DHARMAKĪRTI, DAVIDSON, AND KNOWING REALITY

LAJOS L. BRONS

ABSTRACT: If we distinguish phenomenal effects from their noumenal causes, the former being our conceptual(ized) experiences, the latter their grounds or causes in reality ‘as it is’ independent of our experience, then two contradictory positions with regards to the relationship between these two can be distinguished: either phenomena are identical with their noumenal causes, or they are not. Davidson is among the most influential modern defenders of the former position, metaphysical non-dualism. Dharmakīrti’s strict distinction between ultimate and conventional reality, on the other hand, may be one of the most rigorously elaborated theories of the opposite position, metaphysical dualism. Despite this fundamental difference, their theories about the connection between phenomena and their noumenal causes are surprisingly similar in important respects. Both Dharmakīrti in his theory of ‘apoha’ and Davidson in his theory of ‘triangulation’ argued that the content of words or concepts depends on a process involving at least two communicating beings and shared noumenal stimuli. The main point of divergence is the nature of classification, but ultimately Dharmakīrti’s and Davidson’s conclusions on the noumenal-phenomenal relationship turn out to be complementary more than contradictory, and an integrative reconstruction suggests a ‘middle path’ between dualism and non-dualism.

Keywords: Dharmakīrti, Donald Davidson, apoha, triangulation, reality, meta-ontology, subjectivity, metaphysical dualism

1. INTRODUCTION

What can we know about reality? If two communicating beings are both aware of a certain black queen chess piece as evidenced by their reference(s) to that black queen chess piece in their communication, does that imply that there ‘really’ is (something that is) that black queen chess piece? At least there must be something causing the shared awareness. If we call that ‘something’ a noumenal cause and the shared conceptualized awareness of the black chess queen a phenomenal appearance, then the question could be rephrased as being about the (possibility, necessity, and/or extent of) identity or non-identity between such phenomenal appearances and their

BRONS, LAJOS L.: Researcher, Institute of Humanities and Social Sciences, Nihon University / Center for Advanced Research on Logic and Sensibility, Keio University, Tokyo, Japan. Email: lajos@lc.dds.nl
noumenal causes, or between our phenomenal beliefs and noumenal reality. The extremes on the gamut of answers to this (kind of) question would be (a variety of) *metaphysical dualism* arguing for non-identity of causes and phenomena, and thus for the illusoriness of our phenomenal beliefs; and (a variety of) *non-dualism* identifying phenomena with their causes and rejecting (the possibility of) massive error of our beliefs.

The rejection of metaphysical dualism and the subjectivity it implies is a central theme in the philosophy of Donald Davidson, but his *anti-dualism* is rarely phrased in terms like those employed above. According to Davidson:

> What stands in the way of global skepticism of the senses is, in my view, the fact that we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief. And what we, as interpreters, must take them to be is what they in fact are. Communication begins where causes converge: your utterance means what mine does if belief in its truth is systematically caused by the same events and objects. (1983, 151)

On the other end of the gamut, Dharmakīrti and many other Indian (and Tibetan) Buddhist philosophers argued for the non-identity of conventional (phenomenal) reality (*saṃvṛtisat*) and ultimate (noumenal) reality (*paramārthasat*), and considered the former to be (in some sense) illusory.

Nevertheless, despite that seemingly fundamental difference, there are remarkable similarities between Dharmakīrti’s and Davidson’s ideas about the connection(s) between words and phenomena on the one hand, and the (noumenal) world on the other. Both defended a form of externalism about the content of words or concepts that is both social and physical. That is, the content of words or concepts is determined in a process involving at least two communicating beings (the social aspect), and a shared noumenal stimulus (the physical aspect). In Dharmakīrti’s thought, this idea is part of the theory of *apoha*; in Davidson’s later thought (especially that of the 1990s), it is called ‘triangulation’.

*Apoha* and triangulation are (a.o.) theories about what makes it possible that words can (come to) be used to refer to ‘things’. If those theories are valid, actually having concepts, actually communicating, implies that those necessary conditions (that ‘what makes it possible that ...’) must be actually satisfied as well. In Davidson’s words: “successful communication proves the existence of a shared, and largely true, view of the world” (1977, 201). Reading *apoha* and triangulation in this way, the focus shifts from the prerequisites of language and communication to *meta-ontology*, from phenomenal effects to their noumenal causes, from the observation of the existence of language and communication to Dharmakīrti’s dualism and Davidson’s non-dualism.

Metaphysical dualism, in its strongest form, denies any access to the noumenal world, and thus robs metaphysics of its subject matter – such dualism leads to a kind of ontological deflationism. A radical non-dualism, on the other hand, gives direct access to noumena through language, reducing metaphysics to little more than an
analysis of ordinary language and the findings of physics (e.g. Manley 2009). The debate between ontological deflationism and inflationary (mainstream analytical) metaphysics is the main battlefield of meta-ontology (sometimes called ‘meta-metaphysics’; see Chalmers et al. (2009) for a recent overview). If Dharmakīrti and Davidson have similar theories leading to apparently contradictory meta-ontological conclusions (dualism vs. non-dualism), then an analysis and comparison of those theories, and – if possible – an attempt to integrate them, may be a valuable contribution to this debate.

Not surprisingly, given the distance in time, space, and tradition between Dharmakīrti (7th century, India, Buddhist philosophy) and Davidson (20th century, USA, analytic philosophy), the similarities between their arguments are hidden behind vastly different terminologies, aims, and background assumptions. The combination of similar arguments and (seemingly) contradictory conclusions, however, justifies an attempt to decontextualize, translate, and integrate these theories for the twin purpose of (1) determining the essential difference between apoha and triangulation that causes these two seemingly similar theories about word-world connections to be associated with contradictory meta-ontological perspectives (dualist vs. non-dualist); and (2) analyzing and – if possible – transcending that difference in order to assess the meta-ontological implications of communication. The aim of this paper, therefore, is not just to compare Dharmakīrti and Davidson, but to bring them – or at least (an interpretation of) some of their ideas – together, to integrate and extend. In other words, it aims at ‘constructive engagement’ (Mou 2010) more than at the juxtaposition (or even opposition) inherent in ‘comparison’. And to some extent, given the vast differences in theoretical and traditional backgrounds, such integration or constructive engagement requires deviation from both – the ‘middle path’ between Dharmakīrti and Davidson leads to an answer that is neither dualist, nor non-dualist (and neither strictly Dharmakīrtian nor Davidsonian).

Whatever the intended conclusion of an argument on the identity or non-identity of phenomenal appearances and noumenal causes, these two ‘levels of reality’ must be distinguished (if even as a purely semantic distinction for the sake of rejection). Therefore, before turning to apoha and triangulation, section 2 explores these ‘levels’ – and what is in between – in Dharmakīrti’s and Davidson’s thought. Section 3 briefly introduces Dharmakīrti’s theory of apoha as concept formation and more extensively explains the notion(s) of ‘triangulation’ and the role(s) it play(s) in Davidson’s philosophy, and section 4 argues that apoha in the more restricted sense of ‘negative classification’ is the aforementioned ‘essential difference’ causing the different meta-ontological position, and outlines a variant of triangulation that incorporates such (apoha-ic) negative classification: an apoha-triangulation integration. The role of negative classification is further analyzed in section 5 in terms of the boundaries of types and tokens, which is illustrated by Dharmakīrtian and Davidsonian answers to the question about the black queen chess piece – “no” and “yes” respectively. Section 6 finally, argues that these different answers are really answers to different questions and are, therefore, complementary more than
contradictory; and discusses the consequences of the *apo*ha-triangulation integration presented in section 4.

2. THE NOUMENAL, THE PHENOMENAL, AND THE IN-BETWEEN

*Metaphysical dualism* is the distinction and opposition of a noumenal and phenomenal (level of) reality, the idea that the world as it appears to us may be (or even is) nothing like the world as it ‘really’ is independent of our experience. In Western philosophy, Kant’s distinction between appearances and things-in-themselves, or between *phenomena* and *noumena*, is the paradigm of metaphysical dualism, but Kant scholars disagree about the metaphysical status of the opposites — whether Kant’s noumenal - phenomenal dialectic is an opposition of ‘worlds’ or ‘aspects’. Similarly, the distinction in general can be framed alternatively in terms of ‘two realities’, ‘two worlds’, ‘two levels of reality’, ‘two aspects of reality’, and so forth. However, unless (or until) these terms are clearly defined, any one of them could be a metaphor for any of the others; and any one of them is equally metaphorical (and ambiguous) in itself.

Setting aside metaphors and Kantian connotations, ‘phenomena’ or ‘phenomenal appearances’ will be interpreted here as referring to the things, stuffs, events, and so forth as we consciously and conceptually experience them, and the ‘phenomenal world’, ‘phenomenal reality’, or the ‘phenomenal level of reality’ as the world of phenomena, the world or reality as it appears to us (as conscious observers and participants in that world/reality). Hence, phenomenal experience is conscious, conceptualized experience, experience of cows as cows, water as water, and weddings as weddings. These two predicates, ‘conscious’ and ‘conceptualized’ are not independent — neither Dharmakīrti, nor Davidson allows for conscious, non-conceptual experience for reasons to be explained below. However, there may be room for the converse, for some kind of un- or semi-conscious conceptual experience, although this would be mostly a derivative of conscious experience with little relevance for the subject matter of this paper.

The ‘noumenal world’, ‘noumenal reality’ or the ‘noumenal level of reality’ is the world or reality independent of our experience (or of how we experience it). There are not necessarily ‘noumena’ as counterparts of phenomena, however. It can be argued that phenomena have ‘noumenal causes’ (and perhaps coincide with (or are) those if non-dualism is correct), but it is not necessarily the case that there are

---

1 By way of advance summary: Dharmakīrti distinguishes conceptualized (‘determinate’) phenomenal experience (*pratītī*-pratibhāsa) from non-conceptual and non-conscious raw sensory data (*pratibhāsa*), with external (‘ultimate’/noumenal) reality as the only other category. Hence, ‘conceptual’ and ‘conscious’ come together. Davidson famously rejected ‘non-conceptual content’, *i.e.* something non-conceptual in the conscious mind (see below in this section), and argued that thought requires concepts (see next section).

2 The existence of such (derivative) unconscious but conceptual experience and thought is supported by research in cognitive neuroscience on non- (or un-) conscious processes. (*e.g.* Hassin *et al.* 2005; Tamietto & de Gelder 2010)
discrete, individuated noumena as the count noun character of that term suggests. A further asymmetry between the noumenal and the phenomenal is that while the phenomenal world (or phenomenal reality) is the sum total of phenomena, the noumenal world (or noumenal reality) includes the sum total of ‘noumenal causes’, but there may also be a part of noumenal reality without phenomenal effect(s) or counterpart(s).

*Metaphysical non-dualism* then, can be roughly defined as the idea that in normal circumstances phenomena necessarily are (identical with) their noumenal causes – ‘things’ as they appear to us are the (noumenally) real things(-in-themselves), and therefore, (phenomenal) appearances are non-deceptive. *Metaphysical dualism*, on the other hand, rejects the necessary identity of phenomena and noumena and argues either for possible or actual difference. Consequently, according to dualism, ‘things’ as they appear to us are or may be nothing like things-in-themselves, and therefore, (phenomenal) appearances are (actually or potentially) deceptive (or illusory).

The distinction between metaphysical dualism and non-dualism hinges on ‘identity’ (and its opposite, ‘difference’), which is a rather ambiguous notion here. A theory claiming complete qualitative identity could be easily qualified as non-dualist, and another theory claiming significant qualitative and numerical difference (in which ‘significant difference’ means something like ‘below any reasonable threshold of fuzzy identity’) would be dualist, but a theory claiming numerical identity and qualitative non-identity would be harder to classify, and it is not difficult to imagine further cases that seem to fall somewhere in between dualism and non-dualism. Rather than as a dichotomy, the distinction between metaphysical dualism and non-dualism could be perceived as a scale of some kind with intermediate positions between extremes. Resolving the issue of (non-) dualism – if possible – then, would not be a choice between these extremes, but an assessment of the extent to which phenomena and their noumenal causes can be meaningfully said to be (non-) identical.

Questions like this, and theories and debates about noumena and phenomena and (the) relationship(s) between them in general, are more commonly framed in terms of ‘realism’ and ‘anti-realism’. Depending on which interpretation or definition of these terms one adheres to, the crux of the realism - anti-realism polarity is either the existence of, or the possibility of making truth-apt statements about noumenal reality. The focal point here, however, is neither existence, nor describability, but identity or non-identity of noumena and phenomena, which is much more accurately captured in the terms ‘dualism’ and ‘non-dualism’ than in variants or extensions of ‘realism’ and ‘anti-realism’. Moreover, neither Davidson nor Dharmakīrti (at least in the context of

---

3 The clause ‘in normal circumstances’ is needed to exclude hallucinations and other extraordinary situations in which some experiencer (temporarily or not) experiences things that are not ‘real’. However, defining such ‘normal circumstances’ (and/or the contrary) without reducing the definition of non-dualism to ‘phenomena are noumena except when they are not’ may not be easy, and consequently, this clause is a serious weakness, but because presenting and defending a theory of non-dualism is not the purpose of this paper, as a characterization of the general idea this rough definition should suffice.
his epistemological writings) can be classified unambiguously and uncontroversially as either a realist or a non-realist (except perhaps, after another redefinition of those terms). (Davidson rejects both realism and anti-realism in (1998b), for example.) There have been attempts to recruit them for either camp, but both have also been interpreted (more appropriately) as somehow transcending the realism - anti-realism opposition. Dharmakīrti’s and Davison’s approximate positions on a dualism - non-dualism scale, on the other hand, are reasonably easy to determine – the former explicitly adheres to dualism, while the latter equally explicitly rejects it. Nevertheless, although there should be little doubt that Dharmakīrti is closer to the dualist end of the scale and Davidson closer to the non-dualist end, because of the aforementioned ambiguity of ‘identity’ and its opposite in this context, and because of interpretive ambiguities, it is difficult, perhaps even impossible, to pinpoint the exact position of either Dharmakīrti or Davidson.

One of the defining characteristics of Buddhist philosophy is a form of epistemological dualism, a distinction between conventional (or phenomenal) truth (saṃvṛti) and ultimate truth (paramārtha). Knowledge beyond the mere conventional or phenomenal (hence, beyond saṃvṛti), knowledge of ultimate truth, is (a requirement for) liberation (from suffering), and consequently, in Buddhist philosophy, epistemology is soteriology. Ultimate truth was often conceived as truth about (a.o.) the nature of the relationship between noumena and phenomena, but theories about that relationship differed between schools – some considered ultimate truth to be that phenomena and noumena are identical, while others argued for metaphysical dualism. The most rigorously developed version of such dualism can be found in the epistemological writings of the Yogācāra ‘logicians’ Dignāga (5-6th ct.) and (especially) his indirect student Dharmakīrti (7th ct.).

Like many other Buddhist philosophers, Dharmakīrti distinguished two ‘realities’ (or two levels of reality), ultimate reality (paramārthasat) and conventional reality (saṃvṛtisat) (e.g. Dunne 2004). Ultimate reality is the noumenal world, or the noumenal level of reality, and is composed of svalaḳṣana, unique spatio-temporally non-extended part-less noumenal ‘atoms’; conventional reality is the phenomenal world, or the phenomenal level of reality, the world as it appears to us. These two realities (or levels of reality) are connected through perception (pratyakṣa) and conceptual construction (kalpana). The connection is indirect, however. Central idea in Dignāga and Dharmakīrti’s epistemology is the distinction of and relationship between two kinds (or levels) of ‘appearance’ – indeterminate and determinate appearance (pratibhāsa and pratibhāsa-pratīti).

4 Dharmakīrti, Nyāyabinduprakaraṇa. – There are a number of different terms used for the distinction between determinate and indeterminate appearance, and as is the case with many Sanskrit philosophical terms, these are translated differently by different interpreters. ‘Pratibhāsa’ is an abstraction of ‘pratibhā’, which most literally means ‘reflection’, but which in a philosophical context more commonly meant ‘thought’, ‘appearance’, ‘mental image’ or something similar. A pratibhā is an ‘appearance’ in the sense of something appearing in, to, or before the mind. Furthermore, in the conceptual opposition with ‘pratibhāsa-pratīti’, pratibhāsa is not-pratīti. ‘Pratīti’ means ‘approach’, ‘apprehension’, ‘distinction’, ‘conviction’, or ‘belief’ (or something similar) and has a Vedic
particulars, while in the latter we directly grasp mentally constructed universals, but thereby indirectly grasp particulars” (Siderits 2004, 371). The notion of ‘determination’ (pratīti) separating these two kinds (or levels) of appearance refers to verbal determination or conceptual construction (kalpana). Indeterminate appearance is the result of direct perception (pratyakṣa), which was defined by Dignāga as “free from conceptual construction / nonconceptual” (kalpanāpodha) (Pramāṇasamuccaya 1.3c); and determinate appearance is (conscious) phenomenal perception – perception of something as something. In other words, determinate appearance is phenomenal appearance (or phenomenal experience), and indeterminate appearance is a pre-phenomenal intermediary between the noumenal world and our phenomenal experience.

Davidson famously argued against (certain kinds of) such intermediaries in “On the very idea of a conceptual scheme” (1974) and “The myth of the subjective” (1988a). He opposed what he called the ‘third dogma of empiricism’, the idea that there is “an element in the mind untouched by conceptual interpretation” (1988a, 40), “an ultimate source of evidence the character of which can be wholly specified without reference to what it is evidence for” (p. 42), and the related idea that our ‘conceptual schemes’ somehow ‘organize’ this ‘element’ or ‘fit’ this ‘source of evidence’ (1974). His primary target was the empiricist notion of ‘sense data’ (and the subjectivity it implies), and these short quotes summarize his main objections against that notion. The first objection is closely related to Sellars’s (1956) argument against ‘the myth of the given’, the idea that there are given in the mind independent of, and prior to, conceptualization. The second objection takes issue with the disconnection of phenomenal experience from its noumenal causes.

This latter objection is related to a more general rejection of metaphysical dualism, a recurrent theme throughout many of Davidson’s philosophical writings. According to Davidson, “words and thoughts (...) are necessarily about the sorts of objects and events that commonly cause them” (1988a, 45). It is the distal stimulus, the (noumenally) real world object or event, that determines the content of words and thoughts, not the proximal stimulus of sense data (see also section 3). More than in the arguments against sense data and subjectivity, this anti-dualism finds clear and explicit expression in Davidson’s rejection of the possibility of ‘massive error’. In a number of papers (e.g. 1977; 1982a; 1983), he argued that communication (or interpretation) is only possible if communicators have (shared) largely true beliefs about the world; and since we are able to communicate (often quite successfully), our phenomenal experience cannot be completely false (in ‘massive error’; i.e. different from noumenal reality) (see also the two quotes in the introduction of this paper). In the course of the 1990s, many of these arguments would be integrated within the ‘triangulation’ framework that will be the topic of most of the next section (see also Malpas 2009; 2011b).

*connotation of self-evident clearness. In this context, ‘pratīti’ refers to the determinateness of some pratibhā, to their appearance as something. ‘Pratibhāsa-pratīti’ then, is determinate appearance, and in contrast, ‘pratibhāsa’ is indeterminate (or perhaps pre-determinate) appearance.*
The same non/anti-dualism lays at the foundation of Davidson’s argument(s) against sense data, subjectivity, and conceptual schemes. In the aforementioned two papers, Davidson concluded that “the idea that there is a basic division between uninterpreted experience and an organizing conceptual scheme is a deep mistake, born of the essentially incoherent picture of the mind as a passive but critical spectator of an inner show” (1988a, 52); and “in giving up the dualism of scheme and world, we do not give up the world, but re-establish unmediated touch with the familiar objects whose antics make our sentences and opinions true or false” (1974, 198). Nevertheless, Davidson’s rejection of metaphysical dualism does not imply a rejection of the conceptual distinction between the noumenal and the phenomenal (as defined in this paper) itself, although he did not use those (or related) terms. This is evident, for example, in his philosophy of mind, ‘anomalous monism’, which depends on a distinction between noumenal events and their phenomenal descriptions (e.g. 1970; 1993).

At first glance, there seems to be a fundamental disagreement here between Davidson (and Sellars) on the one hand, and Dharmakīrti (and Dignāga) on the other. However, whether the refutation of (given) sense data is also a refutation of pratibhāsa (indeterminate, pre-phenomenal appearance) depends on whether the latter is (a kind or variety of) the former. And whether Davidson’s criticism of conceptual schemes also applies to kalpana (conceptual construction) depends primarily on whether the source material of kalpana is the same as (or essentially similar to) that what is ‘organized’ or ‘fitted’ in Davidson’s interpretation of conceptual schemes, hence again, to whether pratibhāsa is (given) sense data.

Davidson and Sellars object to (given) sense data as some kind of non-conceptual mental content. In its epistemic role (as ‘evidence’) that mental content must have propositional form. It must consist of identifiable, more or less discrete, and determinably (inter-) related chunks of sensory information that map (almost) directly to concepts. For words (or ‘concepts’) to ‘fit the evidence’ of sense data, that evidence must consist of such (interrelated) sense data chunks. As the source material of the ‘organizing’ role of conceptual schemes (in Davidson’s reconstruction thereof), sense data is similarly structured. Davidson’s interpretation of the metaphor of ‘organization’ is really re-organization – conceptual schemes re-organize what is already in some sense organized (i.e. structured in discrete chunks of sensory information) (1974, 192ff). Hence, ‘non-conceptual’ is a bit of a misnomer here because, although such sense data is non-verbal (or better: not-yet-verbal), it is not entirely non-categorized (but note that the categorization is (supposed to be) provided by noumenal reality, not by a categorizing agent). Davidson and Sellars’s interpretation of sense data is better described as propositional, non-verbal mental content, as something in the mind that has propositional form, but (still) lacks verbalization. Such sense data should be distinguished from ‘sensory experience’ (1988a, 45), the pre-mental, uninterpreted signal of the sense organs, which is non-propositional (in addition to being non-verbal), which is more before than in the mind, and which is causally, but not epistemically, related to conceptual(ized) beliefs (in the mind). Davidson rejects sense data, but not such ‘sensory experience’ (his
term) (but he does consider the latter without philosophical significance). “Of cause there are causal intermediaries. What we must guard against are epistemic intermediaries” (Davidson 1983, 144).

In Nyāya epistemology, the main philosophical opponent of Buddhism in Dhammakīrti’s time, perception is a two stage process involving bare, undifferentiated (nirvikalpaka) perception, followed by perceptual judgment resulting in differentiated (savikalpaka), but still non-verbal perception; and conceptualization is nothing but the attachment of linguistic labels to the propositional content resulting from that second stage of perception. In other words, the Nyāya theory of perception involved a notion of propositional, non-verbal mental content, hence (given) sense data. It was exactly that notion that Dhammakīrti rejected (in Nyāya epistemology). According to Dhammakīrti, “perception is necessarily non-propositional” and “limited to a bare sensing which does not directly produce any useable information” (Dreyfus 1996, 213).

Pratībhāsa (indeterminate, pre-phenomenal appearance) is not (given) sense data. Like Davidson’s ‘sensory experience’, it is pre-mental and uninterpreted, and causally rather than epistemically related to phenomenal experience. There is a further similarity between the two concepts in their ‘innocence’: Davidson argued that the form and nature of ‘sensory experience’ (as it is produced by particular sense organs) is mere empirical accidence without philosophical significance (1988a, 45); and Dhammakīrti added the predicate ‘free of error’ (abhrānta) to Dignāga’s definition of pratībhāsa as non-conceptual.

In other words, despite the fundamental difference with regards to the (non-) identity of noumena and phenomena, there appears to be a lot of agreement between Davidson and Dhammakīrti – both argue for a causal (rather than epistemological) connection between the noumenal and phenomenal levels of reality through an innocent intermediary (pratībhāsa or ‘sensory experience’). Moreover there are striking similarities in their respective theories on the noumenal - phenomenal connection, which is the subject matter of the next section.

3. APOHA AND TRIANGULATION

In Dignāga’s and Dhammakīrti’s philosophical system(s), phenomenal experience and conceptual knowledge are connected (through perception) to ultimate (noumenal) reality by means of apoha. This concept of ‘apo[a]’ plays a role in two different but related theoretical contexts – conceptual construction (kalpa) and concept formation. In the former we construct particular determinate phenomena in opposition

---

5 Davidson’s term for this innocent intermediary, ‘sensory experience’ is not a particularly good term because the word ‘experience’ suggests something in the mind, perhaps even something conscious, and something pre-conceptual in the (conscious) mind is exactly what Davidson objected to, and which is exactly what ‘sensory experience’ is not.

6 Because of the epistemological purpose of Dhammakīrti’s theory, his rejection of an epistemic intermediary lead to theoretical difficulties that kept many generations of (especially Tibetan) Scholastics occupied (see Dreyfus 1996), but that is of no concern to us here.
to what they are not; we perceive a particular cow as not non-cow; in the latter we create (pseudo-) universals rather than particular determinations (phenomena). Dignāga introduced the concept in the former context. Conceptual construction (kalpana) is a form of inference (Pramāṇasamuccaya 5.1) which works by means of exclusion: *apoha* (id. 5.17). The closest Western equivalent to this idea is Derrida’s concept of ‘*différance*’, which similarly expresses a deferring of meaning (or determination) by reference to or embedding in a network of different, but related, concepts. Dharmakīrti further elaborated Dignāga’s idea (and philosophy in general) and applied it to concept formation (mostly in the auto-commentary to his Pramāṇavārttika; Dreyfus 1997; 2011; Tillemans 2011). Making Dharmakīrti’s ideas “clearer than they are” (p. 278), Dreyfus summarizes *apoha* as concept formation as follows:

[O]ur starting point is our experience of things and their mutual resemblances. These experiences give rise to a diffuse concept of similarity. To account for this sense of similarity, we construct a more precise concept by correlating conceptual representations with a single term or sign previously encountered. This creates a more precise concept in which the representations are made to stand for a commonality that the objects are assumed to possess. (…) In this way experiences give rise to mental representations, which are transformed into concepts by association with a linguistic sign. The formation of a concept consists of the assumption that mental representations stand for an agreed-on imagined commonality. Two points must be emphasized here regarding concept formation. First a concept, which is nothing but an assumption of the existence of a fictional commonality projected onto things, comes to be through the conjunction of two factors: the experience of real objects and the social process of language acquisition. (…) Second, a concept is mistaken. (Dreyfus 1997, 227)

Further summarizing Dreyfus’s summary: concepts are formed in communication about shared experiences of real things. Hence, *apoha* is (or involves) a kind of social and physical externalism about language. In (recent) Western philosophy, externalism about language and/or mental content is usually either social, as in Kripkenstein’s or Burge’s externalism, or physical, as in Putnam’s, but rarely both. The most important exception is Davidson, particularly his theory of triangulation, which has much in common with *apoha*. Dreyfus (2011) writes that according to *apoha* “thought and language are causally related to our experiences of things and hence are grounded in reality” (p. 209), but the exact same could be written about triangulation. Compare also the first half of the above quote with Davidson’s assertion that “all creatures classify objects and aspects of the world in the sense that they treat some stimuli as more alike than others. The criterion of such classifying activity is similarity of response” (1991, 212). And compare the above quote’s conclusion with the central idea of triangulation: concepts (and therefore the possibility of communication and thought) depend “on the fact that two or more creatures are responding, more or less simultaneously, to input from a shared world, and from each other” (1997a, 83).

Davidson first introduced the notion of triangulation in the second last paragraph of “Rational animals” (1982b) in the context of an argument for the necessity of
objectivity for thought. “Our sense of objectivity is the consequence” of a kind of triangulation involving two creatures, each of which “interacts with an object, but what gives each the concept of the way things are objectively is the base line formed between the creatures by language” (p. 105). During the 1990s, Davidson employed the notion in a variety of related contexts, and consequently, triangulation became gradually associated with more and more parts of his philosophical system (see also Amoretti & Preyer 2011a; Malpas 2009; 2011b).7 Partially because of that, “Davidson’s triangulation figure is notoriously difficult to explicate” (Fergestad & Ramberg 2011, p.221), and different interpreters understand the notion (sometimes subtly) differently and position it differently in the wider context of Davidson’s philosophy (or even ignore that wider context, as Malpas (2011a) points out). For example, Føllesdal (1999) suggests that the term was introduced “for the process of language learning” (p. 724), specifically for the role of perception in word learning; Sosa (2003) associates it with the principle of charity and radical interpretation in the context of the refutation of skepticism about external reality (hence, in the context of metaphysical anti-dualism); Carpenter (2003) interpretes triangulation as a model of language learning and as such as the ‘transcendental argument’ for externalism about mental content at the very core of Davidson’s philosophical system; Lepore and Ludwig (2005) locate it in the context of the ‘third person perspective’ and Davidson’s arguments for the necessity of language for thought; and Glüer (2006) disentangles the two aspects of triangulation suggested in Davidson (1997b): “without the triangle, there are two aspects of thought for which we cannot account. These two aspects are the objectivity of thought and the empirical contents of thoughts about the external world” (p. 129). All of these interpretations (and probably a few more that are not mentioned here) are correct; the apparent differences are merely the result of differences in focus on particular aspects of the notion’s application.

“In its simplest form” or “its pure state” (e.g. Davidson 1999, 731), triangulation is a singular occasion of pointing out some object by one communicant to another by means of some ad hoc sign. These two communicating creatures and the shared stimulus are the three corners in the triangle. Many of Davidson’s papers make use of a notion of triangulation in a less ‘simple’ or ‘pure’ form, however: as (a model of) a process of word learning by means of repeated similar signs in the (repeated) presence of similar stimuli (e.g. 1990a; 1992; 1994; 1997c; 1998). In the second half of the 1990s, the notion of ostensive learning appeared as a variant denotation of such triangulatory word learning (my term) (e.g. 1997c; 1998), and in that form, the idea made its final appearance in the last pages of the posthumously published Truth and Predication (2005a).8

7 In (1990a), discussing externalism, Davidson wrote that “for some thirty years [he had] been insisting that the contents of our earliest learned and most basic sentences (...) must be determined by what it is in the world that causes us to hold them true” (p. 200); in other words, that the basic idea of triangulation had already been a central part of his philosophy for some thirty years at that point.

8 In (2001a), Davidson wrote that (the importance of) ostensive learning was his original inspiration for the idea of triangulation. Consequently, the later references to ‘ostensive learning’ are a return to the source of triangulation more than a substantial change or innovation.
Either in its ‘simple’ or in its less simple form, triangulation involves the same triangle graphically represented in figure 1. It should be immediately obvious that the term ‘triangle’ is a misnomer if triangles are considered to consist of nothing but three corners and three sides – triangulation involves a fourth element: language (S in figure 1), which Davidson called the ‘base line’ on a number of occasions, but which is more properly characterized as the triangle’s pivot. These four elements are connected by six lines as shown in the figure and explained in the following. A communicating creature U utters (wavy line) some sound S (which could in principle also be another kind of sign) in reference9 (double line) to some object, occurrence, or state of affairs O perceived (single line) by U. Triangulation further involves a second creature I that also perceives (single lines) the object (etc.) O, the sound S, and its utterer U, and the relationships between these three (uttering: U-S, reference9: S-O, perception: U-O; but note that the awareness of these relationships is not graphically represented in the figure). By ‘correlating’ these incoming lines, I finds or creates the meaning (or ‘content’) of S. Words (as a kind of sign S; but not necessarily all words) are learned by repetition of this process (often with different people in the roles of U and I at different occasions, and/or the same people switching roles): by repeated correlation of sufficiently similar verbal signs S, sufficiently similar objects O, and sufficiently similar utterers U (that belong to the same species and seem to speak a similar language, for example).10

Triangulation explains how (it is possible that) words get (empirical) content or meaning. Triangulation “is necessary if there is to be any answer at all to the question what [a creature’s] concepts are concepts of” (1992, 119); without triangulation, words could not have content, and (therefore) (propositional) thought would be

---

9 ‘Reference’ here, is nothing but the act (by the utterer) of referring to something; it should not be confused with the technical term.
10 The repetition of such correlations is essential. Malpas (2011b) correctly points out that there is a “need for triangulation to operate through repetition” (p. 259). This, however, seems to be ignored by some (perhaps even many) critics of triangulation (see Amoretti & Preyer 2011a for an overview). The common criticism that triangulation does not fix a connection between a word and a noumenal kind seems to be based on (a) the assumption that according to Davidson, such a connection can be made in a single triangulatory event (rather than in repeated triangulations), and (b) the neglect of Davidson’s holism and coherentism. The latter fits in with a general tendency observed by Malpas (2011a) to ignore these and other less orthodox aspects of Davidson’s (especially later) philosophy. See further section 4.
impossible (1982b; 1991; 1997b). The question how (it is possible that) words get meaning (or content) should be distinguished, however, from the questions what ‘meaning’ is and how to determine (a) meaning, although neither these questions, nor their answers are always easily kept apart. For example, Davidson’s answer to the last of these questions is his theory of interpretation (radical interpretation) (1973) in cases with the least information to start with) but interpretation (of meaning) is based on the same triangle: interpreter $I$ determines the meaning of $S$ by correlating the ‘incoming lines’.\footnote{On the relation between triangulation and (radical) interpretation, see also Carpenter (2003), Goldberg (2011), and Malpas (2011b).} As his closing statement in a paper rejecting Quine’s dependency on proximal stimuli (see also section 2), Davidson writes: “the active role of the interpreter (...) requires that the interpreter correlate his own responses and those of the speaker by reference to the mutually salient causes in the world of which they speak” (1990b, 62). The terminology here seems to be that of interpretation, but the idea expressed is that of triangulation: speaker, interpreter, mutually salient cause (the distal stimulus); these are the three corners ($U$, $I$, $O$) of the triangle. The apparent incongruity is partially explained by the strong similarities between the two ideas (interpretation and triangulation), which differ more in background and purpose than in process or effect; and partially by the fact that, aside from the singular mention of ‘triangulation’ in (1982b), it was only around this time that Davidson started to apply the term in a wider range of contexts.

Davidson answered the third question, what meaning is, by analyzing meaning in terms of truth (e.g. 1967; 2005a) – ‘$s$ means in $L$ that $p$’ is equivalent to ‘$s$ is true-in-$L$ iff $p$’ where $s$ is some statement, $L$ is a language, and $p$ is a description of the truth condition of $s$ in the meta-language. If $S$ in figure 1 would be a simple constative rather than a word – ‘ame ga futte iru’ (‘it is raining’ in Japanese), for example – then the meaning of $S$ in the language spoken by $U$ is that, if $S$ is true, then $O$, which is described in the meta-language (English in this case) as ‘it is raining’. By defining ‘meaning’ in this way, Davidson avoided an ontological commitment to facts as truth-makers and meanings as abstract entities. ‘Meaning’ in itself is of little relevance in the context of this paper, however; what concerns us here is the implication of the possibility of meaning something – not so much that $S$ is true iff $O$, but that without $O$ (along with some other conditions), there could not be an $S$.

Perhaps confusingly, Davidson sometimes seems to coalesce the questions what meaning is and what makes meaning possible; that is, ‘meaning’ is sometimes understood in terms of triangulatory word learning history, rather than in terms of truth, particularly in the context of metaphysical anti-dualism. Two speakers “may mean different things by the same words because of differences in the external situations in which the words were learned”, and therefore, “the correct interpretation of what a speaker means is not determined solely by what is in the head; it depends also on the natural history of what is in the head” (1988a, 44). Which leads to the conclusion that “in the simplest and most basic cases, words and sentences derive their meaning from the objects and circumstances in whose presence they were
learned” (p. 44), in other words, that meaning depends on triangulatory word learning history; and consequently, “words and thoughts (...) are necessarily about the sorts of objects and events that commonly cause them” (p. 45), and thus not subjective (or illusory).

Although this argument predates Davidson’s explicit usage of the term ‘triangulation’ in the context of word learning (1990a; 1992; 1994; see also above and footnote 7), and the term, therefore, does not occur in the paper, the idea that “words and sentences derive their meaning from the objects and circumstances in whose presence they were learned” is the same idea elaborated in the word learning version of triangulation, thus connecting Davidson’s arguments against subjectivity or metaphysical dualism with triangulation. His (1988a) argument also connects triangulation and anti-dualism with the ‘primacy of the idiolect’, the idea that ‘a language’ is a collection of idiolects (1986; 1994), which also applies to truth-conditional meaning (see above) – the L in ‘true-in-L’ is an idiolect, and ‘true-in-L’ is ‘true-in-the-idolect-of-U’ (U is the utterer in figure 1), where ‘idioclect’ is, moreover, continuously changing under the influence of new triangulatory experiences (thus: ‘true-in-the-idolect-of-U-at-t’).

Aside from the somewhat indirect connection between triangulation and metaphysical anti-dualism through (meaning as) triangulatory word learning history, there are two more direct connections. The first of these can be found in (1983) and (1990a): Davidson’s argument against ‘skepticism of the senses’ (i.e. metaphysical dualism) in (1983) quoted in the introduction of this paper (“What stands in the way of global skepticism of the senses is, in my view, the fact that we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief.”; p. 151), is quoted by Davidson himself in (1990a) in the context of externalism about mental content, i.e. the origination of the content of words and thought – through triangulation – in external reality. The second of these more direct connections was already hinted at above: the stimulus in triangulation (O in figure 1) is shared (e.g. 1991; 1997a), and a shared stimulus can only be a distal stimulus. Using the term ‘interpretation’ rather than ‘triangulation’, Davidson pointed this out in (1990b), as mentioned above, but the argument of that paper partially overlaps with that of (1997a), which explicitly mentions “the triangular nexus of causal relations” (p. 83). There cannot be language, communication, and thought without triangulation, and there can be no triangulation without (shared) distal stimuli (along with the other necessary conditions) – we speak and think, therefore our words are grounded in distal stimuli (i.e. in the noumenal world).

4. TRIANGULATING DIFFERENCE

Triangulation as word learning and apoha as concept formation, while similar in their world - word connection through social learning processes, differ in two respects.

---

12 On the connections between triangulation, (radical) interpretation, anti-dualism, and externalism about mental content, see also Carpenter (2003), Sosa (2003), and Malpas (2011b).
First, and least importantly, the idea and implications of triangulation are much more extensively, and more rigourously explored than *apoha*, which is partly due to the theories’ embeddedness in different philosophical traditions, and partly to the obscurity of many of the key philosophical texts on *apoha*. The second, more important difference is expressed in the characterizations as *word learning* and *concept formation* respectively – triangulation merely names (or assigns words to) the objects and occurrences given by the noumenal world (in a sense, Davidson relocates the ‘given’ from sense data to noumenal reality), while *apoha* simultaneously *constructs and names* phenomena. Assuming that conceptual construction is always and necessarily construction out of sense data, Davidson rejected construction with the rejection of sense data (see section 2). That rejection may have been premature, however. Conceptual construction (*kalpana*) by means of *apoha* in its more narrow sense as *anya-apoha*, (classification by) exclusion of what is different, does not depend on sense data or similar epistemic intermediaries, and is therefore, at least in that respect, not inconsistent with triangulation. And incorporating *apoha* in triangulation – that is, adapting triangulation to allow for construction by exclusion, or extending a more rigorous reconstruction of *apoha* with an exploration of consequences as in triangulation – does neither lead to metaphysical dualism, nor to anti-dualism, as the final sections of this paper will show. Furthermore, similarity-based (positive) classification of experiences is considered a weakness of Davidson’s theory of triangulation by some critics (*e.g.* Fennel 2000; Ludwig 2011), and therefore, the substitution of (negative) classification by exclusion may lead to a stronger version of triangulation; a version that would, moreover, be tied in more closely with Davidson’s holism.

Consider the following very abstract example: a perceiver has become aware of 9 particulars, which all have one and only one characteristic such that these characteristics are values on a single dimension, and each particular has a unique value on that dimension. figure 2 shows the positions of the 9 particulars on that single dimension (the x-axis, ranging from 0 to 1). (An example of such a dimension could be (the noumenal ground or cause of) the range from dark (0) to light (1).) The perceiver perceives the difference between the 9 particulars, but at the same time notices that some are less different (or more similar) than others. The pair \{d,e\} is considerably less different than \{c,h\}, for example.

![Figure 2](image)

Next, assume a second perceiver who, in the presence of e (and nothing but e) utters ‘def’, and does so again in the presence of d (and nothing but d), but utters ‘bac’ in the presence of b (and nothing but b). With this information alone (the utterings and the perceived relative differences), the first perceiver is able to construct two working concepts, ‘bac’ and ‘def’, such that the first refers to \{a,b,c\}, and the
second to \{d,\ldots,i\}. Further communication may refine these concepts – despite the
gap between \(c\) and \(d\), it is in principle possible that \(c\) is classified as ‘\text{def}’ rather than
‘\text{bac}’, for example; and there may be a third concept ‘\text{hig}’ that refers to \(\{g,h,i\}\) and
that restricts ‘\text{def}’ to \(\{d,e,f\}\). (Note that there are many more possible classifications
of the particulars in figure 2, including various discontinuous ones.)

Let us assume that further triangulations result in the latter restriction of ‘\text{def}’
indeed. The formation of this concept makes those particulars seem even more
similar, even identical in their ‘\text{def}’-instantiating capacity, but that is mere
phenomenal illusion. Having a concept ‘\text{def}’ does not mean that the apparently
similar elements of \(\{d,e,f\}\) are ‘\text{def}’, that they share a property ‘\text{def}’-ness, that such a
universal ‘\text{def}’-ness exists, or that \(\text{defs}\) (ultimately/noumenally) exist (as \(\text{defs}\)). All
that the ‘things’ we call call ‘\text{def}’, the phenomenal \(\text{defs}\), share is that they are ‘not
\text{non-def}’.

This ‘not non-...’ is \textit{the} key notion in \textit{anya-apoha} (exclusion of difference). It
should not be interpreted as a double classical negation ‘not not’, but available
sources do not make clear how ‘not non’ should be understood exactly either. Siderits
(\textit{e.g.} 1991) proposed an interpretation based on Matilal’s (1971) distinction between
‘nominally’ and ‘verbally bound negation’ found in Hindu philosophy, but as Siderits
admits himself (2011), there is no clear evidence for that distinction in Buddhist
philosophy (and it is, moreover, not immediately clear how Siderits’s two kinds of
negation relate to ‘exclusion’ and ‘difference’). Nevertheless, the idea that ‘not non’
should be interpreted as two different negation-like operations is not implausible,
provided that it can be made to work such that it coheres with the basic idea of
excluding what is different (\textit{anya-apoha}), while not collapsing into identity (as is the
case with a double classical negation).

The point of the formula ‘\(x\) is not non-\(X\)’ is avoiding the ontological commitment
to an universal involved in ‘\(x\) is (an) \(X\)’. \(X\) marks a class, membership of which is
determined by an universal, meaning or intension, or something similar; and ‘\(x\) is \(X\)’
denotes membership of that class (\(x \in X\)), which implies that there is an \(X\) (that the
universal \(X\) (or \(X\)-ness) exists). Any acceptable interpretation of ‘not non-...’,
therefore should (a) not collapse into \(x \in X\), and (b) not commit to the existence of \(X\)
(or other universals or classes determined thereby) by existentially quantifying over \(X\)
/etc.). Three interpretations will be briefly outlined here, but the first two violate
either (a) or (b) and are, therefore, rejected. In all three, it is assumed that ‘not’ stands
for classical negation, and that it is, therefore, the ‘non’ operation that needs
elucidation.

(1) ‘\(\text{Non-X}\)’ is the absolute complement of \(X\) (\(X^C = \text{def.} U \setminus X\)) and ‘\(x\) is non-\(X\)’
means \(x \in X^C\) (\(x\) is an element of the absolute complement of \(X\); that is, it belongs to
the class of ‘things’ that are not \(X\)\). ‘Not’ as classical negation merely negates \(\epsilon\), and
therefore, ‘\(x\) is not non-\(X\)’ means \(x \notin X^C\) (\(x\) is not in (or is excluded from) the class of
things that are different from \(X\)\). However, this interpretation (obviously) fails
because \(x \notin X^C \iff x \in X\), collapsing the two negations in ‘\(x\) is not non-\(X\)’ into ‘\(x\) is \(X\)’,
thus violating (a). Nevertheless, the implicit identification of ‘not’ with ‘exclusion’
Comparative Philosophy 3.1 (2012)

BRONS

and ‘non’ with ‘difference’ in *anya-apoha* (exclusion of what is different) seems correct.

(2) Rather than to an absolute complement, the prefix ‘non-’ is a negative reference to an open class of mutually exclusive contextually appropriate alternative classifications (where ‘open class’ is a class with undetermined and unfixed membership). If $D = \{X, Y, Z, \ldots\}$ is that open class of such contextually appropriate classifications $X$, $Y$, $Z$, and so forth, then ‘non-$X$’, here symbolized as $D^X$, is $D$ minus $X$: $D^X = D \setminus \{X\}$. However, because the elements of $D$ are classes themselves, the relationship between $D^X$ and any particular is indirect: ‘$x$ is non-$X$’ means that $x$ is an element of a class $D$ that is an element of $D^X$, formally: $\exists D \in D^X [x \in D]$. And since ‘not’ is just classical negation, ‘$x$ is not non-$X$’ is nothing but the negation thereof: $\neg \exists D \in D^X [x \in D]$. In words: there is no contextually appropriate class other than $X$, such that $x$ is a member of that class; or shorter: $x$ is not in a different class (than $X$).

Although interpretation (2) avoids (direct) commitment to (the existence of) $X$, it does so by positing a series of alternative classifications $Y$, $Z$, ..., and in (indirect) violation of (b), is ontologically committed to those: $\exists D \in D^X [x \in D]$ ($x$ is non-$X$) quantifies over elements of $D^X$, which are classes determined by universals (above described as ‘contextually appropriate alternative classifications’). Moreover, *anya-apoha* similarly applies to $Y$, $Z$, ..., and thus, if negative classification assures that $X$ has no members, then these have no members either. By implication, it is true that $\neg \exists D \in D^X [x \in D]$ for *any* $x$ and $X$: if ‘$x$ is not non-$X$’ it is also ‘not non-$Y$’, ‘not-non-$Z$’, and so forth. Therefore, (2) fails, but the reasons for its failure point at the third interpretation.

(3) The source of (2)’s failure is $D$, which is effectively a set of universals. This third interpretation keeps the basic framework of (2), but substitutes $\hat{D}$ for $D$. The members of $\hat{D}$ are not universals, but non-overlapping loose collections of ‘things’ or experiences deemed not-non-something in previous triangulation events, such that both the members of $\hat{D}$ and their members are salient to the interpreter in the occasion of the new triangulation event.\(^\text{13}\) Thus $\hat{D}$ is the collection of ‘things’ considered ‘not non-$Y$’ before, and remembered, reconstructed, or otherwise (sub-)consciously activated in the context of the (new) triangulation of $x$. There may or may not be an element $X$ in $\hat{D}$, depending on whether something was triangulated as ‘not non-$X$’ before. In either case $\hat{D}^X$ excludes that (if $X \notin \hat{D}$ then $\hat{D}^X = \hat{D} \setminus \{X\} = \hat{D}$; if $X \in \hat{D}$ then $\hat{D}^X = \hat{D} \setminus \{X\}$). The members of $\hat{D}^X$ are collections of remembered or reconstructed past triangulations (or strictly speaking, the particular objects or features involved therein), and consequently, the object or feature $x$ in the new triangulation can (barring exceptions such as (partial) loss of memory) not be a member of any of these. This means that the membership relation in the right-hand part of

\(^{13}\) In actual learning situations, relevant members of $\hat{D}$ are often made salient. For example, when offering a new word to a child, parents often follow up the usage of the new word “with additional information in the form of a selection of adjacent terms – neighbors from the relevant semantic domain that help delimit the probable meaning of the new term” (Clark 2010, p. 257).
∃D ∈ D⁻¹[X \{x \in D\}] in (2) needs to be replaced (in addition to dotting the Ds). In the process of triangulation, x is compared to contextually salient previous triangulations, the elements of elements of D, and found to be similar and/or dissimilar to some. The relationship between x and some D (D ∈ D⁻¹[X]) is that of (subjective and contextually salient) similarity: sim(x, D). With those changes to (2), ‘x is non-X’ becomes ∃D ∈ D⁻¹[X] [sim(x, D)], and ‘x is not non-X’ the classical negation thereof: ¬∃D ∈ D⁻¹[X] [sim(x, D)] (which can be read something like: ‘there is no ♦ collection of previously triangulated not-non-somethings other than not-non-X, such that x is ♦ similar to the ♦ members of that collection’; in which each ♦ can be replaced with ‘subjectively, contextually salient(ly)’).

Contrary to (2), (3) is not ontologically committed to universals (or universal-like classes), but merely to collections of (remembered or reconstructed) particulars loosely bound together by that interpreter’s history of triangulations. An obvious objection to (3) would be that ‘similarity’ (sim(x, D)) is too vague and subjective to guarantee success (see also footnote 10), but there are two counterarguments against that objection. Firstly, repeated triangulations (of similar particulars) progressively disambiguate (dis-)similarity and reduce subjectivity. Secondly, the whole process embeds any singular triangulation in the whole triangulation history of that interpreter, locking all words and concepts together in a single, (more or less) coherent whole. Coherence and incoherence with that history and its result (the web of concepts) also disambiguates (dis-)similarity and reduces subjectivity. The essence of anya-apoha, or at least this triangulation-infused interpretation thereof, is that the meaning of a word or concept is not some kind of universal, but its embedded triangulation history, a history that, moreover, never stops. Rather, speakers/interpreters (one has to be both to be either) continuously further ‘atune’ their words and concepts in further communication. In ‘atunement’, at least some vagueness clears up.

There is a second, perhaps more fundamental type of vagueness involved in concept formation: the vagueness of what counts as one particular. In addition to the negative classification (anya-apoha) of ‘things’ (as not-non-something), apoha similarly constructs (at least some of) the boundaries of (at least some of) the ‘things’ classified. In the above example, illustrated in figure 2, there were nine clearly distinguishable and discrete particulars, and the process of concept formation merely added a convenient (negative) classification. However, as Dharmakirti pointed out, singular concepts are either grounded in (noumenally caused) singular effects, or are purely conventional (saṃketa). There are many cases in which particular-hood itself

---

14 In their Thinking through Confucius, Hall and Ames’s (1987) use the term ‘atunement’ for a similar process. Their Derrida-inspired interpretation of Confucius has interesting similarities with the apoha-triangulation integration suggested here. “Classical Chinese is a system of differences, (...) The meaning of a given sign is (...) determined by its active and passive difference, and that meaning is never altogether present but deferred” (p. 293).

15 “Ekavacanam api tad-eka-sakti-sūcana-arthaṃ saṃketa-paratantram vā” (autocommentary (PVS) at Pramāṇavārttika 1.141-142). Note that the expression ‘(noumenally caused) singular effects’ is not an accurate translation of “eka-śakti-sūcana-artha” but a summary of what this fragment refers to: the
is determined by (negative) classification, where the objects we (phenomenally) perceive are constructed (*kalpana*) out of non-discrete indeterminate appearance or sensory experience, and are therefore conventional (*saṃketa*). To illustrate this, consider again a rather abstract example. Figure 3 shows a continuum on one spatial dimension (the *x*-axis), and with one type of characteristic with infinite values between two extremes (0 and 1; the *y*-axis) – a bit like one-dimensional pumpkin soup with solidity on the *y*-axis.

![Figure 3](image)

Given the right circumstances, an agent could form a concept ‘gu’\(^{16}\) (as ‘not non-gu’), but in the same way that particulars could be grouped together differently in case of the previous example, in this case particular-hood itself can be constructed differently. For example, there may be an absolute threshold, the dotted line in the figure, making the concept of ‘gu’ applicable to a, b, \{c,d\}, and e; or a relative threshold, which might exclude a, but include f; and depending on the convention constructed, different bumps in the line are called ‘gu’, and c and d may be considered one *gu* or two. This is, of course, a very abstract example, but it can be easily made less abstract – distinguishing ‘gu’ from ‘non-gu’ could be distinguishing chunks (from non-chunk) in pumpkin soup; or mountains (from non-mountain) in a landscape (on a somewhat larger spatial scale). Or perhaps, if the *x*-axis would represent time rather than space, distinguishing ‘gu’ could be distinguishing specific types of events. Whatever the rule for distinguishing ‘gu’ from ‘non-gu’, it is nothing more than a (constitutive) rule, a convention (*saṃketa*). There are no real (discrete) noumenal *gus* (as a plural of *gu*); it is that convention that creates phenomenal instances of ‘gu’.

In other words, the thing-ness (or particular-hood) of (at least) some perceived ‘things’ is itself phenomenal. Indeed, the 9 particulars in figure 2 – as individual, discrete particulars – may be phenomenal. Furthermore, not just their particular-hood, but also the ‘one characteristic’ that differentiates them (the *x*-axis in figure 2) and that is the basis for classification and concept formation, and similarly, the characteristic expressed on the *y*-axis in figure 3, may be phenomenal (as they are in the explanatory examples). The direct source material for concept formation (and/or conceptual construction) does not (always) need to be the indeterminate sensory experience or *pratibhāsa* that is directly caused by noumenal reality, but may be –

\(^{16}\)Not coincidentally, ‘gu’ (具) is the Japanese word for (among others) (more or less) solid things in soup.
and most likely often is – already otherwise determined, hence phenomenal. In such cases, concept formation adds a conceptual layer, a further classification. Nevertheless, below (or behind) such ‘layers’ of phenomenal description, there must be a level of basic concepts (Davidson’s ‘plainest and methodologically most basic cases’ in the quote in the introduction above, see also 1988a, 44) that is created or learned in triangulation of shared indeterminate appearance.

5. BOUNDARIES OF TYPES AND TOKENS

Apotha in the narrow sense is classification by exclusion of what is different, negative classification, to be contrasted with positive classification, which groups things together on the basis of (perceived) class-internal similarity rather than class-external difference. Nevertheless, all classification, either negative or positive, either conceptual or otherwise, is drawing (class) boundaries; drawing boundaries between ‘gu’ and ‘non-gu’ and between ‘defs’ and ‘non-defs’, for example.

The two examples in the previous section illustrate the two kinds of (class) boundaries involved in conceptual construction and classification: extensional boundaries and object boundaries, or boundaries of types and tokens. Extensional boundaries are the outer limits of the extension of a conceptual class or type, such as the boundaries between ‘defs’ and ‘non-defs’. Object boundaries are the outer limits of what is constructed and classified as one instantiation, case, or token of a certain conceptual type, such as the boundary between one specific instantiation (etc.) of ‘gu’ and the surrounding ‘non-gu’. In other words, extensional boundaries delimit types, and object boundaries delimit tokens.

Both kinds of boundaries can be conventional or intrinsic. Conventional boundaries are more or less arbitrary boundaries that are ‘drawn’ in areas of gradual rather than discrete difference (between noumenal causes) in the social process of concept formation. Non-conventional or intrinsic boundaries, on the other hand, match ‘real’, intrinsic ‘joints’ in noumenal reality and are therefore not created, but discovered or confirmed. In meta-ontology a similar distinction is often phrased in terms of the ‘cookie-cutter’ metaphor (concepts cut up noumenal reality into objects and categories in the way cookie-cutters cut up amorphous dough) versus ‘cutting nature at its joints’. Metaphysical dualism assumes that (at least a significantly large share of) boundaries are conventional; non-dualism assumes that they are intrinsic.

Boundaries of types and tokens do not play a role in Davidson’s philosophy because he implicitly assumed intrinsic boundaries, which – together with the equally implicit assumption of positive classification – became self-confirming in his philosophical system. Conventional boundaries are the result of conceptual construction, and because Davidson assumed that conceptual construction can only take sense data as its source material, with the rejection of sense data Davidson rejected conceptual construction (see also sections 2 and 3). Among the many different names and metaphors used for the notion of ‘conceptual construction’, Davidson (1974) focused on the rather ambiguous ‘organization’ (aside from ‘fitting’, which is mostly irrelevant in the present context), but interpreted it as re-organization...
– concepts re-organize what is already (otherwise) (conceptually) organized. Consequently, he failed to make sense of the notion: “how could something count as a language that only organizes experiences (...). Surely knives and forks, railroads and mountains, cabbages and kingdoms also need organizing” (p. 193). Underlying Davidson’s misunderstanding is the aforementioned implicit (prior) assumption of intrinsic extensional and object boundaries – ‘things’ are already ‘organized’ (that is: intrinsically bounded and classified), and therefore, further ‘organization’ can only be re-organization (or re-classification). However, if ‘things are already organized’, and thus non-constructed, then all that is left for triangulation is the assignment of names to things, and triangulation as the necessary link between words and things circularly confirms their ‘thingness’ (i.e. their intrinsic boundaries). In other words, if ‘defs’ are grouped together in a triangulatory process based on the unconstructed similarity between the elements of \{d,e,f\} (hence, by positive classification), then there really are (intrinsically bounded, noumenal) ‘defs’. Apoha as negative classification avoids this conclusion without ending up at the other extreme – according to Dharmakīrti, at least some, but not necessarily all extensional and object boundaries are purely conventional (saṃketa).

There seems, moreover, no prima facie reason for Davidson to object to the apoha-ic extension (or insertion) suggested in the previous section – apoha-ic construction does not assume sense data (Dharmakīrti rejected that notion; see section 2), which is Davidson’s ground for objection to conceptual construction. Rather in the contrary, the above interpretation of the apoha-ic notion ‘not non-...’ not only coheres with Davidson’s holism and rejection of meanings (universals; see sections 3 and 4), but implies those, and in that way strengthens the ties between triangulation and other core elements of Davidson’s philosophical system (holism/coherentism particularly).

Let us return to the question that opened this paper: If two communicating beings are both aware of a certain black queen chess piece as evidenced by their reference(s) to that black queen chess piece (in their communication), does that imply that there ‘really’ is (something that is) that black queen chess piece? The two communicating beings each ‘have’ a type (concept) ‘black queen chess piece’, such that these types overlap and the shared token is a token of that type-overlap. The shared-ness of that token implies that the indeterminate appearance (‘sensory experience’ or pratibhāsa) classified as a ‘black queen chess piece’ is (ultimately) noumenally caused. Whether that implies that there ‘really’ is (something that is) that black queen chess piece depends on the nature of the boundaries of type(s) and token(s), however. The questionned implication can be split up in accordance with the distinction between those two kinds of boundaries: (1) is there ‘really’ (something that is) that, or: is there ‘really’ some singular and discrete noumenal ‘thing’ that corresponds or coincides with that (singular and discrete) black queen chess piece; and (2) is that something a black queen chess piece (or at least, can that something be meaningfully be considered to be a black queen chess piece)? The first sub-question concerns the nature of object (token) boundaries; the second of extensional (type) boundaries.
An intrinsic object boundary implies that there is a singular and discrete noumenal ‘thing’, ‘object’ or ‘entity’ (if such terms apply to the noumenal) causing and coinciding with the appearance that is classified as a black queen chess piece. A conventional object boundary, on the other hand, does not have such an implication – there may be a singular and discrete noumenal cause, but it may also be the case that more or less arbitrary boundaries are drawn in a ‘landscape’ of non-discrete and heterogeneous noumenal cause(s) (as in the case of the ‘gu’ example in section 4). Consequently, the answers to the first sub-question (is there ‘really’ some singular and discrete noumenal ‘thing’ that corresponds or coincides with that (singular and discrete) black queen chess piece?) are “yes” if one – like Davidson – assumes the intrinsic object boundaries that come with ‘non-constructivism’, and “no, not necessarily” if one – like Dharmakīrti – assumes that at least some object boundaries are conventional (and that this may be a case of such conventional boundaries).

The second sub-question concerns extensional boundaries. An intrinsic extensional boundary means that for each of the tokens \(x_1, x_2, \ldots, x_i\) of a certain type \(X\) there is a noumenal cause \(nc_{x_i}\), such that these noumenal causes \(nc_{x_1}, nc_{x_2}, \ldots, nc_{x_i}\) are identical, or at least very similar, to each other in the relevant (x-token-causing) respect, and (significantly) different (in the same respect) from noumenal causes of phenomena that are not tokens of \(X\). In other words, the clear and unambiguous phenomenal boundary between \(X\) and non-\(X\) coincides with the equally clear and unambiguous noumenal boundary that causes it. And if that is the case, and there are only innocent intermediaries in between the phenomena and their noumenal causes (see section 2), then there seems little reason to not identify the noumenal causes of black queen chess pieces as black queen chess pieces – the object of a belief is the cause of that belief (Davidson 1983, see the quote in the introduction).

A conventional extensional boundary, on the other hand, denies the necessary existence of a noumenal class boundary coinciding with the phenomenal boundary and only grants that the noumenal causes of the appearances that in their (linguistic, situational, etc.) contexts are phenomenalized as tokens \(x\) of a certain type \(X\) are insufficiently different to be classified differently (are not non-\(X\); see section 4). There is a noumenal basis for the classification, but boundaries are drawn by convention, are context dependent, and (potentially) conceal real differences. Hence, if there is a class of noumenal causes of tokens \(x\) of type \(X\), than that class ‘exists’ only by virtue of the phenomenal distinction between \(X\) and non-\(X\), and because the noumenal was defined above (see section 2) as non-phenomenal, as independent from our (phenomenal) experience, there is no purely noumenal (ly determined) class corresponding with (phenomenal) \(X\).

Consequently, the answers to the second question (can the cause of the phenomenal black queen chess piece be meaningfully considered to be a black queen chess piece) are “yes” if one – like Davidson – assumes the intrinsic extensional boundaries associated with positive classification, and “no” if one – like Dharmakīrti – assumes that extensional boundaries are conventional. Combining the answers to the two sub-questions results in two answers to the original question: Yes, the shared experience of a black queen chess piece implies there really is a black queen chess
piece. Or: No, that shared experience merely implies that it is noumenally caused, but its specific phenomenal (or conceptual) effect is or may be largely due to (shared) convention.


Either by means of apoha or triangulation, words and phenomena are grounded in the noumenal world through a conjunction of a number of conditions – there must be at least two agents that have the ability to communicate, and that are perceiving the same ‘things’ and the same similarities and differences between those ‘things’, and at least one of those agents uses similar signs in the presence of (or to refer to) perceived to be similar objects. Where apoha and triangulation fundamentally differ is in the nature of classification of (relevant) experiences (of ‘things’) – while apoha as exclusion is negative and thus constructive, Davidson implicitly assumes a positive, non-constructive form of classification. Consequently, ‘def’ is not defined by difference from ‘non-def’, but by similarity between the things labeled as ‘def’, which suggests perceiving them as instantiations of ‘def’, as defs. In other words, contrary to apoha, (positive) classification based on similarity invites (but not justifies) the idea that having a concept ‘def’ implies that there are defs, and therefore, that language gives us more or less direct access to the (real) world. And it is exactly this illegitimate conclusion that Dignāga and Dharmakīrti wanted to avoid (e.g. Matilal 1971).

On the other hand, having a concept of ‘def’ does imply that there are shared indeterminate appearances (‘sensory experience’ or pratibhāsa) in which that concept is grounded, and the ultimate cause of such shared indeterminate appearances can only be noumenal (the distal stimulus). Furthermore, the rejection of a noumenal category corresponding with (phenomenal) ‘def’ because of the non-identity of its supposed members implies that we are or can become aware of that non-identity. And if that is the case, we can triangulate and conceptualize the difference(s). Perhaps every ‘thing’ that is classified as (not-non-) ‘def’ is uniquely different, and every ‘thing’ constructed as ‘gu’ is merely a conventional construction, and therefore, there really are no defs and gus, but we can be (or become) aware of those differences, conventions and constructions and describe those. Hence, conceptual categories are deceptive or illusory only to the extent of our (contingent) inability or unwillingness to ‘see’ beyond (ordinary) words. Aside from that contingency (or unwillingness), phenomenal perception can only be subjective, deceptive, or illusory in the minimal sense regarded ‘mere empirical accidence without philosophical significance’ by Davidson (see section 2).

In the previous section it was shown that his rejection of conceptual construction and implicit assumption of positive classification would lead Davidson to answer “yes” to the black chess queen question. That answer, however, does not depend thereon – an apoha-enriched Davidson would still answer “yes”. Davidson’s key point is that “successful communication proves the existence of a shared, and largely true, view of the world” (1977, 201), that shared phenomena are necessarily
noumenally caused (or that a shared stimulus is a distal stimulus, in more Davidsonian terminology), and that, given the innocence of causal (rather than epistemic) intermediaries, phenomena are non-deceptive (or only ‘deceptive’ in a very limited sense) and that we can in principle become aware of the ‘real nature’ of things. The abandonment of positive classification results in the recognition of the difference of (identically named) ‘things’, but the noumenal causation of those differences implies that those can be triangulated and conceptualized. The black chess queen piece may be a phenomenal construction, and there may thus not strictly speaking be a noumenal black chess queen, but the shared phenomenon’s cause does (noumenally) really have certain (triangulable) properties that make us perceive it as such (in the given context). That is enough to answer “yes”.

Davidson and Dharmakīrti largely shared the idea that words or concepts are (ultimately) noumenally grounded conventions, but differed in focus and attributed implication. While for Davidson the noumenal grounding of words (and thought) implied that they somehow stand for noumenally real things (ignoring the intricacies of Davidson’s thought about that relationship), thus downplaying their conventionality, it is that conventionality which Dharmakīrti stressed, and which for him implied that conceptual categories are (noumenally) unreal. The difference seems largely ideological – Davidson wanted to prove ‘unmediated touch’ with noumenal objects as much as it was an indubitable premise for Dharmakīrti that the phenomenal or conventional world is an illusion. This ‘ideological difference’ reveals a difference in purpose (of the argument), and a difference between the questions Dharmakīrti and Davidson intended to answer. Because of that difference, for Davidson the question about the black queen chess piece would be a question about necessary noumenal causes, about subjectivity and about the possibility of deception by the senses; for Dharmakīrti it would be a question about the nature of things independent from our experience. Hence, the different answers: “no, the black queen chess piece is not (really) a black queen chess piece independent of our experience, but yes, there is a noumenally real cause of our shared experience of that black queen chess piece, and that experience is no deception”. These answers, however, are complementary rather than contradictory.

These answers are also somewhat unsatisfactory because they remain two separate answers. By integrating apoha and triangulation, section 4 transcended the apparent contradiction between Davidson’s argument that our (phenomenal) beliefs about the (noumenal) world cannot be massively wrong, and Dharmakīrti’s belief in

---

17 That concepts are conventional, as argued here, does not imply that such – or other – conventions are necessary for communication, which Davidson rejected in (1982c) and (1986) (and which also more or less follows from the theory of triangulation itself). Concepts are conventions in the sense that they are created in a social process involving shared objects of beliefs (the real ‘things’ referred to); they are conventions between the participants in triangulation (hence, involving those shared objects). Davidson’s critique of conventions in (1982c) and (1986) rejects the idea that language and communication necessarily involve more or less fixed conventions between all speakers of a language, that social conventions are constitutive of meaning. It is a rejection of social externalism about mental content (see also 1990a) in favor of an externalism that is both social and physical, i.e. triangulation.
the (ultimate) illusoriness of the phenomenal world. The key point of that integration is that our grounding of words and phenomena in the (noumenal) world may be imperfect in the sense that words (may) conceal real differences, but that a shared noumenal world implies that those differences can be triangulated and conceptualized, and therefore, that conceptual categories are deceptive or illusory only to the extent of our (contingent) inability or unwillingness to ‘see’ beyond (ordinary) words.

This metaphor of ‘seeing beyond words’ is a central theme in many currents of Chinese (and Japanese) Buddhist philosophy (most famously, Chán/Zen), which mostly rejected metaphysical dualism (but not the epistemological dualism at the core of Buddhist thought). Such ‘seeing beyond words’ is often interpreted as conscious, non-conceptual perception (or experience), which was rejected in section 2, but whether this interpretation is correct is not the issue here. Above, the bracketed word ‘ordinary’ was added: ‘seeing beyond (ordinary) words’. The point here (and possibly in (some) Chinese/Japanese Buddhist philosophy as well) is not the impossible goal of conscious, non-conceptual experience, but a critical examination of the categories given in our (ordinary) language. ‘Ordinary’ is again bracketed here, because the construction of arbitrary conventional boundaries (as in both examples in section 4) is not necessarily limited to ordinary concepts. However, such ‘critical examination’ necessarily takes place in language – determination of non-identity of the members of a conceptual class requires triangulation and conceptualization or description. Hence, ‘seeing beyond (ordinary) words’ extends rather than rejects language, and it is still ‘seeing through words’, but ‘seeing through’ in both senses of ‘through’.

‘Seeing beyond (ordinary) words’ denotes a ‘middle path’ between two conflicting dogmas: the illusoriness of conventional (phenomenal) reality (and the associated rejection of the categories of (ordinary) language as noumenally true) in Buddhist philosophy, and the authority of ordinary language in analytic philosophy: between the Buddhist wholesale rejection of language as a guide to the (noumenally) real world, and the common ‘analytic’ belief in ordinary language as truly and objectively representing reality. (Ordinary) language is to some extent deceptive, but only to some extent, and the necessary (social) grounding of words in the real world through *apoha* or triangulation also implies the possibility of uncovering ‘deception’: of uncovering the conventional nature of (at least some) conceptual (class) boundaries and their real grounds.

The *apoha*-triangulation integration suggested in this paper (see sections 4 and 5) neither guarantees the identity of phenomena and noumena as assumed in metaphysical non-dualism, nor implies the dualist conception of the phenomenal -

---

18 Most of Buddhist philosophy involved the claim that phenomenal (conventional) reality is illusory, but not all. And most of analytic philosophy believes in the authority of ordinary language, not all. Hence, these two ‘dogma’s’ are not universal.

19 This leaves metaphysics with the question of how to determine what is conventional/constructed, and what is (noumenally) real. If at least some, but not (necessarily) all, boundaries are conventional, then how can one know which boundaries are intrinsic and which are conventional, or which phenomena are constructed and which coincide with discrete (intrinsically bounded) noumena, hence ‘are real’?
noumenal divide as an unbridgeable chasm, but rather suggests another ‘middle path’—a middle path between dualism and non-dualism. There is a gap between the noumenal and the phenomenal, but it is not an unbridgeable gap. Communication (triangulation) continuously creates bridges, but also enables us to build further bridges, and that possibility of building bridges means that the gap, rather than irrelevant because it is (believed to be) either non-existent or too wide, is an invitation for further inquiry.

ACKNOWLEDGMENTS

This paper resulted from a research project supported by the Japan Society for the Promotion of Science (JSPS). I would like to thank the three anonymous referees of *Comparative Philosophy* for their insightful comments and suggestions on an earlier version of this paper.

REFERENCES


