morality found in sympathy. As the judoka example shows, sympathy requires that an agent fulfill the needs for mutual flourishing. However, an action promoting mutual flourishing still does not suffice to explain morality. For instance, apparently, gang members can mutually flourish when they mutually satisfy their needs.[38] Therefore, the idea of the coordinated whole must be investigated. Nishida claims:

> Clearly, a particular demand becomes good only when it is related to the whole. For example, the good of the body derives not from the health of one of its parts but from the harmony of the body as a whole. . . . The good is primarily a coordinated harmony.[39]
>
> The demands of the personality are the unifying power of consciousness and, at the same time, an expression of the infinite unifying power at the base of reality. And so, to actualize and fulfill our personality means to become one

[36] "Itsusu no Kata", United States Judo Association, accessed March 18, 2020, http://www.judomjcnarbonne.fr/pdf/itsutsu_no_kata_guide_lines.pdf
[37] Ibid.
[38] See, for example, Gary Watson, "On the Primacy of Character", in *Identity, Character and Morality*, eds. O. Flanagan and A.O. Rorty (London: MIT Press, 1990), 449–483, 462–3. Also see Rosalind Hursthouse, *On Virtue Ethics* (Oxford: Oxford University Press, 1999), 192–3.
[39] Nishida, *An Inquiry into the Good*, 128.

> with this underlying power. If we construe the good in this way, we can determine the nature of good conduct.[40]

In these passages, following the Platonic–Aristotelian notion of the good as *eudemonia* (or [human] flourishing), Nishida locates the end of a "unifying power" (found in sympathy) in achieving a coordinated harmony. For him, the good is a sort of proper relation (such as health and *eudaemonia*) exhibited in a "whole" (such as the body and a community) sustained by its properly functioning "parts" (such as body parts and community members) in which each "part" has its own history (though in pure experience, the "part–whole" distinction vanishes for all parties). Based on my previous discussion, coordinated harmony as the final end of moral development can be initially characterized as follows. (1) In sympathy which, as previously mentioned, is a unifying power, an agent seeks a "larger unity". As the above passage shows, Nishida sets no limit for this power: potentially, *anything* (living and nonliving things) and *anywhere* (a group, a community, an ecosystem, and a universe) could be a part of a coordinated whole. (2) As discussed in the first section, Nishida holds that the manifestation of this power is *agent*-relative: each and every agent has his/her *own* way to manifest sympathy (i.e., they play a *unique* role in the coordinated whole). (3) Sympathy can be understood as a *placeholder* for a set of dispositions (in the case of humans, "individualized virtues", that is, pure experience wherein a set of virtues is manifested as a particular whole) that enables each and every agent to play his/her unique role in the coordinated whole.[41] (4) This coordinated whole has its own *history* that encompasses all the histories of the members inside the whole. (5) All agents in a community share one reality (pure experience) by manifesting sympathy, though from their own perspectives, in the ways described in (1)–(4).

Consider the following community. (1) This is the *most comprehensive* community (hereafter, Community) in which each and every agent fully and constantly manifests sympathy with one another—no sympathy is manifested to impede anyone from mutually flourishing; any sympathy is manifested to promote someone in mutually flourishing. (2) In the Community, any sympathy is manifested in ways *relative to* each and every agent: relative to their nonmoral dispositions and to the segments of their developments. That is, for most people, given their nonmoral dispositions, there would be circumstances in which they could have

[40] Ibid., 132.
[41] Ibid., 22, 70.

behaved such that they mutually flourish with others. In the Community, these circumstances keep arising, so that no one fails to fully and constantly manifest sympathy. (3) Let us assume then that each (but not every) agent in the actual world has his/her counterpart who shares the same set of nonmoral dispositions: in addition to physical dispositions, each counterpart shares the same personality type—how he/she tends to react to the environment, whether he/she is introverted/extroverted, neurotic/stable,[42] etc.; nonetheless, each counterpart is capable of staying in the Community because the circumstances allow him/her to continue developing a set of virtues relative to his/her nonmoral dispositions. (4) Furthermore, consider the history of an agent (*A*) who has fully developed sympathy, as he/she has developed intelligence and emotion in such a way that he/she could live in the Community *throughout his/her life*.

Let us (roughly) sketch what it is like to live in the Community throughout one's life using Piaget–Kohlberg's theory of moral development.[43] Let us *not* assume that this theory explains the rightness of a moral development: how we *ought* to develop our morality.[44] Instead, for the sake of sketching the Community's membership, let us assume that along with our cognitive development, we typically undergo the following stages of development: (i) the preconventional, (ii) conventional, and (ii) postconventional stages. Based on how we reason against a set of hypothetical moral dilemmas (such as Heinz's dilemma[45]), it is shown that as we grow cognitively, we tend to shift how we deal with the needs of ourselves and others. That is, at stage (i), an agent tends to care for and serve only his/her own

[42] For this, we might facilitate some theories of personality. For instance, see Hans Eysenck, *Dimensions of Personality* (Routledge, 1997).

[43] Lawrence Kohlberg, "Moral Stages and Moralization", in *Moral Development And Behavior*, ed. Thomas Lickona (New York: Holt, Rinehart, And Winston, 1976), 31–53, 34–35.

[44] For instance, see Carol Gilligan, *In different voice* (Cambridge MA: Harvard University Press, 1982), where Gillian provides a feminist criticism of Kohlberg's theory of moral development. Later, Kohlberg criticizes Aristotelian virtues as ineffective in moral education (see L. Kohlberg "Education for Justice – A Modern Statement of the Socratic view", in *The Philosophy of Moral Development* (San Francisco: Harper and Row, 1981), 30–31. However, recently, Aristotle's account of moral development is discussed by Albert Silverstein and Isabel Trombetti, "Aristotle's Account of Moral Development", *Journal of Theoretical and Philosophical Psychology* 33, no. 4 (2013).

[45] This dilemma describes that Heinz, whose wife is on her death bed and needs a drug to survive, must decide, after using all means to collect money for the drug but falling short by half, whether he steals the drug from a druggist who refuses to discount it, Kohlberg, "Moral Stages and Moralization", 44–46.

needs. At stage (ii), he/she comes to tend to care for and serve the needs of the individuals he/she encounters and of a specific group of people (such as his/her society and institute). At stage (iii), he/she comes to tend to care for and serve people's needs in general.[46]

Now consider how sympathy could be manifested at each stage of (i)–(iii) so that an agent (*A*) could stay in the Community. For instance, at stage (i), in a self-effacing action, sympathy could be manifested in such a way that only *A*'s own needs are focused on (as in the above infant and artist examples) yet *in ways that do not impede anyone from his/her mutual flourishing*. For this to be possible, known, and practiced, *A* would need to have some sort of *self*-regarding virtues, such as temperance, self-esteem, and prudence (where these virtues help *A* avoid conflicts with the needs of any others, and in which *A* is incapable of having other-regarding virtues [because his/her needs are self-centered]). At stage (ii), in a self-effacing action, sympathy could be manifested such that only *A*'s own needs and those of a specific group of people are focused on, yet so that *not only* does this not impede anyone from mutually flourishing *it also promotes someone* in mutually flourishing in that group. For this to be possible, known, and practiced, in addition to self-regarding virtues, *A* would need to have some sort of (nonuniversal) *other*-regarding virtues, such as kindness and generosity (where these virtues help *A* avoid conflict with the needs of people outside the group, and in which *A* is incapable of having universal virtues). Finally, at stage (iii), in a self-effacing action, sympathy could be manifested so that the needs of people in general (including those of *A* him/herself) are focused on so that not only does this not impede anyone from his/her mutual flourishing, it also promotes *anyone* mutually flourishing. For this to be possible, known, and practiced, in addition to self-regarding virtues and (nonuniversal) other-regarding virtues, *A* would need to have some sort of *universal* virtues, such as justice and benevolence.

Let us assume thus that the Community's history includes the histories of each and every agent who develops intelligence and emotion, where each and every agent fully and constantly manifests sympathy in ways relative to their nonmoral dispositions and to their segments of development. Moreover, the manifestation of sympathy for a given agent, at a given time, in given circumstances, could be analyzed (though only after deliberation) as a set of individualized virtues: pure

[46] Kohlberg, "Moral Stages and Moralization", 34–35. For the sake of argument, I limit my discussion about the Community membership to people, though Nishida's theory might allow much wider membership such as sentient beings in general (or even things in general), as he does not limit the object of sympathy.

experience from this agent's perspective analyzed as a manifestation of a combination of virtues, which enables the agent, who has a unique set of nonmoral dispositions, to stay in the Community.

4. Conformity as Moral Rightness: "Self-Realization"

In this section, based on Nishida's account of the good discussed so far, I shall argue that the rightness of an action, for a given agent, at a given time, in given circumstances can be explained based on whether the action *conforms* to the sympathy manifested by a virtuous version of this agent who shares the same set of nonmoral dispositions and who is in the same stage of moral development (see stages [i]–[iii]), though who successfully lives in the Community by constantly manifesting his/her individualized virtues.

Johnson criticized the type of approach that attempts to explain the rightness of an action based on a virtuous person's history: how a virtuous person comes to be virtuous. No virtuous person is virtuous by birth; he/she comes to be virtuous through moral education. Then, at earlier stages of moral development, he/she has moral obligations that indicate he/she ought to pursue self-improvement, self-control, and advice from people in morally better positions. Nonetheless, there seems to be innumerous ways to describe such development and, with some cases described below, this type of approach seems to fail.

Sean McAleer argues that there are four ways to solve the difficulties suggested by Johnson,[47] and one of these suggestions is somehow similar to the DBA. In this interesting suggestion, McAleer indicates that the rightness of an action could be explained based on the manifestation of some of the following Mengzian virtues: benevolence (*ren*), righteousness (*yi*), appropriateness (*li*), and wisdom (*chi*).[48] Among those, McAleer interprets that Mengzi's righteousness can be understood as a sort of situational appropriateness, whose appropriateness is agent-relative.[49] That is, he mentions that "possessing the virtue of righteousness ensures that one's conduct conforms to the way", yet, this is also "a disposition to accord with agent-relative prohibitions involving the expression and preservation of one's own ethical character". With some other points, he suggests:

[47] Sean McAleer, "Four Solutions to the Alleged Incompleteness of Virtue Ethics", *Journal of Ethics and Social Philosophy* 4, no. 3 (2010): 1–20.
[48] Ibid., 10.
[49] Ibid., 13.

> The subvirtuous person is not doing what the virtuous person would do in those circumstances, but this is because the same virtue, righteousness, makes different demands on people at different stages or levels of moral development. What is righteous or appropriate depends not only upon the circumstances, but upon the nature of the agent, as well.[50]

In terms of this passage, McAleer mentions that if an agent has an upward disposition to become virtuous, he/she ought to challenge him-/herself toward virtue; if an agent has a downward disposition to vice, he/she ought to stop him/herself from being vicious. For instance, "if one steals a chicken every day from a neighbor, cutting down the thievery to one chicken a month would be a step in the right direction, but appropriateness requires that one stop stealing chickens altogether".[51] According to this interpretation, one ought not to steal a chicken *at all*. Nevertheless, given that an agent has a downward disposition to vice, it is right for one to reduce the frequency of stealing from daily to monthly. However, this seems to imply that it is morally *permissible* for this agent (who has certain degrees of downward dispositions to vice) to steal monthly, and this would still apply even this agent had a tendency to steal (or even kill) humans instead of chickens. Thus, it is unclear how this agent-relative obligation is consistent with the agent-neutral obligation of appropriateness: he/she ought not to steal chickens (or humans) at all.

More seriously, it is unclear what it is to have an upward disposition to virtue and a downward disposition to vice. Consider the case of Oskar Schindler, a member of the Nazi party who pursued material happiness for most of his life, but who during World War II saved more than a thousand Jews by sacrificing his fortune and risking his life. Suppose it were true that before his moral conversion, he went to Poland and made his fortune on the black market, building relationships with the local Gestapo. However, suppose that his character traits significantly changed in an altruistic way during/after the course of these activities. Suppose then that all his activities as a Nazi somehow affected him and caused him to experience a moral conversion (perhaps he regretted what he had done) so that, in the end, he came to manifest sympathy in an altruistic way. In this example, for Schindler to manifest downward dispositions to vice is somehow a part of his upward dispositions to virtue: when he was a Nazi, manifesting downward dispositions to

[50] Ibid., 12.
[51] *Mengzi*, 3B8.

vice disposed him to developing upward dispositions to virtue. Thus, in his case, using the above Mengzian account, it seems that Schindler ought to increase the number of vicious deeds to enable him to be a virtuous person.

This suggests that for any account of right action in favor of virtue ethics based on some sort of moral development, the alleged development enabled by virtues must be properly specified. With the understanding of Nishida's sympathy and the coordinated harmony in the previous section, the DBA will suffice for this. Let us analyze the case of an agent who has a disposition to steal a chicken daily. Consider whether it is morally permissible for this agent to reduce the frequency of stealing a chicken from daily to monthly. To evaluate this, we must first look at the stage of moral development this agent is in: either sympathy is manifested (1) only for satisfying his/her own needs, or (2) for satisfying the needs of a specific group of people, or (3) for satisfying the needs of people in general (namely whoever is involved in the action). Then, we must look at whether this action could be performed by an agent who shares the same set of nonmoral dispositions and who is in the same stage of moral development, though who manifests individualized virtues: either (a) self-regarding virtues or (b) (nonuniversal) other-regarding virtues or (c) universal virtues (where the manifestation of [a]–[c] must enable that agent to live in the Community). Thus, let us assume that stealing a chicken harms someone's property, and this is against the owner's needs to mutually flourish (unless, for instance, the owner wanted to allow another to steal his/her chicken if the stealer were starving to death). Then, this action could not be said to be, say, tempered, self-esteemed, and prudential so that this should be morally impermissible for all stages of development. Moreover, the stealer who tends to steal a chicken daily has an obligation to improve him/herself by developing self-regarding virtues suitable to his/her nonmoral dispositions. This would include pursuing self-improvement, self-control, and advice from people who share the same set of nonmoral dispositions, though who are successfully developing self-regarding virtues.

Notice that this would *not* imply that the above agent ought to reduce the frequency of stealing. This is precisely because cultivating virtues does not necessarily mean changing a course of action *in terms of* a single type of behavior.[52] There are quite different ways in which that agent could reduce the need to steal by reinforcing his/her (individualized) self-regarding virtues. For instance, his/her temperance, self-esteem, and prudence would have told him/her this: stealing would

[52] For instance, see Hursthouse, *On Virtue Ethics*, 10–11.

cause him/her to gain a bad reputation so that to live comfortably, he/she would have to redirect his/her needs. Rather than stealing a chicken, if the agent needed a thrill, he/she could go on an adventure; if he/she needed some food, he/she could inquire about food aid. Moreover, any of these improvements, self-controls, or pieces of advice could be better provided by a counterpart who would share the same set of nonmoral dispositions (which would dispose him/her to steal a chicken if the opportunities that the above agent faced had arisen in the actual world), yet because of the lack of such opportunities, who could stay in the Community. This would be especially important because if an agent had an entirely different set of nonmoral dispositions, he/she would have proceeded on an entirely different course of moral development.

Now let us resume Schindler's example. Suppose that Schindler is in stage (ii) assuming that he is a dedicated Nazi member. Then, it is morally *im*permissible for him to act *only* for the sake of his own flourishing—at the same time, he ought to promote the needs of others in his group (say, the local Gestapos) as, say, his kindness and generosity would have disposed him to do (where he might have no such virtues). From this, nonetheless, it does *not* follow that it is morally *permissible* for him to join in any activities as the local Gestapos do *for the sake of the Nazis*. This is because Schindler is also under the obligation not to impede anyone (including Jews, as they are a part of the Community) from mutually flourishing as his self-regarding virtues could have told him (given that joining in any activities as the local Gestapos do harms his reputation among Jews). In the Community, by virtue of having individualized virtues relative to each and every agent, there is no such manifestation of sympathy that impedes anyone from mutual flourishing. Hence, even if an action would please the local Gestapos, it ought not to be done. This does not prevent Schindler from doing something kind and generous for the local Gestapos—on the contrary, he is obliged to promote their flourishing as a kind and generous version of him would do. Thus, he has the moral obligation to pursue self-improvement, self-control, and advice from a counterpart who is in a morally better position: who is developing (nonuniversal) other-regarding virtues (though, again, relative to his nonmoral dispositions). Furthermore, in stage (ii), it *may* be morally permissible for Schindler not to promote the flourishing of Jews (though, it is impermissible to impede Jews from mutually flourishing), depending on what kind of nonmoral dispositions he/she has.[53]

[53] As I have discussed above, Nishida's account of the good can provide us with an insight specific enough to sketch the DBA in favor of virtue ethics. The advantage of the DBA is in

Conclusion

In this paper, I have defended a version of the virtue-ethical account of right action that I call the "development-based account" (DBA), in which the rightness of an action is determined by the action's conformity to the individual's development, wherein an agent manifests sympathy in ways relative to his/her nonmoral dispositions, stage of moral development, and final purpose. I developed this account based on Kitarō Nishida's virtue-ethical account of the good. For this, I analyzed the problem with virtue-ethical accounts, as presented by Robert Johnson. In "Virtue and Right", his critique is that no virtue-ethical account of right action can sufficiently explain the idea that "we ought to become better people" such that we (ordinary, nonvirtuous people) ought to pursue self-improvement, self-control, and advice from people in morally better positions. I demonstrated how the DBA addressed and resolved this problem.

References

Carter, Robert E. (2013) *The Kyoto School: An Introduction*. (New York: State University of New York Press).

Eysenck, Hans (1997) *Dimensions of Personality.* (London: Routledge).

Fukasawa, Koyo (2014) The Potentiality of Empathy with Others in Competitive Sport: A Suggestion from Nishida's 'Pure Experience' and 'I' and 'Thou'. *International Journal of Sport and Health Science*, 12, 47–52.

Gilligan, Carol (1982) *In different voice.* (Cambridge MA: Harvard University Press).

Grote, George (1875) *Plato, and the other companions of Sokrates.* (London: John Murray).

---

its connectiveness to further researches in the fields of psychology, such as theories of moral development, of moral education, and of personality. For instance, in this paper, I discussed only three stages of moral development without articulating many traits that comprise personality. However, with further researches, various stages of moral development with detailed analyses of nonmoral dispositions could be discovered, and this could assist the DBA in overcoming one of the major difficulties usually ascribed to virtue ethics, namely, the lack of action-guidance (see Hursthouse, *On Virtue Ethics*, 35–42).

Honda, Masaya (2016) Individualizing Virtues: Comparing Kitaro Nishida's Normative Naturalism with Neo-Aristotelian Naturalism, *International Philosophical Quarterly*, 56. 1, 57–76.

Hursthouse, Rosalind (2000) *On Virtue Ethics.* (New York: Oxford University Press).

Johnson, Robert N. (2003) Virtue and Right. *Ethics*, 113, 810–834.

Kohlberg, Lawrence (1976) Moral Stages and Moralization, in T. Lickona (ed.) *Moral Development And Behavior*, (New York: Holt, Rinehart, And Winston), 31–53.

Kohlberg, Lawrence (1981) *The Philosophy of Moral Development.* (San Francisco: Harper and Row).

Krueger, Joel W. (2006) The Varieties of Pure Experience: William James and Kitaro Nishida on Consciousness and Embodiment. *William James Studies,*1, 1–37.

McAleer, Sean (2010) Four Solutions to the Alleged Incompleteness of Virtue Ethics. *Journal of Ethics & Social Philosophy*, 4. 3, 1–20.

Nishida, Kitarō (1990) *An Inquiry into the Good.* (New Haven: Yale University Press).

Plato (1992), *Republic*. translated by G.M.A. Grube revised by C.D.C. Reeve (Indianapolis: Hackett Publishing Company).

Reeve, C. D. C. (1988) *Philosopher-Kings.* (Indianapolis: Hackett Publishing Company).

Silverstein, Albert and Trombetti, Isabel (2013) Aristotle's Account of Moral Development. *Journal of Theoretical and Philosophical Psychology*, 33. 4, 233–252.

Slote, Michael (1995) Agent-Based Virtue Ethics. *Midwest Studies in Philosophy*, 20, 83–101.

Swanton, Christine (2001) A Virtue Ethical Account of Right Action, *Ethics*, 112. 1, 32–52.

Tiberius, Varlerie (2006) How to Think about Virtue and Right, *Philosophical Papers*, 35. 2, 247–265

Watson, Gary (1990) On the Primacy of Character, in O. Flanagan and A.O. Rorty (eds.) *Identity, Character and Morality*, (London: MIT Press), 449–483.

Meta-Cognition Without a Cognizer:
Buddhist Non-Self and Awareness of Awareness

Mark Siderits
Seoul National University / Kyoto University

***Abstract**: The doctrine of non-self is central to the theory and practice of classical Indian Buddhism. One source of recent interest in this doctrine is the widely shared sense that belief in a self may be an obstacle to naturalizing the mind and mental phenomena. While many Buddhists might reject such a project, Buddhist efforts to explain how mental processes function in the absence of an enduring subject of experience may still be worth exploring. One issue about which Buddhist philosophers had much to say is how self-knowledge or meta-cognition (our ability to cognize our own cognitions) is possible if there is no self. The example used by Indian philosophers is that of seeing blue color: typically one is able to report not just the presence of blue, but also that one sees blue. I seem to be aware not only of the things that I see and feel and think about, but also of my seeing and feeling and thinking. Descartes took this as proof of the self. The question is whether Buddhists can account for the ability given their allegiance to non-self.*

I begin by sketching the non-self doctrine as articulated in Buddhist reductionism. This will introduce the problem of self-knowledge, followed by discussion of two Buddhist responses to the problem: the claim that cognition is reflexive in nature (that cognitions cognize themselves), and a higher-order-thought account. I claim that the second response is better suited to the Buddhist project than the first. I then investigate the evidence that can be marshalled in its defense, some of which will be drawn from current work in developmental psychology and philosophy of mind.

1.

It is well known that Buddhists deny the existence of the self. What is not well known is that Buddhists distinguish between the concept of the self and the concept of the person, and that most Buddhists take not an eliminativist but a reductionist stance toward the latter. According to the mainstream Buddhist view, while persons are not to be found in our ultimate ontology, the concept of the person plays a

sufficiently important place in our cognitive economy that it should be granted the second-tier status of 'conceptual construction'. Their view is that while strictly speaking there are no persons, it is still perfectly understandable, given our interests and cognitive limitations, that the concept of a person should play a central role in the conceptual scheme used by most people most of the time. In order to see how this view arose, we need to begin by looking at the soteriological project that Buddhism shares with most other Indian philosophical schools.

Indian philosophizing grew out of the project of seeking release from *saṃsāra* (the round of rebirth). The basic idea behind the project is that life as ordinarily lived is unsatisfactory because goals are chosen based on false beliefs about who we truly are. A life devoted to sensual pleasure, for instance, is based on the assumption that the self is the sort of thing that can be made better or worse through the presence of pleasure or pain. What sets Buddhism apart from other Indian liberation projects is its rejection of the idea of a self as an essence or pole of identification. The crucial mistake we make is not that we identify with the wrong sort of thing, but that we identify at all. Buddhists claim that what they call the 'I'-sense, the sense of being a persisting subject of experience and agent of action, is misleading and the source of existential suffering.

The Buddhist strategy for overcoming this mistake begins by distinguishing between two possible referents for the 'I' of the 'I'-sense: a *self*, understood as the one part among all the psychophysical elements that grounds diachronic personal identity; and a *person*, understood as the whole that is composed of the many psychophysical elements. The best-known Buddhist argument for the nonexistence of a self is the argument from impermanence.[1] It uses a taxonomy of five kinds of psychophysical elements: one corporeal and four mental. The argument is simply that since none of these elements is permanent, and a self would have to be permanent, it follows that there is no self. That the self must be permanent follows from the rebirth assumption plus the claim that the self is the essence of a sentient being. If the sentient being persists through uncountably many lives, then the entity that is its essence must likewise persist.

Many Buddhist philosophers recognized that the refutation of the self left the 'I'-sense largely unscathed.[2] They explained the cognitive impenetrability of the

[1] A clear statement of the argument is to be found at *Majjhima Nikāya* III, 15–20.

[2] Thus Candrakīrti likens someone who thinks refuting this 'philosopher's self' will extirpate the 'I'-sense to one who, knowing there are snakes living in the walls of their house, comforts themselves with the thought that there is no elephant inside (*Madhyamakāvatāra vṛtti* 6, 141).

'I'-sense through their account of the *person*, understood as the concept of the mereological sum of causally connected sets of appropriately assembled psychophysical elements. As mereological nihilists, they naturally deny that there is any such mereological sum. But after giving their argument against mereological realism,[3] they have some important things to say concerning the semantics of enumerative expressions.

One might take the lesson of Buddhist mereological nihilism to be that we should be eliminativists about persons. As the mereological sum of mereological sums—the assembled elements at one moment in causal connection with the assembled elements at the next moment, etc.—the person looks to be ripe for elimination from our ontology. But, says the Buddhist, to eliminate the person from our ontology is to eliminate what for most people grounds appropriation: identification by one set of elements with other sets occurring earlier and later in the causal series. As we have already seen, Buddhists do claim that appropriation is a mistake. Still, simply jettisoning the practice would have bad consequences, as illustrated with the examples of the pregnant woman, the student and the convicted criminal. For instance, eliminativism means the criminal would not appropriate and thus identify with those earlier parts of the causal series that committed the crime, would hence see their present punishment as undeserved, and would thus not be deterred from future crime. The concept *person* plays an important role in our cognitive economy. Yet Buddhists do claim that its use lies behind the problem of existential suffering. The solution is not to eliminate persons but to *reduce* them to causal series of psychophysical elements.

This is explicated through their claim that 'person' is an opaque enumerative term. By contrast, the word 'pair' is a *transparent* enumerative term. It would be at best a bad joke to say that my dresser drawer contains three things: two matching socks and a pair of socks. 'Pair', like 'dozen', 'gross', 'multitude', and 'heap', is merely a useful way to refer to a multiplicity. For creatures with our interests and cognitive limitations, it is a useful cognitive shortcut. That this semantic role is transparent is what explains the fact that use of the term is not taken to commit us to the existence of such things as pairs, heaps, or multitudes. At least such use does not generate a *serious* ontological commitment. It does, we might say, generate a *casual* ontological commitment. Buddhists express this by saying that things of this sort are conventionally, but not ultimately, real.

[3] For details see Siderits 2015, 100–103.

Our use of 'person' is said to be like this, except for the crucial difference that the term's being enumerative is opaque to us. This is, they think, the source of the 'I'-sense, and so of existential suffering. Their recommended solution is to render the term transparently enumerative. Reduction of persons to psychophysical elements in causal succession is their strategy. The idea is that we can then retain the benefits conferred by this useful cognitive shortcut without paying the price of existential suffering.

I shall not go into the details of that strategy here. For our purposes it suffices to say that the proposed reduction involves a sort of dualist ontology in the reduction base. This is what is hinted at in the term 'psychophysical elements': the ultimate simples are sorted into two categories, the physical and the mental. The two categories are distinct insofar as only the physical elements are said to have such spatial properties as location. The dualism here is not substance dualism but trope dualism. Buddhists take their mereological nihilism to rule out not just composite physical objects, but, more generally, substances of any sort; these are understood to reduce to bundles of momentary trope occurrences. Chief among the mental tropes allowed in their ultimate ontology are momentary occurrences of consciousness, understood as bare awareness or registry. And if one were to try to analyze 'thinking substance' into its ultimate mental constituents, this looks like one likely candidate. It seems intuitively plausible that consciousness is qualitatively simple in character and is thus not reducible to anything both mental and yet simpler. Its inclusion does, however, raise considerable difficulties.

2.

These difficulties grow out of the problem of explaining how meta-cognition is possible. Not only are we (at least sometimes) conscious, we may also be aware of our being conscious, cognize our own cognitions.[4] The simplest way of accounting for this fact involves positing a subject with cognizing as one of its modes.

[4] Unlike recent discussions of self-knowledge, the classical Indian debate over how self-knowledge is possible does not start from the presupposition that a mental state is conscious only if that state is itself represented. While some Indian accounts of self-knowledge do have as a consequence that consciousness requires self-consciousness, it is understood that this result requires argument. What is agreed by all is that some cognitions are themselves cognized; this is the phenomenon that is thought to require explaining.

Buddhists reject such an account. They thus owe us an explanation of what it is that cognizes a given cognition. To say that it is that very cognition itself—that a cognition may function as both subject and object simultaneously—seems implausible. For it violates the widely accepted principle of irreflexivity: a simple entity cannot operate on itself. The illustrative examples are legion: the knife that cannot cut itself, the fingertip that cannot touch itself, and the like. In the absence of a valid counter-example, the principle appears to hold, and to rule out reflexive self-cognition as an account of self-knowledge.

Buddhists still needed an account of self-knowledge. They were committed to the claim that one attains enlightenment by coming to somehow *directly* grasp the fact that all existents are characterized by impermanence, suffering and non-self. Since this 'all existents' must include any cognition that grasps these characteristics in other things, it is not clear how such a cognition could be directly grasped as bearing the marks if it could not cognize itself. A variety of possible solutions was discussed. One such solution—Dignāga's reflexivity thesis—came to be particularly influential. It is the view that every cognition cognizes not only its object but also itself. It grows out of Buddhist embrace of a representationalist view of perception, according to which the direct object of perceptual cognition is not an external object but instead the form borne by the cognition due to sensory contact with the external object. A cognition must therefore have two forms: that of its intentional object and that of itself as that which cognizes. Dignāga then argues that the possibility of meta-cognition can only be explained by supposing that these two forms are in fact one, that cognition illuminates its object by illuminating itself. To the objection that this is ruled out by the irreflexivity principle, his commentator Dharmakīrti proposes, as a counter-example to the principle, the case of the light that illuminates itself as well as other things in the room.

In propounding his reflexivity account of self-knowledge, Dignāga seems committed to what Carruthers (2011) calls the transparency thesis—the thesis that the mind has transparent access to its own states. While Carruthers does not himself accept this thesis, he says it may be something that humans at all times and in all cultures are strongly disposed to accept as true. Bogdan (2010) likewise rejects the thesis; his developmentally based account of self-knowledge claims instead that a child's awareness of its own mental states develops through deployment of a theory the child first acquires in order to explain and predict the behavior of others. But he nicely illustrates the transparency thesis in his formulation of the objection he thinks will naturally arise in response to his own account of self-knowledge:

> Doesn't one first need to illuminate one's mind and register what is there before—and in order to—recognize, classify and make use of what is there? Don't we first turn on the lights before seeing and recognizing what is in the room? And isn't consciousness this light that illuminates the mind, before anything else is figured out there? Isn't it the case that in the morning, when awakening, the brain first turns on the consciousness switch before we become aware—and in order to become aware—of what is going on around us and inside our minds? And when consciousness weakens or is switched off, so is the mind worth talking about? In short, mustn't one become introvertly conscious of one's own thoughts *before* determining their attitudinal profile? (Bogdan 2010, 70)

Dignāga would agree.

I shall return to Carruthers' and Bogdan's defense of their alternative opacity thesis, the thesis that the mind is opaque to itself and its states. First, though, I should point out that while Dignāga's reflexivity thesis was influential in Buddhist philosophical circles, it was not the consensus view. Some Buddhists accepted a higher-order perception (HOP) account that sees meta-cognition as introspection or 'looking within'.[5] Still others propounded a sort of higher-order thought (HOT) account, according to which meta-cognition arises out of an abductive inference. The latter account requires some explanation. The classical Indian formulation of HOT was first developed by the non-Buddhist philosopher Kumārila. He starts with a simple but compelling attack on the claim (shared by reflexivists and HOP theorists) that one can be directly aware of one's own mental states. This is not possible, since these are mental states, like those of another person. The idea here is that if we are using a single concept when we attribute mental states to ourselves and to others, then the criteria of application should be uniform across the allegedly huge gap between the first-person case and the third-person case. Since the criteria in the

[5] Here I am using 'higher-order' with reference to theories of meta-cognition. The term is also used in connection with theories concerning what property a mental state must have to be a conscious state. Typically, higher-order accounts of state consciousness claim that a mental state is a conscious state only if it is itself represented by some distinct ('higher-order') state. The best-known alternative first-order account claims instead that a mental state is conscious just in case it has the dispositional property of making its content available in the global workspace, i.e., to mental modules for functions like memory, speech and action. The HOT account of self-knowledge I shall discuss begins from such a first-order account of state consciousness.

third-person case are behavioral (e.g., direction of another's gaze is a sign of their attending to an object in that direction), the same should hold in the first-person case as well. What behaviors might reveal that an object such as a fruit has been cognized? Action, such as reaching for the fruit, and verbal conduct, such as saying that a fruit is present. The key point here is that the mental state that is thereby cognized is a theoretical posit. Cognitions, whether another's or one's own, are never directly grasped. The abductive inference whereby cognitions are cognized is an inference to the best explanation. And the explanation in question is one that uses a theory of mind, a theory first developed in order to help explain and predict the behavior of others in a shared physical and social environment.

Those Buddhists who adopted this account of self-knowledge rejected Kumārila's assumption that the cognitions thus inferred are states of a self. Of course, it was still open to them to hold that consciousness itself is ultimately real. A bundle theorist need not believe in an immaterial self in order to believe that cognitions have non-physical existence. And their existence might still be consistent with their never being directly grasped. But it is also open to a Buddhist who holds the opacity thesis to claim that consciousness is no more than a theoretical construct, and so is reducible to entities of some other sort entirely. This possibility was glimpsed early on by Śrīlāta, the Sautrāntika who pointed out that if there is no direct acquaintance with consciousness, its nature can only be specified in functional terms, and that functionalization is an open invitation to reduction (Dhammajoti 2007, 163–4). The conclusion to be drawn is that consciousness is not ultimately real.[6] And it is, once again, difficult to imagine what non-physical trope consciousness might be reducible to, given its apparent simplicity in character. There is, moreover, a specifically Buddhist motivation at work here as well. Not only does Dignāga's reflexivity thesis violate the irreflexivity principle, it also threatens to reinforce the notion of an inner subjective realm accessible only through first-person modes of awareness. Some modern scholars (e.g., Dreyfus 2011) take Dignāga's claim that every cognition is reflexively self-aware to be true to the

[6] The Buddhists who followed Kumārila in explaining first-person attribution of cognition as abductive inference were global anti-realist Mādhyamikas. As such, they deny that anything is ultimately real, including the physical as well as the mental. They would thus find suspect the physicalism of those who wish to naturalize the mental. Both parties can, though, agree on the negative thesis that consciousness is not ultimately real. Whether a naturalizing project can withstand Madhyamaka critique of all ultimate ontologies is a question that lies beyond the scope of this paper.

phenomenology of our experiential lives. Many Buddhists, though, see it as just another reinscribing of the 'I'-sense.

Why, though, should consciousness be reduced and not eliminated? In a word, because the theory that posits conscious mental states, our 'folk psychology', is useful. Even as we have begun to discover the nature of the underlying neurophysiological states on which conscious mental states supervene, we find it difficult to imagine how knowledge of those states, considered strictly as neurophysiological states, could ever play the role that our folk psychology plays in everyday life. Our folk psychology reliably facilitates complex feats of self-regulation and social interaction. The computational demands would be overwhelming for systems like ours were we to jettison the theory and try to use a completed brain science instead. The states that folk psychology attributes to us are, then, conventionally real. What they reduce to, however, are physical states, and this probably accounts for the fact that Buddhists did not embrace this reductionist strategy. Physicalism makes the karma/rebirth ideology appear implausible, and Indian Buddhists took that ideology quite seriously. Further exploration of this way of accounting for meta-cognition requires that we turn to more recent discussions.

3.

Both Carruthers and Bogdan espouse versions of the opacity thesis that are compatible with the view of some Buddhists that consciousness is a conceptual construction. Bogdan builds his case on the basis of what we now know about human cognitive development. The opacity thesis posits an asymmetry between self-knowledge and other-knowledge: the ability to reliably attribute mental states to ourselves is built up out of tools that were first acquired for the purpose of explaining and predicting the behavior of others.[7] Bogdan supports the claim that there is such an asymmetry by examining the record concerning cognitive development in early childhood. Classical Indian Buddhist philosophers were unaware of many of the facts that make up this record. It would clearly be a mistake

[7] Proponents of the transparency thesis typically posit a reverse asymmetry: subjects are directly acquainted with their own conscious states, and only subsequently come to infer such states in others by analogy. The difficulties with the analogical approach are well known. These are captured in the thought that attributions of mental states must then employ two distinct concepts, one for the first-person case and another for the third-person case. This strikes many as tantamount to solipsism.

to suppose that the case Bogdan makes is what these Buddhists 'really had in mind'. Their concern was just to develop an account of meta-cognition that was consistent with Buddhist non-self. Our question is whether there is now any evidence supporting their account.

Bogdan's story goes roughly like this. Because of the human obstetrical dilemma—bipedalism leads to selection for greater brain size, leading to a requirement for earlier birth—human infants are born premature (relative to other primates) and are thus dependent on adult caregivers. This leads in turn to selection for innate mechanisms that promote attachment and ensure that the infant's subsistence needs are met. These mechanisms include such things as differential attention to faces (present at birth) and imitation of facial gestures like smiling (present soon after birth).[8] It is of course tempting to see this repertoire as part of a deliberate strategy on the infant's part. (More about this temptation later.) But it is no more appropriate to attribute intention here than it is in the case of the car that only fails to start on rainy weekday mornings. These are indeed reflexes. Bogdan sees in them, though, the materials out of which is constructed the full-blown self-consciousness that the transparency thesis claims is innate. The key point for our purposes is that the kind of self-knowledge reflected in the 'I'-sense is a construction: 'self-consciousness appears to be a *by-product assembled* out of unrelated ontogenetic adaptations with their own histories of selection' (Bogdan 2010, 165). Its being so opens the door to the possibility that the sense of an inner first-person realm is, as Buddhists claim, deceptive.

Three phenomena will help explain the opacity thesis and its counter-intuitive take on self-consciousness: infant amnesia, absentmindedness, and blindsight. The first of these should be puzzling if we take the transparency thesis at face value: if our own conscious states are transparently given to us, and experiences are conscious states, why do we lack episodic memory for experiences had prior to about age three? Other sorts of long-term memory, e.g., that involved in aversive conditioning, seem to function properly below that age. The suggestion is that

[8] The selectionist explanation of such phenomena helps answer the objection to theory-of-mind accounts of cognition of others' mental states raised by Krueger and Overgaard 2012, 248. In support of their claim that we are able to perceive the mental states of others (and not merely cognize them by rapidly performing an automatic inference), they point to the manifest difference between looking at a picture of a smiling face and looking at the same picture inverted. The selectionist explanation of the infant's differential attention to even the most schematic depictions of faces helps us see why the same visual processing cannot be expected to function when the image is inverted.

full-blown autobiographical memory requires possession of a particular kind of self-concept, and this is something that is only fully developed in the child around age five.[9] Of course we can say that infants do distinguish between self and other, and so do have some sort of a sense of self; the mastery of motor control in very early childhood depends on this. Goal-directed motor activity involves a feed-forward mechanism: the anticipated sensory input consequent on success is represented in neural form at the time of initiation of activity (the infant forms a neural image of the sensory input that would be present were the nipple in the mouth). This representation is then canceled on success; non-cancellation signals the need for further refinement of the motion already undertaken.[10] The mechanism involved here clearly depends on the distinction between the feed-forward representation generated at the time of initiation and the feedback representation obtained upon completion of the action. And we might call this a distinction between self and other. So we might think that the infant does, after all, possess a rudimentary sense of self. But then we should have to attribute the same sense of self to the motion detector that turns on the outdoor lights when a raccoon enters the backyard. For precisely the same sort of feed-forward mechanism is involved there as well. An inverted form of the representation of the detector's radar readout at one moment is superimposed on the uninverted readout obtained at the next moment. If these do not cancel one another out, the lights are turned on; if they do, then nothing happens. It would clearly be a mistake to attribute a sense of self to the motion detector. This despite the fact that the mechanism depends on distinguishing between an internal state of the detector (its 'memory' of the preceding moment) and its representation of the current state of the external world.

The phenomenon of absentmindedness poses an equal challenge to the transparency thesis. Take the common instance of resolving to stop at the dry cleaner's on the drive home, only to find oneself pulling into the driveway without the cleaning and equally without any recollection of having driven straight home. One was presumably aware of the road, other cars, traffic signals and the like during the drive, and yet none of those experiences seems to have registered. Some of the oddity of the phenomenon can be dispelled if we take self-knowledge to be the result not of introspection but of an abductive inference. We are sure we were aware of the road, other cars, etc., during the drive home, despite our not recalling any of these

[9] For some of the details of development of episodic memory in childhood see Busby Grant and Suddendorf 2005.

[10] For details see Jeannerod 2006, 23–44.

experiences, because this best explains our having arrived home intact. And we take this to best explain the present state of affairs because we have seen what can happen when one performs an attention-demanding task like texting while driving. Still we will wonder how we might have experiences that we cannot recall immediately afterward, if the mind is indeed transparent to itself.

The case of absentmindedness involves performances that we take to involve consciousness. The case of blindsight involves doing things that we think require conscious awareness, but in fact cannot; the lack of awareness cannot be blamed on mere lapses in memory. The blindsight patient is someone whose visual organs are intact and functioning but who lacks conscious awareness of anything visual. The condition is typically brought about by injury to a part of the brain involved in visual processing. What is remarkable about blindsight patients is that they are nevertheless able to perform actions that require visual input, such as walking down a corridor strewn with obstacles. Indeed, after weaving around a piles of books and the like, the blindsight patient will report that they walked straight down the hall. Dual systems theory can be invoked to explain the phenomenon. What is presumably missing in the blindsight patient is System 2 (ventral) visual processing, which makes vision input globally available and thus available for speech processing: it is what enables one to report on what one sees. The patient's System 1 (dorsal) visual system is still intact, though. System 1 processes take afferent input directly to efferent output without routing through the global workspace. This may show that they do not involve consciousness of sensory stimulation. The flinch response is a clear-cut example: visual input as of some object coming at one's head triggers ducking before one is aware of seeing the object. Because the routing goes directly from sensory input to efferent output, it is faster than processing that went via the global workspace and thus resulted in possible registry as a conscious experience. The value to the organism of such a mechanism is obvious. It is, though, relatively inflexible and so prone to error: we duck in 3-D movies too.

What the case of blindsight brings out is the point that consciousness may best be thought of as the property a mental state has when it makes its information content available to other processing modules through presentation in the global workspace. This view of consciousness as global availability can likewise help explain the phenomenon of absentmindedness. Tasks that are routinized, such as driving the same route one has taken every day for years, no longer require System 2 resources, so that one's sensory and motor representations are not globally broadcast. If episodic memory is only of conscious experiences, it is then no mystery that one

does not recall the experience of driving straight home—since there was no such conscious experience. This view does, though, make being conscious an extrinsic property of mental states, one possessed in dependence on the state's relations to other entities. This in turn bolsters the claim that self-knowledge—cognition of one's own conscious mental states—comes about through an abductive inference from action and speech.

Dretske makes the following remarks concerning the consequences of this view of consciousness for the development of self-knowledge:

> [H]ow do I know I have a mind? If introspection tells me only what I think and feel, not that I think and feel, how do I discover that I think and feel, that I'm not a zombie? I am tempted to reply, I learned this the same way I found out a lot of other things—from my mother. She told me. I told her what I thought and experienced, but she told me that I thought and experienced these things….Three-year-olds know, and they are able to tell you, authoritatively, what they think and see (e.g., that there are cookies in the jar, that Daddy is home, etc.), before they know, before they even understand, that this is something they think and see. Somehow they learn they can preface expressions of what they think (Daddy is home) with the words "I think," words that (somewhat magically) shelter them from certain forms of correction or criticism. Parents may not actually tell their children that they think—for the children wouldn't understand them if they did—but they do teach them things (language must be one of them) that, in the end, tell them they think. Children are, at the age of two or three, experts on what they think and feel. They have to learn—if not from their mothers, then from somebody else—that they think and feel these things. Nonhuman animals never learn these things. (Dretske 2003, 140–41)

Bogdan would agree. There is, Bogdan holds, a crucial step in the transition from reliably reporting on one's beliefs and other attitudes, to representing oneself as having beliefs and other attitudes. The latter requires mastery of reflexive meta-representation: representing what are one's representations *as* representations.[11] It is not clear how the child can accomplish this without using resources only available through linguistic abilities. Dretske may be right that this is something the child learns to do through learning certain locutions. But this opens

[11] For a useful discussion of meta-representation see Dretske 1997, 43–55.

up the possibility that the self-ascription is not reliably reporting on something that was already there before the child mastered the locution. Bogdan says this about the propositions that serve as contents of beliefs and other attitudes that are self-ascribed: 'as contents ascribed to attitudes, propositions are *constructs* of commonsense psychology. In other words, propositions are how the contents of attitudes *look* from the vantage point of commonsense psychology' (Bogdan 2010, 134). The same should go for the attitudes that are said to have propositions as content.

The evidence Dretske cited does not, though, fully support his claim that children of two or three are 'experts on what they think and feel'. At least they are not reliable reporters of their own past beliefs—when those beliefs have undergone change. As Bogdan points out, when three-year-old Child_1, has witnessed another child, Child_2, first observe a toy being put inside one of two boxes, then leave the room, after which Child_1 sees the toy moved to the other box, Child_1 will say that when Child_2 returns to the room they will believe the toy is in the second box. And it is only when the child has mastered the task they have failed here, attributing false beliefs to others, that they then become able to self-attribute past false beliefs under similar circumstances. The order of mastery is important. It is the child's mastery of a theory that attributes beliefs and other attitudes to others that comes first; only afterwards does the child learn to apply that theory to themselves. This is the essential asymmetry of the developmental formulation of the opacity thesis.

This asymmetry is important to the account of self-knowledge that is compatible with Buddhist non-self. What the developmental record reveals is that the infant begins with a set of other-directed reflexes, which serve as a foundation for the emergence of a mind-reading faculty that, when enriched by conceptual resources made available through language, enables the child to explicitly attribute intentional states to others. And only then is the child able to turn this proto-theory on themselves and self-attribute. This is how the capacity for meta-representation, crucial for the child's 'knowing *that* they think and feel', emerges. Employment of this folk psychology then endows the child with capacities for self-governance and self-revision that over time foster the development of a narrative self. Use of this theory thus comes to be second nature to us. And this makes us forget that the conscious mental states we attribute to ourselves and to others are the posits of a useful theory, not things with which we are directly acquainted.

We are sure we are not zombies. But what is the evidence that assures us of this? The developmental story we have been discussing claims we rely on a theory

designed, in the first instance, to help us meet our needs by helping us explain and predict the actions of others. When we apply that theory to ourselves, we deploy a model that generates precisely the intuitions that make zombiehood inconceivable. First-person reporting of an experience or an attitude is, the model tells us, possible precisely because experiences and attitudes are real denizens of an inner realm and objects of direct acquaintance by the master of that realm. We know that false theories are sometimes useful. Celestial navigation, which is fairly reliable, deploys the model of Ptolemaic astronomy. Still we tend to find it inconceivable that folk psychology could be no more than a useful model. Why is this, though? Bogdan suggests it is because we are being asked 'to *conceive*, from felt experiences, of phenomenal consciousness as fully reducible, without residue, to physical matter and its functional arrangements. Since the challenge cannot be met, the conclusion is that phenomenal consciousness cannot be explained in such terms, which are the terms of science' (Bogdan 2010, 170). But of course, if we use the model to test the hypothesis that we might be zombies, the outcome is guaranteed. This is hardly a fair test.

Here is one last piece of evidence for the possibility that folk psychology is a useful but nonetheless deceptive theory. When infants interact with their adult caregivers, we often treat the infant's behavior as indicating intention. And we talk to infants, and take their babbling and their later attempts at verbal imitation, as efforts to communicate. It is difficult to resist viewing the infant as having a rich subjective realm that is still hidden to us but transparent to the infant. But it is worth considering the following possibility: that our intuitions here are just the ones we would expect if the process of turning infants into persons employed the powerful mechanism of the self-fulfilling prophecy. We do tend to become what we are taken by others to be. We know, for instance, that acquisition of its first language requires that the child be spoken to. What better way to facilitate the process than to inculcate the belief that the child already possesses a rich inner life and simply lacks the means to communicate it?

4.

There are two possible ways for Buddhists to account for meta-cognition in the absence of a self: claim that cognitions are reflexive, and that the noetic and

noematic poles of a cognition are somehow strictly identical;[12] or claim that such self-knowledge is the result of an abductive inference rendered rapid and automatic due to our routine use of the model of the mind captured in folk psychology. Those who champion the second option must somehow face down withering stares of incredulity. But the first has its own challenges. There is no non-controversial counter-example to the principle of irreflexivity; and it must somehow be explained how the distinction between a cognition's noematic and noetic poles could be illusory.[13] These are formidable hurdles, and the Buddhists who propounded the first solution to the problem of self-knowledge struggled to overcome them. Less effort was expended on trying to make the second option more plausible. I would suggest, though, that recent work in developmental psychology and philosophy of mind shows how one could begin to counter the incredulous stare. Buddhists might want to reconsider their opposition to a naturalistic project.

References

Bogdan, Radu (2010) *Our Own Minds: sociocultural grounds for self-consciousness*. (Cambridge, MA: MIT Press).

Busby Grant, Janie & Suddendorf, Thomas (2005) Recalling yesterday and predicting tomorrow. *Cognitive Development*, 20.2, 362–372.

Carruthers, Peter (2011) *The Opacity of Mind*. (Oxford: Oxford University Press).

Dhammajoti, K.L (2007) *Abhidharma Doctrines and Controversies on Perception*. (Hong Kong: The Buddha-Dharma Centre of Hong Kong).

Dretske, Fred (1997) *Naturalizing the Mind*. (Cambridge, MA: MIT Press).

Dretske, Fred (2003) Externalism and Self-Knowledge, in Susan Nuccetelli (ed.), *New Essays on Semantic Externalism and Self-Knowledge*, (Cambridge, MA: MIT Press), 131–42.

[12] The identity of noetic and noematic poles, of 'grasped' (*grāhya*) and 'grasper' (*grāhaka*), is demanded by the claim that cognitions are ultimately real and irreducible, plus Buddhism's mereological nihilism. For an interesting recent attempt to defend this thesis see Strawson 2008.

[13] MacKenzie tries to solve the second problem by claiming that while cognitions are not mereologically complex, they are 'aspectually complex' (2017, 351). The difficulty with this is that, as Kumārila pointed out long ago, it is not clear how something that is in no sense composite can give rise to the attribution of different aspects. Dharmakīrti's example of the man who may equally be thought of by the boy as 'teacher' and as 'father' depends on the man's carrying out distinct functions at different times.

Dreyfus, Georges (2011) Self and Subjectivity: A Middle Way Approach, in Mark Siderits, Dan Zahavi and Evan Thompson (eds.), *Self, No Self? Perspectives from Analytical, Phenomenological and Indian Traditions*, (Oxford: Oxford University Press), 114–55.

Jeannerod, Marc (2006) *Motor Cognition: what actions tell the self*. (Oxford: Oxford University Press).

Krueger, Joel and Overgaard Søren (2012) Seeing subjectivity: Defending a perceptual account of other minds. *ProtoSociology: Consciousness and Subjectivity*, 47, 239–262.

MacKenzie, Matt (2017) Luminous Mind: Self-Luminosity and Other-Luminosity in Indian Philosophy of Mind, in Joerg Tuske (ed.), *Indian Epistemology and Metaphysics*, (London: Blooomsbury Academic), 335–53.

Siderits, Mark (2015) *Personal Identity and Buddhist Philosophy: Empty Persons*, 2nd edition. (Farnham, Surrey: Ashgate).

Strawson, Galen (2008) What is the Relation Between an Experience, the Subject of the Experience, and the Content of the Experience?, in *Real Materialism and Other Essays,* (Oxford: Oxford University Press), 151–87.

Beyond Monism and Pluralism
On Dōgen's *Pan-Self-Ist* Turn of Metaphysics

MITANI Naozumi
Associate Professor, Shinshu University

Abstract: *This paper aims to take up the philosophy of Dōgen and delineate it as a contemporary piece of philosophy that merits our serious attention. To be more specific, this paper adopts Jonathan Schaffer's path-breaking work on philosophical monism as a platform to illuminate Dōgen's metaphysical positioning. As we shall see, Dōgen will be depicted as a "pan-self-ist metaphysician" who annihilates the whole debate between monism and pluralism.*

Schaffer proposes that we try a new interpretation of monistic philosophy, which he calls "priority monism". Schaffer insists "priority" rather than "existence" should be the topic of ontological debate between monism and pluralism.

According to Schaffer's argument, existence monism is said to be the "world only view", which claims "exactly one thing exists". However, as Schaffer says, "monism would deserve to be dismissed as obviously false, given this interpretation".

In contrast, priority monism, which Schaffer says is more defendable, claims that "[t]he world has parts, but the parts are dependent fragments of an integrated whole". That is, priority monism admits existence pluralism with a caveat that the whole is prior to its parts. So interpreted, "[t]he historical debate is not a debate over which objects exist, but rather a debate over which objects are fundamental".

It is in view of this proposal that I will depict Dōgen as a metaphysician of a new kind. Schaffer's "priority turn" of ontology leads us to one philosophical possibility that Schaffer dimly mentions but does not delve into any detail. That is, pace Schaffer's strenuous effort to revitalize the monistic tradition, the upshot of his priority turn might be an annihilation of the whole debate between monism and pluralism. Seen from Dōgen's perspective, not only at the level of existence but also at the level of priority, the monism/pluralism debate is to be annihilated.

Introduction

This paper aims to take up the philosophy of Dōgen and delineate it as a contemporary piece of philosophy that merits our serious attention. To be more specific, a recent resurgence of monistic philosophy composes the background for this paper. Above all, this paper adopts Jonathan Schaffer's path-breaking work as a platform to illuminate Dōgen's metaphysical positioning. As we shall see, Dōgen will be depicted as a "*pan-self-ist metaphysician" who annihilates the whole debate between monism and pluralism*.

Schaffer proposes that we try a new interpretation of monistic philosophy, to which he gives a name of "priority monism". Schaffer insists "priority" rather than "existence" should be the topic of ontological debate between monism and pluralism.

According to Schaffer's argument, existence monism is said to be the "world only view", which claims "that exactly one thing exists. . . . On such a view there are no particles, pebbles, planets, or any other parts to the world. There is only the One. Perhaps monism would deserve to be dismissed as obviously false, given this interpretation".[1]

In contrast, priority monism, which Schaffer says is more defendable, does not insist that the whole has no parts. Rather, it claims that "[t]he world has parts, but the parts are dependent fragments of an integrated whole". That is, *priority monism admits existence pluralism* with a caveat that the whole is prior to its parts. In this sense, it annihilates the monism/pluralism controversy at the level of existence. So interpreted, "[t]he historical debate is not a debate over which objects exist, but rather a debate over which objects are fundamental".[2]

It is in view of this proposal offered by Schaffer that I will depict Dōgen as a metaphysician of a new kind. Schaffer's "priority turn" of ontology leads us to one philosophical possibility that Schaffer dimly mentions but does not delve into any detail.[3] That is, pace Schaffer's strenuous effort to revitalize the monistic tradition, the upshot of his *priority turn* might be an annihilation of the *whole* debate between monism and pluralism. Seen from Dōgen's perspective, not only at the level of *existence* but also at the level of *priority*, the monism/pluralism debate is to be annihilated.

[1] Schaffer 2010, 32
[2] Schaffer 2010, 33
[3] Schaffer 2010, 35; 45

Dōgen on the Whole and its Parts

As a starter, let us take up the phrase of "*Issai-Shujō Sitsu-U Busshō*" (一切衆生悉有佛性) and explore how this illuminates Dōgen's take on "the one as the whole" and "the many as its parts". In the fascicle of *Busshō* (*buddha-nature*), Dōgen says this:

> [In Buddhism, traditionally, when] one speaks of "*Shujō* (衆生)", [the term is used as] "sentient beings (*Ujō* 有情)". . . . [However, in Dōgen's view,] saying of "the entirety of being (*Shitsu-U* 悉有)" refers to "*Shujō*". . . .[4]

In this citation, Dōgen is talking about "*Shujō*" as a component of "*Issai-Shujō*". "*Shujō*" is an old translation of the original Sanskrit word "*Sattva*", which was replaced by a new translation of "*Ujō*", the literal meaning of which is "that which possesses mind". Following this, in the former part of the citation, Dōgen is talking about "*Shujō*" or "*Sattva*" as "minded/sentient creatures".

However, the latter half of the citation defies this gloss. As is evident from Dōgen's equation of "*Shujō*" with "the entirety of being" (*Shitsu-U*), Dōgen suggests that we take the word "*Shujō*" to have a broader meaning than the traditional understanding of it. He replaces the meaning of "*Shujō*" as "sentient beings" with that of "all beings that exist".

Having offered his interpretation of *Shujō* and *Shitsu-U*, Dōgen proceeds to the next component of "*Issai-Shujō Sitsu-U Busshō*", i.e., the term of "*Busshō*". As he says,

> In a word, the entirety of being (*Shitsu-U*) is the buddha-nature (*Busshō*); One entirety of the entirety of being (*Shitsu-U no I-Shitsu* 悉有の一悉) is called *Shujō*.[5]

This is a perplexing remark. First, Dōgen (quite abruptly) declares that "the entirety of being is the buddha-nature (*Busshō*)", but how are we to understand this crucial vocabulary of "*Busshō*"? Dōgen does not give us any clue about this. Second, Dōgen further proclaims that "one entirety of the entirety of being" is called "*Shujō*" but how are we to make of this?

---

[4] Dōgen 1990, vol. 1, 73
[5] Dōgen 1990, vol. 1, 73

According to Dōgen's *un*traditional interpretation, "*Shujō*" was equated with "the entirety of being". That is, "*Shujō*" à la Dōgen is a notion that refers to the whole as a group of individual beings. To our concern, "*Shujō*" is a word that can be paraphrased as "the one as the whole".

However, at the same time, Dōgen observes that "one entirety of the entirety of being (*Shitsu-U no I-Shitsu*) is called *Shujō*". In this case, the word "*I-Shitsu*" can be decomposed into "*Ichi*" and "*Shitsu*", meaning "one" and "all" respectively. In short, "*Shitsu-U no I-Shitsu*", the translation of which is "one entirety of the entirety of being", can be paraphrased as "one of the individual members of the whole taken in its entirety". That is, "*Shujō*," when it appears as "*Shitsu-U no I-Shitsu*", refers to "one of the many that compose the whole".[6]

Putting these together, the notion of *Shujō* is given a double meaning; "the one as the whole" and "one individual part of the whole". This might be a welcoming move for those who seek a reconciliatory approach to the monism/pluralism debate. For, this seems to suggest that being the whole and being a part of it might turn out to converge under the umbrella concept of *Shujō*. However, this *surely* is a perplexing remark as well. Our task is to put forward an intelligible exposition of this enigmatic worldview proposed by Dōgen.

Inmo **(Such)**

At one point in the fascicle of *Busshō*, where Dōgen gives a further comment on "*Issai-Shujō Sitsu-U Busshō*", he mentions that "the essential point" of this World-Honored One's saying can be expressed by another Ancestor's saying of "*Ze Shimo Butsu Inmo Rai* (是什麼物恁麼來; what is it that comes like this?)".[7]

This phrase, "*Ze Shimo Butsu Inmo Rai*", is one of Dōgen's favorite phrases and makes recurring appearances in the entire book of *Shōbōgenzō*. Thus, in the fascicle of *Henzan*, Dōgen tells us one anecdotal dialogue between Nangaku and Enō, in which the essential point of "*Ze Shimo Butsu Inmo Rai* (what is it that comes thus?)" is illuminated.

The story, which is a famous one in itself, is simple. When Nangaku visited Enō, Enō asked him, "What is it that comes thus?", and to this Nangaku replied, "Were I to try to put [this thing that comes thus] into words, they would miss the

[6] On this reading of "*Shitsu-U no I-Shitsu*", See Nishitani 1991, 130f.
[7] Dōgen 1990, vol. 1, 73

mark (*Setsu Ji Ichi Motsu Soku Hu Chu* 説似一物即不中)". To repeat, Dōgen's suggestion is that this dialogue between Nangaku and Enō leads us to "the essential point" of "*Issai-Shujō Sitsu-U Busshō*".

Let us proceed by focusing on the notion of *Inmo* as appears in "*Ze Shimo Butsu Inmo Rai*". As a glossarial note, it can be pointed out that the word "*Inmo*" has a double meaning. On the one hand, it means "this" or "this such". On the other, it means "what" or "what sort of". Given this double meaning of *Inmo*, the entire phrase of "*Ze Shimo Butsu Inmo Rai*" implies both "What is it that comes like this?" and "Such a thing comes here in this way" at one time.[8]

Keeping this double meaning of *Inmo* in mind, what needs to be emphasized is that *Inmo* is categorially blank as to what it is. All that can be said about *Inmo* is that "something comes somehow" and it always comes with the query—"what is it that comes like this?" In this sense, the something-ness or somehow-ness of *Inmo* is always elusive or ineffable. In this way, *Inmo* always follows Nangaku's saying of "Were I to try to put this such into words, they would miss the mark". For, *Inmo* does not say what the essence of the referent is.

In the fascicle of *Inmo*, Dōgen gives a rather straightforward expression of this elusiveness or the unfathamability of *Inmo*, saying:

> Such [*Inmo*] can't be got; not such can't be got; both such and not such can't be got. What will you do?. . . We should study "such" in "can't be got"; we should question "can't be got" in "such".[9]

In addition, also in the fascicle of *Inmo*, Dōgen makes the following comment.

> Daikan Enō once addressed Diae of Nangaku, saying, "What sort of thing is it that comes in such a way?" These words mean that "such" is beyond doubt, because it is beyond understanding (恁麼はこれ不疑なり、不會なるがゆゑに). . . . "What sort of thing" is not subject to doubt; it is coming of such a thing (什麼物は疑著にはあらざるなり、恁麼來なり).[10]

[8] The word "*Shimo*" (什麼) that appears in "*Ze Shimo Butsu Inmo Rai*", though different in notation, has the same meaning and usage as "*Inmo*" (恁麼).
[9] Dōgen 1990, vol. 1, 416f.
[10] Dōgen 1990, vol. 1, 417f.

Dōgen insists that *Inmo* is not to be doubted because it is unfathomable. This phrase, without doubt, is enigmatic and needs glossing. Dōgen tersely affirms that *Inmo* is undoubtable because it is unfathomable but he says nothing about how this can be the case. We need further clarification as to how the undoubtable because unfathomable character of *Inmo* holds.

Three comments before proceeding:

1) It should be emphasized that the unfathomability of *Inmo* does not lead us to skepticism. The unfathomability of *Inmo* does not mean that it is unacceptable or dubious. On the contrary, Dōgen suggests that the unfathomable character of *Inmo* should be considered vis-à-vis the undoubtability of it. That said, one possible approach to the *undoubtable because unfathomable* character of *Inmo* will be to get the undoubtability aspect into sharper focus and explore whether it might shed some light on the unfathomable side of it.

2) It is worth restating that *Inmo* is a word that has a "double meaning". As we saw, *Inmo* means "this" or "this such" and "what" or "what sort of" at one time. Otherwise put, when you raise a query of *Inmo* (asking "what?") you are, in a sense, already answering it. Take the case of "*Ze Shimo Butsu Inmo Rai*" as an example, the interrogative utterance of "what is it that comes like this?" always comes as a package with an affirmative version of it, "Such a thing comes in such a way".

3) It should be a matter of evidence that "what sort of thing?" cannot be asked unless something is given somehow. *That is, the coming of such a thing must precede the question of "what sort of thing?"* This, I assume, is the sense in which Dōgen says "what sort of thing" is not subject to doubt. When you say "what sort of thing?", it cannot be an expression of your doubt. For it always comes with an affirming experience of the coming of such a thing. This "something-somehow-ness" of *Inmo*, as we shall see, when coupled with Dōgen's *ineffabilist understanding of "self"*, leads us to the identification of *Inmo* with the *ultimate truth of Buddha* that is "beyond predication".[11]

Inmo as the Ultimate Reality of Buddha

To repeat, Dōgen's reference to *Inmo* is made in close connection with the ineffabilist understanding of "*Ze Shimo Butsu Inmo Rai*". That is, if we try to say anything further than "something comes somehow", it always misses the mark. In a

[11] I owe this clarificatory point to a comment by an anonymous referee for the journal.

word, the prominent characteristic of *Inmo* can be found in its ineffability as to what it is.

So, as long as *Inmo* is concerned, any further query beyond "something, somehow" is blocked. *Inmo* refers to that which is just what it is; that which is beyond further predication.[12] In this sense, *Inmo* is not only undoubtable but also infallible. (You can't err as to what the target of your talk is when you give no predicative specification of your referent.) It should be this "infallible because ineffable" aspect of *Inmo* that Dōgen had in mind when he says the following:

> . . . if you wish to get such a thing, you should be such a person. Since you are such a person, why worry about such a thing? The point here is that, for the time being, he is speaking of "proceeding directly to supreme bodhi" as "such" (この宗旨は、直趣無上菩提、しばらくこれを恁麼といふ).[13]

In the first two sentences, Dōgen repeats essentially the same point about *Inmo* that we reviewed above; the always-and-already-got-ness of *Inmo*. *Inmo* is not to be doubted as *Inmo* is always the coming of such a thing. Slightly adjusted to the case in hand, you do not have to worry about *Inmo* as you are always and already such a person and thus *Inmo* is always manifested in you.

What's remarkable is the move that Dōgen makes in the third sentence. Dōgen, getting beyond the undoubtability, goes so far as to claim that when you speak of *Inmo*, you are speaking of "proceeding directly to supreme bodhi". Dōgen is now making a much stronger claim than the assertion of undoutability, declaring that *Inmo* be regarded as a direct expression of the ultimate truth of Buddha.

Nyoze and Existing in Suchness

Let us proceed by making a further exploration into *Inmo*. Again, as a glossarial note, it can be pointed out that *Inmo* is a word that colloquially/informally represents one traditional term of Buddhism: *Nyoze* (如是). *Nyoze* is a word that is conventionally translated as "such"; "suchness" or "thus"; "thusness".

[12] On this point, see Bielefeldt 2017, 30
[13] Dōgen 1990, vol. 1, 402

For example, D. T. Suzuki gives his comment on this "suchness" of *Nyoze* as the following: "*Aru ga mama no aru* means to "exist in suchness." Plants and trees exist in suchness, as do dogs and cats and mountains and rivers and streams".[14]

According to Suzuki, *Nyoze* (as a formal counterpart of *Inmo*) is a word that might as well be translated as "being (existing) as it is by its own nature". That is, the suchness of *Nyoze* is a notion that can be paraphrased as "as-it-is-ness" ("*aru ga mama*").[15]

On the part of Dōgen, this suchness of *Nyoze* is discussed in the fascicle of *Shohō-Jissō*, where he says this:

> The presencings of all the buddhas and ancestors are [the presencings of] real forms in its ultimacy. The real forms are the myriad of things. The myriad of things is their form just as it is (*Nyoze-Sō*), their true nature just as it is (*Nyoze-Shō*), their body just as it is (*Nyoze-Shin*), their mind just as it is (*Nyoze-Shin*) (佛祖の現成は究盡の實相なり。實相は諸法なり。諸法は如是相なり、如是性なり。如是身なり、如是心なり)[16]. . .

This is a dark passage. But with the help of the glossarial note above, it should be allowed to read this passage as "every *Nyoze* as "existing in suchness" is an expression of Buddha's truth". To put it less bluntly, when we speak of the myriad of things in this world just as they are, we are speaking of those things in their real forms or their ultimate reality. That is, everything in the world, when they exist just as they are, i.e., when they exist in suchness, are already manifestations of the ultimate truth of Buddha.

One passage from *Shohō-Jissō*, where Dōgen refers to the idea of "*Honmatsu Kukyō Tō*" (本末究竟等), can be cited as textual support for this reading:

> The Tathagata says "all things, from top to bottom, from beginning to end, are ultimate reality themselves, and this applies equally to all things" (*Honmatsu Kukyō Tō*). This saying of the Tathagata is a self-expression of *Shohō-Jissō* in itself.[17]

---

[14] Suzuki 2016, 111
[15] Suzuki 1997, 219
[16] Dōgen 1990, vol. 2, 432
[17] Dōgen 1990, vol. 2, 432f.

What Dōgen aims to convey should be obvious: he is insisting that all things are manifestations of ultimate reality in themselves and are all equal on this count. To quote yet another expression, "when all things are nothing but those things, we call them the buddhas. When all things, just as they are, are manifestations of their real forms, they are what we call the buddhas".[18]

So the core thought found in Dōgen's view on the undoubtable character of *Inmo* can be expressed as "all things, when they exist in suchness (*Nyoze* or *Inmo*) are equally manifestations of ultimate reality". It is in this sense that Dōgen equates the Tathagata's saying of "every entity is equal in their ultimate reality" as a self-expression of *Shohō-Jissō*. *Inmo*, as a colloquial counterpart of *Nyoze*, is a word that verifies the ultimate truth of Buddha: all things, just as they are, are manifestations of ultimate reality themselves.

Being Such and Being a Self

To move on, let us introduce the notion of "self" as a helping guide to approach the undoubtable because unfathomable trait of *Inmo*.

First, it needs to be noticed that "everything exists as itself" is a primitive, self-evident affair. For it is an undeniable truth that "everything is equal to itself". Otherwise put, it is undoubtable that "nothing can differ from itself". If something differs from itself, that something should be something other than the original self. Or it will be a new self that is distinct from the original self.

Second, it should be noticed that this primitive, undoubtable character of being a self can be taken as an expression of another aspect of being a self: the ineffability of X's being a self. The background assumption for this is simple: X's being a self is a fact or a state that can never be an objective target of theoretical reflection. For example, even as we try to describe the ultimate reality of a flower's being that flower, we simply have to admit that the ultimate truth of that flower's being itself is something that cannot be verified from sideways-on.

"X's being X" is realized only through the existential state of "X's being itself". That is, an adequate account of "X's being itself" can only be given through the primitive fact of "X's being itself". To give it a Dōgenian expression, *aru ga mama* is just *aru ga mama* (*Inmo* is *Inmo*), and we can't say otherwise.

[18] Dōgen 1990, vol. 2, 433

Now, the idea to be joined to this is that *Inmo* is expressive of the truth of something's existing as self. That is, being just as they are is expressive of the fact that X exists as X.

To review the argument offered by Dōgen, it is claimed that *Inmo* is an expression of Buddha's truth in that it is expressive of the state in which the entirety of being exists as what they are. All things, just as they are, are already manifestations of ultimate reality. This, conjoined with the idea that *Inmo* is expressive of the truth of a self's being itself, should imply that the undoubtability of *Inmo* can be illuminated by the undoubtable character of being a self. For it can be argued that X's being such (*Inmo*) is undoubtable because it is expressive of the fact that X exists as itself.

With these considerations in view, we now can see what the undoubtable character of *Inmo* amounts to. As long as the undoubtability of being a self is reflective of the ineffability of being itself, talking of the ineffability of being a self is at one time talking of the ineffability of *Inmo*. What follows from this should be clear. *Inmo*'s undoubtability is explained in terms of the ineffability of being a self *qua* the ineffability of *Inmo*. This should be the sense in which Dōgen says *Inmo* is undoubtable because it is unfathomable.

Dōgen's Pan-Self-Ist Turn of Metaphysics

To wrap up, the central tenet of Dōgen can be condensed to the thesis that all beings, just as they are, are manifestations of ultimate reality. Also, this means that the ultimate reality of things tells us that the entirety of being exists as self. In Dōgen's voice, "[a]rraying the selves in order and we get the whole universe" (われを排列しおきて尽界とせり).[19]

Now, from a meta-metaphysical point of view, it should be allowed to characterize the ontological commitment of Dōgen as a move from *existence* to *self*. That is, the teachings of Dōgen can be interpreted as urging us to set the notion of *self*, not *existence*, as the target unit of ontology. Dōgen suggests we transpose the question of ontology from "what is it that exists?" to "how are we to understand the ultimate reality of everything's existing as self?". The following remark of Dōgen can be read along these lines:

[19] Dōgen 1990, vol. 2, 47

> We should realize that the "being" that is here made the "entirety of being" by the buddha-nature is not the being of being and non-being. (仏性に悉有せらるる有は有無の有にあらず)[20]

The phrase "the "being" that is made the entirety of being by the buddha-nature" is a difficult one. However, it should be permissible to read this as "when the entirety of being verifies the ultimate truth of buddha nature"; i.e. "when the entirety of being verifies the ultimate reality".

Granting this to be the case, the entire sentence should be read as "when talking about the ultimate reality, the being or non-being is not a target question to be asked". That is, the target of our ontological inquiry should not be the question of *existence* that deals with the being or non-being. Instead, Dōgen asserts that we shift the focus of attention from the concept of existence to that of self.

To our concern, this move of Dōgen will have grave consequences. For as the previous considerations have shown, the essential point about the entirety of being's existing as self is that the fundamental structure of self forbids us to take them as a target of theoretical reflection. That said, what is important is to get clear about what follows from this: the ineffable character of self deprives our metaphysical inquiry of its qualification to take the *number of the beings* as its target question.

Given that the selfhood of X (X's being itself) cannot be a target of *numerical* reflection, the question that asks whether that which exists should be counted as *one* or *many* no longer holds. In other words, as long as the ultimate reality of things should neither be regarded as the One nor the many, Dōgen's take on the ultimate reality of things annihilates the debate between *existence* monism and pluralism. As Dōgen impressively notes:

> Hearing the phrase of "the myriad of things", we should not explore it as being the One, nor should we explore it as being the many[21].

Two things are worth mentioning. First, as long as the notion of "existence" is concerned, Dōgen will be depicted as a metaphysician of rather a queer kind: 1) He is not an existence monist, for he denies that that which exists is one. 2) He is not an existence pluralist, for he denies that that which exists is many. 3) He is not an

[20] Dōgen 1990, vol. 1, 73
[21] Dōgen 1990, vol. 2, 433

existence nihilist, for he is committed to the existence of self as the ultimate reality of the world.

In a word, none of the traditional categorizations of ontology fits the positioning of Dōgen. Considering his insistence that self is the topic of metaphysics although it is something whose being or non-being is not a matter of philosophical inquiry, he might as well be called a *pan-self-ist* at the level of the existence talk.

Second, Dōgen's annihilation of the monism/pluralism debate might raise the following concern. To review the argument that Dōgen makes about the notion of "*Issai Shujō Shitsu-U Busshō*", he was committed to the view that "*Shujō*" is "the one as the whole" and "one of the individual parts of the whole" at one time.

However, given the annihilation thesis mentioned above, this does make a queer situation. On the one hand, when discussing the phrase of "*Issai Shujō Shitsu-U Busshō*", he says that *Shujō* refers both to "the one as the whole" and "one individual part of the whole". On the other, he also says that "the myriad of things" (those which exist in this world) should not be explored "as being the One" nor "as being the many".

The One and the Many

To proceed, let us introduce two concepts both of which play a central role in the Buddhists' understanding of self; undeterminedness (無自性) and emptiness (空).[22] Put in advance, the idea is that the notion of self, when given a deeper examination in terms of undeterminedness and emptiness, will shed new light on the issue of the One and the many.

As a first step, let us consider the teaching of "one water, four views (一水四見)", in which Dōgen talks about an idea that the thing we call water appears in various ways according to the type of being that sees it. As Dōgen says, "[i]n seeing water, there are beings who see it as a jeweled necklace, as miraculous flowers, as raging flames or as pus and blood".[23]

By referring to these examples, Dōgen makes us see the process by which entities with determinate character emerge from a world of emptiness where there is no determinate essence that makes each entity the entity that it is. Otherwise put, Dōgen is trying to give his answer to the question: "what makes an object token the

[22] For an excellent review of these core Buddhist ideas, see Garfield 2015, chap 2 and 3.
[23] Dōgen 1990, vol. 2, 192f.

object token that it is?" For example, "what is it that makes water what it is (i.e., water)?"

Dōgen's answer to this question is that it is our practical confrontation with the things that makes each being the particular being that it is. Water's being water (and not a jeweled necklace) is dependent on human's seeing it as water. As Dōgen says, water is "given life" when humans see it as water. Conversely, water is "killed" when different types of being see it as something other than water such as a jeweled necklace or miraculous flowers.[24]

What has to be noticed is that Dōgen's take on "one water, four views" cannot hold if water has its self-determined essence that makes water what it is. For there's no denying that what is essentially water cannot be something other than water. To put it in the other way around, what we call water can become water (or anything other than that) just because it carries no determinate essence in itself. That is, the thing that we regard as water can be a necklace just because the thing that we regard as water is empty in itself.

However, at the same time, it has to be noticed that Dōgen's rejection of the self-determined character of things does not commit him to any sort of metaphysical nihilism. True, "X's being X" is not sustained or realized by independent, essential selfhood of X, but this does not mean that X is nothing in itself. Dōgen admits that X, though empty in itself, emerges and exists as X that is distinct from other elements of the world, say, Y or Z.

In the fascicle of *Zenki*, Dōgen explains this moment of emergence from the world of emptiness in an illuminating way:

> Life can be likened to a time when a person is sailing in a boat. On this boat, I am operating the sail, I have taken the rudder. Whereas I am pushing the pole, the boat is carrying me, and there is no I beyond the boat. My sailing in a boat is what makes this boat be a boat. . . .[25]

Obviously, Dogen is talking about the practical dependence of what-it-is-ness of things. Again, it is argued that the determinate characters of things become present from the world of emptiness through our practical confrontations with them. A boat that is not a boat in itself becomes a boat by my confronting it as a boat.

---

[24] Dōgen 1990, vol. 2, 193
[25] Dōgen 1990, vol. 2, 83f.

In addition, what has to be remarked is that Dōgen seems to suggest a *reciprocal* portrayal of this practical confrontation. As will become clearer in due course, Dōgen seems to be saying that the boat is made the boat by the person's using it as a boat, and at the same time, it is being carried by the boat that makes this person the person that he is.

Also, importantly, this reciprocal relationship between the self and the boat can be extended to include the relationship between a thing and the entire universe. As Dōgen's wording of "life can be likened to" implies, what is talked about here is our life at large and not the individual contexts where only the local interdependences matter. Thus, says Dōgen;

> At this very moment, the entire universe is nothing but the boat itself: the sky, the water, the shore, have all become the moment of the boat. . . . In this way, life is what I am making it, and I am what life is making me. While I am sailing in the boat, my body and mind and circumstances are all essential parts of the boat; and the whole Earth and the whole of space are all essential parts of the boat. All of this applies to self as life and life as self in general.[26]

In this passage, Dōgen quite explicitly endorses the idea of universalized reciprocal dependency. It is not the boat that is in operation at a localized time and place, but the boat that contains the entire universe as its moment that is "given life" by my using it as a boat (Life is what I am making it). Conversely, this self that sails in the boat is "given life" by being carried by the boat that contains the whole earth and the whole space as essential parts of it (I am what life is making me).

To recall, Schaffer's argument for *priority* monism was that both the whole and its parts exist but the whole is prior to its parts (The parts are dependent fragments of the whole). In contrast, Dōgen insists that it is reciprocal—"simultaneous", to use Dōgen's original wording—dependency between the whole universe and each of its parts that makes each being the thing that it is. For Dōgen, it is not the *priority* relation but the *reciprocal-cum-simultaneous* dependence that obtains between the whole and its parts.

Beyond Monism and Pluralism

[26] Dōgen 1990, vol. 2, 84

At one point, where the annihilation of *existence* monism/pluralism debate was discussed, Dōgen mentioned: "[h]earing the phrase of 'the myriad of things', we should not explore it as being the One, nor should we explore it as being the many". Also, in the context of discussing the phrase of "*Issai Shujō Shitsu-U Busshō*", Dōgen made a remark to the effect that "*Shujō* as the multitude of beings is the One as the whole and one individual part of the whole at one time".

Putting these two thoughts together, Dōgen is eventually holding that the entirety of being is "neither the One nor the many" and "both the One and (one of) the many" at one time. Now, having the thesis of universal reciprocity/simultaneity between the whole and its parts in view, how are we to make of this peculiar commitment made by Dōgen?

Getting back to the case of the boat, Dōgen said that the moment at which my sailing in a boat makes the boat be a boat, the sky, the water, the shore—in short, the entire universe—have all become the moment of the boat. It is in this sense that "the whole Earth and the whole of space" are said to be "all essential parts of the boat".

However, at the same time, what needs to be remembered is that when I am using the boat as a boat, I am using it as a boat and not as anything else. As long as I am confronting it as a boat, the boat exists as a boat; i.e. as an individual part of the whole that retains a character that is distinct from other elements of the universe.

So, in sum, Dōgen's thought on the One and each part of it can be compressed as the following: a boat is given life and *simultaneously* the whole universe becomes its essential parts when a boat is used as a boat (i.e., as an individual part of the whole). That is, "X's existing as X (i.e., as an individual part of the whole)" is always concatenated with its *being one with the whole*. That being said, what are the consequences of these considerations for the issue of *priority*?

One the one hand, when asked whether X is a part or the whole, it will not be allowed to claim that "X is a part [full stop]". As long as "X's being an individual part of the whole" comes as a package of its "being one with the whole", the claiming of "X is a part" should always be concatenated with the saying of "*simultaneously* X is nothing but the whole". Thus, Dōgen's saying of "X should not be explored as being [one of] the many [full stop]"; the part is always one with the whole.

On the other, the same sort of consideration can be applied to the myriad of things as "the One". That is, it will not be allowed to claim that "the entirety of being as *totum simul* is the One [full stop]". For, as long as the One as the whole is

always the one that each of the many is one with, the saying of "the One" should always be concatenated with the saying of "*simultaneously* this One is the One with which each of the many is one with". Thus, Dōgen's saying of "the myriad of things should not be explored as being the One [full stop]"; this One's being the One is made possible by its concatenation with the part.

From above follows the verdict on the *priority* debate by Dōgen. Based on his *pan-self-ist* understanding of the world, he proclaims that the monism/pluralism debate is annihilated even at the level of priority; it is always the *simultaneous-cum-reciprocal* interdependence, not the *priority*, that holds between the whole and each of its parts. For Dōgen, the moment at which something exists as an individual part of the world is, *simultaneously*, the moment at which it is one with the entire universe.

References

Bielefeldt, Carl (2017) Introduction [to his translation of Such (*Inmo*)]. *Soto Zen Journal: Dharma Eye*, 40, 30.

Dōgen (1990–1993) *Shōbōgenzō 4 vols.* Nahoko Mizuno (ed.) (Tokyo: Iwanami Shoten) *Translations by Carl Bielefeldt (#1) and Hubert Nearman (#2) have been referenced but the citations in this paper do not necessarily follow them.

(#1)https://buddhiststudies.stanford.edu/research/hcbss-research-projects

(#2)http://www.thezensite.com/ZenTeachings/Dogen_Teachings/Shobogenzo_Complete.html

Garfield, Jay (2015) *Engaging Buddhism: Why It Matters to Philosophy.* (Oxford: Oxford University Press).

Nishitani, Keiji (1991) *Shōbōgenzō Kōwa II*, *Nishitani Keiji Chosaku-Shu*, vol. 23. (Tokyo: Sobunsha).

Schaffer, Jonathan (2010) Monism: The Priority of the Whole. *Philosophical Review*, 119.1, 31–76.

Suzuki, D. T. (1997) *Toyo Teki Na Mikata.* (Tokyo: Iwanami Shoten).

Suzuki, D. T. (2016) *Selected Works of D.T. Suzuki, Volume III: Comparative Religion.* (Oakland: University of California Press)

Dōgen on Time and the Self

MORIYAMA Shinya
Professor, Shinshu University
SAKON Takeshi
Associate Professor, Osaka City University

***Abstract**: In this article, we examine Dōgen's paradoxical theses concerning the passage of time. In one instance, he denies the commonsense view that time passes but, in another, accepts time's aspect of coming and going. Although the two theses look incompatible, there is a way to come to a coherent interpretation. The key is Dōgen's peculiar concept of the self. Unlike the other no-self doctrine of Buddhism, his theory of no-self indicates not only the refutation of self as the independent subject but also the acceptance of SELF, which is equal to the myriad things existing as the manifestation of the Buddha-nature. When we look at our initial problem on the paradoxical theses based on the above observation of Dōgen's theory of the self, we can understand why he needs to describe time as possessing two different aspects.*

1. Introduction

This article presents preliminary research for our future study of an Asian analytical-philosophical approach to Dōgen's theory of time.[1] The aim of this article is to examine several important passages from the fascicle "Uji" of the *Shōbōgenzō*[2] to give a plausible explanation of his two paradoxical theses with respect to the passage of time:

(T1) Time does not pass.
(T2) Time has the aspect of coming and going.

[1] This article is partially based on Moriyama 2019, a critical review of Vorenkamp 1995.
[2] There is a number of studies on the fascicle "Uji", for instance, Abe 1992, Chaps. III & IV; Kopf 2001; Tsujiguchi 2012, Chap. 4; Yorizumi 2014; Izutsu 2015; and Roberts 2018. Our reading in the following is mostly influenced by Izutsu's analysis.

It seems that T1 and T2 are incompatible because T2 is naturally read as saying that time passes. If T1 is true, T2 must be false and vice versa. Nevertheless, Dōgen seems to assert the two theses, especially when he states that time has both characteristics of coming and going and neither-coming-nor-going. Although one might think that such a paradoxical expression is a usual practice of Zen, we can try to offer a reasonable account by assuming the different purposes of the two theses: while T1 aims to reject the commonsense view of time's passage, T2 aims to establish the ultimate nature of time that is discovered in Zen meditation, in which things interpenetrate and exhibit real transition without the distinction of past, present, or future.

2. Time Does Not Fly Away

Let us start with T1, namely, the denial of the passage of time. In the commonsense view, the passage of time is expressed using various phrases, like "time flows", "time flies away", etc. According to Dōgen, however, this notion of the passage of time is untenable. He begins with the following observation:

> Text 1: 十二時の長遠短促，いまだ度量せずといへども，これを十二時といふ。去来の方跡あきらかなるによりて，人これを疑著せず，疑著せざれどもしれるにあらず。(SG II: 47)
> Even though one does not measure the length of twelve hours, one calls it "twelve hours". Since the mark of coming and going [of time] is obvious, one does not doubt it. Although one does not doubt it, one does not know it.[3]

In the above passage, Dōgen raises an objection against the commonsense view that time passes at a certain rate. People simply assume that 12 hours (i.e., 24 hours in the modern calculation of time) pass per day without questioning what "time" denotes or how long (or short) 12 hours is. Why do they tend to hold such an ungrounded belief regarding the passage of time? To answer this question, Dōgen points out the following two fallacies in the commonsense view: First, the passage of time can be considered analogous to movement in space; second, time is supposed to exist independently of its observer.

[3] Cf. Raud 2012, 159f.; Roberts 2018, 57; Uchiyama 2018, 188f.

> Text 2: しかあるを，仏法をならはざる凡夫の時節にあらゆる見解は，有時のことばをきくにおもはく，あるときは三頭八臂となれりき，あるときは丈六金身となれりき。たとへば、河をすぎ，山をすぎしがごとくなりと。いまはその山河，たとひあるらめども，われすぎきたりて，いまは玉殿朱楼に処せり，山河とわれと，天と地となりとおもふ。(SG II: 48–49)
>
> Nevertheless, the understanding of an ordinary man who has not studied the Buddhist teachings is such that on hearing the word 'at one time' (*uji*), he thinks: 'At one time someone had become a demon [three-heads-eight-arms], at another time he had become an enlightened one [six-jo-eight-shaku]'. This [passage of time] is just like crossing a river, passing a mountain. Even if the mountain and the river still exist, I have passed them and now stay in this jewel palace and vermilion tower. Me and the mountains-rivers are like heaven and earth to each other".[4]

An ordinary person misunderstands the term *uji* as merely indicating "(at) one time". Such a misunderstanding would lead us to think that just as Kyoto and Tokyo exist as two different places, one time (e.g., the moment of crossing a river) exists independently of another (e.g., the moment of staying at a palace). Depending on which position in space one occupies, it may be said that such-and-such is the case at one place and so-and-so at another. Similarly, one tends to say that such-and-such is the case at one time and so-and-so at another according to the passage of time. This is how time is treated analogously to space, as the first fallacy shows.

Dōgen also remarks that the commonsense view presupposes "I", or a self, as a subject of experience existing against the objective background of time and space. Thus, one firmly believes that the self moves from one place to another or from one time to another. Dōgen's conclusion is that the ordinary conception of time's passage presupposes the distinction between *what passes* and *what is passed*: Time (or place) and I (i.e., the subject of an experience) are distinct, as the second fallacy shows. Against such a misconception, he presents the following argument:

> Text 3: 時は飛去するとのみ解会すべからず，飛去は時の能とのみは学すべからず。時もし飛去に一任せば，間隙ありぬべし。有時の道を経聞せざるは，すぎぬるとのみ学するによりてなり。(SG II: 50)

[4] Cf. Raud 2004, 39; Raud 2012, 162; Roberts 2018, 86; Uchiyama 2018, 190.

> One should not understand that time just flies away. One should not learn that it is time's capability that it flies away. If time [really] retains [the capability of] flying away, there must be a gap. One does not know the true doctrine of being-time (*uji*) because one has learnt only time's passage.[5]

Here is our reconstruction of what Dōgen suggests in the above text: First, for *reductio ad absurdum*, suppose that (i) time flies away, as the ordinary expression states. Dōgen then points out that (ii) if time flew away, there would be a gap between one time and another. Since (iii) there is no such gap, he concludes that time does not pass.

The absurdity might be obscure at first. As for premise (ii), the analogy of time and space will do the expected job. If something (say, a bird) flies from one place to another, there must be a gap between the two places. Likewise, if time "flies away", there must be a gap between a certain time and another.[6] Dōgen suggests that if the passage of time is treated analogously to movement in space, a gap between two given times should be unavoidable.[7] Premise (iii) may be more difficult to grasp, but we suggest that Dōgen's phenomenological tendency plays a crucial role here. In our view, Dōgen means to say that if one looks carefully at what s/he really experiences, s/he will know that no gap is *experienced* at any time and that, therefore, time must be seamless. This phenomenological interpretation of Dōgen is surely controversial. What is crucial for our purpose is whether Dōgen

---

[5] Cf. Roberts 2018, 107; Uchiyama 2018, 191f.

[6] There are mainly two possibilities to consider when interpreting the "gap". While some previous studies (Shaner 1985, 150; Tanahashi 1995) show that the gap is between oneself and time, others assume that the gap is between two times, as Raud (2012, 164) states: "[I]f we would, indeed, against the text's admonition, presume that moments fly past, one after another, like the stages of the moment of Zeno's arrow, it would be very logical to ask what is present during the almost imperceptible interval when one moment has already passed and another one is still not yet here". Although we follow the latter in our interpretation of the text, we do not deny that, as its implication, the former's view is also concluded.

[7] This argument reminds us of Nāgārjuna's famous refutation of motion in the second chapter of his *Mūlamadhyamakakārikā*, where the concept of motion is rejected because, by grammatically analyzing the verb $\sqrt{}gam$, i.e., to go, it is revealed that the object of the present-tense motion (e.g., the path to be traversed) is impossible for both the path already traversed and the path not yet traversed. This is because there is no third place that is neither the path already traversed nor the path not yet traversed. As such, there is no possibility of establishing the object of the motion. Without the object, the motion itself is also unestablished. To explain the third place that is neither the past place nor the future place, the later commentator, Candrakīrti, describes the place's infinite divisibility, like Zeno's paradox of flying arrows at rest. Cf. Katsura & Siderits 2013, 32.

would accept or reject the idea of an independent reality. To answer this question, it is necessary to analyze what Dōgen thinks of the self and the world.

3. SELF as the Entire World

As seen in the previous section, the commonsense view of the passage of time rests upon the distinction between *what passes* and *what is passed*. Once the distinction disappears, the passage of time is experienced in a completely different manner. Then, how does the distinction disappear?

On this point, as several previous studies have already clarified, Dōgen's notion of self (*jiko*, *ware*) gives us a clue regarding its answer.[8] To put it briefly, we should distinguish two different meanings of self in Dōgen's usage: while the term denotes an independent entity as the subject or agent, the same term also signifies the true self as the Buddha-nature (*busshō*). For convenience, we shall write the former as *self* and the latter as SELF. Dōgen denies *self* and affirms SELF. It should be noted here that Dōgen's idea of no-self differs from the orthodox Buddhist Reductionist view that *self* is reducible to five aggregates of physical and mental elements (*pañcaskandha*). While this view still allows distinctions between my body and mind and another's body and mind or between internal elements and external elements, Dōgen radically removes such distinctions by referring to "body-and-mind-dropping-off" (*shinjin datsuraku* 身心脱落). Note, however, that the negation of body and mind does not mean mere nothingness, because it implies the "dropped-off-body-and-mind" (*datsuraku shinjin* 脱落身心), namely, the affirmation of body and mind that newly emerges through discarding wrong conceptions concerning the dichotomy of body and mind.[9] Likewise, when the so-called *self* is denied, SELF newly emerges.[10] According to Dōgen, this SELF is equated with the entire world as being-time. Let us look at the following passage from the fascicle "Uji":

---

[8] Most recently, Mitani 2019 clearly presents Dōgen's position as the "pan-self-ism" through his fusion-philosophical reading of Dōgen, Nishida, and Sellars. Cf. also Kopf 2001, Chap. 2.

[9] See Izutsu 1982, 5.

[10] In this connection, Dōgen's famous phrase of Genjōkoan, "To learn SELF is to forget *self*. To forget *self* is to be illuminated by myriad things" is well understood. To be aware of SELF, it is necessary to forget *self* that is grasped mistakenly as being restricted to a certain body and mind. To forget such *self* is nothing but to emancipate oneself from the limitation of a certain body and mind.

> Text 4: われを排列しおきて尽界とせり，この尽界の頭頭物物を時時なりと見すべし。物物の相礙せざるは，時時の相礙せざるがごとし。このゆゑに同時発心あり，同心発時あり。および修行成道もかくのごとし。われを排列してわれこれをみるなり。自己の時なる道理，それかくのごとし。(SG II: 47–48)
>
> Having unfolded SELF, the entire world occurs.[11] One should regard each different thing of the entire world as each different time. Just as different things do not obstruct each other, different times do not obstruct each other. Therefore, the resolution of each mind [for Buddhist practice] occurs simultaneously. The occurrence of each [different] time is shared by the same mind (i.e., SELF). The practice and the enlightenment are also considered in the same manner. Having unfolded SELF, I see this [SELF].[12] The principle "self is time" is as such.

According to the oldest commentary of *Shōbōgenzō*, Kyōgō's *Shōbōgenzōshō*, the "*ware*" in this passage means the "SELF as the Buddhist entity" (*buppō no ware*), which is likely the same as the Buddha-nature.[13] Against our commonsense view that *self* exists as a subject in front of the world as an object, this SELF as the Buddha-nature becomes manifest as the entire world.[14] Dōgen describes the world

---

[11] On the translation *hairetsu* 排列, several translations are possible. Cf. Raud 2012, 160: "The I unfolds and become the world in its entirety, ..."; Roberts 2018, 61: "We set the self out in array and make the whole universe"; Mitani 2019, 14: "The self extends itself and [as a result] it contains the whole universe." We follow Mitani's direction but choose "(to) unfold" for *hairetsu*'s translation.

[12] In this peculiar worldview, a situation where I see a flower can be paraphrased as "the entire world sees the entire world" or as "time sees time." Cf. SGS 10: 排列の様如前云，尽界が尽界を見る程の道理なるべし，時が時を見る心地也。

[13] SGS 10: 此の我は仏法の我也。…唯我が尽界なる所を排列とは云なり。

[14] The implication of SELF as the entire world is to deny an opponent's view that the Buddha-nature is hidden inside oneself. Shirō Matsumoto (2000, 192–193) named it *busshō naizai ron* 仏性内在論, which presupposes the independent agent distinct from its environment. In Genjōkōan, their opinion is: "To carry *self* forward to undertake the practice and to attain the enlightenment of myriad things" (自己をはこびて万法を修証する). To this, the sentence "Myriad things go forward and undertake the practice and the enlightenment of SELF" (万法すすみて自己を修証する) represents Dōgen's position. Contrary to the opponent's view, this view, *busshō kenzai ron* 仏性顕在論, accepts that myriad things already exist as the manifestation of the Buddha-nature. Cf. also Sueki 2009, 248–250.

as having no obstruction, namely, no boundary between the self and the other, between the subject and object, and between one moment in time and another. However, it does not refer to a genuinely monistic worldview. To accept SELF as the entire world is nothing but to accept a seamless connection of different entities (*dharma*s) of different times in *one* SELF. Therefore, it is said: "[T]he resolution of each mind [for Buddhist practice] occurs simultaneously. The occurrence of each [different] time is shared by the same mind (i.e., SELF)". If *one* SELF makes the resolution, other *SELVES* make the resolution simultaneously. Thus, the entire world makes the resolution. If the resolution occurs at *one* time, it occurs at other times, too. Thus, the resolution always occurs. This co-relation of "oneness in multiplicity"/"multiplicity in oneness" forms the core of Dōgen's philosophy.

4. Time Recurs

On the basis of the above understanding of Dōgen's SELF, we can begin to comprehend the reason why he describes the nature of time using such paradoxical expressions. In the following section of "Uji", we find that he allows two different aspects (*sou*, 相) of time, namely, coming and going (*kyorai*, 去来) and its negation. The following is the subsequent paragraph from the previous Text 2:

> Text 5: しかあれども，道理この一條のみにあらず，いはゆる山をのぼり河をわたりし時にわれありき，われに時あるべし。われすでにあり，時さるべからず。時もし去来の相にあらずば，上山の時は有時の而今なり。
> 時もし去来の相を保任せば，われに有時の而今ある，これ有時なり。かの上山渡河の時，この玉殿朱楼の時を呑却せざらんや，吐却せざらんや。
> 三頭八臂はきのふの時なり，丈六八尺はけふの時なり。しかあれども，その昨今の道理，ただこれ山のなかに直入して，千峰万峰をみわたす時節なり，すぎぬるにあらず。三頭八臂もすなはちわが有時にて一経す，彼方にあるににたれども而今なり。丈六八尺も，すなはちわが有時にて一経す，彼処にあるににたれども而今なり。(SG II 49–50)
> [1.] Nevertheless, the principle of time is not only this one. At the time of climbing the mountain and crossing the river, SELF has already been. Time belongs to SELF. Inasmuch as SELF has already been, time does not pass.

> As time does not retain the aspect of coming and going, the time of mountain-climbing is the "immediate now of being-time" (*uji no nikon*).
> [2.] If time retains the aspect of coming and going, the "immediate now of being-time" belongs to [my] SELF. This is [my] being-time.
> [3.1] Does not the time of climbing the mountain or crossing the river shallow up the time of the jewel palace and vermilion tower?
> [3.2] Does not that time spit out this time?
> [4] The demon with three heads and eight arms is yesterday's time. The Buddha with six-jo and eight-shaku is today's time. Nevertheless, the principle of naming "today" and "yesterday" is used at the time when one enters directly into mountains and look at the myriad mountain summits. [Such events] do not pass away [in its ordinary sense]. [The event of] having three heads and eight arms recurs in *my* being-time (*waga uji*). Although it looks to have occurred in a remote time, it recurs at the immediate now. [The event of] being with six-jō and eight-shaku also recurs in my being-time. Although it looks to have occurred in a remote place, it recurs at the immediate now.[15]

Two metaphorical expressions, "swallowing up" and "spitting out", are important here. While the former indicates time in the past being integrated with time in the present in one's experience (i.e., oneness in multiplicity), the second indicates that time in the past is distinguished from that of the present in one's being-time (i.e., multiplicity in oneness). While the former is described in paragraph [1], the latter is described in paragraph [2].

Paragraph [1] is summarized as follows: Unlike the commonsense view that the moment in time of crossing a river (i.e., practice's time) differs from the time of living in palace (i.e., goal's time), Dōgen claims that the two times are connected seamlessly because they share the same SELF. In this case, since practice's time

[15] For other translations, see Raud 2012, 162-163; Roberts 2018, 91, 96, 103; Uchiyama 2018, 191. Among previous studies, the important point of the relationship between Dōgen's self (i.e., the selfless self) and time as the immediate now has been explained several times. Cf. Kopf 2015, 184: "In contrast to inauthentic experience of time, which Dōgen compares to the experience of an individual who "passes mountains and rivers," authentic experience of time is characterized by an internal relationship between the (selfless) self and time and in the sense that time functions as the internal negation of self and vice versa. In short, time temporalizes itself as the internal self-negation of the self qua from the present to present." Our analysis repeats the same point by emphasizing the two paradoxical aspects of time, namely, coming and going and neither-coming-nor-going.

does not pass away, it is called the "immediate now", just as goal's time is called the "immediate now". Thus, each time of each entity of the entire world is called the "immediate now".

In paragraph [2], on the other hand, the same situation is analyzed from a different angle. Dōgen states that the idea that time has the aspect of coming and going is also acceptable. Of course, its meaning is not same as the commonsense view of time's passage, which was already refused in the first section. What Dōgen claims here is a different kind of "passage of time", which should be called more correctly by the term *kyōryaku* 経歴 as we will see later. Broadly speaking, in meditation, *one* SELF (*waga uji*) reexperiences what the Buddha and ancient Zen masters experienced in the exact same way. Dōgen says elsewhere, "Qingyuan 青原 is time, Huangbo 黄檗 is time. Mazu 馬祖(江西) and Shito 石頭 are also times" (SG II: 51). Since the empirical *self* is already removed, one SELF is not restricted to a certain body and mind. It contains each different time of each different Zen master. Thus, at the "immediate now" of *my* being-time (i.e., SELF), each different time becomes manifest.

Thus, it is now clear that two paragraphs [3.1] and [3.2] metaphorically express [1] (oneness in multiplicity) and [2] (multiplicity in oneness), respectively. The image is like Indra's net, where each jewel reflects other jewels in the net, and in its totality, every jewel reflects every other jewel. In the same manner, one being-time is reflected in all other being-times, and all other being-times penetrate into one being-time.

Moreover, paragraph [4] again explains [2]'s point. If one is not yet familiar with Zen meditation, such an unskilled practitioner might think that s/he experienced a demon's image yesterday and experienced the Buddha's image today; when Buddha's image is experienced, the yesterday's demon image has already gone. To this, Dōgen claims that yesterday's demon is not gone; it recurs in one's being-time. In the same manner, today's Buddha also recurs in one's being-time.

To understand time's aspect of coming and going in its special sense, we need to turn to Dōgen's technical term *kyōryaku*, which indicates a recursive time in our interpretation.[16] One usually considers that an event that happened at one time

[16] On *kyōryaku* 経歴, there have been several translations in previous studies. For instance, Raud (2012, 164-165) proposes "shifting"; Roberts (2018, 243, n. 1) summarizes previous translations of *kyōryaku* and proposes "seriatim passage"; in Uchiyama (2018, 192f., fn. 172), a translator D. T. Wright comments: "Translators have variously translated *kyōryaku* as "flowing", "passing", "continuous existence", "changing", "moving", and so forth. However, all of these words only articulate the seemingly moving aspect of time. This is

is now passing away and never again recurs because we believe that the subject "I" moves along an objective timeline. This is an illusion. Through Dōgen's lens of meditation, all events recur without passing away in one direction. On the basis of the Huayan thought of the gate of ten different times,[17] he states that a time and another time are connected in multiple manners: from today to tomorrow, from today to yesterday, from yesterday to today, from today to today, from tomorrow to tomorrow, etc.[18] However, notice that such multiple connections of times do not commit the fallacy of the gap involved in time's flying away, as pointed out in the criticism of the commonsense view of time's passage. In the Huayan manner, times are connected in the "interpenetration" (相入), a seamless connection without any gap between two times.

With respect to this *kyōryaku*, we should not misunderstand that it denotes a successive occurrence of different times one after another. According to Dōgen, spring does not pass away winter, nor does summer pass away spring. Spring only passes away spring; that is, spring recurs without presupposing the distinction of *what passes* and *what is passed.* Dōgen states:

---

why Uchiyama Rōshi uses the term *kokkoku* (刻々), or "moment-by-moment", to express the nuance of *kyōryaku* in ordinary language". Although those previous translations each have merit and demerit, we do not have any decisive idea for a better translation here. However, the striking aspect of *kyōryaku*, on which we would focus in our context, is its nature of "recurrence"; that is, time does not fly away, but it recurs again and again in a Zen practitioner's pure experience. Note, however, that time's being recursive does not imply time's one-way recurrence; rather, it expresses that all moments in time penetrate into the religious time (i.e., *my* immediate now) in multiple ways. At any rate, for a more precise understanding of *kyōryaku*, we need to examine Dōgen's other arguments on time, especially the notion of the "discontinuity of before and after" (*zengo saidan* 前後裁断). However, this will be a future task.

[17] In the theory, nine kinds of time, i.e., past in past, present in past, future in past, past in present, present in present, future in present, past in future, present in future, and future in future, are said to interpenetrate. In addition, there is also one all-inclusive time as the 10th time. The 10th time stands as the ground upon which past, present, and future times are mutually connected to each other in multiple directions. Cf. *Huayan wujiao zhang*, T. 1866 Vol. 45.506c16-22: 八者十世隔法異成門。此上諸雑義遍十世中同時別異具足顯現。以時与法不相離故。言十世者。過去未来現在三世。各有過去未来及現在。即爲九世也。然此九世迭相即入故，成一總句。總別合成十世也。此十世具足別異同時顕現成縁起故。得即入也。

[18] SG II: 50: 有時に経歴の功徳あり。いはゆる今日より明日に経歴す，今日より昨日に経歴す，昨日より今日に経歴す。今日より今日に経歴す，明日より明日に経歴す。

> Text 6: 経歴といふは、風雨の東西するがごとく学しきたるべからず。尽界は不動転なるにあらず、不進退なるにあらず、経歴なり。経歴は、たとへば春のごとし。春に許多般の様子あり、これを経歴といふ。外物なきに経歴すると参学すべし。たとへば、春の経歴はかならず春を経歴するなり。経歴は春にあらざれども、春の経歴なるがゆゑに、経歴いま春の時に成道せり。審細に参来参去すべし。経歴をいふに、境は外頭にして、能経歴の法は東にむきて百千世界をゆきすぎて、百千劫をふるとおもふは、仏道の参学、これのみを専一にせざるなり。(SG II: 53–54)
>
> *Kyōryaku* 経歴 should not be understood like wind and rain moving from east to west. The entire world is neither changeless nor motionless. It recurs. It is like spring. That many things (e.g., birds, flowers, glasses) appear in spring is called *kyōryaku*. One should learn that it recurs without external objects (i.e., *what passes* and *what is passed*). For instance, spring's *kyōryaku* definitely means that [spring] passes away spring (i.e., spring recurs). The recursive passage itself is different from spring. However, since it is the recursive passage of spring, the recursive passage is established at the time of spring. This point should be examined in detail. Hearing the term *kyōryaku*, people think that external objects exist and *what passes* moves toward the east and passes a hundred thousand worlds and a hundred thousand *kalpa*s away. It is because they have not been devoted to learning only the Buddhist path.[19]

It is difficult to imagine that something passes without *what passes* and *what is passed*. However, if we remember that Dōgen's SELF is no longer an agent and that a simple grammatical structure of S-V-O is paraphrased in many manners, we will notice that his idea of *kyōryaku* is also understood in a similar manner. For instance, when one says, "Spring comes", since neither *what passes* nor *what is passed* exists, there is only a situation that should be called *spring-ing* or something similar. Birds fly, flowers blossom, plum trees boom, etc. That the entire world is full of such spring things and events may be called *spring-ing*. Or, it can be said that everything is integrated into spring. By using the metaphor of spring's *kyōryaku*, Dōgen teaches us how to understand time's recursive passage in Zen meditation. It is wrong to consider that time's passage happens outside oneself. It is also wrong to consider that it happens inside oneself. It should be a pure experience in which the seer and

[19] Cf. Raud 2012, 166; Roberts 2018, 172; Uchiyama 2018, 197.

the seen (or *what passes* and *what is passed*) are interpenetrated. It might be helpful to present Toshihiko Izutsu's explanation to understand how such a religious experience beyond subject-object dichotomy works in the Zen tradition:

> Zen argues as follows. One cannot *become* water because one is observing it from outside, that is to say, because the ego is, as an outsider, looking at water as an 'object'. Instead of doing so, Zen continues to argue, one must first learn to 'forget one's ego-subject' and let oneself be completely absorbed into the water. One would then be flowing *as* the flowing river. No more would there be any consciousness of the ego. Nor would there be any 'consciousness of' the water. Strictly speaking, it is not even the case that one *becomes* the water and flows on as the water. For in such a dimension there would be no ego existent to *become* anything. Simply: The water flows on. No more, no less. (Izutsu 1982, 81)

In the pure experience of looking at water, *self* as the subject or ego disappears, and only flowing water remains. In our context of being-time, it can be said that in *becoming* water, one experiences water's time. Likewise, when Dōgen states, "A pine is time, a bamboo is time" or "The ocean is time, a mountain is time", in *becoming* a pine, a bamboo, the ocean, or a mountain, one experiences a pine's time, a bamboo's time, the ocean's time, and a mountain's time. Dōgen's notions of being-time and recursive passage *kyōryaku* show us how it is possible to think of the radical transformation of time's experience.

5. Is It a Non-Experience View or an Error Theory?

To repeat, Dōgen denies time's flying away, which involves a distinction between what passes and what is passed. On the other hand, he endorses time's coming and going, which does not involve such a distinction. How can this line of thought be justified? Our proposal is that it is supported by the phenomenological fact that the latter is really experienced, whereas the former is not: *A fortiori*, one never experiences time flying away (from the future to the past), but one does experience (or live in) some transition that Dōgen expresses as "coming and going" (*kyorai*) or "recurrence" (*kyōryaku*). If Dōgen is a radical phenomenologist or idealist who holds both that true experience itself constitutes reality and that there is no

independent reality, he would think that time's coming and going is real, whereas time's flying away is not.

Is this a realist view of time's passage or not? If our interpretation is correct, it is clear that Dōgen holds a realist view with respect to time's coming and going. What about time's flying away? To examine the implication of his philosophy of time, let us consider this question: Does Dōgen hold a non-experience view or an error theory with respect to time's flying away? We suggest that he could go either way.[20]

If Dōgen holds a non-experience view, he would argue as follows: (i) Time is real only if it is experienced as it appears; (ii) Time is not experienced as flying away while it is experienced as coming and going. Therefore, time's flying away is not real, whereas time's coming and going is real. This form of a non-experience view regarding time's flying away would be a natural way of thinking that fits nicely with our interpretation of Dōgen's philosophy of time.

However, Dōgen could also be an error theorist. In this case, he would argue as follows: (i) It is *believed* that time is experienced as flying away; (ii) In reality, for the Buddhist practitioner, time is not experienced as flying away while it is experienced as coming and going. Therefore, the belief that time is experienced as flying away is false. It seems that this is also a natural way of thinking if our interpretation is correct.

It is important to note that this form of error theory is very different from usual error theories. They typically assume an independent reality and argue that it falsifies certain human beliefs: Things are not as they appear. Dōgen's error theory would lead us in the opposite direction. It says that there is no independent reality. Instead, it points to the fact that there are two layers in one's experience: pure experience and belief. It is beyond the scope of the present paper to explain what "pure experience" means, except to say that it is supposed to be achieved through

[20] According to a preceding survey by Miller (2017, 772), Dōgen is regarded either as a "hard-nosed" realist who is called an A-theorist in contemporary metaphysics after McTaggart (1908) or as a "middle-way" realist who is called a B-theorist. Miller attributes the latter reading to Vorenkamp (1995) and the former to Heine (1985) and Stambaugh (1990), although she also points out Vorenkamp's hesitation in classifying Dōgen as a simple B-theorist. Miller's own diagnosis is that Dōgen is a moving-spotlight theorist who belongs to the hard-nosed category, contrary to what we have argued in the present paper. In our view, what was missing in the preceding discussion is the phenomenological tendency found in Dōgen's writings and his rejection of an independent reality. If our interpretation is correct, Dōgen is neither an A-theorist nor B-theorist; at best, such a classification is not clear in his case. Cf. also Moriyama 2019.

practice (Zen meditation). In ordinary life, one's pure experience and beliefs may affect each other, as what one perceives may affect what one believes, or vice versa. Nonetheless, the core of experience is pure experience. What Dōgen's error theory implies is that one may mistake what s/he believes for what s/he really experiences in Zen meditation.

This error theory is not a metaphysical error theory that assumes an independent reality but, so to speak, a phenomenological error theory in which pure experience plays a crucial role. Regardless of whether Dōgen actually adopts an error theory or a non-experience view, his message is now clear: The truth consists not in an independent reality but in what is really experienced. Insofar as how the world really appears to be *is* how it is, time does come and go (i.e., time recurs). However, this is not to say that time flies away. Time is not a thing that flies away according to Dōgen. This is either because such temporal passage is not truly experienced or because the belief in flying away is falsified by the pure experience of time's coming and going or recurrence.

6. Concluding Remarks

So far, we have seen the main argument of Dōgen's theory of time. The results can be summarized as follows:

1. Dōgen criticizes the commonsense view of time's passage because the view is based on preconceptions of time's being measurable independent of its observer and analogous to space. In particular, since the view presupposes a subject-object dichotomy, Dōgen presents an alternative, innovative view of time beyond this dichotomy.
2. To go beyond the dichotomy, Dōgen deconstructs another common view of the self and demonstrates the entire world as the unfolding of SELF. In this view, the true self, or the Buddha-nature, is manifest in each and every entity of the world, and no space remains for *self* as an independent subject.
3. From 1, it is concluded that time does not pass (T1). Dōgen explains that even though a past event contains all other events in different times, it remains as the immediate now.
4. Since each time interpenetrates, time's recursive nature (*kyōryaku*) is also admitted as an alternative conclusion (T2). Unlike our ordinary view that everything

happens only once and passes away, Dōgen's claim is based on his meditative experience, in which everything recurs and never passes away. The recurrence of religious events in one's meditative state can be best explained by *kyōryaku*, which is said to be possible only when one abandons the preconception of the subject-object dichotomy.

5. If our interpretation of Dōgen's philosophy of time is correct, he is classified as a realist with respect to "coming and going" (*kyorai*) or "recurrence" (*kyōryaku*), while he may adopt a non-experience view or an error theory with respect to time's flying away. In any case, the truth consists not in an independent reality but in what is really experienced—that is, pure experience.

References

*Citations from Dōgen's *Shōbōgenzō* 正法眼蔵 (SG) are taken from the four-volume *Shōbōgenzō* edited by Yaoko Mizuno, Tokyo: Iwanami shoten, 1990 (SG I, II), 1991 (SG III), 1993 (SG IV). Kyōgō's *Shōbōgenzōshō* 正法眼蔵抄 (SGS) is cited from *Shōbōgenzō chūkai zensho* 正法眼蔵註解全書 vol. II edited by Nyoten Jinpo & Bunei Andō, Tokyo: Nakayama shobō, 1956.

Abe, Masao (1992) *A Study of Dōgen: His Philosophy and Religion*. (Albany: State University of New York Press).

Heine, Steven (1985) *Existential and Ontological Dimensions of Time in Heidegger and Dōgen*. (Albany: State University of New York Press).

Izutsu, Toshihiko (1982) *Toward a Philosophy of Zen Buddhism*. (Boulder: Prajñā Press). [First published in 1977, The Imperial Iranian Academy of Philosophy].

Izutsu, Toshihiko (2015) Sōzō fudan – tōyōteki jikan ishiki no genkei, in *Izutsu Toshihiko Zenshū* 9, (Keiō gijuku daigaku shuppankai). (The article first published in 1986).

Kopf, Gereon (2015) *Beyond Personal Identity: Dōgen, Nishida, and a Phenomenology of No-Self*. (New York & London: Routledge). [First published 2001, Curzon Press].

McTaggart, J. E. M. (1908) The Unreality of Time. *Mind,*17, 457–474.

Matsumoto, Shiro (2000) *Dōgen shisō ron*. (Tokyo: Daitō shuppansha).

Miller, Kristie (2017) A Taxonomy of Views about Time in Buddhist and Western Philosophy. *Philosophy East and West*, 67, 763–782.

Mitani, Naozumi (2019) The World in Which Everything Is the Self: The Philosophy of the Original Image and Pan-Self-Ism, in Jay L. Garfield (ed.), *Wilfrid Sellars and Buddhist Philosophy: Freedom from Foundations*, (New York & London: Routledge), 3–31.

Moriyama, Shinya (2019) Uji kanken: Dōgen no jikan ron to B keiretsu no jikan ron. *Shinshū Daigaku Jinbunkagaku Ronshū*, 6, 19–31.

Raud, Rein (2004) 'Place' and 'Being-Time': Spatiotemporal Concepts in the Thought of Nishida Kitarō and Dōgen Kigen. *Philosophy East & West*, 54/1, 29–51.

Raud, Rein (2012) The Existential Moment: Reading Dōgen's Theory of Time. *Philosophy East & West*, 62/2, 153–173.

Roberts, Shinshu (2018) *Being-Time: A Practitioner's Guide to Dōgen's Shōbōgenzō Uji*. (Somerville, MA: Wisdom Publications).

Shaner D. E. (1985) *The Bodymind Experience in Japanese Buddhism: A Phenomenological Study of Kūkai and Dōgen*. (New York: SUNY Press).

Siderits, Mark & Katsura, Shoryu (2013) *Nāgārjuna's Middle Way*. (Somerville, MA: Wisdom Publications).

Stambaugh, Joan (1990) *Impermanence Is Buddha-nature: Dōgen's Understanding of Temporality*. (Honolulu: University of Hawaii Press).

Sueki, Fumihiko (2009) *Butten o yomu – Shi kara hajimaru bukkyō shi*. (Tokyo: Shunjūsya).

Tanahashi, Kazuaki. et al. (1985) *Moon in a Dewdrop: Writings of Zen Master Dōgen*. (New York: North Point Press).

Tsujiguchi, Yuichiro (2012) *Shōbōgenzō no shisō teki kenkyū*. (Tokyo: Hokuju shuppan).

Uchiyama, Kosho (2018) *Deepest Practice, Deepest Wisdom: Three Fascicles from Shōbōgenzō with Commentaries*. Trans. D. T. Wright & A. Okumura. (Somerville, MA: Wisdom Publications).

Vorenkamp, Dirck (1995) B-Series Temporal Order in Dōgen's Theory of Time. *Philosophy East & West*, 45/3, 387–408.

Yorizumi, Mitsuko (2014) *Shōbōgenzō nyūmon*. (Tokyo: Kadokawa shoten).

A Contrarian View of the Philosophy of Zhuangzi, "Seeing *Zhuangzi* as Confucian"

Liu Kuan-ling
National Tsing Hua University

***Abstract**: Seeing Zhuangzi as Confucian 莊子儒門說, as will be developed in this paper,[1] is highly controversial because of the huge difference between Confucianism and the classical Daoist interpretation of Zhuangzi. On the other hand, such controversy could be seen as a noteworthy starting point for clarifying Zhuangzi's philosophical structure because of its disclosure of some crucial incompatibility with Daoism inherent in Zhuangzi that remains obscure if seen only within a Daoist context. Seeing Zhuangzi as Confucian argues that Zhuangzi shares the same concern as Confucianism with the Yi Jing (The Book of Changes) and it inherits its tradition in a genuine sense. Those Confucians who pose a challenge to the traditional Daoist interpretation of Zhuangzi further propose suggestions about how to face Zhuangzi's intrinsic "incompatibility" in a positive way. My focus in this paper will be not on a choice between Confucianism and the Daoist school, or between the philosophy of "being" in a Confucian sense and the philosophy of "nothingness" with its influential Daoist background, but instead on the coexistence of conflicts that have been highlighted by the debates between a Confucian approach and the Daoist school's viewpoint. However, these also imply an essential characteristic of the Zhuangzi philosophy that is worthy of inquiry. Although a controversial issue, "seeing Zhuangzi as Confucian" all the same raises a fundamental question concerning the logic of Zhuangzi's philosophy. By investigating the philosophical underpinnings, it could also provide an opportunity for advancing the philosophy of Zhuangzi in a contemporary context.*

[1] I would like to express my gratitude to all reviewers for their valuable suggestions, some of which could not be dealt with in this paper but do open a new horizon for my future research, especially concerning the issue of *qi*. And also, I would like to thank Prof. Uehara Mayuko and Prof. Linda Gail Arrigo for their generosity in supporting and advising.

1. Is *Zhuangzi* a Classic of the Daoist School, or Is It a Specific Commentary on Confucian Teachings 儒門別傳?[2]

The classical view of *Zhuangzi* philosophy is that it was derived from or at least connected with the philosophy of *Laozi*, and there is evidence to support such an argument. Equally there is evidence that has come to light to suggest that this classical view may need to be reconsidered, for instance, in the discussion of "nothingness". Nothingness is regarded as one of essential notions in *Zhuangzi*, one which mirrors the understanding of *Zhuangzi* as under the influence of the Daoist interpretative tradition. In this tradition nothingness was the core of *Laozi*'s thought and it was further developed by *Zhuangzi* based on his understanding of *Laozi* or that of the concept of *Dao*. Nothingness, interpreted through a Daoist approach, has continued to be a primary issue until the present in the discussion of *Zhuangzi*, whether one's focus is on its close relation to or subtle differences from *Laozi* philosophy.

Essential doctrines of the Daoist school have also been developed in the field of religion. Daoism reveals the potentiality of such doctrines in the sense that some aspects of "body" and its related "practice" should be taken into serious consideration when elucidating these Daoist classics. It does appear that *Zhuangzi* therefore seems to possess a dual characteristic: as a philosophical text or as a religious classic. It can be observed that fundamental concerns still remain in such religious interpretation, and these have continued to influence and reinforce those traditional understandings of *Zhuangzi* in their particular ways. "Reversion 逆反" is a central tenet of Daoism, a practice whose focus is on reversing the course of creation for the purpose of returning to the primordial perfection. It is the thought that still echoes through the typical attitudes already emphasized in *Laozi* and *Zhuangzi,* through their well-known concepts such as "weakness 弱", "softness 柔", "emptiness 虛", and "non-doing 無為", basically showing the stance of passivity and an ultimate pursuit of simplicity.

All of those notions or nuances can be understood against a broader background of "nothingness" defined in a Daoist way: the practice of negation or the

[2] "A specific commentary on Confucian teachings" is the term used by Yang Rur-bin, see Yang Rur-bin 楊儒賓, 〈儒門別傳——明末清初《莊》《易》同流的思想史意義〉 [Rumen Biezhuan: Mingmoqingchu Zhunagyitongliu de Sixiangshi Yiyi], in Chung Tsai-chun 鍾彩鈞 Yang Chin-lung 楊晉龍 eds.,《明清文學與思想中之主體意識與社會——學術思想篇》 [*Mingqing Wenxue yu Sixiang zhong zhi Zhutiyishi yu Shehui: Xueshusixiang Pian*], Taipei: Institute of Chinese Literature and Philosophy, Academia Sinica, 2004): 245–289.

realm achieved through such practice. Both of them are fundamentally associated with each other and should be recognized as different dimensions to be fully aware of in regards to the discussion on nothingness. For Daoist believers, even if they are not firmly disengaging from the mundane world, they are apparently indifferent to that world because of their belief that the perfect state of real human beings is never accomplished by any practice of following the course of creation. Rather, the solution lies exactly in the reversion of such a course. If using *Laozi*'s terms, the crux of the matter is in returning to the original state, like an "uncarved block 樸", in *Zhuangzi*'s terms; the suggestion would be to "sit and forget 坐忘", an essential notion highlighting the necessity of "leaving form and discarding intelligence 離形去智", both of which have entailed the connotation of "reversion", whether the word they have used is exactly "reversion" or not and whether the field they are discussing is purely oriented towards metaphysics or essentially involving "bodily practice". Such theoretical and practical concern has led both the Daoist school and Daoism into a stance that is counter to Confucianism.

One of the fundamental Confucian doctrines is "ceaseless procreating 生生", which is derived from the *Yi Jing* and fundamentally associated with its principle of "interaction and resonance 交感", emphasizing a metaphysical foundation shared by "myriad things 萬物" in this world. While the tradition of the *Yi Jing* highlights the importance of the creation of things and how such creation is possible, in contrast these themes have hardly appeared or been seen as a major question within a traditional Daoist tradition. This suggests a view that the concern with "creation", a concern which reveals the incessant habits of humanity that have been maintained by Confucianism in various aspects, is just a secular fallacy or a tiresome treadmill that from a classical Daoist viewpoint betrays the sacred truth about life itself.

Nevertheless, there are some ambiguities in *Zhuangzi* that are quite noticeable, through which another interpretative choice could be understood and some challenges could be posed that would be meaningful to a Daoist approach. The first example I would like to touch upon is "wandering 遊", a concept not only sufficient to represent the core of *Zhuangzi's* spirit, but also one encompassing a wide spectrum of the philosophy of *Zhuangzi*, one which includes aspects of metaphysics, "body", "practice", and "subject", and also extends its influence beyond the realm of philosophy or religion to the domains of art and literature. The significance of "wandering" has been elevated to be seen as the basis of a Daoist understanding, because wandering is easily related to a tradition of "wandering with immortals 遊仙", or regarded as a certain spiritual development, the nuance of

which is implied in the term "mind-wandering 遊心" in *Zhuangzi*. These interpretations each belong to their own contexts, but they have conveyed a shared message: the concept of "wandering" suggests or at least implies disengaging from the turbid mundane world. Such a passive attitude could be demonstrated based on *Zhuangzi*'s text, but it would also face a potential challenge from another quote, "wandering amidst the mundane world 遊世俗之間", mentioned in Chapter 12, "Heaven and Earth 天地", under the discussion on "chaos".

It could be disputed as to how representative the text is, primarily because it belongs to the outer chapters, a less important part of *Zhuangzi* from a traditional viewpoint, or disputed because such a kind of meaning is emphasized by two Confucian personages: Confucius himself and Zi-gong, whose philosophical concerns appeared to be incompatible with those of the typical interpretation of *Zhuangzi* or the classical Daoist school. A noteworthy philosophical relationship between *Zhuangzi* and Confucianism will be discussed later, but it is still worthy of note here that the mention of such "active" wandering, not suggesting disengagement from, but rather engagement with, the turbid mundane world, does pose a question as to whether the traditional approach is persuasive enough to respond to such dimensions of *Zhuangzi*. It further provides an opportunity to contextualize other important issues in an atypical way.

Another related issue worthy of much notice is language. "Forgetting words 忘言" or "without words 無言" has been seen as the basic attitude of *Zhuangzi* towards the issue of language. This interpretation focuses on the danger accompanying the usage of words, and it is concerned about a necessary fragmentation or unavoidable distortion of "true language"; it suggests a passive or negative attitude towards words, as if based on a Daoist approach, and this has evidential support from *Zhuangzi*'s text, yet it also leaves some ambiguity for further reconsideration. Besides the emphasis on "forgetting words", the importance of which deserves serious attention since its insight is still enlightening today, "goblet words 卮言", another term mentioned in Chapter 27, "Parable 寓言", occupies a key position as well in a further investigation of *Zhuangzi*'s attitude towards language. It would arouse debate, not only because of its connotations concerning the emergence of words, but also because the chapter it is located in is

the one whose theme is language and the focus there is on what true language means.[3]

Examining these issues and inquiring into such inherent incompatibility through an atypical lens does not intend to underestimate the decisive role of Daoist interpretation, but, rather, it attempts to advance this inherited tradition in a different way. These conflicts could be regarded as an obstruction to qualifying *Zhuangzi*'s philosophical identity, or, on the other hand, as a starting point from which to further develop the logical structure of *Zhuangzi,* to the extent that the entangled passages could be articulated. The apparent ambiguity in turn may serve as the key to penetrating into *Zhuangzi*'s specific philosophical concern.

For revealing such a possibility, it will be crucial to introduce an unorthodox and controversial approach called "seeing *Zhuangzi* as Confucian". This choice, however, does not imply an intention of replacing authoritative exegeses with deliberate misunderstanding. On the contrary, it is important to advance this philosophical tradition by confronting the ambiguity within it that would not present a problem within its old tradition. An active, constructive side of *Zhuangzi* has already been highlighted against the horizon of Confucianism, and it supports the supposition of "seeing *Zhuangzi* as Confucian", in the view of which conflicts inherent in *Zhuangzi* are increasingly clarified and interpreted in a more positive way, an attitude that is not only staying vigilant to the discontinuous nature of the mundane world and its relevant dangers, but one also facing such an existential situation and still waiting for some accounting of it.

With respect to "seeing *Zhuangzi* as Confucian", one could associate it with some statements like "the reconciliation of Confucianism and Daoism 儒道會通", a major issue of Neo-Daoism 玄學 or "the harmony of the Three Teachings 三教合一" in the Ming dynasty. My focus here, however, is on a specific interpretative approach shedding light on the *Yi Jing*'s origin in *Zhuangzi*, claiming that *Zhuangzi* shares the same metaphysical structure with the *Yi Jing* and penetrates deep into the profundity of the *Yi Jing*'s thought. Based on such philosophical reasoning, advocates like Wang Fu-zhi 王夫之 (1619–1692) argued that *Zhuangzi* should not be regarded as a Daoist classic, but rather known as a noteworthy commentary on Confucian teachings.

[3] For more discussion on "goblet words", see Yang Rur-bin 楊儒賓, 〈莊子的卮言論〉[Zhuangzi de Zhiyan Lun], 《儒門內的莊子》[*Rumennei de Zhuangzi*], Taipei: Linking Publishing, 2016: 225–264.

2. "Emptiness *ji qi* 虛空即氣" and "the Centre of the Ring 環中"

What interests me here is the interpretation given by Wang Fu-zhi, a representative figure of such philosophical trends as "seeing *Zhuangzi* as Confucian". Wang Fu-zhi's interpretation of *Zhuangzi* is fundamentally against a peculiar background of conceptualization of *qi* that was mainly philosophized by Zhang Zai 張載 (1020–1077), a Neo-Confucian philosopher in the Northern Song dynasty. Zhang Zai delved into a deeper nuance of *qi*, a signature Chinese philosophical term that had been evaluated before and was mainly understood from the viewpoint of the "primordial *qi* 元氣" but one that rarely shows any connection with "emptiness", the concept that aroused Zhang Zai's attention and became one constitutive element in his interpretation of *qi*. "Emptiness *ji qi*", which has the sense of "emptiness is the same as *qi*", is the term Zhang Zai made use of to formulate his theory of *qi*, highlighting an inherent connection between *qi* and emptiness. They are not opposed to each other, he proposed, unlike previous thinkers suggesting that *qi* is absolutely non-emptiness because of its undoubtable actuality and profundity, but he saw that rather they are inclusive of each other, meaning only *qi* with its inherent emptiness is the one fundamentally possessing an identity of "actuality-being 實有". *Ji* illuminates the relationship that *qi*/actuality and emptiness should hold, though such an inherent relationship, however, is not contained in the discussion on *qi* and emptiness; but it extends to every relationship between the one and the other no matter whether or not they are opposed to each other in appearance.[4]

A more detailed account of emptiness *ji qi* and how such logic might work in Wang Fu-zhi's interpretation of *Zhuangzi* will be mentioned later. But it is still important to note in advance that such a background of the usage of *qi* is highlighted here not for the use of proving again the philosophical succession between Zhang Zai and Wang Fu-zhi, but rather for clarifying the strategy Wang Fu-zhi employed in his interpretation of *Zhuangzi*. Zhang Zai's influence mainly shows in two

[4] Although it is possible and would become more clear if "emptiness *ji qi*" is translated into "emptiness is the same as *qi*" or "emptiness as *qi*", but I would like to use the term of *ji* for clarifying that sense in which "emptiness *ji qi*" should be taken into more careful consideration. That is primarily because, even as a conjunction, *ji* is an essential concept representing a specific logic in an eastern context. The following discussion surrounding *qi* and its inherent connotation of paradoxical unity attempts to reveal the meaning of *ji* and its intrinsically paradoxical state against a specific Confucian background. It entails meanings both of "the same" and "not the same"; the former is one aspect of *ji*, but *ji* goes beyond the definition of "the same", at least in our present context.

aspects: the theory of *qi* prepares the ground for arguing the plausibility of "seeing *Zhuangzi* as Confucian", the reason for which is that *Zhuangzi* shared the same metaphysical concern with the *Yi Jing*, and that has been much clarified and shows its logic through Zhang Zai's elucidation of emptiness *ji qi*; furthermore, emptiness *ji qi* provides insight into how to reconsider apparent conflicts from a profound horizon. Such conflicts are also found in *Zhuangzi* and might pose critical challenges to Wang Fu-zhi's interpretation if his approach were still based on the philosophy of "being" that is exclusive of its opposing side, sides of "emptiness" and "nothingness".

It would not, to a large extent, have been necessary for most researchers to identify *Zhuangzi* as a Daoist or a Confucian, at least for those who focus on *Zhuangzi* for its own philosophical identity. Nevertheless, it would be essential to face up to the fact that two opposing interpretative approaches are fundamentally elicited by the incompatibility intrinsic to *Zhuangzi* as long as further development of *Zhuangzi*'s thought is necessary in a contemporary philosophical sense. In line with this, the perspective of "seeing *Zhuangzi* as Confucian" should not and cannot be underestimated because inherent conflicts have emerged, whether they are acknowledged or denied in a positive or negative way.

Wang Fu-zhi's interpretation is important not because it seems to represent an available approach much different from the traditional one, but because it shows a potential for communicating with the other school that is seemingly counter to the one it belongs to. The importance of Wang Fu-zhi's interpretation in the context of the philosophy of *Zhuangzi* lies in the fact that he confronts the challenge of conflicts, not only those inherent in *Zhuangzi*'s text but also those that have emerged and are in heated dispute because of his unorthodox position. His interpretation took a new step by penetrating into a profound logical structure that is adequate to explain those conflicts of text, thought, and interpretation, as well as the sense in which his view on "conflict" deserves to be regarded as a starting point to develop *Zhuangzi* philosophy and advance its tradition in a distinguished way.

My focus here is on the concept of "the centre of the ring 環中", a key notion in *Zhuangzi* but also one fraught with ambiguity. This concept is traditionally viewed as a symbol relating to "nothingness" and its relevant notions, all of which are basically understood in a Daoist context; on the other hand, it is used by Wang Fu-zhi as well to demonstrate its metaphysical concern as already embedded in "Taiji 太極", the concept in which *Zhuangzi*'s Confucian identity will be justified along with *Zhuangzi* through its own text. The question concerning "the centre of

the ring", as developed in this paper, is not on which perspective will be more proper for interpretation of such a notion, but rather on whether implicit nuances of the centre of the ring could be revealed progressively, and if so, how, and also how through such clarification to face up to *Zhuangzi*'s conflicting dimensions. In this sense, I would like to focus on the structure of the centre of ring, which has been explicated in Wang Fu-zhi's commentary on a dialogue between two scholars, Shao Zhi 少知 and Da-gong Diao 大公調, in Chapter 25, "Zeyang 則陽".[5]

There are two notions running through such discussion: "doing 或使" and "non-doing 莫為". The question posed by Shao Zhi is: Which is biased and which is the proper way? It seems that there should be no hesitation to give a clear answer in a typical Daoist way; however, one will realize its complexity and difficulty when further looking to the way in which Da-gong Diao answered such a question. In the present context, there are two points worthy of note. Based on Da-gong Diao's response, first, neither "doing" nor "non-doing" are accurate enough to be representative of the exact practice of Dao, and second, questions relating to doing and non-doing are fundamentally associated with being/actuality and nothingness/emptiness. It is against this background that Wang Fu-zhi goes further in the clarification of such emerging issues through pointing out a constitutive feature of the centre of the ring.

As Wang Fu-zhi argues, the centre of the ring is a key notion with two fundamental elements: "a ring 一環" and the "emptiness within the ring 中虛". The former highlights the dimension of actuality/being, the later emphasizes the part relating to emptiness/nothingness. It would be important to shed light both on what is the connection between them and on how to reveal their meanings within such a specific context. In Wang Fu-zhi's view, the real being is not located in the center, which implies a "pivot" or "axis", but rather it only presents itself as a ring with consistent movement. Furthermore, such a ring is the one fundamentally consisting of its inherent emptiness. There could be no ring without its inherent emptiness, which does also entail that there is no actual being that could exist without its profound nothingness. Similarly, in his thought, emptiness should not be interpreted as one philosophical concern opposed to the other concern for actual being – it rather serves as an essential element constituting actual being by participating in its consistent movement.

[5] See Wang Fu-zhi 王夫之, 《老子衍 莊子通 莊子解》 [*Laoziyan Zhuangzitong Zhuangzijie*], Beijing: Zhonghua Book Company, 2009: 310–311.

It would be more obvious that Wang Fu-zhi takes a positive attitude towards "emptiness" to a great extent, instead of underestimating its significance that has been demonstrated in a traditional context. His approach to the term of the centre of the ring shows a clear indication that issues surrounding being and nothingness, actuality and emptiness, should be discussed on a new horizon: the Great Dao is represented as such a kind of "a ring", the ring with its inherent emptiness. And the emptiness is ultimately emptiness itself that could not be replaced by "actuality" in any sense. Accordingly, debates surrounding being and nothingness, actuality and emptiness, could not be retained at the level of apparent difference because their irreplaceable unique identities are only revealed based on such disclosure of their profound relationship. The centre of the ring is full of such potentiality for reformulating relevant essential issues in *Zhuangzi*, and Wang Fu-zhi makes further efforts to develop its insights, demonstrating a logical structure that is familiar to him: the logic of *ji*.

As mentioned before, Wang Fu-zhi is much influenced by Zhang Zai's thought, especially his argument concerning emptiness *ji qi*. My aim in introducing Zhang Zai's articulation of *qi* is not intended to discuss Zhang Zai's philosophy against the horizon of Confucianism as a whole, but rather to point out the importance of *ji* in its providing a unique perspective on the issue of "conflict".

Emptiness *ji qi* is highly controversial especially within a Confucian context, mainly because of the criticism concerning the insufficient understanding of true "being" in using the term "emptiness", as applied by other Confucian masters such as the Cheng brothers 二程 or Zhu Xi 朱熹. Their views are that such a highlight on emptiness makes Zhang Zai incline to the dimension of emptiness/nothingness, or causes his theory to just remain at the level of "below form 形而下" rather than achieve the level of "above form 形而上".[6] Their concern is Zhang Zai's usage of "emptiness", the term that was perceived as a very strong Buddhist-Daoist concept, one that is fundamentally in conflict with Confucian philosophical concerns. In addition to such debates surrounding "emptiness", it would be also crucial to mention the position that *ji* occupied in a Chinese philosophical context. As already shown in Tiantai 天台 Buddhism, the meaning of *ji* is much more complicated and

[6] See Mou Tsung-san 牟宗三, 《心體與性體（一）》 [*Xinti yu Xingti (I)*], Taipei: Cheng Chung Bookstore, 1968: 455. A different perspective on Zhang Zai's interpretation, see Zhang Heng 張亨, 〈張載「太虛即氣」疏釋〉 [Zhang Zai Taixujiqi Shushi], *Bulletin of the Department of Chinese Literature, National Taiwan University* 臺大中文學報, vol. 3 (Dec., 1989): 1–44.

with multidimensional aspects,[7] and its meaning will differ depending on different viewpoints, which reminds us of the need to interpret *ji* against our present horizon: emptiness *ji qi*.

Following the interpretation given by Tang Chun-i 唐君毅(1909–1978),[8] an important figure of New Confucianism in modern Chinese philosophical movement, emptiness *ji qi* represents the way that Zhang Zai inherits the tradition of the *Yi Jing*, a very influential book for him, and the way in which Zhang Zai interprets *qi* as the metaphysical foundation of "interaction and resonance", the principle formulated by the *Yi Jing* and seen as its major subject. Such *qi* is the ultimate actuality—being only in the sense of its inclusiveness of emptiness, as opposed to negating it, in Zhang Zai's thought. Q*i* is the foundation making it possible to consider that myriad things are interacting with, but simultaneously opposing, each other internally, which means each thing is capable of interacting with the other based on its inherent emptiness, while each thing is containing inherent emptiness, because each one is actually one with the identity of *qi*, the *qi* fundamentally inclusive of emptiness.

Tang focuses on the structure of emptiness *ji qi*, both its vertical and horizontal dimensions, the clarification of which is to show how the logic of *ji* functions both inside one thing and in between the one and the other, and the key to understanding lies in a shift of horizon from the level of form and matter to the level of *qi.*[9] In Zhang Zai's view, based on his understanding of the *Yi Jing*, the innermost relation between myriad things should not be described as one conflicting with the other, but rather as one interacting with the other. A view of the former is based on the horizon limited by "form and matter 形質", while the latter is against a deeper horizon opened by *qi*, the actual being that makes all changes of form and matter possible. Indeed, one thing could not be viewed as the thing without its specific form and matter, and in this fact any implication of "conflict" is unavoidable because each thing possesses its form and matter which is different

---

[7] See Chan Wing-cheuk 陳榮灼, 〈「即」之分析——簡別佛教「同一性」哲學諸型態〉 [Ji zhi Fenxi: Jianbie Fojiao Tongyixing Zhexue Zhuxingtai], *The Annual of International Buddhistic Studies* 國際佛學研究年刊, vol. 1 (Dec., 1991): 1–22.

[8] The understanding of Zhang Zai's theory of *qi* is basically depending on the interpretation from Tang Chun-i, see Tang Chun-i 唐君毅, 〈張橫渠之以人道合天道之道〉 [Zhang Heng-Qu zhi yi Rendao he Tiandao zhi Dao], 《中國哲學原論—原教篇》 [*Zhongguo Zhexue Yuanlun: Yuanjiao Pian*], Taipei: Student Bookstore,1990: 72–120.

[9] See Tang Chun-i 唐君毅, 〈張橫渠之心性論及其形上學之根據〉 [Zhang Heng-Qu zhi Xinxinglun ji qi Xingshangxue zhi Genju], 《哲學論集》[*Zhexue Lunji*], Taipei: Student Bookstore, 1990: 219–224.

from the other one's; but on the other hand, a "thing" is not just the thing limited by its form and matter, but also the one with the identity of *qi*, *qi* with its inherent emptiness, such nature making it possible to encounter and interact with the other one, both of which exceed the limitation of their form and matter based on their own emptiness. Furthermore, if mutual-inclusiveness is revealed as a more profound relation between myriad things, differences of forms and matter should no longer be seen as the root of a sequence of conflicts, but rather as luxuriant expressions of *qi* with its consistent movement.

It is such a mutual-inclusiveness that *ji* implies against the background of *qi*, which has been philosophized by Zhang Zai and succeeded to by Wang Fu-zhi. What *ji* shows is not a kind of naivete or negligence of difference to avoid any potential conflicts. It emphasizes a way in which addressing challenges of conflict is possible and probably meaningful. In Zhang Zai's case, he uses emptiness *ji qi* to clarify conflicting notions especially concerning being and nothingness, actuality and emptiness and to reveal their shared metaphysical foundation; in Wang Fu-zhi's, he found that *ji* is the thread running through *Zhuangzi*'s thought as well, by means of which the existing ambiguous texts and conflicting dimensions could be further investigated depending on a deeper logical structure. The specific strategy has emerged through his interpretation of the centre of the ring, and such clarification not merely focuses on a concept called the centre of the ring or only aims to understand the true meaning of the sayings of Da-gong Diao, but also endeavors to head towards the horizon that supports different interpretations but belongs to neither of them. In the discussion on doing and non-doing, Wang Fu-zhi reminds us that both of them are right and wrong: neither of them achieves the marvelousness of the centre of the ring, but either of them still sheds light on one of the two irreplaceable elements.

As shown in his analysis of the centre of the ring, "the actuality of a ring" and "the emptiness inherent in the ring" are two fundamental elements, which means it would be difficult to grasp the essence of the centre of the ring if only focusing on one element but excluding the other. It is the same with tangled issues concerning being and nothingness, actuality and emptiness, doing and non-doing, and other conflicting aspects in or perspectives on *Zhuangzi.* Their differences contain more meaning than conflict, and such nuances would be revealed based on "mutual-inclusiveness" rather than "mutual-exclusiveness". Such an interpretative approach adopted by Wang Fu-zhi is illuminating. Instead of arguing that he, as a Confucian intellectual, interprets *Zhuangzi* in an ordinary Confucian way, it may be

more accurate to suggest that he focuses on such a possibility of furthering *Zhuangzi* with insight into *ji*, which is deeply associated with the philosophical tradition of *qi* and significant influence from Zhang Zai.

3. Paradoxical Unity

Rur-bin Yang 楊儒賓 makes use of a term "paradoxical unity 詭譎的同一", also within the specific context of *qi* formulated by Zhang Zai, to elaborate the logic of *ji* for the purpose of discovering the contemporary position that the philosophy of *Zhuangzi* should hold.[10] With respects to the legitimacy of arguing that *Zhuangzi* inherited the thought of the *Yi Jing*, Yang articulates a basic structure supporting *Zhuangzi*'s thought and enunciates it as "change and non-change present simultaneously 化與不化的同時具足".[11] This view is derived from a conversation between Confucius and Yanyuan in Chapter 22, "Zhi's Wandering in the North 知北遊",

> *The ancients, amid external changes, did not change internally; now-a-days men change internally, but take no note of external changes. When one only notes the changes of things, himself continuing one and the same, he does not change.* (James Legge trans.)
>
> 古之人，外化而內不化，今之人，內化而外不化。與物化者，一不化者也。[12]

In Yang's view, a question raised by this text is how to demystify the relation between change and non-change, external and internal, in a metaphysical sense. "Change and non-change present simultaneously" is not an answer to reply to a given question, i.e. a question such as whether the main focus should aim at "internal changes" or at "external changes", or which attitude is better than the other.

---

[10] For more discussion, see Yang Rur-bin 楊儒賓, 〈《易經》與理學的分派〉 [Yijing yu Lixue de Fenpai], 《從《五經》到《新五經》》 [*Cong Wujing dao Xinwujing*], Taipei: National Taiwan University Press, 2013: 279–322.

[11] See Yang Rur-bin 楊儒賓, 〈儒門內的莊子〉 [Rumennei de Zhuangzi], 《儒門內的莊子》 [*Rumennei de Zhuangzi*]: 154.

[12] See Guo Qing-fan 郭慶藩, 《莊子集釋》 [*Zhuangzi Jishi*], Beijing: Zhonghua Book Company, 2007: 765.

Instead, it is an attempt to reformulate such questions by means of inquiring into their constitutive structure that makes them possible and reasonable. Also, it is this fundamental structure that makes it possible to make a statement about "changing internally rather externally" or vice versa, statements that are made according to their approaches to a certain question that has determined by a specific horizon. "Change and non-change present simultaneously" is an articulation of this structure, not only focusing on a certain relation between two apparent conflicting aspects but also highlighting such a relation as fundamentally constituted by two aspects that are opposed to each other. Two conflicting aspects are, accordingly, not only contradictory to but also complementary to each other, with such an inherent relationship that makes their apparent conflicts possible and guarantees their irreplaceable individual identities at the same time. In other words, the situation of conflict is not sufficient to eclipse their inherent relationship, and such relationship never ignores one another's difference nor intends to reduce any potential conflict caused by differences. It is "paradoxical unity" that is used to explain such an implicit relationship.

This kind of unity does not show a tendency to equate one with the other; it rather implies a specific state of "neither the same nor the other".[13] The negation of "the same" emphasizes that the genuine unity needs going further into the depth of *qi* inherent in and shared by myriad things, as opposed to some statements that argue the meaning of unity can be well explained at the level of forms, or argue that *qi* can be clarified from a purely empirical viewpoint. "Not the other" describes such a profound relationship between things based on the horizon opened by *qi*, the principle of which is mutual-inclusiveness; more importantly, it further suggests that the genuine unity is the one acknowledging all differences, regarding them as concrete expressions of *qi,* not as conflicts from abstraction to be transcended. It is the mutual-inclusiveness that makes it possible for myriad things to be presented as such with their own form and matter. "Not the other" is fundamentally associated with "not the same", a unity with such a paradoxical condition that it neither stops at the level of apparent conflicts nor remains limited to a narrow interpretation of "actual being". This unity entails the logic of *ji*, which is the teaching embedded in the *Yi Jing* and furthered by Zhang Zai's clarification of *qi*.

According to Yang's understanding, paradoxical unity is the main thread running through all major issues of *Zhuangzi*, an obvious example is his particular

[13] See Yang Rur-bin 楊儒賓, 〈《易經》與理學的分派〉 [Yijing yu Lixue de Fenpai]: 296–297.

attention to a series of notions related to a metaphor of the "potter's wheel 陶均". He is of the view that the potter's wheel is the fundamental metaphor of *Zhuangzi* and intrinsically related to its expressions of other essential concepts, especially "heavenly wheel 天均", "undifferentiated heaven 渾天" and "the centre of the ring", all of which share its round shape, symbolizing completeness, embrace, and wholeness, and contain its theme of driving all movements of the universe. The theme of such internal energy could find its possible origin against a mythological background, but it also can be developed within a philosophical context in which posing questions on its own logic and structure would be essential.

Paradoxical unity could be seen as the thread to clarify questions arising from the metaphor of such forms that are "undifferentiated round 渾圓". Yang mentions such a fundamental metaphor and its relevant notions "with the essence that is the coalescence of permanence and changeableness, of absolute and relative",[14] they are in the relation of mutual-inclusiveness, not in the relation that one could be replaced by the other, the sense in which such wholeness is the one with paradoxical identity, and paradoxical unity aims to shed light on such heterogeneousness intrinsic to such an undifferentiated round form.

The focus of paradoxical unity is on its character of "round interpenetration", rather than on the "flattened sameness".[15] The former highlights irreplaceable individual identities and their vertical structure within as opposed to the latter's purely empirical horizon, and such an approach is basically related to Yang's understanding about *qi* for further examining the issue of "essence 體" and "function 用", an essential issue in the context of Confucianism and still under discussion today. It seems irrelevant to mention such an issue if one's concern is on *Zhuangzi* and with no interest in statements about "seeing *Zhuangzi* as Confucian"; however, if an approach to such an essential Confucian issue is associated with an insight inherent in *Zhuangzi*, whether *Zhuangzi* belongs to Confucianism or not, it already shows a possibility that *Zhuangzi* does not necessarily play a passive role in the prospect of "seeing *Zhuangzi* as Confucian". It could and would in turn occupy a constructive position in a contemporary dialogue with Confucianism.

Wang Fu-zhi's exegesis has revealed such a possibility and suggests some complexity *Zhuangzi* already possessed. Based on his interpretation in the last chapter "All Under Heaven 天下", the genuine relationship between essence and

[14] See Yang Rur-bin 楊儒賓, 〈莊子的卮言論〉 [Zhuangzi de Zhiyan Lun]: 245.
[15] See Yang Rur-bin 楊儒賓, 〈檢證氣學——理學史脈絡下的觀點〉 (On the Classification of "*Qixue*"), *Chinese Studies* 漢學研究, vol. 25.1 (Jun., 1997): 264.

function has been indicated by *Zhuangzi* through its fundamental notions, and such a relationship should not be understood as something like "following function yet abandoning essence 捐體而狥用", or "demonstrating essence yet making function void 立體以廢用", or like "analyzing essence and function as two independent parts 析體用而二之", or "identifying essence and function as the same 槩體用而一之", but rather it should be constructed that "essence is embodied in function and as non-essence 寓體于用而無體以為體". [16] The term *ji* does not appear in these sentences; however, it is obvious that all words are used to clarify the accurate meaning of *ji* that he grasps depending on his understanding of *Zhuangzi*.

Whether the philosophy of *Zhuangzi* would be detained at the horizon revealed by Confucian philosophers is an open question; nevertheless, *ji* with such paradoxical identity represents an approach to address the challenge of the conflict facing *Zhuangzi*, both in the aspects of text and of interpretation, in a positive way; it also, probably, serves as a starting point from which the meaning of *ji* could be deepened through consistent disclosure of *Zhuangzi*'s logical structure, the step from which a more profound conformation between *Zhuangzi* and Confucianism will emerge again.

4. Two Models of *qi*

In this paper, my focus is on a possibility of developing *Zhuangzi*'s logical structure by an approach surrounding the concept of *qi* and its logic of *ji* that has been demonstrated through Wang Fu-zhi's interpretation of *Zhuangzi*, and such an approach is deeply influenced by Zhang Zai's interpretation and has received further clarification from a contemporary philosophical view. A crucial issue on whether "seeing *Zhuangzi* as Confucian" can justify its argument with convincing reasons and evidence, however, would not be discussed in this paper, mainly because some preliminary issues are necessary while entering into such a discussion, which means some clarification of the question itself is crucial as well. More attention needs to be paid to issues such as which kind of Confucianism is discussed, which philosophical system of Confucianism is referred to, and which approach is adopted by such a system to claim the validity of "seeing *Zhuangzi* as Confucian".

[16] See Wang Fu-zhi 王夫之, 《老子衍 莊子通 莊子解》 [*Laoziyan Zhuangzitong Zhuangzijie*]: 353–354.

The approach to *qi* mentioned in this paper is such a kind of preliminary research, emphasizing that it would not be very appropriate to only regard it as the *qi* in a universal meaning or as the *qi* broadly defined by Confucianism as a whole, but rather, it contains a specific meaning to be revealed only following the interpretation given by Zhang Zai and his followers. And also, this clarification does not aim at proving or disapproving its legitimacy, but rather at introducing such an insight to open up the horizon already inherent in *Zhuangzi*. The horizon in which the logical structure of *Zhuangzi* could be reinvestigated, and the relation to *Laozi* philosophy could be subject to further critical inquiry, and the place in which the creativity of *Zhuangzi* philosophy could be discovered in a more profound way.

The legitimacy concerning the approach to *qi* is not further examined in our context, but a criticism of such an approach, however, is still mentionable as one of opportunities for elucidation of the types of *qi*: what is and is not the *qi* used here to formulate the statement on "seeing *Zhuangzi* as Confucian". A good place to open such discussion would be through the debates between Jean François Billeter and Rur-bin Yang on the issue of the following question: Is it appropriate to interpret *Zhuangzi* from the view of *qi,* suggesting *Zhuangzi* shared the same philosophical concern with the *Yi Jing*, which is the evidence for claiming "seeing *Zhuangzi* as Confucian"?[17] According to Billeter, the position that *qi* occupies within *Zhuangzi*'s context is questionable. First, although the term of *qi* does exist in *Zhuangzi*, this fact does not mean that it can be regarded as a key notion formulating a philosophical system, mainly because *Zhuangzi* philosophy is not a philosophy with continuity that allows us to evaluate it in historical perspective. Second, and more important, is the problem of the continuity of *qi*.

Billeter further explains, against the traditional background of *qi*, that since the Song dynasty the concept of *qi* has played an important role in providing a theoretical foundation for traditional Chinese philosophy, and it serves as the origin of each phenomenon in the universe. As the shared foundation, *qi* promises a fundamental continuity between all phenomena, the sense in which there is no true difference between the one and the other, the sense that everything is changeable

[17] Jean François Billeter (1939–), trans. Song Gang 宋剛, 〈莊子九札〉 (Nine Notes on Zhuangzi and Philosophy), *Newsletter of the Institute of Chinese Literature and Philosophy* 中國文哲研究通訊, vol. 22.3 (Sep., 2012): 13–15, and Yang Rur-bin 楊儒賓, 〈莊子與儒家——回應《莊子四講》〉 (Zhuangzi and Confucianism—A Response to Billeter's Lectures on Zhuangzi), *Newsletter of the Institute of Chinese Literature and Philosophy* 中國文哲研究通訊, vol. 22.3 (Sep., 2012): 137–141.

and reversible. In this context, it is impossible to allow any sense of "break", and it is therefore much more difficult to find a real starting point for creation. Owning to such intrinsic defect or imperfection, the theory of *qi* would not provide a convincing argument on issues concerning the freedom of subject or the potentiality of creativity, key issues that have been discussed and developed in the context of the philosophy of *Zhuangzi*. Given this background, Billeter centers the essentiality of paradox and discontinuity in *Zhuangzi*, regarding them as key elements for disclosing a new model of subject specifically from the perspective given by *Zhuangzi*. Issues of creation would have more appropriate discussion after such clarification of this new subject.

To address such challenges, Yang emphasizes a crucial distinction of two types of *qi* within Confucianism: *qi* with a pre-created model and with a post-created model,[18] basically related to their respective understandings of *qi*: the pre-celestial *qi* and the post-celestial *qi*, both of which are different from those meanings elaborated in a Daoist context. The focus of the post-created model is on its empirical characteristic of *qi*, showing a tendency towards naturalism or materialism, owing to which *qi* has been regarded as a less important issue for discussion in Neo-Confucianism. In contrast, the pre-created model of *qi,* the main topic of this paper, dedicates its effort to revealing the depth of myriad things within, and the depth we shared with others is the metaphysical foundation, making our individual identity possible. In Yang's view, the pre-created model of *qi* indicates a particular sense of "continuity". Such continuity is not claiming "undifferentiation"; instead, it claims the fundamental unity that necessarily includes and authenticates the specific differences belonging to each one, the sense in which without focusing on such continuity, a complete explanation of individuality or creativity would not be possible. Differences in appearance do not suffice to state individual identity, and also sameness in appearance do not illustrate any concrete unity. The continuity of *qi,* under the disclosure of its pre-created mode, suggests that kind of paradoxical unity, the unity which covers two dimensions both of "the one" and "the other" but encompasses neither of them. Its paradox presents as neither the same nor the other, as mentioned before. Such interpretation echoes Wang Fu-zhi's approach to *Zhuangzi* to a greater extent, and with respect to relations between myriad things he

[18] Yang Rur-bin 楊儒賓, 〈兩種氣學，兩種儒學〉 (Two Kinds of Ch'i Philosophy, Two Kinds of Confucianism), *Taiwan Journal of East Asian Studies* 台灣東亞文明研究學刊, vol. 3.2 (Dec., 2006): 1–39.

leaves an instructive comment: "the myriad things are uniting with others and none are independent of others 萬物合一而莫非獨".[19]

One further question would arise against the context of debates as to the appropriateness of the theory of *qi* while approaching *Zhuangzi* philosophy. Besides the possible diversity inherent in Confucianism and different connotations implied in the concept of *qi*, various understandings on continuity deserve more attention as well. Before making a judgment on whether continuity is the main concern of *Zhuangzi* or on whether *Zhuangzi* represents a philosophy with continuity, the meaning of continuity and its structure should be taken into careful consideration. A successor to the philosophical thought formulated by Zhang Zai's theory of *qi* has appeared and has kept developing understanding of his concepts, aside from such a Chinese philosophical context.

Nishida Kitarō 西田幾多郎 (1870–1945), the founder of the Kyoto school, also expressed views particularly on the "continuity of discontinuity 非連続の連続", one of his fundamental concepts to demonstrate the logic of "self-identity of absolute contradiction 絶対矛盾的自己同一". Their specific meanings and related approaches towards *Zhuangzi* are beyond the scope of the present paper,[20] but this case might contribute to discovering the value of "seeing *Zhuangzi* as Confucian". These concepts of Nishida mean dialogues related to *ji*, paradox, and continuity are no longer limited to the field of Chinese philosophy. And so *Zhuangzi* further extends its possibility to an encounter with Japanese philosophy, not based on any possible relevance traced to seemingly historical factors, but mainly based on the capabilities for developing their own concepts consistently. In the case of *Zhuangzi*, such possibility is revealed and much clarified with the help of a "Confucian" interpretation, the sense in which the view of "seeing *Zhuangzi* as Confucian" would be essential for the furtherance of *Zhuangzi* philosophy and as one constitutive element of such furtherance.

References

[19] See Wang Fu-zhi 王夫之, 《老子衍 莊子通 莊子解》[*Laoziyan Zhuangzitong Zhuangzijie*]: 253.

[20] I have dealt with these issues by investigating a profound meaning of "nothingness". See Liu Kuan-ling 劉冠伶, 《《莊子》思想的「環中」構造——「絕對無」的線索》 (*The structure of Huanzhong in Zhuangzi's thought: in view of absolute nothingness*), (Ph. D. diss., National Tsing Hua University, 2016): 70–94.

Billeter, Jean François (2012) 〈莊子九札〉 (Nine Notes on Zhuangzi and Philosophy). Trans. Song Gang 宋剛, *Newsletter of the Institute of Chinese Literature and Philosophy* 中國文哲研究通訊, 22.3, 5–39.

Chan, Wing-cheuk 陳榮灼 (1991) 〈「即」之分析——簡別佛教「同一性」哲學諸型態〉 [Ji zhi Fenxi: Jianbie Fojiao Tongyixing Zhexue Zhuxingtai] *The Annual of International Buddhistic Studies* 國際佛學研究年刊, 1, 1–22.

Guo, Qing-fan 郭慶藩 (2007) 《莊子集釋》 [*Zhuangzi Jishi*] (Beijing: Zhonghua Book Company).

Liu, Kuan-ling 劉冠伶 (2016) 《《莊子》思想的「環中」構造——「絕對無」的線索》 (*The structure of Huanzhong in Zhuangzi's thought: in view of absolute nothingness*). Unpublished doctoral dissertation, National Tsing Hua University, Taiwan.

Mou, Tsung-san 牟宗三 (1968) 《心體與性體(一)》 [*Xinti yu Xingti (I)*] (Taipei: Cheng Chung Bookstore).

Tang, Chun-i 唐君毅 (1990) 《中國哲學原論—原教篇》 [*Zhongguo Zhexue Yuanlun: Yuanjiao Pian*] (Taipei: Student Bookstore).

Tang, Chun-i 唐君毅 (1990)《哲學論集》 [*Zhexue Lunji*] (Taipei: Student Bookstore).

Wang, Fu-zhi 王夫之 (2009)《老子衍 莊子通 莊子解》[*Laoziyan Zhuangzitong Zhuangzijie*] (Beijing: Zhonghua Book Company).

Yang, Rur-bin 楊儒賓 (1997) 〈檢證氣學——理學史脈絡下的觀點〉 (On the Classification of "*Qixue*"). *Chinese Studies* 漢學研究, 25.1, 247–281.

Yang, Rur-bin 楊儒賓 (2004) 〈儒門別傳——明末清初《莊》《易》同流的思想史意義〉 [Rumen Biezhuan: Mingmoqingchu Zhunagyitongliu de Sixiangshi Yiyi], in Chung Tsai-chun 鍾彩鈞 and Yang Chin-lung 楊晉龍 (eds.) 《明清文學與思想中之主體意識與社會——學術思想篇》 [*Mingqing Wenxue yu Sixiang zhong zhi Zhutiyishi yu Shehui: Xueshusixiang Pian*] (Taipei: Institute of Chinese Literature and Philosophy, Academia Sinica), 245–289.

Yang, Rur-bin 楊儒賓 (2006) 〈兩種氣學，兩種儒學〉 (Two Kinds of Ch'i Philosophy, Two Kinds of Confucianism). *Taiwan Journal of East Asian Studies* 台灣東亞文明研究學刊, 3.2, 1–39.

Yang, Rur-bin 楊儒賓 (2012) 〈莊子與儒家——回應《莊子四講》〉 (Zhuangzi and Confucianism—A Response to Billeter's Lectures on Zhuangzi). *Newsletter of the Institute of Chinese Literature and Philosophy* 中國文哲研究通訊, 22.3, 137–141.

Yang, Rur-bin 楊儒賓 (2013) 《從《五經》到《新五經》》 [*Cong Wujing dao Xinwujing*]. (Taipei: National Taiwan University Press).

Yang, Rur-bin 楊儒賓 (2016)《儒門內的莊子》[*Rumennei de Zhuangzi*]. (Taipei: Linking Publishing).

Zhang, Heng 張亨 (1989)〈張載「太虛即氣」疏釋〉[Zhang Zai Taixujiqi Shushi]. *Bulletin of the Department of Chinese Literature, National Taiwan University* 臺大中文學報, 3, 1–44.

The Similarity between Buddhist Logic and Assertion Theory: Exclude Pakṣa and Context

Liu Chi Yen
National Yang Ming University / Kyoto University

***Abstract**: When scholars try to compare Dignāga's Buddhist logic with Western logic, most of them take Aristotle's syllogism as the paradigm—since they propose that the canonical argument for Dignāga is a deductive argument. However, some scholars argue against this interpretation, they claim that the canonical argument cannot be a deductive one because of exclude pakṣa. In this paper, I suggest that exclude pakṣa of Dignāga's Buddhist logic are compatible with deduction from the contextual point of view.*

The canonical argument for Dignāga is:

Thesis:	*Sounds are impermanent.*
Reason:	*Because of being produced.*
Examples:	***Similar corroboration & instance:*** *what is produced is observed to be impermanent, like a pot.*
	Dissimilar corroboration & instance: *what is permanent is observed not to be produced, like space.*

The reason explains why the thesis should be accepted, and to be the right reason requires 3 specific criteria, which are called "tri-rūpa-hetu". It is believed that after satisfying tri-rūpa-hetu, we would get two universal statements as premises to derive the thesis. Thus, the canonical argument is usually reconstructed as

Major premise: *All things that are produced are impermanent.*
Minor premise: *All sounds are produced.*
Conclusion: *All sounds are impermanent.*

Then it seems that the canonical argument is deductive.

But exclude pakṣa makes things complicated. Pakṣa is the subject of the thesis, and exclude pakṣa means that when providing examples to satisfy the 2nd and 3rd rūpa, instances and corroborations cannot contain pakṣa. This provision triggers

a major disputation among contemporary Chinese Buddhist logicians. Some scholars therefore claim that exclude pakṣa keeps Dignāga's Buddhist logic out of deductive arguments.

This paper argues for the opposite. By demonstrating how to formalize Buddhist logic with the use of symbolic logic, particularly predicate logic, I explain why some scholars claim that exclude pakṣa would keep the canonical argument out of deduction. To solve the problem, I reveal the similarity between Buddhist logic and Stalnaker's assertion theory, in which exclude pakṣa relates to the domain of discourse. Finally, I provide more detail about the role of exclude pakṣa and explain why it does not compromise the deductive power of the canonical argument.

1. Introduction

Consider a canonical argument for Dignāga mentioned by S. Katsura below (cf. Katsura 2004, 143):

Thesis:	Sounds are impermanent.
Reason:	Because of being produced.
Examples:	**Similar corroboration & instance:** what is produced is observed to be impermanent, like a pot.
	Dissimilar corroboration & instance: what is permanent is observed not to be produced, like space.

Table 1.

This way of argumentation is also called "3-membered argument", for it has 3 components: thesis, reason, and example. The thesis is the conclusion of the argument, which can be divided into a subject (*pakṣa*) and a predicate (*sādhyadharma*). Theoretically, every statement can be divided into a subject and a predicate, and for predicate logic, we can also formalize a subject as a predicate. But for Buddhist logic, the thesis's truth and falsity is the most important concern, so it has a special terminology "pakṣa" for the subject of the thesis, and another terminology "sādhyadharma" for the predicate of the thesis. This paper uses these two terminologies to avoid misunderstandings.

The reason explains why the thesis should be accepted, and to be the right reason requires three specific criteria, which are called "*tri-rūpa-hetu*". Tri-rūpa-hetu is a requirement for a proper *reason* in Dignāga's Buddhist logic, which are necessary conditions for a good argument. Hetu means reason, rūpa can be understood as form, and tri-rūpa-hetu denotes three forms that a reason should have. The first rūpa is that the reason all occurs in the pakṣa. As we see in the canonical argument, the reason "being produced" satisfies this rūpa, because sounds are all produced. To understand the other two rūpa further, it would be helpful to know what example is. Before Dignāga, examples are just instances to support the reason, and they can be divided into similar ones and dissimilar ones. Similar instances are those which have the same reason and sādhyadharma as the pakṣa. In the canonical argument, a pot is a similar instance because it is both produced and impermanent. Conversely, dissimilar instances do not have the same reason and sādhyadharma as the pakṣa. Space is usually used as a dissimilar instance in the canonical argument because it is permanent and not produced as Abhidharma school claims.

As we see in the canonical argument, besides instances, Dignāga adds two more elements into the category of examples, the universal statements "similar corroboration" and "dissimilar corroboration". In the canonical argument, the similar corroboration means all produced things are impermanent, and the dissimilar corroboration means all permanent things are not produced. This move is considered as a substantial improvement in Buddhist logic in that it makes the canonical argument appear to be deductive. But how can we get these two corroborations? The answer lies in the other two rūpa.

The second rūpa is that the reason must occur in a similar kind of sādhyadharma. Here, similar kind refers to things having the same sādhyadharma as the pakṣa. Thus, this rūpa requires that something possesses the same reason and sādhyadharma as the pakṣa must exist, that is to say, at least one similar instance must be provided. As we already have seen, a pot is a similar instance of both being impermanent and produced, so the 2nd rūpa is satisfied in the canonical argument.

The third rūpa is that the reason cannot be found in the dissimilar kind of the sādhyadharma. This means all things which do not possess the same sādhyadharma as the pakṣa cannot have the reason either. Now we can see that to provide the dissimilar corroboration is to satisfy the 3rd rūpa.

So far, we have shown how a similar instance and dissimilar corroboration relate to the 2nd and 3rd rūpa, but how about the dissimilar instance and similar

corroboration? Intriguingly, Dignāga explicitly says that it is unnecessary to provide the dissimilar instance for the argument, but he does not provide enough explanation for this. For the similar corroboration, it is believed that Dignāga is inclined to derive the similar corroboration from the dissimilar corroboration, and this means that Dignāga needs at least one direction of contraposition (Matilal 1986).

Whether Dignāga should admit contraposition or not is still debatable, but this paper focuses on one special provision Dignāga seems to require: exclude pakṣa (from examples).[1] Exclude pakṣa says that when providing examples to satisfy the 2nd and 3rd rūpa, our instance and corroboration cannot contain pakṣa—the subject of the thesis, and this provision provokes a major disputation among contemporary Chinese Buddhist logicians (Shen 1994; Yao 1990; Zheng 1990). Some scholars claim that exclude pakṣa keeps Dignāga's Buddhist logic out of deductive arguments. The reasoning is that under the constraint of exclude pakṣa, similar corroboration cannot be a universal statement. When discussing Dignāga's Buddhist logic, Richard Hayes distinguishes two domains: the domain of the subject and the induction domain (Hayes 1988). The domain of the subject is the domain of pakṣa; the induction domain is the domain without pakṣa. Hayes' question is: while both similar corroboration and dissimilar corroboration apply to the induction domain, is this sufficient to make a universal statement including pakṣa? Hayes says: "To this question it is clear that we must give a negative reply" (Hayes 1988, 122). Zheng further argues that since exclude pakṣa makes it impossible for similar corroboration to be a universal statement, the canonical argument cannot be a deductive one (Zheng 1990).

Despite of Zheng's' argument, this paper suggests the opposite, which is that exclude pakṣa and deduction are compatible from the contextual point of view. Therefore, it is crucial to show that exclude pakṣa does not prevent the similar corroboration from being a universal statement. To achieve my goal, I would demonstrate how to formalize Buddhist's logic from symbolic logic in section 2, particularly from predicate logic. As soon as a precise formalization is obtained, I would discuss why some scholars claim that exclude pakṣa would keep the canonical argument out of deduction. In section 3, I reveal the similarity between Buddhist logic and Stalnaker's assertion theory and how exclude pakṣa relates to it, particularly to the domain of discourse. Finally, in section 4, more details about the

[1] The original text is Chinese "除宗有法". Interestingly, Dignāga never explicitly mentions this provision. As far as we know, it was first mentioned in a Chinese text—因明入正理論疏(*Commentary on Nyāyapraveśa*)—written by Wengui 文軌.

role of exclude pakṣa and the reason why it does not compromise the deductive power of the canonical argument would be provided.

2. Formalization of Dignāga's Buddhist Logic

In an attempt to compare Buddhist logic with Western logic, most scholars take Aristotle's syllogism as the paradigm. Thus, they reconstruct the canonical argument as:

> **Major premise:** All things that are produced are impermanent.
> **Minor premise:** All sounds are produced.
> **Conclusion:** All sounds are impermanent.

As we have shown in the introduction, to satisfy the 1st rūpa is to give us the minor premise. Satisfying the 3rd rūpa gives us the dissimilar corroboration, and applying contraposition to it can derive the similar corroboration, that is, the major premise. From this point of view, Dignāga's tri-rūpa-hetu tries to give us a deductive argument, or more precisely, a valid and sound argument. Astute readers may wonder about the role of 2nd rūpa now. I discussed this issue in another unpublished paper, but due to being irrelevant to the main point of this paper, we need to skip this.

However, some scholars may complain about the above reconstruction because of the limited expressive power of Aristotle's logic (Ho 2002). So, I suggest a reconstruction of the canonical argument from predicate logic, and this can help us to formalize it under the constraint of exclude pakṣa latter, which Aristotle's logic cannot. Let S_x means x is a sound; E_x means x is permanent; P_x means x is produced. In this setting, we can reconstruct the canonical argument roughly as follows:[2]

1. $(x)(S_x \supset P_x)$	premise, by 1st rūpa
2. $(x)(E_x \supset \neg P_x)$	premise, by 3rd rūpa
3. $(x)(P_x \supset \neg E_x)$	from 2, by contraposition
4. $(x)(S_x \supset \neg E_x)$	from1 & 3, by hypothetical syllogism

[2] Some inference rules relating to quantifiers are omitted from this proof: universal instantiation and universal generalization.

To make things simpler, I use material conditionals to formalize general statements. Thus, $(x)(S_x \supset P_x)$ means "sounds are produced", $(x)(E_x \supset \neg P_x)$ means "permanent things are not produced", $(x)(P_x \supset \neg E_x)$ means "produced things are impermanent", and $(x)(S_x \supset \neg E_x)$ means "sounds are impermanent.

But things get more complicated if we consider the requirement of exclude pakṣa that Dignāga asks. Exclude pakṣa demands that when testing whether the reason satisfies the 2nd and 3rd rūpa, pakṣa cannot be taken as the similar or dissimilar kind. Thus, some scholars claim that exclude pakṣa makes both similar and dissimilar corroborations not general statements anymore, and that would compromise the deductive reasoning of the canonical argument.

If exclude pakṣa is a semantic requirement for the 3rd rūpa, the meaning of the 3rd rūpa would be: the reason cannot be found in things which are the dissimilar kind of the sādhyadharma but not pakṣa. Hence, even though satisfying the 3rd rūpa, the dissimilar corroboration itself would not be related to pakṣa at all. The similar corroboration, which is derived from the dissimilar corroboration, is not related to pakṣa either, and it also means that the similar corroboration in the canonical argument cannot be a general statement.

Following this line of thought, the canonical argument should be formalized as

1. $(x)(S_x \supset P_x)$	premise, by 1st rūpa
2*. $(x)((E_x \wedge \neg S_x) \supset \neg P_x)$	premise, by 3rd rūpa
2.1. $(x)((\neg S_x \wedge E_x) \supset \neg P_x)$	from 2*, by commutation
2.2. $(x)(\neg S_x \supset (E_x \supset \neg P_x))$	from 2.1, by exportation
2.3. $(x)(\neg S_x \supset (P_x \supset \neg E_x))$	from 2.2, by contraposition
3*. $(x)((\neg S_x \wedge P_x) \supset \neg E_x)$	from 2.3, by exportation
4. $(x)(S_x \supset \neg E_x)$	from?

Now the dissimilar corroboration is reformulated as 2*, which means "permanent things except sounds are not produced". As I show, the similar corroboration 3*, which means "produced things except sounds are impermanent", can be derived from 2*.[3] Hence, the problem of exclude pakṣa is not that Dignāga cannot derive the similar corroboration from the dissimilar corroboration (Zheng 1990), but rather that we can easily conceive that even if premises 1 and 2* are true, conclusion 4 could still be false.

[3] Claus Oetke has a similar formalization (cf. Oetke 1996, 472–473).

This means that under the constraint of exclude pakṣa, to satisfy tri-rūpa-hetu would not make the canonical argument a deductive one. This is a fair objection if exclude pakṣa is a semantic constraint. However, inspired by Stalnaker's account of assertion, I argue that the exclude pakṣa is not a semantic constraint but rather a pragmatic one. The similarity between Stalnaker's assertion theory and Buddhist logic is hardly noticed by contemporary Buddhist logicians (Chen 2006; Ho 2002; Katsura 1996; Matilal 1986; Mohanty 1992; Shen 1994; Yao 1990; Zheng 1990), and I would explore this in the next section.

3. Assertion

The role of the thesis in Buddhist logic reminds us of the role of assertion elaborated by Stalnaker (Stalnaker 1999). For Buddhist logic, the thesis should be accepted by the proponent but not accepted by the opponent.[4] During the debate, the proponent provides the reason which satisfies tri-rūpa-hetu to persuade the opponent of the thesis. This kind of process was surprisingly captured by Stalnaker's assertion theory.

For Stalnaker, to make an assertion is to assert a proposition against a context. A proposition is a set of possible worlds, that is, the set of possible worlds where the proposition is true. Here, let us interpret possible worlds as our epistemic states: the current state of our knowledge about the actual world. The context is a set of possible worlds in which we make assertions as the background, and it has two parts: presupposed propositions and worlds which are compatible with presupposed propositions. Presupposed propositions are known, believed, or assumed for the conversation by participants, and it means that they are true in all the possible worlds in the context set.

In other words, a context is speakers' presuppositions, which can be represented as follows:

[4] This is what Dignāga calls "不顧論宗 (thesis regardless of the opponent's stance)" in his *Hetuvidyā-nyāya-dvāra-śāstra* 因明正理門論.

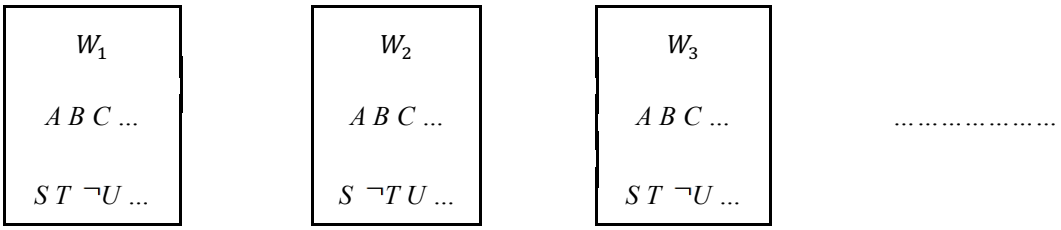

Figure 1.

The middle line in figure 1 is the presupposed propositions: *A*, *B*, *C*, etc., and as we see it, they should be true in all worlds in the context. The bottom line is propositions that are compatible with the presupposed ones, and their truth values may be different in different worlds. The function of making an assertion is "to reduce the context set in a particular way, provided that there are no objections from the other participants in the conversation" (Stalnaker 1999, 86). Ideally, Figure 1 can be represented as our knowledge states, and to reduce the context set means that we eliminate our ignorance and know more about the actual world.

To fulfill the goal of assertion, as Stalnaker demands: "A proposition asserted is always true in some but not all of the possible worlds in the context set" (Ibid., 88). This means that we should not assert propositions which are incompatible with presupposed propositions, because "one wants to reduce the context set, but not to eliminate it altogether" (Ibid., 89). We should not assert presupposed propositions either, because "to assert something which is already presupposed is to attempt to do something that is already done" (Ibid.). Interestingly, these accounts of assertions perfectly explain the requirement of the thesis in Buddhist logic.

For Buddhist logic, the thesis is accepted by the proponent but rejected by the opponent. The purpose of the proponent to establish the thesis is to persuade the opponent to accept it, and in that sense, they are just like participants in a conversation with their context. Just like an assertion, the thesis should be true in some but not all of the worlds in the context of a conversation, and that is why to establish a thesis agreed or rejected by both parties is to commit a fallacy for Buddhist logic.

To persuade the opponent, the proponent must show the reason he provides satisfying tri-rūpa-hetu. To satisfy a rūpa is to make an assertion to successfully eliminate possible situations which are not compatible with it. In the end, ideally, the thesis would be true in all of the worlds in the context set. We would see this in more detail in the next section. This is a dynamic process, as Stalnaker says: "A

conversation is a process taking place in an ever-changing context" (Ibid., 86); "The context—what is presupposed in it—is constantly changing as things are said" (Ibid., 101). This idea of "ever-changing context" may also relate to exclude pakṣa, because it relates to universal statements.

It is well known that the interpretation of a universal statement depends on the domain of discourse. When I say "The burglar took everything", I do not mean everything in the world, of course; only a specific domain is relevant, say, all the valuable objects in my house. Likewise, the subject in "all produced things are permanent" may refer to different things based on different domains. François Recanati proposes that we can also apply the ever-changing aspect of context to the domain of discourse. He says:

> It is therefore to be expected that the domain of discourse itself can change in mid-utterance. This means that there can be more than one domain, more than one 'situation', corresponding to a given utterance. (Recanati 1996, 454)

This gives us a hint that during the reasoning process of Buddhist logic, we do not have to stick to a fixed domain of discourse. This opens up the possibility to interpret exclude pakṣa differently from previous research.

As we have discussed in section 2, it seems that to satisfy the 3rd rūpa will not give us a universal statement due to exclude pakṣa. Now, we can see that this comes from the assumption that pakṣa is in the domain of discourse which is fixed during the reasoning process; however, Recanati reminds us that it may not be the case. In other words, we do not have to suppose that pakṣa is always in the domain of discourse when considering different rūpa. Thus, it is arguable that to satisfy the 3rd rūpa could give us a universal statement, which is true relative to a domain without pakṣa.

The reader may wonder that if the domain of discourse is not always the same for different universal statements, how can they be put together to form an argument? After all, we need to evaluate an argument based on the same domain. Consequently, we need a new perspective to see the reasoning process in Buddhist logic. Let us see how this can work in the next section.

4. What Kind of Reasoning?

What is the role of exclude pakṣa? While Daqi Chen thinks it is to avoid circular argument (Chen 2006), I suggest it is to avoid begging the question. Buddhist logic requires that what is similar and dissimilar kind must be commonly recognized by both parties concerning the topic in question, therefore, to which kind pakṣa belongs cannot be commonly recognized by default. Under this constraint of exclude pakṣa, Dignāga tries to establish a universal statement without mentioning pakṣa. I try to explain how Dignāga can do this.

Let us consider the canonical argument again. We can image an initial situation such that the proponent believes that sounds are impermanent, but the opponent holds that sounds are permanent. Suppose that the proponent knows that all produced things are impermanent and sounds are produced—but his opponent is not aware of this yet—the proponent provides "being produced" as the reason. To convince his opponent of the similar corroboration, the proponent needs to start from a smaller domain of discourse without sounds, otherwise the opponent would reject the universal statement immediately. Intriguingly, Buddhist logicians propose to achieve this goal by satisfying the 3rd rūpa instead.

As more objects are demonstrated to satisfy the similar corroboration, it becomes more and more plausible. Finally, after investigating all the objects in the domain, the proponent can convince the opponent that all produced things are impermanent. Now, by reminding the opponent that all sounds are produced, since the 1st rūpa is satisfied, the proponent brings sounds into the domain of discourse. Should the opponent still believe the similar corroboration after this expansion? Normally he should, unless he can provide a really good reason rejecting it. This may look like Katsura's idea (cf. Katsura 1996, 12), but since Katsura does not consider the problem of exclude pakṣa, it is unlikely he would consider the domain of discourse may change during the reasoning process.

Another important aspect requires to be clarified is the role of premises in a deductive argument. When Katsura discusses the canonical argument, he claims that it is "fundamentally the results of an Inductive Reasoning" (Ibid., 8). We should say that the premises are fundamentally the results of induction instead. To determine whether an argument is deductive or not, our concern is only whether the truths of the premises can guarantee the truth of the conclusion. How the truth of a premise is obtained is irrelevant to its role in a deductive argument.

For contemporary Chinese Buddhist logicians, exclude pakṣa is the main reason why the canonical argument is not a deductive one. According to Zheng the similar corroboration cannot be a universal claim because of exclude pakṣa, thus the

canonical argument cannot be a deductive argument (Zheng 1990). On the contrary the opposite camp insists that the similar corroboration is a universal statement (Shen 1994; Yao 1990), since Dignāga never explicitly claims that we should exclude pakṣa from the similar corroboration. However, they both presuppose a fixed domain of discourse and this is what this paper tries to point out: We do not have to consider the whole reasoning process based on a fixed domain. This may help Shen and Yao to defend their position.

In other words, exclude pakṣa is just a pragmatic strategy. It suspends pakṣa in the beginning of the conversation for the proponent to make the reasoning easier or even possible so as to establish a general statement to convince the opponent in the second stage. In the final stage, the proponent can then bring the main subject into our conversation, and in this sense, the two premises—sounds are produced and all produced things are impermanent—share the same domain of discourse now.

Some scholars suggest Indian logic is a kind of logic of cognitions (Mohanty 1992, 130), or logic of knowledge acquiring (Ho 2002, 32). But how does it work? Unfortunately, the details are never provided. Here I try to connect Buddhist logic to context, and by that we can see how it relates to epistemic states of participants in a debate or conversation. This gives us a new perspective to see the relation between Buddhist logic and contemporary Western logic, and that may help us to understand Mohanty and Ho's perspective.

For example, Stalnaker notices that if we take context into consideration, we may evaluate inferences from a different angle. Stalnaker has an interesting account for this:

> An inference from a sequence of assertions or suppositions (the premises) to an assertion or hypothetical assertion (the conclusion) is reasonable just in case, in every context in which the premises could appropriately be asserted or supposed, it is impossible for anyone to accept the premises without committing himself to the conclusion. (Stalnaker 1975, 138)

Though this account is proposed to deal with inferences involving indicative conditionals, it can be seen as a general account of our reasoning regarding contexts.

According to Stalnaker, some reasonable inferences involving conditionals are not deductive arguments, but almost all deductive arguments are reasonable inferences. In this paper I argue that if we analyze the canonical argument from the contextual point of view, exclude pakṣa just indicates that the domain of discourse is

flexible. Thus, exclude pakṣa would not keep the canonical argument out of deductive one.

5. Conclusion

This paper reveals that the interpretation of Dignāga's exclude pakṣa in Buddhist logic relies on how we consider the domain of discourse. If we interpret it based on a fixed domain of discourse as some scholars does, we cannot insist that the similar corroboration is a universal statement anymore. The spirit of Buddhist logic has a surprising similarity to Stalnaker's account of assertion, which is ignored by contemporary scholars. This reminds us that Buddhist logic actually evolves from debating strategy, and we should take context into consideration. By doing this, we should interpret exclude pakṣa based on a flexible domain of discourse changing with context, as Recanati suggests.

This paper suggests the role of exclude pakṣa is to suspend the main subject from the debate to avoid unnecessary dispute in the beginning. Dignāga never used the term "exclude pakṣa" in his remarkable work *Hetuvidyā-nyāya-dvāra-śāstra*, though this idea occurred in his discussion of tri-rūpa-hetu. However, this could only indicate that Dignāga does not consider pakṣa at this stage at all, because it is completely removed out from the domain of discourse. But after showing the reason can satisfy *tri-rūpa-hetu*, the pakṣa can be brought back into the domain. And this kind of everchanging context in the conversation may just be too common in their practice, so they do not have to mention or explain it at all.

Before taking context into account for exclude pakṣa, it is difficult to see the dynamic aspect of Buddhist logic. It is believed that dynamic semantics is motivated by Stalnaker's assertion theory, and I try to point out that this kind of idea is hidden in Dignāga's Buddhist logic. The role of exclude pakṣa is also a noteworthy aspect that may present real situation when we try to convince people who disagree with us. How to develop a logic model to capture this process would be very interesting. In sum, I hope that this would indicate the strong connection between Buddhist logic and contemporary Western logic, and by that they may benefit from each other.

Acknowledgments

I would like to thank three anonymous reviewers and the editor for their careful reading and inspiring comments. I also thank Ministry of Science and Technology of Taiwan (108-2911-I-010-507) and Major Program of National Social Science Foundation of China (18ZDA031) for support.

References

Chen, Daqi 陈大齐 (2006) *Observations on the Great Commentary on Hetuvidyā* 因明大疏蠡测. (Bejing: Zhonghua Book Company).

Hayes, Richard (1988) *Dignāga on the interpretation of signs.* (Studies of Classical India: 9) (Dordrecht: Kluwer).

Ho, Chien-hsing 何建興 (2002) On Dignāga's logic. *Journal of the Center for Buddhist Studies* 佛學研究中心學報, 7, 27–47.

Katsura, Shoryu (1996) *How did the Buddhists prove something? The nature of Buddhist logic. The Numata Yehan Lecture in Buddhism.* (Calgary: The University of Calgary).

Katsura, Shoryu (2004) The role of dṛṣṭānta in Dignāga's logic, in Shoryu Katsura and Ernst Steinkellner (eds.) *The role of the example (dṛṣṭānta) in classical Indian logic*, (*Wiener Studien zur Tibetologie und Buddhismuskunde*: 58) (Vienna: Arbeitskreis für Tibetische und Buddhistische Studien), 135–174.

Matilal, Bimal K. (1986) Buddhist Logic and Epistemology, in Bimal K. Matilal and Robert D. Evans (eds.) *Buddhist Logic and Epistemology*, (D. Reidel, Dordrecht), 1–30.

Mohanty, Jitendra N. (1992) *Reason and Tradition in Indian Thought*. (New York: Oxford University Press).

Oetke, Claus (1996) Ancient Indian logic as a theory of non-monotonic reasoning. *Journal of Indian Philosophy*, 24, 447–539.

Recanati, François (1996) Domains of discourse. *Linguistics and Philosophy*, 19 (5), 445–475.

Shen, Jian-ying 沈劍英 (1994). *Buddhist logic* 佛家邏輯. (Taipei: Shinning Culture).

Stalnaker, Robert (1968) A theory of conditionals, in Nicholas Rescher (ed.) *Studies in Logical Theory, American Philosophical Quarterly monograph series, 2*, (Oxford: Basil Blackwell), 98–112.

Stalnaker, Robert (1975) Indicative Conditionals. *Philosophia*, 5, 269–286.

Stalnaker, Robert (1999) *Context and Content: Essays on Intentionality in Speech and Though*. (Oxford: Oxford University Press).

Yao, Nan-qiang 姚南強 (1990) On the essence of Dignāga's 3-membered logic 论陈那三支式的逻辑本质. *Fudan Journal (Social Sciences Edition)*, 6, 40–43.

Zheng, Wei-hong 鄭偉宏 (1990) Is Dignāga's new Hetuvidya deductive argument? 陳那新因明是演繹論證嗎？ *Nei Ming* 內明, 216, 27–34.

Hetuvidyā-nyāya-dvāra-śāstra《因明正理門論》, CBETA, Vol. 32, no. 1628.

Commentary on Nyāyapraveśa《因明入正理論疏》, CBETA, X53

One Negation, Two Ways of Using It: Prasajyapratiṣedha in Bhāviveka and Candrakīrti's Argumentation[1]

Chen Hsun-Mei
National Taiwan University / Kyoto University / Harvard Yenching Institute
Wang Wen-Fang
Professor, National Yang Ming University

***Abstract**: The famous Madhyamaka philosopher Nāgārjuna has a very special way of arguing against his opponent: he often argues and concludes that a certain thesis of his opponent should be rejected while at the same time denies that he has therefore endorsed the negation of the thesis of his opponent. This special way of argumentation had a tremendous influence upon later Indian Buddhist philosophers but has invoked two different interpretations, the Prāsaṅgika and the Svātantrika, about what exactly this special way of arguments is. While Svātantrika proposes that one should establish full hetuvidya syllogisms in accord with the ultimate truth to rebut opponents, Prāsaṅgika proposes that the only legitimate Madhyamaka way of argumentation is to deny opponents' theses by merely indicating their absurdities. Strikingly, both Svātantrika and Prāsaṅgika emphasize that the negation used in the Madhyamaka arguments should be* prasajyapratiṣedha*. In this paper, the authors explore how Bhāviveka and Candrakīrti, two representors of Svātantrika and Prāsaṅgika respectively, had different understandings of* prasajyapratiṣedha*. Based on textual evidences and philosophical analysis, the authors argue that while Bhāvivek's way of using* prasajyapratiṣedha *in a Svātantrika argument would commit him to a certain conclusion, Candrakīrti's way of using prasajyapratiṣedha in a Prāsaṅgika argument would not have the same effect.*

From a rational reconstruction, the authors then propose that a Mādhyamika, one who advocates Madhyamaka ideas, should simply reject his opponent's thesis by drawing absurdity from it and should at the same time refrain from making any conclusion, and therefore the Prāsaṅgika interpretation is the right Nāgārjuna way of argumentation as shown in Mūlamadhyamakakārikā*. In the end, based on Prāsaṅgika way of argumentation, the authors will provide a formal regimentation of Nāgārjuna's rebuttal.*

[1] This paper was supported by Harvard Yenching Institute and through New Partnership Program for the Connection to the Top Labs in the World by Ministry of Science and Technology, grant no. 106-2911-1-010-512.

Introduction

Prāsaṅgika and Svātantrika, two major Madhyamaka schools, advocated separately two distinct ways of rebutting an opponent, and each of them claimed that their way was the approach that Nāgārjuna used in his masterpiece *Mūlamadhyamakakārikā* (hereafter: MMK*)*.[2] Ostensibly, the main difference between the two schools is about whether a Mādhyamika should argue for and assert his/her affirmative/negative thesis either before or after repudiating his/her opponent's thesis. While the Prāsaṅgika uses a method that we will call "*prasaṅga*"[3] (simply pointing to a consequence of the opponent's thesis that the opponent is unwilling to accept without arguing for or making a conclusion) as a special way to rebut the opponent, the Svātantrika, on the other hand, always argues for and asserts the conclusion in accord with *emptiness* (*śūnyatā*), the ultimate truth (*paramārtha-satya*) in the Madhyamaka school of thought. For the Prāsaṅgika, however, making any assertion about the conclusion, no matter a positive one or a negative one, would disobey Nāgārjuna's "no-thesis" position. Because the Prāsaṅgika thinks that Nāgārjuna does not intend to propose any statement whatsoever on the level of the ultimate truth in view of the "ineffable" nature of *emptiness*, *prasaṅga* is the only

[2] According to this Prāsaṅgika/ Svātantrika distinction, Bhāviveka (500–570 CE) and his successor Śāntarakṣita (725–788 CE) represent the Svātantrika school while Buddhapālita (470–540 CE) and his successor Candrakīrti (600–650 CE) represent the Prāsaṅgika one. The Svātantrika interpretation of Madhyamaka thoughts were very influential both in China and in Tibet before the 9th century, but the Prāsaṅgikā school has gradually evolved and finally become the authoritative Madhyamaka interpretation in Tibet (not in China) since then. Moreover, modern studies and formal regimentations of Madhyamaka philosophy have largely been influenced by the Prāsaṅgikā tradition. For the historical background, please see Hsu, *Bhāviveka's Jewel*, 25–34. For modern studies, please see: Garfield, The Fundamental Wisdom of the Middle Way; Napper, Dependent-Arising and Emptiness; Hopkins and Napper, Meditation on Emptiness, and Priest, "The Structure of Emptiness", 468–480.

[3] Briefly, the name of Prāsaṅgika is derived from the term prasaṅga, a method similar to the classical *Reductio ad Absurdum.* However, it is controversial whether the Prāsaṅgika way of argumentation, like *Reductio ad Absurdum,* implies a conclusion that is a negation of the original propositions. Because of this possible difference, we will not translate Prāsaṅgika as *Reductio ad Absurdum* in English, but just use the original Sanskrit term. For a good discussion of reductio ad absurdum or indirect proof, see Gasser, "Argumentative Aspects of Indirect Proof", 41–49. We will have more to say about this method below and in Section 4.

adequate method used in a debate for the Prāsaṅgika on issues about the ultimate truth. Hence, for a Prāsaṅgika philosopher, the difference between the Prāsaṅgika and the Svātantrika ways of argumentation is not only methodological, but also substantial: it reveals two different understandings of Nāgārjuna's doctrine of emptiness.[4]

These two ways of argumentation can be roughly characterized by the following two argumentation schemes (where "A" stands for the thesis of the opponent, "⊥" stand for a falsity or an absurdity, and *deleting* "A" and "~A" indicates, respectively, the *denial* and the *negation* of A):[5]

The Svātantrika	The Prāsaṅgika
Conclusion: ~A Reasons Examples	Assume A; Derive ⊥; ~~A~~ (also ~~~A~~).

From the schemes, one can clearly see that the Svātantrika way of argumentation meets Dignāga's basic criteria for a valid hetuvidya syllogism: the thesis to be proved (siddhānta, pratijña), reason (hetu), and example (dṛṣṭānta). Moreover, the thesis has the subject (pakṣa) and the target property to be proved (sādhya-dharma). The main purpose of hetuvidya syllogism is to prove that the subject in the thesis does (or does not) possess the target (possibly negative) property by the tight connection between the reason and the property. Note that it is a requirement of Dignāga's hetuvidya syllogism that the arguer must explicitly assert the conclusion, so Svātantrika is actually proposing that a **Mādhyamika should always assert his/her conclusion when s/he argues with his/her opponents.** This trait will be

[4] Many later Indo-Tibetan scholars, such as Tsong Kha Pa, endorse Candrakīrti's critiques on the Svātantrika and takes the Prāsaṅgika to be the highest and the correct Madhyamaka school. See: David Ruegg, *The Buddhist Philosophy of the Middle*, 4. In contrast to Tsong Kha Pa, some Tibetan Buddhist scholars, such as Go Rams Pa (1429–1489 CE) and Bu Ston Rin Chen Grub (1290–1364 CE), argue that the differences between the Prāsaṅgika and the Svātantrika are merely methodological and insubstantial; see: Mi Pham, *L'opalescent Joyau*, 6. Debates over the question on whether the difference between the Prāsaṅgika and the Svātantrika is merely methodological but not substantial continues between modern Buddhist scholars, including Ames, Huntington, and McClintock. According to Dreyfus's report, Ames and McClintock reply to the question with a positive answer, while Huntington gives a negative reply. No consensus about this question has been reached so far. See: Dreyfus and McClintock, *Svatantrika-Prasangika Distintion*, 6.

[5] The exact meaning of the deletion will be explained in Section 4.

crucial for our later discussion. On the contrary, as mentioned above, the Prāsaṅgika way of argumentation only rebuts the opponents' theses by indicating the absurdities of the theses without making any assertion.

Intriguingly, both the Svātantrika and the Prāsaṅgika, represented in this paper by Bhāviveka and Candrakīrti respectively, claim that the negation used in their arguments/rebuttal is *prasajyapratiṣedha*, usually translated as non-implicative negation, which syntactically is construed with a verb which negates the action or purported fact without further implication. However, even if they adopt the same term "*prasajyapratiṣedha*" for the type of negation that should be used in Madhyamaka arguments/rebuttal, one plausible question to be asked would be: does *prasajyapratiṣedha* used in both schools have the same connotation? If the negations used in Svātantrika and Prāsaṅgika arguments are the same *prasajyapratiṣedha*, do these two schools just arrive at the same end by different means? After all, both schools eventually apply the same non-implicative negation to the opponents' theses. On the other hand, if *prasajyapratiṣedha* has different meanings in these two schools, how should we clarify the air of this discrepancy? In the following sections, by examining the textual evidences provided by Bhāviveka and Candrakīrti and their models of argumentation, we will argue that the negations used in the two schools are actually different: while the *prasajyapratiṣedha* used in Prāsaṅgika is a "non-committal" negation, in Svātantrika it is a non-implicative negation. Moreover, by our rational reconstruction, we suggest that Prāsaṅgika's interpretation is more faithful to Nāgārjuna. In the final section, we will review previous logical formalization of Nāgārjuna's philosophy and provide a formal regimentation based on our understanding of *prasajyapratiṣedha* in Prāsaṅgika.

1. Two Negations and Bhāviveka's Way of Argumentation

In the Indian tradition, grammarians often distinguish two usages of negation: *prasajyapratiṣedha* and *paryudāsapratiṣedha* (implicative negation).[6] Syntactically, this is a distinction about to which part of a sentence the negative word is attached. *Prasajyapratiṣedha* denotes a negation which refers to the predicate or the verb of a

[6] While prasajyapratiṣedha is usually translated as "non-implicative negation", paryudāsapratiṣedha is "implicative negation". However, since we argue that prasajyapratiṣedha is used differently in Svātantrika and in Prāsaṅgika, we keep prasajyapratiṣedha untranslated in this section, and adopt non-implicative negation for paryudāsapratiṣedha.

statement (e.g.: "something is not X"), while the implicative negation designates a negation referring to another word that combines with it to form the predicate (e.g.: "something is a non-X").[7] This syntactic difference corresponds also to a semantic difference according to Bhāviveka:

> When one says "[There is] no white silk fabric", one is only to negate the white silk fabric, but without any power to imply black, red, or yellow silk fabric.[8]

This example by Bhāviveka shows how people may be confused about implicative negation and *prasajyapratiṣedha*. From the semantic point of view, the implicative negation implies other possible affirmations of a positive sentence, while the *prasajyapratiṣedha* simply negates the entire sentence without implying any such possible affirmation, as used in the example.

Given this distinction, Bhāviveka proposes that when a Mādhyamika, such as Nāgārjuna, uses a negation to rebut a proposition of his/her opponent from the perspective of the ultimate truth, that negation s/he utters must be the *prasajyapratiṣedha*.[9] For example, in his commentary on Nāgārjuna's *MMK*, *Prajñāpradīpa*, Bhāviveka proposes that the negation used in *MMK* Chap.1 v.1 to reject all the possibilities of the arising of things should be understood as *prasajyapratiṣedha*. Thus, Bhāviveka states:

> Moreover, [when Nāgārjuna says that a thing is] not produced from itself, he just means [that it is] not produced by itself [without any affirmative implication]. If [one has a] different understanding and says that it is not produced by itself [but by other things], that understanding is not correct, since that would imply that it is produced by other things. [10]

---

[7] For example, in Āpadeva's Mimāṃsānyāyaprakāśa 330f: "Where the negation particle is connected with the word that follows it, this is to be understood as an implicative negation (paryudāsā). While the negation particles are associated with an activity it is to be understood as [prasajya-] pratiṣedha".
[paryudāsaḥ sa vijñeyo yatrottarapadena nañ iti, pratiṣedhaḥ sa vijñeyaḥ kriyayā saha yatra nañ iti.]

[8]《大乘掌珍論》, CBETA, T30, 270c12–15

[9] Kajiyama, "Three Kinds of Affirmation", 161–175.

[10]《般若燈論》, CBETA, T30, 52b26–28

Here, Bhāviveka's point is clear: if the negation used to rebut a thesis in *MMK* is understood as other than the *prasajyapratiṣedha*, i.e. as the implicative negation, Nārgārjuna's verse would imply that a thing is produced by other things. This last implication, however, is not what Nārgārjuna or a Mādhyamika would want to say from the perspective of the ultimate truth, because, for Nārgārjuna in particular and for the Madhyamaka school in general, a thing is neither produced by itself nor by others from that perspective; the production of a thing is simply empty from the perspective of the ultimate truth.

Having this distinction in mind, readers are now in a position to understand Bhāviveka's own way of argumentation, which can be exemplified by the beginning verse of *Ta-sheng chang chen lun* (大乘掌珍論):

> In the ultimate level, conditioned things are empty, because they are produced from conditions, like a magical production. The unconditioned is not real because it is not produced, like a sky-flower.[11]

This verse represents the main thesis of *Ta-sheng chang chen lun*, and the rest of the text contains the opponent's counter arguments and Bhāviveka's counter-counter arguments. Each sentence of the verse actually comprises an argument, thus this verse contains two arguments, the first of which is as follows (the second one can be spelled out in a similar way):

> The First Argument:
> Hidden Major Premise:
> *Ultimately, things that are produced from conditions are empty.*
> Minor Premise: *Conditioned things are produced from conditions.*
> Conclusion: *Ultimately, conditioned things are empty.*
> Similar Example: *Magical production.*

In the beginning of this verse, Bhāviveka uses 'in the ultimate level' or 'ultimately' (paramārthataḥ) to make the opponent understand that the two arguments in this

[11] Here we use Ruegg's translation with some adjustment. Originally, Ruegg uses "in reality" to translate the Sanskrit term "paramārthataḥ"; however, since there are two levels of truth in the Madhyamaka school, i.e. the conventional truth and the ultimate truth, here we use 'in the ultimate level' instead of 'in reality' for the sake of clarity. For the original translation, please see: Ruegg, *The Literature*, 63. For the original Chinese verse, see《大乘掌珍論》, CBETA, T30, 268b21–22.

verse are restricted to the perspective of the ineffable ultimate truth, not to that of the conventional truth which is how things appear to us and what we take for granted. The subject, the target property to be proved, and the reason in the first argument are, separately, "conditioned things", "empty", and "produced from conditions", with "magical production" as a similar example. One can clearly see that the structure of this argument meets Dignāga's criteria for a valid hetuvidya syllogism.[12] Similarly, the second argument also contains the subject (unconditioned things), target property to be proved (not real), the reason (not produced) and the similar example (sky-flower).

It is clear from the above example that Bhāviveka adopts Dignāga's hetuvidya syllogism to establish a Madhyamaka's thesis about the ultimate truth, and, actually, he is the first one who synthesizes the Mādhamaka thoughts and Indian classical logic (*hetuvidyā*) to rebut the opponent. Note that it is a requirement of Dignāga's hetuvidya syllogism that the arguer must explicitly assert the conclusion, so Bhāviveka is actually proposing that a Mādhyamika should always assert his/her conclusion when s/he argue with his/her opponent. This trait will be crucial for our discussion in section 3. Although *prasajyapratiṣedha* in Svātantrika arguments can indeed prevent the implication of other affirmations, in this sense non-implicative, it cannot prevent the affirmation of the conclusion itself, given the nature of hetuvidya syllogism.[13] In sum, for Bhāviveka, a Mādhyamika should not merely use *prasaṅga* to rebut his/her opponent, s/he should also assert his/her own

[12] In Dignāga's logic system, a hetuvidya syllogism is something like an enthymeme syllogism and should be created with three parts: the thesis, the reason, and a similar example and/or a dissimilar example. Most importantly, the thesis (conclusion) of a hetuvidya syllogism must always be asserted by the arguer. Also of note, in the case of this opening verse, there is no dissimilar example that can be offered, because, for a Mādhyamika, everything is empty. For more information about Dignāga's logic system, please see *Hetuvidyā nyāya dvāra śāstra* (in Chinese:《因明正理門論》).

[13] In Bhāviveka's truth system, he proposes that there are two kinds of ultimate truths: unverbalizable ultimate truth (*aparyāya paramārtha*) and concordance ultimate truth (*paryāya paramārtha*) which is the verbal and conceptual representations of the ultimate truth. According to this classification, the conclusions of Svātantrika arguments belong to concordance ultimate truth which can guide people to attain the unverbalizable ultimate truth. Therefore, despite the ontological status of concordant ultimate truth, the usage of *prasajyapratiṣedha* in Svātantrika cannot dispel all linguistic statements or judgements, but it at least derives the conclusion at the level of the concordant ultimate truth. However, due to limited space, we are unable to provide more discussion on this special classification by Bhāviveka. For more discussion, please see: Lusthaus, *Buddhist phenomenology*, 449; Ruegg, *The Buddhist Philosophy of the Middle*, 170.

thesis in accord with the ultimate truth.[14] Moreover, the negation used in a Mādhyamika conclusion should be understood as the *prasajyapratiṣedha*; other interpretations would be incorrect.

2. Candrakīrti's Way of Argumentation

In *Prasannapadā Mūlamadhyamaka-vṛtti* (hereafter: *PPMV*) Candrakīrti harshly scolds Svātantrika Mādhyamika, such as Bhāviveka and his disciples, as just logicians obsessed with making arguments but not true Mādhyamika at all.[15] Candrakīrti then argues that Mādhyamika should not adopt the Svātantrika's way of arguing because Mādhyamika should have "no view" about the ultimate truth.[16] Due to this tenet, Candrakīrti therefore thinks that the Madhyamaka way of arguing for the conclusion by offering an independent argument violates the core spirit of the Nāgārjuna. For Candrakīrti, this idea echoes the 29th verse in *Vigrahavyāvartanī* (in Chinese:《迴諍論》), where Nāgārjuna affirms:

> If I had some thesis, the defect [just mentioned] would as a consequence attach to me. But I have no thesis, so this defect is not applicable to me.[17]

Following this underlying principle, Candrakīrti proposes that:

> For those who establish the independence inference (argument), the mistake will be on their side. We do not establish the independence inference (argument), since the function of argument is merely to rebut the opponent.[18]

In the quoted passage, Candrakīrti indicates that the correct way of making a Madhyamaka argument about the ultimate reality is *merely* to refute the opponent by *prasaṅga*, as Buddhapālita (and Nāgārjuna) always does.

Candrakīrti also criticizes Bhāviveka that, if he really accepts the Madhaymaka doctrine that everything is ultimately empty, then it is impossible for

[14] Bhavya, *Madhyamakahṛdaya*, 9.
[15] Dreyfus and McClintock, *Svatantrika-Prasangika Distintion*, 81–82.
[16] Candrakīrti, *PPMV*, 16.
[17] Here we adopt Westerhoff's English translation, please see: Westerhoff, "The No-Thesis View", 25–39.
[18] Candrakīrti, *PPMV*, 34.

him and his opponent to have the same understanding about the subject term in the disputed thesis because only Mādhyamika, not his opponent, would take it to be ultimately empty. And, according to the Indian hetuvidya tradition, there can be no argument unless the speaker and his rival have consensus on the meaning of the subject term in the argument. Consequently, Bhāviveka's way of argumentation will violate this basic principle of hetuvidya logic. Moreover, if Bhāviveka agrees that everything, including the subject of the thesis, property to be proved, the reasons, and the similar/ dissimilar cases, is ultimately empty, how could he then make a valid argument that contains a subject, explanandum, etc., which are all empty? If Bhāviveka does accept that everything is ultimately empty, how can he establish a thesis with a subject term that refers to things that do not exist at all? That is to say, if the Mādhyamika uses Dignāga's Indian syllogism to establish his own thesis, he would have to commit to the existence of every component in the thesis, and that will be inconsistent with the Madhyamaka main doctrine that everything is empty *ultimately*. Therefore, Candrakīrti firmly rejects the Svātantrika way of argumentation and proposes that the only suitable way of argumentation that fits the Madhyamaka is *prasaṅga*.

Nevertheless, it is striking that Candrakīrti also interpreted the negation used in *MMK* and Prāsaṅgika inference to be prasajyapratiṣedha:

> Does not "[a thing is] not produced from itself" establish the undesirable [thesis] "from others"? [answer:] It is not the case, since the prasajyapratiṣedha is what wants to be said, and therefore "from others" will be dispelled as well.[19]

It should be noticed that here "a thing is not produced from itself" is not a conclusion derived from reasons and examples as in the argument by Svātantrika, but just a *prasaṅga* refutation of the opponents' thesis that "a thing is not produced from itself". Therefore, for those Mādhyamika making svātantrika inferences, such as Bhāviveka, they have to accept the conclusion derived from the inferences in accord with ultimate truth, and therefore the *prasajyapratiṣedha* used in the inferences will still imply the assertion the conclusion. By contrast, Candrakīrti had no intention of making any assertion when rebutting the opponents. Candrakīrti took

[19] For original Sanskrit, please see: PPMVS on MMK 1-1:
nanu ca naiva svata utpannā ity avadhāryamāṇe parata utpannā ityaniṣṭaṃ prāpnoti | na prāpnoti, *prasajyapratiṣedhasya* vivakṣitatvāt parato 'pyutpādasya pratiṣetsyamānatvāt |.

the prasajyapratiṣedha in prāsaṅgika as merely the denial of the opponents' thesis. Therefore, according to those textual evidence and analysis, we name the *prasajyapratiṣedha* in Svātantrika "non-implicative negation", and the prasajyapratiṣedha in Prāsaṅgika "non-committal negation". This discrepancy, which will be further explored in the following section, is crucial for our rational reconstruction of Nāgārjuna's philosophy.

3. The More Plausible Interpretation

Both the Prāsaṅgika and the Svātantrika ways of arguing intend to interpret *how* Nāgārjuna rebuts his opponent's thesis in *MMK*. The main difference between the two interpretations is whether Nāgārjuna *asserts* his conclusion or his own thesis after he rejects his opponent's thesis by drawing falsity or absurdity from the latter. It seems at first sight that the Svātantrika's interpretation must be right; after all, Nāgārjuna *apparently* makes many assertions, though all of them being negative, after drawing falsity or absurdity from his opponent's theses or assumptions. Especially, Nāgārjuna *apparently* asserts that a thing cannot be self-caused (or that it cannot be caused by other things, caused by both, or uncaused) by drawing absurd consequences from the assumption that it is self-caused (or the assumption that it is caused by other things, caused both by itself and other things, or that it has no cause), and, he *apparently* aims to assert that nothing has an intrinsic nature, which is the negation of his opponent's thesis that at least something, i.e., the production of a thing, has an intrinsic nature. However, closer inspection of Nāgārjuna's two-truth theory shows that facts are not as they appear.

After all, Nāgārjuna is trying to see the whole debate, as Bhāviveka points out, from the perspective of the ultimate truth, and, as mentioned in section 1, everything whatsoever, according to Nāgārjuna, is empty from this ultimate perspective. Yet, if everything is ultimately empty and therefore ultimately does not exist, what we say about anything is ultimately about nothing and therefore cannot be true from this ultimate perspective. Conversely, if there is any "truth" from this ultimate perspective, it must be something ineffable. Thus, any assertion about the ultimate truth, no matter a positive or a negative one, must not be true (and must not be false either, as we shall see below) ultimately. How then can Nāgārjuna assert, from this ultimate perspective, any truth about what is ultimately empty if he is faithful to his own two-truth theory? It seems that Nāgārjuna must not be making

any assertion, either a positive or negative one, when he rebuts his opponent's thesis after drawing falsity or absurdity from the thesis and wants to convey something "true" from this ultimate perspective. That he apparently asserts any statement must only be an appearance.[20]

One can also view things from another angle. *Were* Nāgārjuna to assert, say, "a thing cannot be self-caused" after he draws the absurdity from the assumption that the things are self-caused, he would be forced to accept that one of the other three possibilities mentioned in the verse (that it is caused by other things, caused by both, or uncaused) must be true, for these possibilities jointly exhaust all possibilities in which a thing is not self-caused. Nāgārjuna should not go on to deny all of them. The fact that Nāgārjuna goes on to reject all possibilities shows that he is not, contrary to appearance, asserting the negative sentence "a thing cannot be self-caused" after he draws out the absurdity of the assumption.

Several related issues follow: if Nāgārjuna does not mean to *assert* those negative sentences that he utters after reducing his opponent's theses to absurdity in *MMK*, how should we understand the function of those negative sentences in *MMK*? Especially, how should we understand the function of the negative word "not" used in these utterances? Should we understand it as the prasajyapratiṣedha that Bhāviveka suggests? If not, should we understand Nāgārjuna's apparent assertion of these negative sentences as a "pretension"? We suggest that "no" is the answer to the final two questions. After all, what is the benefit of pretending to make a negative assertion when one's aim is to convey some "ineffable truth" about the ultimate reality?[21] And, after all, using the word "not" in the Bhāviveka's prasajyapratiṣedha, which we call the non-implicative negation, is still making an assertion of a negative sentence with the non-implicative "not",[22] and we have seen

[20] Or it could be a mistake. However, we do not think that it is plausible to attribute such a mistake to Nāgārjuna; here, we adhere to the famous principle of charity proposed by D. Davidson. See: Davidson, *Inquiries into Truth,* Chapter 13.

[21] What is an ineffable truth? It is about the ultimate reality but cannot be expressed in any language. Since it is not expressible in any language, it is perhaps misleading to call it a "truth". However, if one takes a truth to be a mind-independent proposition and allows that there are propositions that cannot be expressed in any language, then it makes perfect sense to talk about an ineffable truth. Does Nāgārjuna intend to convey some such ineffable truth to his readers of *MMK*? On one reading, especially the one hinted by later Chinese Chan philosophers, Nāgārjuna does intend to convey some such ineffable truth to his readers. Is this a possible mission? If so, how? These are questions that we need not answer here (though we think that they can be answered), for our purpose here is merely to point out a way to understand Nāgārjuna's rebuttal, not to defend such a possibility.

[22] For more on this claim, please see the final paragraph of this section.

that this way of understanding *MMK* will wrongly attribute the attitude of disobedience to his own two-truth theory to Nāgārjuna. But, then, to repeat, how should we understand the function of those negative sentences found in *MMK*? Especially, how should we understand the function of the negative word "not" used in these negative utterances?

We believe that the word "not", when uttered by Nāgārjuna in *MMK* to rebut his opponent's thesis, is used, neither as Bhāviveka's non-implicative negation nor as implicative negation, but in a third way as Candrakīrti's prasajyapratiṣedha, which we call "the non-committal negation", referred to by the *deletion* of a sentence mentioned in the introduction. Unlike the non-implicative use of "not", which commits one to the truth of a negative sentence containing a "not" with a wider scope, and unlike the implicative use of "not", which commits one to the truth of a negative sentence containing a "not" with a narrower scope, the non-committal use of the word "not" commits one to neither, but *merely indicates* a speech act of denial or a propositional attitude of rejection to a sentence or a proposition.[23] We emphasize the phrase "merely indicate" because when one uses the word "not" implicatively or non-implicatively, it is used, of course, to deny (or reject) a sentence or a proposition; however, one typically does more than that: s/he also asserts (or accepts) a negative sentence or a negative proposition. The non-committal use of the word "not", by contrast, *merely indicates* a speech act of denial (or a rejection) of a sentence or a proposition *and no more*. Therefore, the denial (or the rejection) of a sentence *p*, indicated by such non-committal use of "not", should not be thought of as accompanied by (or as implying or containing) an assertion (or an acceptance) of the negation (either implicative or non-implicative) of *p*. One does not have to assert (or accept) not-*p* in order to *merely* deny *p* (or *merely* reject *p*), and in some cases (especially when one takes that both *p* and not-*p* are somehow "defective"), one can both rationally deny (and/or reject) *p* and deny (and/or reject) not-*p*.[24]

---

[23] A speech act, also called an "illocutionary act", is what one does with his/her utterance (Austin, *How to Do Things With Words*, 98), while a propositional attitude is one's mental attitude toward a proposition. Speech acts are not propositional attitudes and vice versa. Because they are different, our explanation is actually a disjunctive one: either the word "not" merely indicates a speech act, or it merely indicates a propositional attitude of Nāgārjuna. We do not know which disjunct of the disjunction is true and we do not exclude the possibility that both disjuncts are true either. However, either disjunct is enough to lend its support to our conclusions.

[24] For a good discussion of negation, denial, and speech acts, see J. Moeschler, "The Pragmatic Aspects", 51–76; Ripley, "Negation, Denial, and Rejection", 622-629; and Horn

To support our claims, we cite here a discussion made by H. Field of the distinction between rejecting a proposition p and accepting its negation not-p (similar things can be said about the distinction between the denial of p and the assertion of not-p):[25]

> [One] can . . . distinguish between rejection [of *p*] and acceptance of the negation [of *p*]. Rejection should be taken to involve, at the very least, a commitment not to accept [p]. . . . [A] defender of [this view] should . . . take rejection not to require acceptance of the negation. By doing so, we can allow for the simultaneous rejection of both a sentence and its negation; even the full rejection of both.

The point here is that one should not confuse the rejection of a proposition *p* with the acceptance of its negation *not-p* (implicative or non-implicative); especially, rejecting a proposition *p* does not necessarily imply accepting *not-p* (implicative or non-implicative). Similarly, one should not confuse the denial to a sentence *p* with the assertion of its negation *not-p* (implicative or non-implicative); especially, denying a sentence *p* does not necessarily imply asserting *not-p* (implicative or non-implicative).

Unfortunately, even though the speech act of *merely denying p* (or the propositional attitude of *merely rejecting p*) should be distinguished from the speech act asserting *not-p* (or the propositional attitude of accepting *not-p*), the most ordinary means that we have for *merely denying* a sentence (or for *merely rejecting* it) is still to use the word "not", which may be the main source of confusion.[26] If one asserts that "the round square is round" and we intend merely to deny it (without committing to its negation) because we think that there is no such a thing as the round square, it is very natural for us to respond to such assertion with "No, it is not!" Of course, this may mislead one to believe that part of our what we mean is that "*it* is not round" (where "not" is used non-implicatively) or "*it* is non-round" (where 'non' is used implicatively), but it should be clear from our explanation that, since we believe that there is no such thing as the round square, we do not intend to assert

& Wansing, "Negation".

[25] Field, *Saving Truth from Paradox*, 73–7

[26] We can avoid this confusion by, say, *stipulating* that, whenever one utters a sentence with his/her nose being pulled out by his/her fingers or with his/her head shaking violently, s/he is *merely denying* it without asserting any negative sentence. However, this is not the convention that we currently have.

or accept either one of them. If one goes on to ask us "so it is not round?" or "so it is non-round?", we will certainly reply to both questions by saying "neither!" out loud. Our reply, to be sure, presupposes a Fregean view about "defective sentences" with empty names, according to which a sentence of whatsoever kind is not true (and not false either) if its subject term does not refer, but we think that this view makes perfect sense when applied to sentences about the realm of emptiness, i.e., about the realm of ineffable things.[27]

That being said, we can now go back to see how Nāgārjuna uses the word "not" in *MMK* to rebut his opponent's thesis and why he uses it in that way. When Nāgārjuna takes a state-of-affair (or a proposition, if you like) *p* to be ineffable from the ultimate perspective, it is also reasonable for him to take, just like a Fregean would do, its negation not-*p* to be ineffable from the same perspective. As a result, if Nāgārjuna is faithful to his doctrine that everything is ultimately empty and ineffable from the ultimate perspective, he simply *cannot* be asserting "not-*p*" *even if he utters it* after concluding, by reasoning, that his opponent's claim of *p* leads to absurdity. Therefore, the "not" used by Nāgārjuna in *MMK* actually indicates neither the assertion of a negative sentence nor the acceptance of a negative proposition, but *merely* the denial of the thesis at issue. It is only because our ordinary means to merely deny a sentence (or to merely reject it) is still to use the word "not", the Svātantrika is misled into believing that Nāgārjuna is actually making an (negative) assertion after he draws the absurd consequences of his opponent's thesis. But this is a confusion. In short, the Prāsaṅgika is right: whenever the Mādhyamika successfully reduces his/her opponent's thesis to absurdity, s/he should simply reject or deny his/her opponent's thesis without making his/her own negative conclusion if s/he is to see things from the ultimate perspective.

It may be suggested that Bhāviveka's implicative/non-implicative distinction is nothing but our mere-denial/assertion or mere-rejection/acceptance distinction and therefore the dispute between the Svātantrika and the Prāsaṅgika is nothing but verbal. This suggestion, however, is not plausible. After all, the Svātantrika is trying to argue in the hetuvidya way that Dignāga's demands, and an assertion of the conclusion is always needed in a hetuvidya argument. If the conclusion is a negative sentence, no matter whether the "not" is used implicatively or non-implicatively, an assertion is still made, which makes the conclusion a *mere* denial impossible.[28]

[27] For details of this Fregean view about defective sentences with empty names, see: Gottlob Frege, "On Sense and Reference", 25–50.

[28] It may be argued that non-implicative negation lacks assertoric force, and, if so, how

Besides, an implicative or a non-implicative negation is always, when uttered, a part of a sentence or a part of a proposition, but the mere denial of a sentence or the mere rejection of a proposition is an act or a mental attitude which can never be a part of a sentence or a proposition.[29] Thus, it is nonsensical to equate Bhāviveka's implicative/non-implicative negation distinction with our mere-denial/assertion or mere-rejection/acceptance distinction.

4. A Formal Regimentation of Nāgārjuna's Rebuttal

We have concluded that the Prāsaṅgika is the right interpretation of how Nāgārjuna rebuts his opponent in *MMK*: whenever Nāgārjuna successfully reduces his opponent's thesis to absurdity, he simply rejects or denies his opponent's thesis without making his own conclusion. This also explains why Nāgārjuna claims that he has not proposed any thesis at all. It remains to be explained whether this way of argumentation can make perfect sense from a logical point of view. There are a few contemporary philosophers, especially Garfield, Priest, and Westerhoff,[30] who have addressed this issue. We will briefly review their views before we give our own.

Like Bhāviveka, Westerhoff also distinguishes two kinds of negation and calls them "presupposition-preserving" and "presupposition-canceling" respectively. It is difficult to tell whether Westerhoff's distinction is exactly the same as Bhāviveka's implicative/non-implicative distinction, because there is evidence

---

does its use by Bhāviveka commit him to making an assertion about a negative sentence? But this is exactly what we have been trying to argue in the paper: a non-implicative negation, especially being the conclusion of a hetuvidya argument, still has some assertoric force; it commits one to the acceptance of a negative sentence. The difference between the implicative and the non-implicative negation is not that the former has, while the latter lacks, an assertoric force. The distinction is rather that the former has, while the latter lacks, the implication that some other *positive* affirmation is true. The utterance of a non-implicative negation still commits one to accept a certain negative sentence, thus still having some assertoric force. By contrast, what we call the "non-committal" use of the word "not" does not commit one to any such positive or negative sentence at all.

[29] You may think that, even if an act or mental attitude cannot be a part of a sentence, still it is something that can be truly asserted or correctly described by a true sentence. Our reply to this criticism is this: that is certainly true, but only from the conventional perspective. Viewed from the ultimate perspective, such a description or such an assertion should also be denied or rejected.

[30] See: Garfield and Priest. "Mountains Are Just Mountains", 71–82 and Garfield and Priest, "Nāgārjuna and the Limits of Thought", 1–21. For Westerhoff, please see: Westerhoff, *Nāgārjuna's Madhyamaka: A Philosophical Introduction*.

indicating both that he intends them to be identical[31] and opposite[32]. Fortunately, we do not have to get involved into this interpretational problem; we can simply say two things about Westerhoff's interpretation: (a) as long as he takes Nāgārjuna's rebuttal to be an assertion with a non-implicative or a presupposition-canceling negation, he is wrong about the rebuttal for reasons mentioned in the previous section; (b) on the other hand, if he takes Nāgārjuna's rebuttal to be something like a *mere* illocutionary act,[33] he is right about the rebuttal and agrees with us. The only issue left for (b) is whether Nāgārjuna's way of argumentation can make perfect sense from a logical point of view, which is a topic Westerhoff does not discuss.

More Recently, Garfield & Priest developed a novel but controversial interpretation of how Nāgārjuna rebuts his opponent in *MMK*.[34] They take a hint from Westerhoff and suggest that the rebuttal of a thesis in *MMK* by Nāgārjuna should be taken as "the external negation" which is another name that Westerhoff uses for his presupposition-canceling negation. The reference to Westerhoff's external negation also creates ambiguity between (a) and (b) in the above paragraph, and we will comment similarly as we did for Westerhoff's suggestion. So far, there is nothing very exciting in their interpretation. However, their formal regimentation of Nāgārjuna's way of rebuttal has a few very interesting features that we will explain and comment on in the rest of this section.

Garfield and Priest's formal interpretation of how Nāgārjuna rebuts his opponent in *MMK* consists mainly of two parts: a part about the positive *catuṣkoṭi* and a second part about the negative *catuṣkoṭi*.[35] In the first part, they take Nāgārjuna to be a logician who thinks that all sentences about the conventional

[31] Especially, he identifies his presupposition-preserving/presupposition-canceling distinction with Bhāviveka's prasajyapratiṣedha and paryudāsapratiṣedha.

[32] Especially, he suggests that his presupposition-canceling negation is an illocutionary act.

[33] Cotnoir in 2015 also suggests that we take Nāgārjuna's rebuttal along the line of (b) here.

[34] Not everything they say in that article is relevant to the current issue; especially, the main focus of their joint paper, namely, the logical form of a *catuṣkoṭi* (either a positive one or a negative one) is not our current concern.

[35] 'Catuṣkoṭi' means 'four corners' in Sanskrit. Nāgārjuna (as well as many other ancient Indian philosophers) often divided possibilities about a thing into four kinds and then reasoned about whether any one of them would hold. When he divided possibilities of a thing in this way, we call the four kinds of possibilities a 'catuṣkoṭi' (or four koṭi). One example of such a catuṣkoṭi (self-caused, caused by other things, both, and no cause) can be found in verse 1.1 of *MMK* mentioned in section 1 of this paper. When Nāgārjunas affirms, usually from the conventional perspective, that at least one of a four koṭi holds, such catuṣkoṭi is said to be a 'positive catuṣkoṭi'. Nāgārjuna, however, often denies all four koṭi from the ultimate perspective; in this case, it is said to be a 'negative catuṣkoṭi'.

reality can be divided into four mutually exclusive non-empty categories: true-but-not-false (*t*), false-but-not-true (*f*), both-true-and-false (*b*), and neither-true-nor-false (*n*). This part of their interpretation is rather controversial and we will reject it for reasons discussed in the next paragraph. On the other hand, we think that their interpretation of the negative *catuṣkoṭi* is closer to truth: they take Nāgārjuna to be a logician who adds one more value, *e* (standing for "ineffable"), to the original values for the conventional truth when using a negative *catuṣkoṭi*. However, their interpretation of the negative *catuṣkoṭi* are implausible for two reasons: (a) since a sentence is certainly something effable, it makes no sense to attribute the value *e* ("ineffable") to it; (b) the value *e* should not be an extra value that can be attributed to sentences with conventional values; it is rather, according to Nāgārjuna, the value that everything (including every sentence) ultimately has from the ultimate perspective. These two problems, however, are nicely corrected by Priest (2018) in the following ways: (1) the value-bearers in the negative *catuṣkoṭi* are now assumed to be state-of-affairs rather than sentences; (2) each state-of-affairs is now assumed to have the value *e* and possibly an extra conventional value. Priest calls this final version of the formal semantics for *MMK* "plurivalent' semantics, which is quite inspiring in interpreting Nāgārjuna's rebuttal, though we still believe that their four-valued construal of the positive *catuṣkoṭi* is quite wrong.[36]

There are two main reasons for why we believe that the four-valued construal of the positive catuṣkoṭi is quite wrong. To spell them out in full detail, however, would require another paper, so our comments below will only be brief. First, there is a problem of literature and historical support for their interpretation. By making both-true-and-false (b) a non-empty category for a positive catuskoti, Garfield and Priest are actually interpreting Nāgārjuna's view about the conventional reality as a sort of dialetheism, according to which there are true contradictions in the conventional reality. But it seems to us that such a dialetheist interpretation of Nāgārjuna's view of the conventional reality has neither literature nor historical support. Second, none of the formal systems (FDE, FDEe, and P-FDEe) that they propose seems to be adequate to ground the logic at play in MMK for the simple reason that they are all too weak. As Cotnoir (2015) points out, Nāgārjuna often uses classically valid inference patterns, such as Modus Ponens, Modus Tollens, Hypothetical Syllogism, reductio ad absurdum (RAA) to argue against his opponent in MMK, but none of these inference patterns are valid in any of their proposed logic systems.

[36] See: Priest, *The Fifth Corner of Four*.

Due to the above observations, we therefore suggest the following formal treatment of Nāgārjuna's rebuttal in *MMK*. In Nāgārjuna's conception, each state-of-affairs can be viewed from two perspectives, a conventional perspective and an ultimate perspective. Viewed from the ultimate perspective, every state-of-affairs is ineffable and hence has the value *e* (in agreement with Priest, 2018).[37] Viewed from the conventional perspective, on the other hand, each state-of-affairs is effable and may have one of the classical values *t* and *f* (this is where we differ from Priest, 2018). More formally, a model *v* (which we shall call a "plurivalent Ke-model"[38]) is a function that assigns to each atomic state-of-affairs a one-membered or two-membered subset (a "value-set", so to speak) of $\{t, f, e\}$ that includes *e* and that satisfies the following conditions (where an underlined formula stands for a state-of-affairs):[39]

1. *f* belongs to $v(\underline{\sim A})$ if *t* belongs to $v(\underline{A})$ and *t* belongs to $v(\underline{\sim A})$ if *f* belongs to $v(\underline{A})$, otherwise $v(\underline{\sim A}) = v(\underline{A})$;
2. $v(\underline{A \,\&\, B}) = \{e\} \cup \{t\}$ if *t* belongs to both $v(\underline{A})$ and $v(\underline{B})$, $v(\underline{A \,\&\, B}) = \{e\}$ if $v(\underline{A})$ or $v(\underline{B}) = \{e\}$, otherwise $v(\underline{A \,\&\, B}) = \{e, f\}$;
3. $v(A \text{ v } B) = \{e\} \cup \{f\}$ if *f* belongs to both $v(\underline{A})$ and $v(\underline{B})$, $v(\underline{A \text{ v } B}) = \{e\}$ if $v(\underline{A})$ or $v(\underline{B}) = \{e\}$, otherwise $v(\underline{A \text{ v } B}) = \{e, t\}$.[40]

---

[37] The assertion that "[v]iewed from the ultimate perspective, every state-of-affair is ineffable and hence has the value *e*" *in the meta-language* should not be regarded as violating Madhyamaka's doctrine of ineffable or Madhyamaka's practice. One reason is that such an assertion is still made from the conventional perspective. The main reason, however, is that the model-theoretical semantics is formulated in the meta-language; such meta-talks, including the model itself and all kinds of assertions that can be made in or about it, are only *heuristic tools* for helping us to decide which sentences in the *object-language* can be *asserted*, and, if so, from which perspective. Since every state-of-affair has the ultimate value *e* in a model, this indicates that nothing can really be truly asserted about any state-of-affair from the ultimate perspective. However, sentences can still be asserted from the conventional perspective. If $\underline{A}$'s value-set contains *t*, then the sentence "A" is assertable from the conventional perspective. Similarly, if $\underline{A}$'s value-set contains *f*, then the sentence "~A" is assertable from the conventional perspective.

[38] Plurivalent semantics allows a sentence to have more than one value in a model. The semantic we will give is obviously a plurivalent one. It is called "Ke" because it is essentially a weak Kleen semantics K with one more value *e*.

[39] In terms of truth tables, the three semantic rules can be explained by the following tables:

A	~A
$\{e, t\}$	$\{e, f\}$
$\{e\}$	$\{e\}$
$\{e, f\}$	$\{e, t\}$

&	$\{e, t\}$	$\{e\}$	$\{e, f\}$
$\{e, t\}$	$\{e, t\}$	$\{e\}$	$\{e, f\}$
$\{e\}$	$\{e\}$	$\{e\}$	$\{e\}$
$\{e, f\}$	$\{e, f\}$	$\{e\}$	$\{e, f\}$

v	$\{e, t\}$	$\{e\}$	$\{e, f\}$
$\{e, t\}$	$\{e, t\}$	$\{e\}$	$\{e, t\}$
$\{e\}$	$\{e\}$	$\{e\}$	$\{e\}$
$\{e, f\}$	$\{e, t\}$	$\{e\}$	$\{e, f\}$

[40] It may be questioned why we do not have a connective or an operator, corresponding to the ineffable value e or the non-committal denial (or rejection), in our language. The main

The main ideas behind the model and the value-conditions of various state-of-affairs are actually quite simple: each state-of-affair has an "ultimate" value *e* and perhaps an extra classical "conventional" value within its value-set. The value *e* is "infectious" in the sense that any state-of-affair that has a part with {*e*} as its value-set will also have the value {*e*}. The value-set for a complex state-of-affairs with no part with the value-set {*e*} is then determined in the ordinary classical way. Since the value *e* stands for "ineffable", the stipulation that each state-of-affairs has at least the value *e* means that each state-of-affairs is ineffable from the ultimate perspective. Yet, each state-of-affairs may also have a classical truth value *t* or *f*, which means that each state-of-affairs may also have an effable aspect and is either true or false when viewed from the conventional perspective. This formal regimentation seems to be quite faithful to Nāgārjuna's two-truth theory. As usual, we define an argument to be valid if it preserves the designated value *t* in every plurivalent Ke-model. It can be proved that the resultant logic, while much weaker than the classical logic, still preserves many classical valid inferential rules such as Modus Ponens, Modus Tollens, and Hypothetical Syllogism, but we will not give the proofs here.

How does this semantics verify that the Prāsaṅgika is the right interpretation of how Nāgārjuna rebuts his opponent? Notice that the classical valid rule *reductio ad absurdum* is no more valid in the above plurivalent Ke-semantics (see the Proof below), so that the derivation of an absurdity from the assumption of a state-of-affairs no more guarantees that the negation of it obtains though it still guarantees that the state-of-affairs must not obtain (and hence can be "deleted"). Because of the importance of this conclusive claim, we give its proof here:

> *Proof*: (Here we take reductio ad absurdum as reduction to contradiction, not absurdity in general, for simplicity.) Reduction to contradiction is the rule "if A entails contradiction, then infer ~A". Now this is invalid in Ke-semantics because while (p & ~p) entails (p & ~p), ~(p & ~p) may still not be true (consider a valuation v in which v(p)={e}). On the other hand, this does not mean that we should accept any contradiction (a state-of-affairs of the form "p & ~p") at all or any state-of-affairs that entails a contradiction, because that would require some state-of-affairs to have both t and f in its value-set,

reason is that, as we said in the previous section, the non-committal denial (or rejection) is not a part of our language but a speech act (or a mental attitude). It is, unlike a connective or an operator, unable to be combined with other elements in a language to form a complex whole and therefore should not be taken to be a connective or an operator of our language.

> which is impossible for any Ke-model v. Therefore, if A entails a contradiction, A (and ~A) can never be true according to Ke-semantics, and hence A (and ~A) should be denied (or rejected).

Therefore, even if Mādhyamika successfully draws an absurdity from his opponent's thesis and therefore shows that the thesis cannot be true (or can be deleted), s/he is not thereby warranted to conclude that the negation of the thesis is true. Indeed, s/he should not conclude this if s/he is to view things from the ultimate perspective. This formal regimentation is in the spirit of the Prāsaṅgika and Nāgārjuna, but not with the Svātantrika. It also explains why *prasaṅga*, rather than the classical rule *reductio ad absurdum*, is the right way for a Mādhyamika to rebut his opponent.

References

龍樹造，梵志青目釋，鳩摩羅什譯。《中論》，CBETA T30, no. 1564。
龍樹造，清辨釋，波羅頗蜜多羅譯。《般若燈論釋》，CBETA T30, no. 1566。
龍樹造，毘目智仙共瞿曇流譯。《迴諍論》，CBETA T32, no.1631
清辯造，玄奘譯。《大乘掌珍論》，CBETA T30, no. 1578。
大域龍[=陳那]造，玄奘譯。《因明正理門論》，CBETA T32, no. 1628。

佐々木, 月樵 (1926) *龍樹の中観及其哲学*. (東京: 甲子社書房).
江島, 惠教 (1990) Bhāvaviveka/ Bhavya/ Bhāviveka. 印度學佛教學研究, 38.2, 838–846.
Austin, John L. (2009) *How to Do Things with Words*. (Oxford: Oxford University Press).
Candrakīrti (1903–13) *Prasannapadā Mūlamadhyamaka-vṛtti with Nāgārjuna's MK,* in Louis de La Vallee Poussin (ed.) *Bibliotheca Buddhica IV*. (St Petersburg: n.p.).
Candrakīrti (1970) *Mūlamadhyamakakārikās (Mādhyamikasūtras) De Nāgārjuna Avec La Prasannapadā*. (Osnabrück: Biblio-Verlag).
Conze, Edward (1993) *Perfect Wisdom*. (Leicester: Buddhist Pub. Group).
Cotnoir, Arron. J. (2015) Nāgārjuna's Logic, in Koji Tanaka, Yasuo Deguchi, Jay Garfield and Graham Priest (eds.) *The Moon Points* Back, (Oxford: Oxford University Press), 176–88.

Davidson, Donald (1984) *Inquiries into Truth and Interpretation*. (New York: Oxford University Press).

Dhammajoti, Kuala Lumpur (2009) *Sarvāstivāda Abhidharma*. (Hong Kong: Centre of Buddhist Studies, University of Hong Kong).

Dreyfus, Georges B. J, and Sara L McClintock (2003) *Svatantrika-Prasangika Distinction*. (Somerville, MA: Wisdom Publications).

Ejima, Yasunori (1990) Bhāvaviveka/Bhavya/Bhāviveka, *Journal of Indian and Buddhist Studies (Indogaku Bukkyogaku Kenkyu)*, 38. 2, 846–838.

Field, Hartry H. (2008) *Saving Truth from Paradox*. (Oxford: Oxford University Press).

Frege, Gottlob (1982) Über Sinn Und Bedeutung [On Sense and Reference], *Zeitschrift Für Philosophie Und Philosophische Kritik*, 100, 25–50.

Garfield, Jay L. and Priest, Graham (2003) Nāgārjuna and the Limits of Thought, *Philosophy East and West*, 53.1, 1–21.

Garfield, Jay L. and Priest, Graham (2009) Mountains Are Just Mountains, in Jay L. Garfield, Tom J. L. Tillemans, and Mario D'Amato (eds.) *Pointing at the Moon—Buddhism, Logic, Analytic Philosophy*, (New York: Oxford University Press), 71–82.

Gasser, James (1992) Argumentative Aspects of Indirect Proof, *Argumentation*, 6.1, 41–49.

Horn, Laurence R. and Wansing, Heinrich (2020) Negation in The Stanford Encyclopedia of Philosophy (Spring 2020 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2020/entries/negation/.

Hsu, Chien Y (2013) Bhāviveka's Jewel in the Hand Treatise: Elucidating a Path to Awakening Utilizing Formal Inference. (PhD diss., University of Calgary).

Kajiyama, Yuichi (1973) Three Kinds of Affirmation and Two Kinds of Negation in Buddhist Philosophy, *Wiener Zeitschrift Für Die Kunde Südasiens Und Archiv Für Indische Philosophie*, 17, 161–175.

Lusthaus, Dan (2014) *Buddhist Phenomenology: A Philosophical Investigation of Yogacara Buddhism and the Ch'eng Wei-shih Lun*. (New York: Routledge).

Mœschler, Jacques (1992) The Pragmatic Aspects of Linguistic Negation: Speech Act, Argumentation and Pragmatic Inference, *Argumentation*, 6.1, 51–76.

Nāgārjuna, Siderits, Mark, and Katsura, Shoryu (2013) *Nāgārjuna's Middle Way*. (Boston, MA: Wisdom Publications).

Priest, Graham (2018) *The Fifth Corner of Four: An Essay on Buddhist Metaphysics and The Catuṣkoṭi*. (Oxford: Oxford University Press).

Rgya Mtsho, Mi Pham (2004) *L'opalescent Joyau: Nor-Bu Ke-Ta-Ka (Trésors Du Bouddhisme)*. (France: Fayard).

Ripley, David (2011) Negation, Denial, and Rejection, *Philosophy Compass* 6.9, 622–629.

Robinson, Richard H. (1967) *Early Madhyamika in India and China*. (Delhi: M. Banarsidass).

Ruegg, David Seyfort (1981) *The Literature of the Madhyamaka School of Philosophy in India Vol. 7*. (Wiesbaden: Otto Harrassowitz Verlag).

Ruegg, David Seyfort (2010) *The Buddhist Philosophy of the Middle: Essays on Indian and Tibetan Madhyamaka*. (Boston, MA: Wisdom Publications).

Siderits, Mark, and Westerhoff, Jan (2016) *Studies in Buddhist Philosophy*. (Oxford: Oxford University Press).

Staal, Johan Frederik (1962) Negation and the Law of Contradiction in Indian Thought: A Comparative Study, *Bulletin of the School of Oriental and African Studies*, 25.1, 52–71.

Tanaka, Koji, Deguchi, Yasuo, Garfield, Jay and Priest Graham (2015) Nāgārjuna's Logic, in *The Moon Points Back,* (Oxford: Oxford University Press), 176–88.

Westerhoff, Jan (2009) *Nāgārjuna's Madhyamaka: A Philosophical Introduction*. *1st* ed. (Oxford: Oxford University Press).

Westerhoff, Jan (2009) The No-Thesis View: Making Sense of Verse 29 of Nāgārjuna's Vigrahavyāvartanī, in *Pointing at the Moon: Buddhism, Logic, Analytic Philosophy*, (Oxford: Oxford University Press), 25–39.

Yotsuya, Kodo (1999) *The Critique of Svatantra Reasoning by Candrakīrti And Tsong-Kha-Pa*. (Stuttgart: Steiner).

Tetsugaku
International Journal of
the Philosophical Association of Japan
Volume4, 2020

Article

« Le langage conditionne la pensée » Le son et le signe chez Levinas

HIRAOKA Hiroshi
Associé de recherche, Université de Tokyo

Résumé : *Le présent travail s'emploie à clarifier la théorie sur la raison que soutient Levinas dans* Totalité et infini*. L'un des enjeux de la pensée lévinassienne sur l'éthique consiste à montrer que la pensée raisonnable est conditionnée par le langage conçu comme accomplissant ma relation éthique avec autrui. Celle-ci s'accomplit par le langage dans sa sonorité qui fait retentir la présence de celui qui le profère. Cependant, le langage est aussi un système de signes verbaux. Aux yeux de Levinas, l'usage de ce système contribue au fonctionnement de la pensée raisonnable, mais ce système, tel qu'il est séparé de mon rapport éthique à autrui, équivaut au langage dans sa rationalité transparente qui entraîne le dialogue silencieux avec soi de la raison impersonnelle. En trouvant dans le silence un analogon du son, Levinas reconduit ce dialogue silencieux avec soi au dialogue oral avec autrui, ce qui signifie que la raison est originellement personnelle et le langage originel qui la conditionne est la parole. Mais cette dernière, telle qu'elle constitue une fable mythique du monde, reste une figure du dialogue silencieux avec soi de la raison impersonnelle, ce qui se produit aussi dans la lecture de l'écrit. Tout cela conduit Levinas à voir le langage conditionnant la raison personnelle dans la parole d'autrui qui enseigne le sens du monde. Écouter l'enseignement oral d'autrui implique de reconnaître autrui comme intelligence qui connaît le monde mieux que moi. Et l'enseignement d'autrui se déroule comme dialogue vivant entre autrui et moi à travers lequel le monde acquiert une signification rationnelle, ce qui suppose une mise en commun d'un système de signes verbaux. Dans ce dialogue, j'écoute la parole d'autrui avec une attention extrême afin d'être mieux instruit par lui sur le monde. En ce sens, la fonction originelle de la raison personnelle en moi consiste à écouter attentivement la parole d'autrui.*

Introduction

Le présent travail[1] a pour objectif de clarifier la thèse sur la pensée raisonnable que soutient Levinas dans *Totalité et infini*. Dans cet ouvrage, la question sur la raison, dont l'importance est souvent négligée par les commentateurs de Levinas, est en vérité l'un des enjeux de la pensée lévinassienne sur l'éthique. Ce dont témoigne Levinas lui-même dans sa conférence de 1962, en explicitant que son effort dans *Totalité et infini* vise à montrer comment ma relation éthique à autrui « se dessine à partir du problème même du commencement critique de la philosophie ».[2] Dans cet ouvrage, il écrit en effet que la philosophie, qui est la pensée raisonnable par excellence, « consiste à savoir d'une façon critique, c'est-à-dire à chercher un fondement à sa liberté à la justifier ».[3] Si la pensée tente de fonder une connaissance théorique sur le monde par une autre connaissance, elle mène à une régression à l'infini. Elle n'arrive à justifier la connaissance théorique sur le monde que si, en découvrant le dogmatisme et la spontanéité naïve de cette connaissance, elle se met en question. La philosophie implique ainsi la « mise en question de soi ».[4] Or, l'éthique est, par définition lévinassienne, précisément la « mise en question de ma spontanéité par la présence d'Autrui ».[5] C'est donc l'éthique qui accomplit l'essence de la pensée philosophique : la philosophie commence avec l'éthique. Cependant, il faut souligner que la pensée philosophique ou raisonnable ne se borne pas à cette critique de soi, elle porte aussi sur le monde, parce que la pensée, ainsi mise en question par autrui, est conduite à se justifier et, ce qui revient au même, à fonder la connaissance sur le monde. Dans l'éthique lévinassienne, il ne s'agit pas seulement de reconduire la théorie à l'éthique, mais aussi de déduire à partir de celle-ci la pensée raisonnable.

Il est aisé de constater que la façon dont Levinas remplit effectivement cette double tâche se lit dans sa conception du langage. Dans *Totalité et infini*, en effet,

[1] Ce travail a été soutenu par JSPS KAKENHI Grant Number JP19K20766.
[2] LC, 73. Les ouvrages de Levinas seront désignés par les abréviations suivantes :
« EE » : *De l'existence à l'existant* (1947), Paris, Vrin, 2004.
« TI » : *Totalité et infini. Essai sur l'extériorité* (1961), Den Haag, Martinus Nijhoff, 1984.
« EDE » : *En découvrant l'existence avec Husserl et Heidegger*, 2e éd., Paris, Vrin, 1967.
« HS » : *Hors sujet*, Saint-Clément-de-rivière, Fata Morgana, 1984.
« DQVI » : *De Dieu qui vient à l'idée*, 2e éd., Paris, Vrin, 1986.
« EN » : *Entre nous*, Paris, Grasset, 1991.
« LC » : *Liberté et commandement*, Saint-Clément-de-rivière, Fata Morgana, 1994.
« Œ 2 » : *Œuvres*, t. 2, *Parole et Silence*, Paris, Grasset/Imec, 2011.
[3] TI, p. 59.
[4] TI, p. 53.
[5] TI, p. 13.

après avoir défini le langage comme « mise en question de moi »,[6] c'est-à-dire comme accomplissant ma relation éthique avec autrui, Levinas soutient que le langage ainsi conçu rend possible la raison : « Le langage conditionne la pensée », c'est-à-dire « le fonctionnement de la pensée raisonnable ».[7] Afin de clarifier la théorie lévinassienne de la raison, il importe donc en particulier d'élucider ce conditionnement. Mais il n'est pas si facile d'appréhender les démarches précises de Levinas à ce sujet. Les commentateurs thématisant la conception lévinassienne du langage mettent l'accent tantôt sur son aspect vocatif ou interpellant, c'est-à-dire l'appel à la responsabilité pour autrui,[8] tantôt sur son déroulement dialogique oral sous la forme de questions et de réponses,[9] mais il nous semble que ces interprétations centrées sur l'oralité du langage ne suffisent pas à cette lecture. C'est plutôt sur l'analyse lévinassienne du langage *dans son système de signes* que la lecture attentive de *Totalité et infini* nous conduit à focaliser notre attention. En effet, Levinas établit le surgissement de la pensée raisonnable par le langage, en analysant le signe et sa signification : la fonction du signe se fonde sur la signification, c'est-à-dire « la présentation du sens » qui se produit originellement comme manifestation d'autrui, et cette signification originelle n'est rien d'autre que « le premier intelligible ».[10] La signification originelle est d'ordre théorique et non pas pratique. Cela montre que Levinas voit dans le recours au système de signes verbaux le fonctionnement de la pensée raisonnable : l'usage de ce système rend possible la raison. Mais il ne faut pas négliger le fait que Levinas affirme aussi que ce système de signes verbaux, tel qu'il est conçu en soi (c'est-à-dire séparé de la relation éthique à autrui), ne joue qu'un rôle servile à l'égard de la pensée de la « raison impersonnelle »[11] qui, en absorbant les interlocuteurs en elle, rend inutile le dialogue entre les êtres raisonnables. Il y a donc chez Levinas une tension entre la conception du système de signes verbaux dont l'usage conditionne la pensée raisonnable, d'un côté, et la critique de la raison impersonnelle qui, utilisant ce système, n'a pas besoin du dialogue, de l'autre. Cela suggère que si la pensée véritablement raisonnable surgit dans le langage accomplissant mon rapport éthique

[6] TI, p. 146.
[7] TI, p. 179.
[8] Cf. par exemple, C. Del Maestro, *La métaphore chez Levinas. Une philosophie de la vulnérabilité*, Bruxelles, Lessius, 2012, p. 58–64.
[9] Cf. par exemple, R. Moati, *Événements nocturnes. Essai sur* Totalité et infini, Paris, Hermann, 2012, p. 205–231.
[10] Respectivement, TI, p. 181 et 183.
[11] TI, p. 59.

avec autrui, alors cette raison, tout en utilisant les signes verbaux, n'en demeure pas moins personnelle dans l'oralité. Pour dégager la nature d'une telle raison personnelle, il faut avant tout élucider la façon dont Levinas, dans *Totalité et infini*, caractérise le signe et son rapport avec le son dans le langage.

Néanmoins, l'analyse du signe et de sa signification fournie dans *Totalité et infini* et d'autres ouvrages publiés du vivant de Levinas se révèle trop sommaire. Il est par conséquent indispensable de se référer aux conférences tenues pendant la période préparatoire à l'ouvrage de 1961 et publiées dans le deuxième tome des *Œuvres* de Levinas, surtout à la conférence en 1948, intitulée « Parole et Silence », où l'on trouve une réflexion détaillée sur le signe, sa signification et son lien avec la raison impersonnelle et l'analyse phénoménologique du son, ainsi qu'à la conférence de 1952 sur « L'Écrit et l'Oral » où Levinas approfondit son analyse sur le rapport entre le son et le signe dans le langage oral. À l'aide de la lecture de ces textes, nous nous emploierons à mettre en évidence les caractères fondamentaux de la raison personnelle lévinassienne.

1. Le système de signes verbaux et la raison impersonnelle

Commençons par examiner comment le système de signes verbaux est lié à la raison impersonnelle. Levinas définit le signe en général de la manière suivante : « il ne révèle qu'en cachant ».[12] Cela signifie qu'il se caractérise principalement par sa rupture avec « le signifiant », c'est-à-dire « l'émetteur du signe ».[13] Le signe est essentiellement émis par quelqu'un et par là manifeste dans une certaine mesure son émetteur. Cependant, cette manifestation n'est pas la manifestation de l'émetteur tel qu'il est en lui-même — manifestation *sui generis* que Levinas appelle « expression ». Car si l'émetteur se manifeste par le signe, c'est seulement en tant qu'il est signifié « par un signe dans un système de signes »,[14] en tant qu'il est devenu un signe. En ce sens, le signe cache son émetteur.

Cette détermination du signe en général s'applique également au signe verbal. Nous reviendrons plus loin sur la spécificité de ce dernier. Remarquons ici qu'en se référant à Merleau-Ponty (et indirectement à Saussure), Levinas conçoit le système de signes verbaux sous le principe holiste de la signification selon lequel « un signe

[12] TI, p. 151.
[13] TI, p. 69.
[14] TI, p. 152.

vise latéralement un autre signe et non pas du tout la signification elle-même ».[15] C'est alors le rapport latéral du signe au signe qui fonde le rapport du signe au sens : un signe signifie par sa différence par rapport à un autre signe. De sorte que le signifié d'un signe verbal est toujours à son tour un signe verbal et non pas son émetteur.[16] C'est pourquoi le signe « ferme » un accès à son émetteur « au moment même où il ouvre le passage qui mène au signifié ».[17]

Si le système de signes verbaux ainsi compris s'utilise par la raison impersonnelle, c'est par ses deux traits caractéristiques : 1/ l'idéalité du sens dans ce système et 2/ la rationalité de ce système.

1/ Selon Levinas, le sens du signe verbal ainsi conçu est un objet intentionnel : « Que les signes du langage […] soient diacritiques et signifient latéralement de signe à signe — il n'en reste pas moins vrai que la signification qu'ils établissent est un objet de pensée ».[18] De même que le signe est essentiellement « une identité formelle » et « idéale »,[19] son sens est idéal. Rappelons une constante dans la lecture lévinassienne de l'intentionnalité husserlienne : l'intentionnalité est chez Husserl une donation de sens (*Sinngebung*), c'est-à-dire le « processus d'identification de l'idéal », et par conséquent l'objet de la conscience est « une identité idéale », c'est-à-dire un « pôle idéal » qui « s'identifie à travers une multiplicité de visées ».[20] L'intentionnalité consiste à comprendre ou plutôt à *interpréter* « ceci en tant que cela », à subsumer ceci sous un sens idéal de cela. De sorte que tout comme le signe, les choses ne signifient que latéralement.[21] Le sens du signe et l'objet intentionnel ont en commun cette signification de *ceci en tant que cela*.[22]

2/ Ce système de signes verbaux est le langage saisi dans sa rationalité. Le sens du signe verbal utilisé par la pensée correspond, en vertu de son idéalité, aux objets visés par cette pensée. Les mauvais signes qui font écran sont déjà exclus du système. Levinas dit souvent que le mot joue un rôle de « fenêtre », pour insister sur

---

[15] Œ 2, p. 359. Cf. M. Merleau-Ponty, *Signes*, Paris, Gallimard, 1960, p, 49.
[16] Cf. TI, p. 69.
[17] TI, p. 157.
[18] Œ 2, p. 378.
[19] J. Derrida, *La voix et le phénomène* (1967), 3e éd., Paris, PUF, 2007. p. 56.
[20] EDE, p. 147. Cf. D. Pradelle, « Y a-t-il une phénoménologie de la signifiance éthique ? », in D. Cohen-Levinas et B. Clément (éd.), *Emmanuel Levinas et les territoires de la pensée*, Paris, PUF, 2007, p. 73–98, surtout p. 81–86.
[21] Cf. TI, p. 165–166.
[22] Cf. R. Calin, « Préface », in Œ 2, p. 37.

cette « transparence du langage ».[23] La pensée comprend bien ce que signifient les signes qu'elle utilise et l'objet auquel ils se réfèrent. La rationalité du langage accomplit « l'intimité silencieuse de la pensée avec l'être »[24] et par conséquent actualise la raison impersonnelle en chaque personne. Chaque personne pense ainsi suivant cette dernière, de sorte que les êtres raisonnables n'ont pas besoin de se parler l'un à l'autre. La pensée raisonnable ainsi conçue n'est rien d'autre qu'un « dialogue silencieux de l'âme avec elle-même ».[25]

Comment s'effectue ce dialogue silencieux avec soi ? Nous pouvons le lire dans la description de la représentation fournie dans B de la Section II de *Totalité et infini*. La représentation, c'est le *je pense,* la pensée à la première personne qui est en réalité une pensée universelle de la raison impersonnelle.[26] Nous avons deux points à remarquer. 1/ La représentation, définie par Levinas comme « détermination non réciproque de l'Autre par le Même »,[27] s'accomplit par l'intentionnalité donatrice de sens mentionnée ci-dessus, dont l'élément est la lumière. Cette dernière est, dit Levinas, l'« élément de la pensée où l'objet apparaît, se livre et où le signe verbal le désigne »[28] : dans la clarté et par la vision, le *je* interprète son objet en tant que ceci ou cela. L'objet idéal de la conscience est une « phosphorescence ou luisance »[29] dans le jeu de lumières, selon l'expression utilisée souvent par Levinas. 2/ Mais il faut souligner d'autre part qu'à ce sujet, Levinas évoque le *son*. À travers cette vision dans la lumière, le sujet *s'écoute penser* ou *écoute sa pensée* : « Le sujet qui pense par la représentation est un sujet qui écoute sa pensée : la pensée se pense dans un élément analogue au son et non pas à la lumière ».[30] Selon Levinas, c'est grâce à cet analogon du son que la pensée représentative se hisse à la raison impersonnelle. En pensant quelque chose, le « moi particulier » de la représentation écoute la pensée du « "démon" qui lui parle dans la pensée et qui est la pensée universelle ».[31] Ce « démon » socratique est une figure de la raison impersonnelle qui parle au fond du *je* et le guide.[32] La pensée du *je pense* se confond ainsi avec la

[23] Œ 2, p. 73.
[24] Œ 2, p. 75.
[25] Œ 2, p. 75. Levinas reprend la conception platonicienne de la pensée (cf. *Théétète*, 189 e – 190 a).
[26] Cf. TI, p. 6.
[27] TI, p. 99.
[28] Œ 2, p. 70.
[29] Œ 2, p. 77.
[30] TI, p. 99.
[31] TI, p. 99.
[32] Cf. TI, p. 250.

pensée universelle de la raison impersonnelle qui œuvre en lui, pour se transformer « en un “Je pense” qui ne parle plus ».[33]

Mais comment comprendre cet analogon du son qui est l'élément du dialogue silencieux de la pensée avec elle-même ? Nous pourrions y trouver la voix et l'écoute spirituelles ou spiritualisées, car, comme Derrida le met en évidence, la voix, en s'entendant, assure la conscience de l'objet idéal et surtout la présence à soi de la conscience se produisant comme « *auto-affection pure* ».[34] Nous acceptons dans ce sens la remarque de D. Arbib selon laquelle le moi de la représentation écoute « la voix de l'universel » recueillie « dans une écoute, non pas écoute *de l'autre*, mais écoute *de soi* ».[35] Cependant, il ne faut pas oublier que Levinas voit dans le silence de la représentation l'*analogon* du son. Qu'est-ce que cela signifie ? Remarquons que la représentation ainsi décrite est la représentation telle qu'elle est « Prise en elle-même, en quelque manière déracinée » et détachée « de ses sources ».[36] La suite de *Totalité et infini* montre en effet que la représentation s'effectue sous de multiples conditions : la vie de jouissance,[37] « la présence discrète du Féminin »[38] et notamment la rencontre avec le « visage indiscret d'Autrui qui me mette en question ».[39] C'est ce statut dérivé de la représentation qui est indiqué, nous semble-t-il, par le caractère analogue au son du silence de la représentation. En d'autres termes, et c'est un point capital, le dialogue silencieux avec soi de la représentation et son élément analogue au son marquent précisément *le lien étroit que la pensée raisonnable noue originellement avec le dialogue oral avec autrui*. C'est donc à tort que D. Arbib affirme que l'analogon du son de la représentation décrite dans *Totalité et infini* contredit la sonorité du son qui, dans la conférence de 1948, « Parole et Silence », est analysée comme introduisant une

[33] TI, p. 183. Nous verrons plus loin que cette pensée n'est en vérité qu'une pensée mythique du monde, dans la mesure où elle consiste à interpréter le monde.
[34] J. Derrida, *op. cit.,* p. 89. Ajoutons que l'intention de *La voix et le phénomène* consiste à affirmer que cette auto-affection implique en vérité une différence pure et par là n'est jamais pure (*ibid*., p. 92). Sur la « spiritualisation de la voix et de l'écoute » dans *La voix et le phénomène*, nous nous reporterons à M. Dufrenne, *L'œil et l'oreille*, Paris, Jean-Michel Place, 1991, p. 55.
[35] D. Arbib, *La lucidité de l'éthique. Études sur Levinas*, Paris, Hermann, 2014, p. 32–33.
[36] TI, p. 95.
[37] Cf. TI, p. 101.
[38] TI, p. 145.
[39] TI, p. 145.

altérité dans le monde.[40] La voix spirituelle qui est le silence dialogique de la raison impersonnelle renvoie à une voix réelle que quelqu'un adresse à quelqu'un d'autre.

En comprenant comment le système de signes verbaux sert, grâce à son idéalité et sa rationalité, le dialogue silencieux avec soi de la raison impersonnelle, lequel à son tour renvoie au langage oral, nous arrivons ainsi à penser que si la raison personnelle utilise le système de signes verbaux, ce langage n'en garde pas moins une sonorité. Le langage de la raison personnelle est à la fois transparent et sonore. Il faut à présent approfondir en quel sens la sonorité du langage concerne le surgissement de la raison personnelle. Examinons pour cela la phénoménologie du son de 1948 où Levinas analyse la sonorité du son et, à partir de cette dernière, le signe verbal.

2. L'enchevêtrement du son et du signe dans la parole

Dans cette analyse, il s'agit tout d'abord de décrire le son comme événement introduisant une altérité dans le monde. Tout en venant du dehors, le son, dans la mesure où il se donne comme une sensation parmi d'autres, est « entendu » et par conséquent compris « comme s'il venait de nous ».[41] Cependant, la sonorité du son consiste dans son retentissement, qui annonce à celui qui l'entend un autre « qui ne peut être donné à la lumière ».[42] De plus, la sonorité indique aussi « ce qu'il y a d'événements dans toutes les manifestations de l'être » ou, autrement dit, « ce qu'il y a de verbe dans tous les substantifs »,[43] parce que le son est un bruit qui accompagne toute action. Par le son dont elle s'accompagne, l'action retentit non pas comme un fait passé, mais comme un événement singulier présent. Le son dans sa sonorité révèle ainsi le monde dans son aspect d'événement ou de verbe. En ce sens, il est le retentissement même de « l'être en tant qu'autre »[44] et par là réalise une relation avec l'autre.

Mais ce n'est pas ici que cette phénoménologie du son est menée à terme. Elle est destinée, soulignons-le, « au problème du langage comme signe ».[45] Après avoir décrit la sonorité du son, Levinas reconduit le mot au son : « l'élément naturel

[40] D. Arbib, *op. cit.*, p. 34.
[41] Œ 2, p. 90.
[42] Œ 2, p. 95.
[43] Œ 2, p. 91.
[44] Œ 2, p. 92.
[45] Œ 2, p. 94.

du mot est le son ».[46] Le mot ne se réduit pas pour autant à un son pur, à une voix qui ne dit rien, car « le mot a une signification, […] il est par conséquent aussi signe ».[47] Levinas se voit ainsi conduit à la tâche de « déduire la modification essentielle du son en mot ».[48]

Pour préciser ce dont il s'agit ici, il est crucial de comprendre la distinction entre le langage comme signe et le langage comme son, que Levinas appelle en 1948 « symbole ».[49] Le symbole n'est rien d'autre que le mot dans sa sonorité, le mot tel qu'il retentit, le mot dans son caractère d'événement : « la valeur symbolique de l'expression […] s'accomplit dans l'élément du son ».[50] Grâce à cette fonction symbolique, le langage accomplit une relation avec l'autre réfractaire à la lumière. En revanche, le signe transparent ne permet pas de dépasser le monde de la lumière, parce qu'il « n'est qu'un renvoi à ce qui est absent et non pas ce qui est inaccessible — à la lumière ».[51] Nous voyons ainsi que la différence entre le symbole et le signe est moins substantielle que fonctionnelle. En effet, si l'expression verbale a une valeur symbolique qui s'accomplit dans l'élément du son, « l'expression est aussi signe », dit Levinas.[52] Un mot peut fonctionner dans son symbolisme ou dans son système de signes, comme son ou comme signe. Cela signifie que, par « expression », Levinas pense avant tout à un langage qui possède cette double fonction de son et de signe, c'est-à-dire au *langage parlé*. Le langage originel conçu à partir du son, c'est la parole. Or, la parole n'est bien entendu pas seulement une voix qui retentit ou un simple événement sonore. Une fois proférée, elle se fige inévitablement en signes silencieux, dans la mesure où elle est un ensemble de signes. En ce sens, la déduction de la modification du son en mot consiste en réalité à décrire *un tel enchevêtrement du son et du signe dans la parole*, c'est-à-dire ce qu'apporte l'usage du système de signes verbaux par la parole,[53] ce que Levinas appelle « dialectique de la parole ».[54] Suivons cette dialectique sommairement.

Elle part du langage dans sa sonorité. Si la sonorité du son fait retentir l'être en tant qu'autre, la sonorité du mot consiste à « faire retentir l'altérité même du

[46] Œ 2, p. 92.
[47] Œ 2, p. 92.
[48] Œ 2, p. 92.
[49] Œ 2, p. 92.
[50] Œ 2, p. 95.
[51] Œ 2, p. 95.
[52] Œ 2, p. 95.
[53] Ce que méconnaît D. Arbib, lorsque, par cette déduction, il entend l'intégration immédiate du son au monde de la lumière (cf. D. Arbib, *op. cit.*, p. 24–25).
[54] Œ 2, p. 100.

sujet »[55] qui l'énonce. Par exemple, ma parole fait de moi un être sonore pour moi-même et surtout pour les autres.[56] La parole dans sa sonorité révèle en ce sens que la relation intersubjective se produit originellement comme un événement temporel : « l'accomplissement de l'ordre intersubjectif […] est le temps ».[57] Dans ce temps de la relation intersubjective, selon Levinas, la parole instaure la « simultanéité »[58] du monde de la lumière. Une telle parole, c'est la « fable », qui consiste à *interpréter* une société et un monde pour former une civilisation où la personne est désormais abordée à travers la fable.[59] Au lieu de retentir en tant qu'autre, je m'y manifeste en renvoyant à autre chose que moi, tout comme dans le système de signes : « Le sujet se manifeste […] revêtu de son mythe, dans sa décence. C'est moi-même, mais déjà engagé dans des relations qui m'identifient et que le mot comme un signe évoque ».[60] La parole fabulant, dans la mesure où elle recourt au système de signes transparents et impersonnels, arrive inévitablement à cacher le sujet qui la profère et la relation intersubjective originelle.[61] La fable ne forme en ce sens pas autre chose que le dialogue silencieux avec soi de la raison impersonnelle.

Cette dialectique de la parole nous montre que la parole qui fait surgir la raison personnelle se distingue de la fable. Comme l'indique l'expression « mythe » dans la phrase citée, la fable ne forme en réalité qu'un mythe et non pas une pensée raisonnable. Quelle est alors la parole fondant la raison qui, tout en utilisant le système de signes verbaux transparents, n'interprète pas le monde ? Dans la conférence de 1948, Levinas suggère la réponse : « l'enseignement », c'est-à-dire la parole d'autrui telle qu'elle *enseigne le sens du monde* « *dans l'élément* […] *du son* ».[62] Car l'enseignement d'autrui dans sa sonorité m'annonce « la présence

[55] Œ 2, p. 92.
[56] Cf. HS, p. 221 : « Parler, c'est interrompre mon existence de sujet et de maître, mais l'interrompre sans m'offrir en spectacle, *en me laissant simultanément objet et sujet* » (nous soulignons). À ce sujet, nous nous reporterons aux analyses par M. Dufrenne sur l'ouïe (cf. M. Dufrenne, *op. cit.*, p. 94).
[57] Œ 2, p. 99. Ce que Levinas montre sous la figure de la fécondité, dont il a développé l'analyse dans *Le temps et l'autre*. Le présent travail ne s'attarde pas sur ce point.
[58] Œ 2, p. 99.
[59] Cf. Œ 2, p. 99–100.
[60] Œ 2, p. 99.
[61] Le monde civilisé instauré par la fable est le monde en tant que société décente, décrit déjà en 1947 dans *De l'existence à l'existant* (cf. EE, p. 59–64). Cette description s'approfondira dans *Totalité et infini* où Levinas décrit le monde économique comme marché anonyme des œuvres où tout est échangeable (cf. TI, p. 136).
[62] Œ 2, p. 83 (nous soulignons). À notre avis, dans ce sens, la conception lévinassienne de

d'une raison autre que la mienne ».[63] Écouter l'enseignement d'autrui, cela implique de reconnaitre autrui « comme maître »,[64] c'est-à-dire comme « une intelligence ».[65] L'écoute de l'enseignement d'autrui n'est ni une sensation, ni une interprétation, mais précisément la compréhension raisonnable originelle, c'est-à-dire la compréhension de la signification originelle que nous avons évoquée au début du présent travail.

L'examen de l'analyse lévinassienne du son en 1948 nous conduit donc à dire que le langage transparent et sonore qui fait surgir en moi la raison personnelle, distinct de la fable mythique qui revient au dialogue silencieux avec soi de la raison impersonnelle, s'effectue comme enseignement oral. Mais la parole qui enseigne, comme c'est le cas dans la fable, n'en devrait pas moins recourir aux signes verbaux. Sinon, elle ne m'enseignerait rien sur le monde. Comment alors aboutit-elle à la pensée raisonnable ? Pour répondre à cette question, il nous est à présent nécessaire d'élucider comment l'enseignement d'autrui se déroule effectivement, en nous appuyant sur la conférence de 1952 sur « L'Écrit et l'Oral » et *Totalité et infini*.

3. La raison attentive qui écoute

Examinons en quoi le signe verbal se distingue du signe en général. Nous avons vu que le trait caractéristique du signe en général réside dans sa rupture d'avec son émetteur : ce dernier, à cause de la signification latérale du signe, ne s'y annonce que tel qu'il est un signifié et par conséquent un signe à son tour, et non pas tel qu'il est en lui-même. De sorte qu'il est impossible de remonter à partir du signe vers son émetteur comme tel. Si le langage écrit se détermine par Levinas comme « redevenu signe »,[66] c'est parce qu'il possède lui aussi ce trait. L'écrit signifie latéralement et par conséquent indépendamment de son auteur. Lorsque je tente de comprendre quelqu'un à partir de ce qu'il écrit, j'arrive à le comprendre seulement tel qu'il se

---

la parole n'est pas tributaire de ce que Derrida appelle « phonocentrisme ». La transparence pure du langage devant la pensée qui constitue le trait caractéristique du phonocentrisme ne s'accomplit pas par l'enseignement oral d'autrui, mais par le langage utilisé dans la fable qui est une figure du dialogue silencieux avec soi de la raison impersonnelle. La parole n'est chez Levinas pas purement et simplement transparente grâce à sa sonorité. Examiner en quel sens Levinas s'oppose au phonocentrisme en détail fera l'objet d'un autre travail.

[63] Œ 2, p. 95.
[64] Œ 2, p. 93.
[65] TI, p. 191.
[66] TI, p. 157.

signale par le contenu de son écrit. En se manifestant dans son écrit, cette personne est abordée à la troisième personne, comme c'est le cas dans la fable. Cette personne est donc toujours méconnue, comme le sont les morts évoqués par l'historiographe à partir de leurs œuvres héritées.

Qu'est-ce qui constitue alors la spécificité du signe verbal par rapport au signe en général ? Selon Levinas, ce n'est rien d'autre que l'intention d'exprimer. Le signe non verbal est émis par quelqu'un, mais sans être voulu ; tandis que le signe verbal est impossible sans cette intention : « L'acte libre d'écrire, acte aboutissant à une œuvre est aussi librement voulu comme expression ».[67] L'écrit garde essentiellement une marque de la volonté d'exprimer. En plus, par cette marque, l'écrit renvoie à la parole : « L'Écrit est le seul produit humain […] qui repose sur la parole, animée par l'intention d'exprimer ».[68] L'auteur n'est donc chez Levinas pas totalement absent de son écrit, « Puisqu'il fallait qu'il fût là et qu'il s'exprimât ».[69] Autrement dit, « à partir du passé »,[70] l'écrit me parle maintenant. C'est ainsi que, en lisant un écrit, je suis engagé « dans la situation du dialogue »[71] : je me pose des questions sur tel ou tel contenu du texte. Mais le discours écrit « n'entend pas mes questions ».[72] D'où la nécessité de l'*interprétation* : « Le lecteur se pose une question à laquelle il ne peut trouver de réponses que dans d'autres écrits. On appelle cela interpréter ».[73] La lecture de l'écrit forme donc un dialogue silencieux du lecteur avec lui-même : « En interprétant on fournit soi-même les questions et les réponses — c'est-à-dire on se trouve dans la situation même que Platon décrit pour caractériser la pensée, dialogue silencieux de l'âme avec elle-même ».[74] La lecture de l'écrit est en ce sens une figure du dialogue silencieux avec soi de la raison impersonnelle, tout comme la fable.

En revanche, celui qui parle assiste toujours à sa parole même : « La parole […] déverrouille ce que tout signe ferme au moment même où il ouvre le passage qui mène au signifié, en faisant *assister* le signifiant à cette manifestation du signifié ».[75] Cette assistance consiste à « "porter secours" à sa parole »,[76]

[67] Œ 2, p. 208.
[68] Œ 2, p. 208.
[69] Œ 2, p. 213.
[70] Œ 2, p. 211.
[71] Œ 2, p. 209.
[72] Œ 2, p. 213.
[73] Œ 2, p. 221.
[74] Œ 2, p. 213.
[75] TI, p. 157.
[76] Levinas emprunte cette expression au *Phèdre* de Platon (cf. *Phèdre*, 275 e).

c'est-à-dire à reprendre sans cesse sa parole, « Comme si la présence de celui qui parle inversait le mouvement inévitable qui conduit le mot proféré vers le passé du mot écrit ».[77] Par cette reprise se rétablit sans répit le lien, toujours sur le point d'être rompu, du signe délivré avec celui qui le parle. Enseignant, autrui reprend sans cesse sa parole. Cela signifie que, selon Levinas, en écoutant la parole d'autrui ou, ce qui revient au même, en étant instruit par lui, je ne reçois pas seulement une proposition thématisant le monde, mais aussi « la possibilité de questionner »,[78] c'est-à-dire de *poser des questions à autrui qui me parle*. L'enseignement d'autrui est inséparable des questions que je lui adresse, car c'est autrui lui-même qui me répond : « le contenu qui s'offre à moi est inséparable de celui qui l'a pensé, ce qui signifie que l'auteur du discours répond aux questions ».[79] Cela est rendu possible par la sonorité de la parole : si « Dans la parole vivante, le son […] disparaît derrière la pensée exprimée »,[80] il n'en fait pas moins retentir la présence d'autrui qui parle, parce que la question que je lui pose s'explique « par la présence de celui à qui elle s'adresse »,[81] en l'occurrence, la présence d'autrui. Porter secours à sa parole, reprendre sa parole sans cesse, cela consiste donc à *continuer à répondre aux questions que suscite sa parole*. Il n'y a aucune réponse qui ne soit pas suivie de nouvelles questions de ma part et de nouvelles réponses de la part d'autrui, comme le clarifient E. Féron et récemment R. Moati.[82] Ne pas répondre aux questions équivaut à laisser se figer sa parole en langage écrit. L'enseignement oral d'autrui se déroule donc comme dialogue vivant entre le maître (autrui) et le disciple (moi).

Allons plus loin, et dégageons deux remarques sur la raison personnelle, dont l'une sur son surgissement et l'autre sur son caractère fondamental.

1/ Ce dialogue, qui semble de prime abord réciproque, ne contredit néanmoins pas l'« asymétrie »[83] de ma relation avec autrui. Dans le dialogue oral avec autrui, *je suis toujours enseigné par autrui*. C'est afin d'être mieux instruit sur le monde par autrui que je pose des questions, ce qui n'est possible que si, comme nous l'avons vu, je reconnais autrui comme maître, c'est-à-dire comme *intelligence qui connaît le monde mieux que moi*. De sorte que le dialogue ne suppose jamais une

[77] TI, p. 41.
[78] TI, p. 69.
[79] TI, p. 43.
[80] Œ 2, p. 210.
[81] TI, p. 69.
[82] Cf. E. Féron, « La réponse à l'autre et la question de l'un », *Études Phénoménologiques*, t. 6, n° 12, 1990, p. 67–100, surtout p. 69–73 ; R. Moati, *op. cit*., p. 209–211.
[83] TI, p. 24.

mise en commun préalable des pensées des interlocuteurs, mais seulement *celle d'un système de signes verbaux*. C'est ainsi que se fonde l'intelligibilité du monde : « L'objectivité […] se *pose* dans un discours, dans un *entre-tien* qui *propose* le monde ».[84] Le monde y acquiert « une signification *rationnelle* »,[85] sans devoir être interprété. Il n'est pas un monde des objets idéals phosphorescents, mais l'être : « L'être est un monde *où l'on parle et dont on parle* ».[86] Le dialogue oral avec autrui thématisant le monde en fournit le savoir philosophique par excellence. C'est en ce sens que la parole d'autrui qui enseigne fait surgir la raison personnelle en moi.

2/ L'enseignement oral d'autrui m'engage à lui répondre en lui posant des questions sans tarder. Il exige mon « attention extrême »[87] à l'égard d'autrui et de sa parole. Ou plutôt, l'attention est originellement celle « qui *essentiellement* répond à un appel » d'autrui.[88] La fonction originelle de la raison consiste donc, et telle est l'originalité de la théorie lévinassienne de la raison, à *écouter attentivement la parole d'autrui*. Cette écoute attentive me permet de répondre à autrui en lui posant des questions. Cependant, cela ne me dissout pas dans la pensée universelle de la raison impersonnelle, car l'enseignement oral que j'écoute fait toujours retentir l'altérité de la raison d'autrui qui m'enseigne le sens du monde. *La raison attentive qui écoute la parole d'autrui*, telle est précisément la raison personnelle en moi.

Conclusion

Récapitulons nos réflexions sur la pensée lévinassienne de la raison personnelle : 1/ en analysant comment le système de signes verbaux conçus en soi entraîne le dialogue silencieux avec soi de la raison impersonnelle, nous avons vu que ce dialogue silencieux avec soi renvoie au dialogue oral avec autrui ; 2/ en examinant en quel sens la sonorité du son introduit l'altérité et comment la parole interprétative qui est la fable mythique ne forme en réalité qu'un dialogue silencieux avec soi de la raison impersonnelle, nous avons remarqué que la parole qui fait surgir en moi la raison personnelle est l'enseignement oral d'autrui dont la sonorité me fait reconnaître autrui comme intelligence qui connaît le monde mieux que moi ; 3/ en

[84] TI, p. 68.
[85] TI, p. 184 (nous soulignons).
[86] TI, p. 156 (nous soulignons).
[87] TI, p. 153.
[88] TI, p. 73.

élucidant la façon dont l'enseignement d'autrui se déroule comme dialogue oral entre moi et autrui supposant une mise en commun d'un système de signes verbaux, nous sommes parvenus à clarifier que la fonction originelle de la raison personnelle en moi consiste à écouter attentivement la parole d'autrui.[89]

Dans les textes ultérieurs, Levinas thématisera à nouveau cette conception de la raison personnelle à partir de la phénoménologie husserlienne. La raison personnelle attentive à autrui s'y conçoit comme conscience non-intentionnelle plus originelle que la conscience intentionnelle et qui s'appelle « éveil ».[90] La question fondamentale reste cependant toujours la même : celle de savoir si, en tant que conscience identificatrice, la raison s'éveille suffisamment. D'où la conception de la philosophie comme « vigilance contre l'évidence et contre ses rêves de plein jour », ou, en un mot, contre la « naïveté ».[91] Il s'agit en réalité de considérer l'éthique comme philosophie première. Car pour que cette vigilance soit véritablement éveillée, pour que le sommeil se dissipe, il faut une mise en question de soi par autrui. En ce sens, être éthique consiste à réveiller sans cesse sa raison personnelle de son assoupissement dogmatique toujours réitéré.

Bibliographie

Levinas, Emmanuel (1947) *De l'existence à l'existant*. (Paris: Vrin, 2004). « EE »
Levinas, Emmanuel (1961) *Totalité et infini. Essai sur l'extériorité.* (Den Haag: Martinus Nijhoff, 1984). « TI »
Levinas, Emmanuel (1967) *En découvrant l'existence avec Husserl et Heidegger*, 2e éd. (Paris: Vrin). « EDE »
Levinas, Emmanuel (1984) *Hors sujet*. (Saint-Clément-de-rivière: Fata Morgana). « HS »
Levinas, Emmanuel (1986) *De Dieu qui vient à l'idée*, 2e éd. (Paris: Vrin). « DQVI »
Levinas, Emmanuel (1991) *Entre nous*. (Paris: Grasset). « EN »
Levinas, Emmanuel (1994) *Liberté et commandement*. (Saint-Clément-de-rivière:

[89] L'analyse du présent travail concernant la sonorité du son et celle de la parole nous permettra de comprendre la pensée du dernier Levinas, où on trouve plusieurs expressions concernant le son : résonance, écho, voix, etc. Sans analyser ces expressions, on ne saurait clarifier les concepts du dernier Levinas : trace, passé immémorial, et surtout *se*.
[90] EN, p. 105. Cf. J. Colette, « Lévinas et la phénoménologie husserlienne », *Les Cahiers de la Nuit surveillée*, nº 3, 1984, p. 19–36, surtout p. 27–31.
[91] DQVI, p. 35.

Fata Morgana). « LC »
Levinas, Emmanuel (2011) *Œuvres*, t. 2, *Parole et Silence*. (Paris: Grasset/Imec). « Œ 2 »
Arbib, Dan (2014) *La lucidité de l'éthique. Études sur Levinas*. (Paris: Hermann).
Colette, Jacques (1984) « Lévinas et la phénoménologie husserlienne ». *Les Cahiers de la Nuit surveillée*, nº 3, 19–36.
Del Maestro, Cesare (2012) *La métaphore chez Levinas. Une philosophie de la vulnérabilité.* (Bruxelles: Lessius).
Derrida, Jacques (1967) *La voix et le phénomène*, 3e éd. (Paris: PUF, 2007).
Dufrenne, Mikel (1991) *L'œil et l'oreille*. (Paris: Jean-Michel Place).
Féron, Etienne (1990) « La réponse à l'autre et la question de l'un ». *Études Phénoménologiques*, t. 6, nº 12, 67–100.
Merleau-Ponty, Maurice (1960) *Signes*. (Paris: Gallimard).
Moati, Raoul (2012) *Événements nocturnes. Essai sur* Totalité et infini. (Paris: Hermann).
Pradelle, Dominique (2007) « Y a-t-il une phénoménologie de la signifiance éthique ? ». in D. Cohen-Levinas et B. Clément (éd.), *Emmanuel Levinas et les territoires de la pensée*, (Paris: PUF), 73–98.

Editorial Notes

Volume 4 of this journal is a special issue on “Analytic Asian Philosophy”, a completely new field which was suggested by our colleague on the editorial team, Deguchi Yasuo. The chief editor would note that bringing together this volume by receiving a sufficient number of articles was a difficult challenge for us. But here we have a fascinating collection containing a wide range of subjects relating both to analytic and to Eastern philosophy. We would be grateful if the readers would send us their comments (ejournal@philosophy-japan.org). On-line communication will motivate more and more the “raison d’être” of our journal, particularly in this present confined life with Covid-19.

I wish to express by sincere gratitude to all the authors for their precious contributions to the present issue. I would also like to acknowledge the generous advice and help I received from deputy chief editors, Jeremiah Alberg and Baba Tomokazu, the other editorial members of *Tetsugaku* and all the external collaborators. I should not forget my great thanks due to our young assistant editors, Yamamori Maiko and Shirakawa Shintaro for their regular, careful and efficient help with the editing process.

U. M.
5 May 2020, Kyoto

Call for Papers for *Tetsugaku* Vol.5, 2021 Spring

Special Issue: "Philosophy of Care"

Tetsugaku: The International Journal (e-journal) of the Philosophical Association of Japan, calls for papers for the special issue, "Philosophy of Care" (Vol. 5, 2021).

Philosophy of care was initially proposed by nursing practitioners to grasp their own attitudes in their practices in the 1960s, when paternalism was first being discussed. This line of thought respects the patient's self-decision, and the issues quickly become complicated. Nurses and medical practitioners face the needs and wants of patients whose capacity of consent is limited. They have strived to establish their own guidelines of care according to their choice among a wide range of moral views. Moreover, consideration of the vulnerable must be included in the deliberation. Society of Hospital Medicine characterizes vulnerable populations as "groups who are at increased risk of receiving a disparity in medical care on the basis of financial circumstances or social characteristics such as age, race, gender, ethnicity, sexual orientation, spirituality, disability, or socioeconomic or insurance status."

Furthermore, philosophical concerns around care are cast not just against medical care but also against care given during/after disasters, war, industrial and economical restructuring, technology, and any forms of threats which deprive opportunity or capacity to people to decide their own course of life.

Thus, the special issue covers the following themes (non-exhaustive):

Philosophy of nursing
Philosophy of restorative justice
Productivity and successful aging
Phenomenological approach to care
Philosophy of narratives
Philosophy of Disaster
Philosophy of War and displacement
Feminist ethics
Ethics and care in business
Ethical concerns on care using artificial intelligence
Enhancement vs normalization

Philosophical analysis of participatory research

[Deadline: 31 October 2020]
To submit your paper, please read carefully our Guidelines for Contributors.

Submission guidelines are available at
http://philosophyjapan.org/en/international_journal/guideline/