

Wittgenstein, Phenomenology and Cognitive Linguistics

Yingjin Xu¹

Received: 24 March 2017 / Accepted: 26 May 2017 / Published online: 26 June 2017
© Fudan University 2017

Abstract Although the literature devoted to the naturalization of mainstream phenomenology has been blooming recently, not so many efforts have been made to make the intellectual legacy from Wittgenstein, who could also be viewed as a “linguistic phenomenologist,” accessible to cognitive science. The reluctance of making Wittgenstein naturalized is sometimes backed by the worry that Wittgenstein’s criticism of the notion of “thinking” as some “internal process” is also potentially threatening the computational theory of cognition. But this worry itself is based on some serious misunderstandings of the internal/external dichotomies, the clarifications of which would greatly relieve the tension between Wittgenstein and cognitive science. Moreover, cognitive linguistics could also be viewed as the intermediate theory between Wittgenstein and cognitive science due to the affinities it bears with both Wittgenstein’s later philosophy and cognitive science.

Keywords Cognitive linguistics · Linguistic phenomenology · Schemas · Internalism

1 Introduction

Although the researches of the relationship between phenomenology and cognitive science (hereafter the “P/C relation”) have been blooming recently, the relationship between Wittgenstein and cognitive science (hereafter the “W/C relation”) has been basically neglected in the literature, despite some unobtrusive studies (which would be addressed in this article a bit later). This situation is, nonetheless, not understandable at all. Firstly, Wittgenstein’s rich comments on the nature of

✉ Yingjin Xu
yjxu@fudan.edu.cn

¹ Fudan University, Shanghai, China

sensation and mentality naturally render him fairly relevant to the main concerns of phenomenologists, philosophers of mind as well as cognitive scientists; hence, one may naturally expect that later Wittgenstein's legacy could offer many inspirations for cognitive science just like Husserl and Heidegger's legacies do; secondly, even when the label of "phenomenological reduction" is applied to Wittgenstein's philosophical methodology in a general sense, it is undeniable that his version of "phenomenological reduction" is executed in a fairly loose manner (namely loose enough to accommodate some weak form of naturalism); hence, the "naturalization" of Wittgenstein should be much easier than that of mainstream phenomenologists. Therefore, there appears to be no reason to preclude Wittgenstein from our picture if a broader vision of the phenomenology/cognitive science alliance is pursued.

But what theoretical benefits would be brought to cognitive science if Wittgenstein's legacy is also put on the table? It is widely known that Wittgenstein is first of all a philosopher of language, and his notion of "language," whatever it is, is different from both the behaviorist understanding of it as linguistic behaviors (or dispositions of delivering them) and the nativist understanding of it as encoders of the Fodorrian "Language of Thought" or Chomskyan generative grammar. Hence, in some unstrict term, Wittgenstein's notion of language, or even that of "thinking," is overarching the gap between the "outer" and the "inner." This feature, as I will argue, makes his idea bear some significant affinity with contemporary cognitive linguistics, which is itself both in debt to phenomenological thinking on the one hand and friendly to cognitive modeling on the other. In this sense, the Wittgenstein/cognitive science alliance, which is intended to be built by this article, is actually an alliance between Wittgenstein and *cognitive linguistics*. Such a proposal should be somehow distinct from Varela (1996)'s approach of doing "neuropsychology" and therefore bring more diversity to the current spectrum of "naturalized phenomenology." Moreover, this research would also be beneficial to Wittgenstein studies in the sense that Wittgenstein's notion of "thinking" could be therefore construed in terms of "cognitive schemas," i.e., a core concept of cognitive linguistics.

This article is composed of three parts. Firstly, I will make some general remarks on the triadic relation among Wittgenstein, phenomenology and naturalism and therefore explain why it is easier to naturalize Wittgenstein than to naturalize traditional phenomenologists; secondly, I will criticize some typical arguments preventing one from positively evaluating the W/C relations; and thirdly and most importantly, I will explain why cognitive linguistics should play a pivotal role in the process of reconstructing the so-called Wittgensteinian cognitive science.

2 The First Approximation to the Triadic Relation Among Wittgenstein, Phenomenology and Naturalism

Since the W/C relation is evaluated in this article in the background of the so-called naturalized phenomenology, it would be helpful to assess the W/P relation in the first place as a preliminary discussion.

First of all, we do have some reasons to confirm the existence of the affinity between Wittgenstein's thought and phenomenology. By affirming this point, I do not mean anything related to Hintikka and Hintikka (1986) and Byong-Chul Park (2012)'s bold interpretation of early Wittgenstein's *Tractatus* as a phenomenological system in which "atomic facts" should be construed as entities like unanalyzable Russellian sense data. Rather, my positive view about the W/P relation is mostly motivated by the following observations; some are initially inspired by Herbert Spiegelberg (1981: 202–228)'s discovery of "The puzzles of Wittgenstein's *Phänomenologie*":

- (1) As Spiegelberg (1981: 83–93) insightfully pointed out, J. Austin's treatment of ordinary language, which is undoubtedly within the tradition of later Wittgenstein, is not only labeled by 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" (surely for J. Austin, "the Given" means the linguistic phenomena, whereas for Husserl "the Given" means "the phenomena within consciousness," but 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 rejection of resorting 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 entities like "cones," "rods," "waves").
- (2) The term "phenomenology" or "phenomenological language" was employed by Wittgenstein himself when his philosophical transition from *Tractatus* to *Philosophical Investigation* is being conducted in 1920s–1930s. Although the only existing text in which Wittgenstein commented on Husserl does not show any of his special knowledge about the latter's work (cf. Waismann 1979: 67, where Wittgenstein is more interested in expressing his own ideas about the categorization of judgments when he was asked to make comments on Husserl's notion of "synthetic a priori judgments"), Wittgenstein himself did, maybe during the time when his 1929 paper "Some Remarks on Logic Forms" (Wittgenstein 1929) was being prepared, intend to view "phenomenological language" as a primary linguistic system in which any discourses about physical entities should be suspended. And this position may be viewed as middle Wittgenstein's version of "phenomenological reduction." Although he gradually discarded this idea by asserting that it is misleading to establish "phenomenological language" as an independent linguistic system which can be detached from ordinary language (cf. Wittgenstein 1980: §1), he still insisted on using the term "phenomenology" as a title of the sections in his posthumously published manuscript *The Big Typescript* (Wittgenstein 2005: §94). In this material, although "phenomenological studies" mainly means a systematic survey of grammars already imbedded in ordinary language, such survey is still in the shadow of "phenomenological

reduction” of the Austin type in the sense that it still exclusively appeals to what can be given to language speakers. Even in *Some Remarks on Colours*, a manuscript elaborated when Wittgenstein was very near to the end of his own life, he still acknowledges the meaningfulness of “phenomenological problems” along with his negative attitudes toward “phenomenology” as an independent discipline detached from ordinary language (Wittgenstein 1979 I: §53). Therefore, we may conclude that Wittgenstein’s own philosophy is deeply tangled with phenomenological thinking in this aspect or another.

However, there are also a couple of reasons for distancing Wittgenstein from classical phenomenology. By asserting this, I do not mean that Wittgenstein’s well-known argument against private language would immediately make his philosophical image hostile to Husserl, since such an image may be based on some substantial misunderstanding of Husserl as an advocate of “private language.” Neither do I intend to mean that Wittgenstein’s emphasis on the “form of life” induces a new type of intersubjectivity-oriented philosophical discourse that Husserl cannot fruitfully participate, since the Husserlian notion of “*Lebenswelt*” could also make the issues on intersubjectivity salient. What I mean is rather more related to their different attitudes toward natural languages. As Wataru Kuroda (1979) sharply points out, a careful reading of Husserl’s *Logical Investigations* (Husserl 2001, especially the first study of six studies within Volume Two) will lead one to see how Husserl reduces the communicative function of language (i.e., the function of “indication” or “*Anzeichen*”) to its expressive function (or “*Ausdruck*”) on the one hand and neglects the physical aspect of signs on the other. But for Wittgenstein, the communicative function of language is so fundamental that it should survive any type of phenomenological reduction, and in this sense, Husserl’s neglect of the physical aspect of signs would also appear to be unacceptable since such an aspect is so essential to any possible form of social communications. Therefore, it may be concluded that although Wittgenstein conducts some general form of “phenomenological reduction” in his later philosophy, the “residue” of such reduction does possess some fairly rich social and physical features that its Husserlian counterpart does lack. And the existence of these features, as I believe, would naturally make it easier to naturalize Wittgenstein than to naturalize Husserl, since the whole Wittgensteinian notion of language looks more friendly to some form of physicalism (e.g., Ryle’s behaviorism), which is the metaphysical backbone of naturalism.

However, since “a naturalized Wittgenstein” is not a widely accepted philosophical image yet (due to the reasons that I will mention in the next section), it may be helpful to learn something about the existing methodologies to naturalize Husserl or Heidegger in the first place. Roy et al. (1999) listed five possible strategies for naturalizing classical phenomenology, including: (1) naturalization as reductionism, namely the idea that phenomena could be reduced to logical construction of something non-phenomenal in a manner paralleling to a reversed version of Carnap (2003)’s *Aufbau* project; (2) naturalization as an “as if” strategy, namely the Dennett (2003)-inspired idea that the ontological commitments of

mental phenomena should be made only on the level of predicting the behaviors of other agents, on Dennett's own term, on the level of "heterophenomenology"; (3) naturalization as mutual constraining, namely the idea that neurological findings and phenomenological findings should be supporting each other in this way or another; (4) naturalization as some form of the enlargement of the concept of nature, namely the idea that mental phenomena could be naturalized by being included into a broader scope of "nature"; and (5) functionalist naturalization, namely the idea that phenomenological descriptions are just a "user interface design" expecting to be substantially supported by a "backstage operation," which in turn should be realized via a computational system as what classical functionalists would like to provide.

Among all of the aforementioned strategies, (5) looks most interesting for advocates for the proposed W/C alliance, since it would bring about most fruitful theoretical outcomes to cognitive science on the one hand and be maximally loyal to the literal meaning of Wittgenstein's texts on the other. The reason for making such an assertion is as below.

Although computational functionalism has long been criticized since new models of cognition, like the model of embodiment or enactivism, were developed to compete with it, it is undeniable that it is still the most accepted working principle for cognitive science. Hence, if the desired W/C alliance could be built in terms of the Wittgenstein/computational functionalism alliance, this will convincingly show how solid the alliance itself could be. By contrast, the idea of "naturalization as reductionism," as Roy et al. (1999) themselves admit, is hard to execute no matter whose philosophy is to be naturalized. Insofar as the idea of viewing naturalization as an "as if" strategy is concerned, though it is quite fitting the communicative nature of the Wittgensteinian notion of "language games," it can hardly bring about fruitful instructions on how cognitive science should be done except for providing some form of hindsight. The similar criticism could be applied to the idea of enlarging the nature to accommodate phenomena, since for most cognitive scientists, this seems to be nothing but a terminological trick. Surely the idea of viewing naturalization as mutual constraining between the phenomenological findings and neurological findings is very relevant to the basic concerns of cognitive science, but unfortunately, there is no evidence to show that Wittgenstein himself is interested in neurological underpinnings of linguistic behaviors. More unfortunately, since the traditional neurological approach to cognition mostly focuses on individual subjects, there should be a gap between such an approach and the multiple-agent-involving nature of Wittgensteinian language games. Surely such a gap could be minimized by inventing new experimental instruments to probe into the neural activities of multiple agents, but such an approach would be made less promising when a huge number of language game participants are being considered. By contrast, a relatively high-level functionalist description of how language is used by agents could be more fitting Wittgenstein's own avoidance of resorting to detailed neurological stories on the one hand and easier to be connected to scientific reality via, say, cognitive modeling on the other. Therefore, to appeal to some form of computational functionalism looks like the only game in the town for building the W/C alliance.

However, the consensus among Wittgenstein scholars appears to be that Wittgenstein could only be depicted as an opponent of cognitive science in the general sense and computational functionalism in particular. Now I will immediately show what arguments are supportive of this opinion and why they cannot work.

3 How to Remove the Obstacles for Building the Wittgenstein/Cognitive Science Alliance?

Insofar as I know, the relation between Wittgenstein and cognitive science has been highlighted mainly by a couple of Wittgenstein scholars like P. M. S. Hacker (1996), Stuart Shanker (1998), D. Proudfoot (1997), Roger Teichman (2001) and Robert Arrington (2001). D. Dreyfus (1979), who seems to be the only famous philosopher of cognitive science who treats Wittgenstein seriously, views early Wittgenstein as a spiritual godfather of symbolic AI, and accordingly, he views later Wittgenstein's criticism of his early thought as a departure from the philosophical tradition embedded in symbolic AI. However, Dreyfus' own idea about how cognitive science or AI should be done seems to be more influenced by Heidegger and Merleau-Ponty rather than later Wittgenstein. Even in Margaret Boden (2006)'s epic treatment of the history of cognitive science, Wittgenstein's thought is basically marginalized as part of the fading background of the prehistory of cognitive science. So it is fairly impartial to conclude that Wittgenstein is a "forgotten philosopher" by most, if not all, cognitive scientists.

Surely aforementioned Wittgenstein scholars should be more responsible for this situation than cognitive scientists, since it is their diagnoses that undermine any attempt to build the W/C alliance constructively. To be more specific, diagnoses of this type can be divided into three categories: (i) diagnoses linking certain doctrines of cognitive science, e.g., the Chomskyan notion of language with its Tractarian counterpart, and therefore distancing later Wittgenstein from cognitive science by resorting to how later Wittgenstein criticized his early philosophy (cf. Hacker 1996: 270–271); (ii) diagnoses exploiting Wittgenstein's negative responses to Alan Turing for showing how weak the W/C alliance could be (cf. Shanker 1998, especially chapter 1); and (iii) diagnoses exhibiting the similarity between cognitive scientists' image of mind as neurally instantiated computational mechanism and the Cartesian image of mind as some form of internal reality, and therefore applying Wittgenstein's criticism of the Cartesian doctrines to contemporary cognitive science (cf. Proudfoot 1997).

In this section, due to the following considerations, more attention will be paid to diagnoses of type (iii) rather than (i) and (ii): Firstly, diagnoses of type (i) would not pose a real threat to the desired W/C alliance, since the specific doctrines (which are targeted by the relevant diagnostician), e.g., the Chomskyan doctrine on generative grammars, are not exclusively accepted even in the arena of cognitive science. Hence, advocates of the W/C alliance can still safeguard it by replacing the targeted doctrine by another. Secondly, even though as diagnosticians of type (ii) depict, Wittgenstein did (especially in Wittgenstein 1974 III: § 87) criticize Turing's

mechanical notion of computation as a substitution of meaningful rules by meaningless sub-rules, his criticism is actually irrelevant to contemporary computational theory of mind. Rather, the metaphysical assumption of contemporary cognitive science includes something more than Turing's mechanical notion of computation. To be more specific, the core metaphysical idea for cognitive science is functionalism, according to which high-level meaningful rules are *irreducible* to meaningless sub-rules, although they are supervenient on the latter. And due to the functionalists' tolerance of high-level features of cognition, their doctrine is sometimes even viewed as a disguised form of dualism, and in this sense, the substantial supplement of functionalism to Turing's original idea dramatically covers the reductionist flavor of the latter. And also in this sense, diagnosticians of type (ii) cannot undermine the W/C alliance if they fail to construct an argument for showing the incompatibility between Wittgenstein and functionalism per se, rather than Turing's notion of computation. In contrast, diagnoses of type (iii) deserve a more serious treatment than (i) and (ii) since they are aimed at the relation between Wittgenstein and a *general* feature of classical cognitive science rather than a certain doctrine of it, and hence, if the verdicts of type (iii) are really sound, the W/C alliance would really be in deep trouble.

Now it is the time for showing even diagnosticians of type (iii) are on the wrong track. But before that, it is necessary to uncover the inferential structure of such diagnoses by taking Proudfoot (1997) as a sample case. His argument for undermining the W/C alliance could be reconstructed as below:

- (1) Wittgenstein denies that cognition is some inner process as Cartesians or Lockeans depict (textual evidence provided in Table 1).
- (2) The contemporary cognitive scientists' image of mind as neurally instantiated computational mechanism is similar to the Cartesian image of mind as some form of internal reality, since both of them intend to view cognition as a disembodied internal reality.
- (3) So one may plausibly expect that Wittgenstein would criticize cognitive scientists' image of mind as well.
- (4) Therefore, there cannot be any solid W/C alliance.

But I don't think this is a sound argument. Since textual evidence for supporting premise (1) looks so abundant, I would not doubt (1). The more controversial premise is (2), wherein Proudfoot seems to have misunderstood what machine functionalists actually mean by "internal" or "inner process." The following citation from Proudfoot (1997) would suffice for showing how misleading his notion of machine functionalism is:

Indeed, proponents of "Strong AI" have often envisaged the construction of a *res cogitans* even narrower in conception than Descartes': Descartes' claim is *in thought* thinking is separable from bodily activity—walking, handling or looking at things, and so on—whereas AI researchers have often claimed to be able to build computer systems in which thinking *actually* is separated from any bodily activity (Proudfoot 1997: 190).

Table 1 Paraphrased citations from Wittgenstein's work for denying the thesis that thinking is an process

No.	Typical claims by Wittgenstein (rephrased from his original expressions)	Source
C1	To explain intentionality by asserting to inner processes would assume a further inner representation to fix such explanations, and hence drive one into regress	Wittgenstein (2009: §198)
C2	There is no way to view thinking as a process completely enclosed in the head	Wittgenstein (1981: §606)
C3	Whether someone understands something is determined by what she would do under certain requests	Wittgenstein (1988: §209)
C4	The idea that expressions like "thinking" do denote to same internal processes is very misleading	Wittgenstein (1965 :7)

Proudfoot is simply attacking a straw man here, since no serious AI researchers would claim that it is possible to build computer systems in which thinking is actually separated from any bodily activity. The reason is so straightforward: For AI researchers, the term "body" simply means the hardware for running the program, and no program could be executed without resorting to hardware. In addition, for AI researchers, how to delimit the boundary between the hardware responsible for "pure thinking" and "bodily movements" is a quite trivial issue, since one may quickly redefine such a boundary by connecting the former to varieties of artificial sensors or navigation system, and so on. As a matter of fact, Proudfoot's criticism of the mainstream AI researchers' methodology could be made to be more plausible if his emphasis were on their underestimation of the significance of designing artificial sensors or navigation systems, and it was criticisms like this kind that led Rodney Brooks (1990, 1999) to claim that the development of robotics is a more promising project than "pure AI." However, even if Proudfoot were to criticize traditional AI in this way, this criticism would still be irrelevant to his assertion that traditional AI systems are disembodied, since to be deprived of peripheral sensors is not tantamount to being disembodied, which actually means to be deprived of any spatial features.

Now it is the time for scrutinizing the meanings of the "internal–external" dichotomy, by virtue of which Proudfoot (1997)'s misleading description of the W/C relation is formed. A sad fact about contemporary analytic philosophy is that this dichotomy is often used in different philosophical contexts in different ways; hence, it looks necessary to clarify the meanings of this dichotomy before we provide a competing description of the W/C relation. To be more specific, there are actually four "internal–external" dichotomies, which are sometimes unfortunately confused with each other:

- (1) The ontological version of the "internal–external" dichotomy, according to which to be internal just means to be accessible to the introspection of the cognizer on the one hand and to be deprived of any spatial properties on the

- other. Otherwise, it is external. Cartesian dualists should have endorsed the “internal–external” dichotomy of this type.
- (2) The epistemological version of the “internal–external” dichotomy, according to which to be internal just only means to be accessible to the introspection of the cognizer in the process of justifying a belief. Otherwise, it is external. Epistemological internalists should have endorsed the “internal–external” dichotomy of this type.
 - (3) The semantic version of the “internal–external” dichotomy, according to which meanings’ being internal is tantamount to their being located in the head. Otherwise, it is external. The target theory of Putnam (1975)’s criticism should have endorsed the “internal–external” dichotomy of this type.
 - (4) The cognitive version of the “internal–external” dichotomy, according to which a cognitive process is internal iff it is processed within the head. The opponent theory of Clark and Chalmers (1998)’s “Extended Mind Thesis” should have endorsed the “internalism” of this type.

According to the foregoing analysis, whether Wittgenstein’s criticism of Cartesian internalism could be applied to cognitive science or AI patently depends on what kind of “internalism” mainstream cognitive scientists or AI researchers have endorsed. Surely no AI researchers would endorse (i), which is straightforwardly conflicting with physicalism, namely the metaphysical assumption of nearly all AI researchers. However, whether AI researchers would endorse (ii) is a more complicated issue. It is undeniable that the traditional symbolic approach to AI looks very similar to epistemological internalism in the sense that the process of achieving knowledge for the system is basically a process exclusively resorting to what has been stored in the system’s own memory base. But for a connectionist system (whose working principles are summarized in J. Garson 2016), some form of externalism in the sense of (ii) should be assumed, especially when a “training set” is fed into the system as what the so-called supervised algorithm of the whole network requires. The reason for saying this is as follows: Since (a) the “training set” fed into a connectionist system is consisting of nothing but many exemplary inputs and their desired outputs for a given task and (b) the issues related to which inputs are “exemplary” or which outputs are “desired” are patently hinging on some factors *external* to the system, it is easy to conclude that no connectionist system can do its “reasoning” job without resorting to some *normative* factors *external* to the system. (Surely in the connectionist context, the meaning of the term “reasoning” has been extended to the sub-symbolic level.) As to the internal/external dichotomy of type (iii) and its relation to cognitive science or AI, the situation is also a bit complicated. I believe that most classical cognitive scientists or AI researchers taking the symbolic approach would take the externalist route as granted, since the meanings of, say, the stored axiomatic propositions in a computational system are patently derived from the human programmers who have encoded them, and these programmers are also patently external to the designed system. But how meanings are represented is a quite different issue in the connectionist case, in which the meaning of a term can be transferred into its relative position *within* the so-called activation space or vector space as the result of

the neural learning. Since the “activation space” itself is an *internal* representation, the fixation of a meaning in the connectionist way appears to have assumed internalism of type (iii). As to what most cognitive scientists and AI researchers would say about the internalism–externalism debate of type (iv), I suppose that none of them would like to take the internalist route of type (iv) in a very strict sense, since it looks too easy to redefine the internal/external boundary by spatially rearranging the information-processing channels of the whole cognitive system. Surely the boundary between the internalism/externalism of type (iv) could be made more salient by redefining “internalism” as the thesis that “*most essential* cognitive processes are located in the central part of the cognitive system” and by redefining “externalism” as the denial of the foregoing thesis. But even when such redefinitions are provided, it would be unclear why all cognitive scientists or AI researchers should endorse “internalism” in this new sense, since it is still too easy to imagine a cognitive system works in an “externalist” way, as Brooks (1990, 1999) has suggested. Hence, the upshot of the observation of this paragraph is: Whether cognitivist scientists and AI researchers are internalists heavily depends on what type of “internalism” is being discussed: If it is metaphysical internalists, they are not; if it is epistemological/semantic/cognitive internalism, they may be but may be not. Therefore, there is simply no way to make an oversimplified conclusion that *all* cognitive science or AI parallels the Cartesian notion of mind just because both of them have endorsed “internalism.”

But what about Wittgenstein’s position concerning the four types of the internalism–externalism debate?

Patently Wittgenstein will deny metaphysical internalism, not because he is a self-identified physicalist, but because he is not willing to make ontological commitments of any entities which are not presentable in public space (for instance, his notion of “pain” is ontologically tangled with the features of the actual life in which the pain occurs. Cf. Wittgenstein 1981: §533). Or more precisely, for him, the boundary between the internal side of, say, pain and the external side of it is so fuzzy that any attempt to delimit it would like to distort the very way in which the world is experienced by us in language games. I also believe that Wittgenstein would reject epistemological internalism as well, since according to him, the normative source of the language games is obviously not something exclusively accessible to the introspection of a game participator in solitude. Even when such a participator retrieves a stored belief about the authoritative source of a piece of information, the mental retrieval of this sort, according to Wittgenstein, cannot be ultimately warranted without being compared with the real entities outside the agent’s memory (cf. Wittgenstein 2009: §265). In a similar sense, Wittgenstein should deny semantic internalism as well, according to which meaning is located in individuals’ minds rather in a spatially extended form of social network. As to the internalism–externalism debate of type (iv), I concede that I am not sure about which side Wittgenstein would take. But I am more inclined to believe that he would take the externalist route, since he himself has provided so many inspiring thought experiments for radically rearranging the triadic relationship among the central part of cognitive system, the body and the environment, and cognitive externalism is very likely to be induced by these thought experiments (Cf.

Wittgenstein 1980: §72, where Wittgenstein imagines the visual experience of an eyeball detached from the rest of human body, and Wittgenstein 2009: §253, where the Siamese twins is taken as an exemplary case for showing how mental activities like pain can occur in “another” body).

Hence, briefly speaking, Wittgenstein is an externalist in the ontological/epistemological/semantical sense, and very probably he would also take the cognitive version of externalism. Meanwhile, nearly all cognitive scientists and AI researchers have endorsed ontological externalism just as Wittgenstein does, and at least a part of cognitive scientists and AI researchers would like to endorse epistemological/semantic/cognitive version of externalism just as Wittgenstein does. Therefore, a W/C union is very probably to build if a Wittgensteinian could patiently search for those more Wittgenstein-oriented schools in the thought market of cognitive science and AI rather than thoughtlessly take the whole market as a unified and hostile school.

However, there may be two questions left for building such a union. First, there is still no evidence to prove that Wittgenstein is a full-fledged physicalist, whereas physicalism is the basic ontological assumption of cognitive science. Second, as I have mentioned previously, Wittgenstein’s attitude toward computation in the form of Turing-machine table looks negative, whereas such notion of computation looks so central to the contemporary cognitive science. What I can briefly say about the two questions includes four points: (1) A 100% fitting between Wittgenstein and cognitive science is not only impossible but also undesirable, since the main purpose of building the W/C alliance is just to make Wittgenstein’s legacy as a source for inspiring brainstorming ideas for cognitive science, rather than making it as a new bible that cognitive scientists should follow in a word-by-word manner; (2) to build the W/C alliance does not require Wittgensteinians’ self-identification as physicalists, but only their commitment to a weaker thesis that basic tenets of later Wittgenstein are not conflicting with all versions of physicalism; (3) if one takes the commitment to the “supervenience thesis” as the minimal requirement for any physicalist (for the sake of simplicity, by “supervenience” I basically mean “weak supervenience” defined by Kim 1993 in this paper), then she will immediately take non-reductive physicalism, which definitely satisfies this requirement, as a serious option, and consequently, non-reductive physicalists’ emphasis on the non-reducibility of high-level features of cognition will be perfectly paralleling to Wittgenstein’s emphasis on the non-reducibility of high-level features of language games; and (4) if Wittgenstein’s hostility to the notion of computation in the form of Turing-machine table could be carefully reinterpreted as his rejection of the idea for *reducing* linguistic phenomena to Turing-machine calculation rather than the idea that linguistic phenomena merely *supervene on* Turing-machine calculation, then there will be no reason to make Wittgenstein a foe of cognitive science or AI.

Now I think I have removed those most troublesome obstacles for building the W/C union. In the following section, I will sketch the outline for building the desired union with the aid of some basic ideas borrowed from cognitive linguistics.

4 Cognitive Schemas and Wittgenstein's Phenomenological Notion of "Grammar"

Why do we need to correlate Wittgenstein with cognitive linguistics if the W/C union is the pursuit of our research? And why is cognitive linguistics so special?

The first reason is definitely that both sides share some core tenets. As Geeraerts and Cuyckens (2007) depict, although cognitive linguistics has not been built into a single uniform theory yet, the following features still suffice for distinguishing a cognitive linguist from other linguists who reject cognitive linguistics: (1) the acceptance of the primacy of semantic issue in linguistic analysis (contra Chomskyans' emphasis on the primacy of syntactical analysis and the purported "autonomy" of grammars); (2) the acceptance of the encyclopedic nature of linguistic meanings (contra any attempt to encode meanings in a structural manner which is remote from how meanings are chaotically presenting themselves in the social reality); and (3) the acceptance of the perspectival nature of linguistic meanings (contra the idea that linguistic meanings encode the information about the world in an objective way). Surely a Wittgensteinian would like to accept all of these points made by cognitive linguists. Firstly, later Wittgenstein's notion of "grammar" is nothing but the usages of words rather than some content-free structural description of language of Chomskyan type. Or in another around, one may even say that Wittgenstein's notion of "grammar" is word-centered, in the sense that the typical "grammatical" question raised by Wittgenstein is often in the form of "What is the typical use of word X in context Y?", the formulation of which is apparently X-centered. Hence, he would very likely side with cognitive linguists by claiming that grammars are basically derived from the semantic configurations of meanings of lexicons. Secondly, also paralleled to what cognitive linguists say in point (2), Wittgenstein would admit that the categorization of a concepts is *not* guided by a set of sufficient and necessary conditions for determining, say, how a concept is subject to another, but in a way that could be metaphorically presented as how newly built houses and streets gradually change the original outlook of the downtown (Wittgenstein 2009: §18). Thirdly, although the term "the perspectival nature of linguistic meanings" is not used by Wittgenstein himself, his hostility to Augustine's notion of what language is, namely the view that meanings derive from the object that words stand for, is definitely aiming the same target that cognitive linguists are aiming at.

The second reason for taking cognitive linguistics as an essential integral for the W/C union is that some core tenets shared by both Wittgenstein and cognitive linguistics have already been converted into working theories in cognitive science. The most exemplary case here is Wittgenstein's notion of "family resemblance," according to which members of a category may be related to each other without exclusively sharing a common property. Although the original expression of this idea looks too metaphorical and hence a bit coarse-grained, a more science-oriented Wittgensteinian definitely can use some relevant scientific resources to make it a bit more fine-grained. For instance, Lotfi Zadeh's fuzzy logic (1965, 1972) is a tool that offers the first technical approximation to the very idea of "family resemblance" in

the sense that it can characterize the fuzziness of boundaries between “insiders” of a category and its “outsiders.” But if we want to more carefully scrutinize the internal structure of a category, we may go one step further by resorting to some other “fuzzy” concepts like “centrality” and “gradience,” the characterization of which is provided by Brand Berlin and Paul Kay (1999)’s study of color concepts. That led to Eleanor Rosch (1973)’s well-known theory of “prototype,” according to which whether a linguistic entity can be viewed as a member of a given category depends on how “near” that entity is to the “prototype” of that category. The computational model for measuring such conceptual distances is provided by Smith and Medin (1999) under the title of “Exemplar Theory,” which is derived from prototype theory. Although I do not want to claim that “prototype theory” is the most faithful technical realization of Wittgenstein’s original idea, Rosch and her followers’ work is definitely in the process of realizing it. In addition, since prototype theory now is a received part of both mainstream cognitive linguistics and cognitive science, it is easy to conclude that at least some ideas taken from Wittgenstein’s legacy can be rendered more assessable to the theoretical construction of cognitive linguistics. Hence, one may naturally expect that more of Wittgenstein’s ideas can be treated in the similar way.

The third reason for bringing Wittgenstein and cognitive linguistics together is based on the following observation: Wittgenstein’s notion of “grammar” can be made more understandable in terms of “cognitive schemas,” which is the core concept of cognitive linguistics. As I have mentioned, in *The Big Typescript*, Wittgenstein identifies the goal of constructing a “phenomenological language” with “the investigation of the use of the rules of our language” or “grammar” in short (Wittgenstein 2005: 320). In another paragraph, he even asserts that the identity of a word is determined by the identity of the grammar/usage of this word and hence highlights the grammatical level of a word as something more essential than its physical outlook (ibid.: 206). Hence, we may identify four features possessed by the Wittgensteinian notion of “grammar”: (1) It has a salient phenomenological aspect, namely the usages of words are presented to us as some primarily given “phenomena.” (2) It is word-centered rather than something whose propositional representations are taken as primary. (3) It is so diversified that its diversity normally overrides the relative uniformity of the physical outlooks of corresponding words. (4) It sets some implicit parameters for forming a background knowledge in which certain forms of combinations of words are allowed to appear. I believe that all of the features listed here can be caught by cognitive linguists via the notion of “cognitive schemas.”

According to Ronald Langacker (2008), in the vocabulary of cognitive linguistics, the notion of “cognitive schema” is closely related to that of “rule,” which actually means the characterization of certain linguistic patterns (which are closely related to the semantic contents of certain lexicons rather than content-free). Accordingly, schemas “are abstract templates obtained by reinforcing the commonality inherent in a set of instances” (ibid.: 23). It should be emphasized that the term “abstract templates” used here is not correlated with the Platonic notion of “abstracta” in any sense. Rather, it only means the subordination of a concept to a set of larger “domains” or “domain matrix,” e.g., the subordination of

the concept *SOPHOMORE* to a set of domains like *TWO*, *PERSON*, *KNOW* and *YEAR* (ibid.: 46). There is no conceptual realism-friendly way to characterize these subordination relations due to the following considerations: Firstly, both the subordination relations and the scope of domains merely stand in a degree manner (rather than categorically), and these relations may change when new data about how to use corresponding words are acquired; secondly, as to which subordination relations, among many other competing ones, is highlighted in a certain context, it is an issue mostly dependent on which perspective the agent is taking (rather than in a perspective-free manner). Thus, it looks that a schematicity-based narrative of grammars/concepts can naturally obtain the following features that a Wittgensteinian conception of grammars/concepts should have: The former is as word-/term-/concept-centered (rather than content-free) as the latter (since schematicity is always related to a given concept); the former is as diversified as the latter (since the obtaining of, say, phonological features of a concept does not always suffice for obtaining its essential schematicity); the former is as relevant to the notion of “background knowledge” as the latter is (since the notion of “domain matrix” itself is another way for addressing “background knowledge”); and finally, the former is as phenomenology-oriented as the latter (since the characterization of schematicity is basically dependent on how the conceptualization of a certain linguistic unit is intuitively construed by the cognitive linguists). Therefore, we may conclude that cognitive linguistics makes a perfect bridge for bringing the Wittgensteinian point of view of what language is into a wider picture of “naturalized Wittgenstein” or “naturalized phenomenology” as so on.

However, there are still two questions left concerning the triadic relationship among Wittgenstein, cognitive linguistics and cognitive science. The first one is: What cognitive linguists would say about the “internalism–externalism” debate mentioned in the last section? Is there any guarantee for ensuring that cognitive linguists’ point of view of this debate precisely coincides with Wittgenstein’s? Secondly, is cognitive linguistics itself naturalized in a full-fledged manner?

As to the first question, although I cannot promise that cognitive linguists’ attitude toward the “internalism–externalism” debate is precisely coinciding with Wittgenstein’s, it can be asserted that the former is not so distant from the latter. Since semantics is the primary concern of cognitive linguistics, so for cognitive linguists, the “internalism–externalism” debate should be addressed from the semantic perspective, and the term “internalism” should be accordingly interpreted as the idea that meanings are localized in individuals’ heads, whereas the term “externalism” should be accordingly construed as the idea that meanings are distributed over a speech community. Representative cognitive linguists like Langacker believe that neither the extreme version of “localism” nor that of “distributism” is tenable, and his positive position is bit more near to the pole of “localism” in the sense that “an expression’s meaning resides in the conceptualizing activity of individual speakers” (ibid.: 29). This claim may make cognitive linguists’ position a bit more distant from Wittgenstein’s own, which may be supposed to be more near to the pole of “distributism.” But such a minor divergence from Wittgenstein is still understandable since cognitive linguistics is an empirical science, whose research projects should be executed in some feasible way, and no

executable project in cognitive modeling could take “distributism” as a starter. Moreover, Langacker’s localism-oriented approach is also required by the perspective-centered feature of schematicity, since sensitivity to certain perspectives is always sensitivity to certain perspectives *of individual speakers*. In addition, the Wittgenstein-favored communicative aspect of language would not be sacrificed by taking this localism-oriented approach, since this approach still permits potential meanings of some expressions to exist in a relatively objective manner even though these meanings are beyond the perspective that the agents in question is actually taking. Hence, when the semantic problem is being addressed, it can be generally concluded that both Wittgenstein and cognitive linguists intend to achieve a compromise between extreme internalism and extreme externalism, although the trade-offs that they actually offer may differ from each other in a more trivial sense.

As to the second question, it is undeniable that a full-fledged naturalization of cognitive linguistics is still in its process, if the expression “the naturalization of X” means to covert a high-level story of X either into a neurological story or a Turing-machine-executable story. But even this is true, it would be too arbitrary to verdict that the assimilation of cognitive linguistic into cognitive science would be conceptually impossible. There are two reasons for saying this:

Firstly, it is undeniable that the phenomenological methodology employed by most cognitive linguists is recognized as the most salient obstacle for naturalizing cognitive linguistics, which is supposed to have to take the third-personal or purely objective perspective of cognition. However, in my view, if this obstacle could not be removed simply due to conceptual reasons of this type (rather than some technical problems), then the naturalization of phenomenology would seem to be even more troublesome. But many would like to admit that naturalized phenomenology is a promising project; hence, if they are right, there will be no plausible reason to claim that naturalized cognitive linguistics, among the features of which the employment of the phenomenological methodology is just one among others, is not at least equally promising.

Secondly, if we confine the scope of “cognitive science” merely to the inquiries into the computational models of cognition, then the difficulties of converting any cognitive schema into a corresponding algorithmic description would appear to be a problem related to the conflicts between cognitive linguists’ notion of “concepts” and its counterpart in the mind of traditional representational theorists of cognition. Or to be more specific, for representational theorists of mind, concepts are discrete representational units waiting to be combined by domain-free syntactical entities to form more complexed propositional representations, whereas for cognitive linguists, concepts should be imagistic in the sense that a central component of a concept is the visual/auditory images of it. In addition, since propositional complexes are supposed to be derived from the entrenched schemas of these imagistic pictures rather than an external combination of “form” and “content,” for cognitive linguists, neither the introduction of discrete representations nor that of domain-free syntax looks necessary. However, the sad story about mainstream computational approaches to cognition is that most researchers simply do not know how to achieve cognitive linguists’ theoretical goal, namely to handle concepts in a pre-propositional manner on the one hand and to keep it still minimally concept-like on the

other. For instance, the symbolic approach definitely can treat concepts as high-level linguistic entities, but it handles them precisely on the propositional level; by contrast, the connectionist approach can treat the concepts on the pre-propositionally level, but the cost is that any concept-like entities should be sacrificed when a connectionist network is initially built. Nonetheless, what cognitive linguists want is to have their cake and eat it too.

But even though difficulties of this type cannot be ignored, they are actually not conceptual problems and could be overcome by new approaches in the near future. Optimism of this kind is supported by the following considerations: Firstly, if we introduce non-reductive physicalism as the basic metaphysical framework for treating the relationship between schema-based description of cognition and its algorithmic depiction, then there is no reason to believe that the former's being supervenient on the latter implies that the former is reducible to the latter as well. Hence, it is possible to make cognitive schemas computable if a right algorithm is selected. Secondly, at least some solid part of cognitive linguistics, e.g., the prototype theory, has its computational model; hence, it may be expected that more and more parts of cognitive linguistics can be treated in a similar manner. Thirdly, in AI we actually already have a general-purpose reasoning system the spirit of which is very similar to that of cognitive linguistics, that is, Non-Axiomatic Reasoning System (or "NARS" for short. Cf. Pei Wang 1994, 2001, 2004, 2006, 2009a, 2009b, 2013), whose logic is term based rather than propositional in the Fregean sense. Moreover, similar to cognitive linguistics, NARS also takes categorization or subordination as the most fundamental linguistic phenomena, from which more complicated syntactic features are derived. Moreover, NARS has a broad vision of what a "concept" is by allowing imagistic entities to be taken as "concepts," in case these entities are within certain relations of categorization or subordination that can be characterized by NARS. Although I do not mean that NARS provides the most suitable platform for linking cognitive linguistics to the reality of computer science, it looks like the most suitable one insofar as I do know currently. However, the technical details for building such a link should be left to another paper due to the limitation of space here.

5 Conclusion

The upshot of the foregoing analysis is that the philosophical ideas of later Wittgenstein can be realized via the theoretical construction of cognitive linguistics, whereas the results of cognitive linguistics can be assimilated into cognitive science with the aid of some suitable formal tools, say, NARS. In addition to this point, two philosophical morals for naturalizing Wittgenstein should be highlighted as well: First, the non-reductive physicalism is a necessary metaphysical framework to achieve a compromise between Wittgenstein's phenomenology-oriented approach to language and the physical realization of high-level features of cognition; second, the ambiguity of the "internal-external" dichotomy should be clarified in advance for showing that Wittgenstein's criticism of the Cartesian notion of thinking as an "internal process" is not supposed to be conflicting with a wider naturalistic picture of cognition.

References

- Arrington, Robert. 2001. Thought and its expressions. In *Wittgenstein and contemporary philosophy of mind*, ed. Severin Schroeder, 129–149. Palgrave: Houndmills.
- Austin, J. 1957. A plea for excuses. In *Proceedings of the aristotelian society*, LVII, pp. 1–30.
- Berlin, B., and Kay, P. 1999. *Basic color terms: Their universality and evolution*, revised version, The Center for the Study of Language and Information Publications.
- Boden, Margaret. 2006. *Mind as machine, a history of cognitive science (two volumes)*. Oxford: Clarendon Press.
- Brooks, R. 1990. Elephants don't play chess. *Robotics and Autonomous Systems*. 6: 139–159.
- Brooks, R. 1999. *Cambrian intelligence: The early history of the new AI*. Cambridge: MIT Press.
- Carnap, Rudolf. 2003. *The logical structure of the world and pseudoproblems in philosophy* (trans: by George, Rolf. A.). Open Court Publications.
- Clark, A., and D. Chalmers. 1998. The extended mind. *Analysis* 58 (1): 7–19.
- Dennett, D. 2003. Who's on first? Heterophenomenology explained. *Journal of Consciousness Studies, Special Issue: Trusting the Subject? (Part 1)* 10 (9–10): 19–30.
- Dreyfus, H. 1979. *What computers can't do: The limits of artificial intelligence*. Cambridge, MA: The MIT Press.
- Garson, J. 2016. Connectionism. In *The stanford encyclopedia of philosophy*, ed. Edward N. Zalta, (Winter 2016 Edition). <https://plato.stanford.edu/archives/win2016/entries/connectionism/>.
- Geeraerts, D., and H. Cuyckens. 2007. Introducing cognitive linguistics. In *The Oxford handbook of cognitive linguistics*, ed. D. Geeraerts, and H. Cuyckens, 3–24. Oxford: Oxford University Press.
- Hacker, P.M.S. 1996. *Wittgenstein's place in the twentieth-century analytical philosophy*. Oxford: Blackwell Publishers Inc.
- Hintikka, M., and J. Hintikka. 1986. *Investigating Wittgenstein*. Oxford: Blackwell Publishing Ltd.
- Husserl, E. 2001. *Logical investigations* (Vol. 1, 2), (trans: Findlay, J. N.) New York: Routledge.
- Kim, Jaegwon. 1993. 'Strong' and 'global' supervenience revisited. In *Supervenience and mind: Selected philosophical essays*, ed. Jaegwon Kim. Cambridge: Cambridge University Press.
- Kuroda, W. 1979. Phenomenology and grammar: A consideration of the relation between Husserl's logical investigations and Wittgenstein's later philosophy. In *Japanese phenomenology: Phenomenology as the transcultural philosophical approach*, ed. Yoshihiro Nitta, and Hirotaaka Tatematsu, 89–108. Dordrecht: D. Reidel Publishing Company.
- Langacker, R. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Park, B.-C. 2012. *The Phenomenological Aspects of Wittgenstein's Philosophy*. Springer.
- Proudfoot, D. 1997. On Wittgenstein on cognitive science. *Philosophy* 72 (280): 189–217.
- Putnam, H. 1975. The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science* 7: 131–193.
- Rosch, E. 1973. Natural categories. *Cognitive Psychology* 4 (3): 328–350.
- Roy, Jean-Michel, Jean Petitot, Bernard Pachoud, and Francesco Varela. 1999. Beyond the gap: An introduction to naturalizing phenomenology. In *naturalizing phenomenology*, ed. Jean Petitot, et al., 1–82. Stanford: Stanford University Press.
- Shanker, Stuart. 1998. *Wittgenstein's remarks on the foundations of AI*. Abingdon, Oxan: Routledge.
- Smith, E., and D. Medin. 1999. The exemplar view. In *Concepts: Core readings*, ed. Eric Margolis, and Stephen Laurence, 207–222. Cambridge, MA: The MIT Press.
- Spiegelberg, H. 1981. *The context of the phenomenological movement*. The Hague: Martinus Nijhoff Publishers.
- Teichman, Roger. 2001. The functionalists' inner state. In *Wittgenstein and contemporary philosophy of mind*, ed. Severin Schroeder, 24–35. Palgrave: Houndmills.
- Varela, F. 1996. Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies* 3: 330–350.
- Wang, Pei. 1994. From inheritance relation to non-axiomatic logic. *International Journal of Approximate Reasoning* 11: 281–319.
- Wang, Pei. 2001. Confidence as higher-order uncertainty. In *The second international symposium on imprecise probabilities and their applications*, ed. G. De Cooman, 352–361. New York: Ithaca.
- Wang, Pei. 2004. The limitation of bayesianism. *Artificial Intelligence* 158: 97–106.
- Wang, Pei. 2006. *Rigid flexibility: The logic of intelligence*. Dordrecht: Springer.
- Wang, Pei. 2009a. Formalization of evidence: A comparative study. *Journal of Artificial General Intelligence* 1: 25–53.

- Wang, Pei. 2009b. Analogy in a general-purpose reasoning system. *Cognitive Systems Research* 10: 286–296.
- Wang, Pei. 2013. *Non-axiomatic logic: A model of intelligent reasoning*. Singapore: World Scientific Publishing Co.
- Waismann, F. 1979. *Wittgenstein and Vienna circle*, ed. Brian McGuiness (trans: Joachim Schulte and Brian McGuiness).
- Wittgenstein, Ludwig. 1929. Some remarks on logical form. *Proceedings of the Aristotelian Society* 9: 162–171.
- Wittgenstein, Ludwig. 1965. *The blue and brown books*. New Yorker: Harper Torchbooks.
- Wittgenstein, Ludwig. 1974. *Remarks on the foundations of mathematics*, Revised Edition, ed. G.H. von Wright, R. Rhees and G.E.M. Anscombe (trans: Anscombe, G. E.). Oxford: Basil Blackwell.
- Wittgenstein, Ludwig. 1979. *Remarks on colour*, ed. G. E. M. Anscombe, (tans: McAlister, Linda L. and Schattle, Margarete). Oxford: Wiley-Blackwell.
- Wittgenstein, Ludwig. 1980. *Philosophical remarks*, ed. Rush Rhees and Raymond Hargreaves, (trans: Maximilian A. E. Aue). Chicago: University of Chicago Press.
- Wittgenstein, Ludwig. 1981. *Zettel*, second edition, ed. G. E. Anscombe and G. H. von Wright, translated by G. E. Anscombe, Oxford: Basil Blackwell.
- Wittgenstein, Ludwig. 1988. *Remarks on the Philosophy of Psychology*, Volume 2, edited by G. H. von Wright and Heikki Nyman, translated by C. J. Luckhardt, Chicago: University of Chicago Press.
- Wittgenstein, Ludwig. 2005. *The Big Typescript: TS 213*, German-English scholar's edition, edited and translated by C. Grant Luckhardt and Maximilian A. E. Aue, Malden, MA: Blackwell Publishing.
- Wittgenstein, Ludwig. 2009. *Philosophical Investigations*, 4th edition, trans. by G. E. M. Anscombe, P. M. S. Hacker and Joachim Schulte, Oxford: Wiley-Blackwell.
- Zadeh, L. 1965. Fuzzy sets. *Information and Control* 1965 (8): 338–353.
- Zadeh, L. 1972. A fuzzy-set-theoretical interpretation of linguistic hedges. *Journal of Cybernetics* 2: 4–34.

Yingjin Xu, professor of the school of philosophy of Fudan Universty, specializing in philosophy of cognitive science and Wittgenstein.