

Defense of Putnam's Twin Earth Experiment

Yiliang Wu

School of Fujian Normal University, Fuzhou, China

*Corresponding author: 102062022064@student.fjnu.edu.cn

Abstract:

This essay reviews and examines the controversies surrounding Searle's rebuttal sparked by Putnam's Twin Earth experiment. The article firstly analyzes Putnam's definition of natural kind terms and points out that there is a difference between natural kind terms and descriptivism, which cannot be explained in equal quantities. From the perspective of the Twin Earth experiment, understanding Putnam's definition and interpretation of natural class terms can be the key to refuting Searle's rebuttal, which can prove that the externalist stance of the Twin Earth experiment has not been severely undermined. On this basis, this article provides multiple examples to further explore and demonstrate that even with the use of index words that are consistent with Searle's explanatory approach, the final conclusion is not contradictory to semantic externalism and can still point to the perspective of externalism. At the end of the article, the author attempts to achieve a more reasonable balance in semantic interpretation by explaining the connection between semantic internalist and externalist elements.

Keywords: Twin Earth; natural kind terms; indexicals

1. Introduction

The issue of meaning is fundamental in the philosophy of language, and the debate between internalism and externalism of semantics has never ceased since its inception. Hilary Putnam explicitly proposed the semantic externalism viewpoint that "meaning just ain't in the head" through the Twin Earth experiment in his article "The Meaning of Meaning" in 1975. This thought experiment generated tremendous stimulation and triggered a great response while also giving rise to a large number of thoughts and refutations about the Twin Earth experiment. This article will first introduce Putnam's thought experiment on the twin earth and then mainly focus on the refutation that has had a significant impact on the academic community -- John Searle's refutation of the thought experiment -- to defend Putnam's semantic externalism position. Finally, the essay will introduce the debate between semantic externalism and internalism through the Twin Earth experiment.

In the article "The Meaning of Meaning", Putnam hypothesized that in the 1970s, when chemical science was insufficient to understand the microscopic structure of matter, there was a planet very similar to Earth, which could be called the Twin Earth. On the Twin Earth, everything except for what the Twin Earthers called "water" had a one-to-one molecular correspondence with things on Earth. What the Twin Earthers called "water" was a substance

with a molecular formula that could be represented by XYZ, which had the same or extremely similar surface characteristics and functions as the substance H₂O that constituted what the Earthlings called "water", except for the different internal microscopic structure. Just as there is no XYZ on Earth, there is no H₂O on the Twin Earth. Therefore, when the Earthlings use the word "water", the semantic meaning of "water" points to H₂O; when the twin Earthers use the word "water", the semantic meaning of "water" points to XYZ. From the conditions of this thought experiment, it can be seen that the brains of Earthlings and twin Earthlings are the same (even though due to the difference between H₂O and XYZ, the water in the brains of Earthlings and twin Earthlings is different, but both the influence of H₂O on the brains of Earthlings and the influence of XYZ on the brains of twin Earthlings are consistent, and everything else remains molecularly consistent, so they can be regarded as being similar), that is, their minds are similar. However, the substances referred to by using the word "water" are different things. Putnam thus concluded that "meaning is not in the mind." [1]. The same psychological state can correspond to different references [2].

In explaining the Twin Earth experiment, Putnam proposed the concept of "natural kind terms", which can be roughly equated to Kripke's definition of proper nouns. When people say "water", they refer to a specific thing in this world, just like parents may fix the reference to "Jo-

seph Haydn” by declaring, “By this name, we mean this child in this world”[3]. It can be seen that the reference to natural class words needs to involve the real conditions of the external world. Therefore, Putnam proposed that the traditional internalism theory at least has loopholes, which stem from its cognitive limitation on the methodology of determining semantics, which is solely strengthened by the role of the mind. He used this thought experiment mainly to break the long-term stable rule of internalism on semantics. This move certainly aroused a lot of rebuttals, among which the most influential one was Searle’s rebuttal to Putnam. The discussion will be launched from here.

2. Analysis

2.1 The Debate between Searle and Putnam on Twin Earth

Regarding the experiment on the idea of the Twin Earths, Searle proposed his refutation: firstly, the traditional idea of having a list of features, where each word in language is associated with some features, has not been abandoned but is reflected as an index definition in Putnam’s Twin Earth experiment. Indexing can be understood as individuals receiving stimuli from external objects, utilizing a series of impressions in their minds, namely the feature list mentioned earlier, to query and define the properties of the external object quickly. Due to the limitations of the Twin Earth experiment on the chemical development of both Earthlings and Twin Earthlings, both of them can’t determine the microstructure of H₂O and XYZ. The cognitive path of Earthlings and Twin Earthlings is only composed of macroscopic observations and experiences, and the list of features in their minds does not contain relevant information on the microstructure of H₂O and XYZ. Therefore, in the context of the Twin Earth experiment, there is a proposition that “water is any substance that is structurally equivalent to something that can precisely cause such a visual experience.”[4]. At the same time, it can also be proven that both Earthlings and twin Earthlings have set a condition in their minds when referring to “water” that can be satisfied by anything similar to the substance in terms of relevance, that is, an index definition or index condition. This is also one of the conditions under which Earthlings and twin Earthlings can refer to “water” on their respective planets. Based on the above judgment, Searle believes that the reason why Putnam can conclude that “meaning just ain’t in the head” is because Putnam believes that the extension is determined by the microstructure rather than the index definition in the mind. To better refute Putnam, Searle gave the example of “the person who killed Brown” to illustrate, believing that intentional content determines an extension for even

those who do not know who the killer is, and the true culprit falls under this extension [5]. Overall, Searle believes that regarding the “water” of the Twin Earth experiment, the setting of its index conditions depends on the mind, and whether the external object meets these conditions depends on the world outside the individual. The inherent logic of the Twin Earth experiment essentially does not negate the internalist argument of the index method, so it cannot stand firm on the position of externalism.

Saier’s rebuttal affirms the potential consensus on the definition of indexing in the Earth and twin Earth systems and proposes that the setting of indexing conditions is inevitably dominated by the mind. He believes that meaning and intentionality exist in a pre-linguistic form before language is expressed. What can be agreed is that in the use of language, meaning does exist in the mind before language. The perceptible features of external objects stimulate the brain to judge whether it can fall on the extension of a certain connotation based on the so-called “semantic list” and whether it can be expressed as a word on this basis. At this point, whether a certain language can be used to refer to an external object depends on whether the external object meets the indexing conditions of the language.

However, this process should apply to explaining the use of language, but it is not as applicable to determining the meaning of language words. Moreover, from the perspective of the Twin Earth experiment, Searle’s refutation is essentially based on descriptive theory, and he seems to have not truly grasped the connotation of Putnam’s “natural kind terms”. Gary Kemp further explains Putnam’s natural kind terms, stating that when people use natural kind terms, it refers to a sample of the kind and intends that the word shall denote whatever is the same kind as the sample [3]. Natural kind terms are similar to the proper names described by Kripke in a broad sense. The references to different natural kind terms are mutually exclusive in the same world, and each is used exclusively to describe something. Kripke distinguishes between the contingent truth (which is true in some possible world) and inevitable propositions and indicates that the descriptive view on the reference of proper names seems to be incorrect [3]. Firstly, define a concept: A proposition is necessary if it is true in every possible world; a proposition is possible if it is true in some possible world [3]. Subsequently, Kripke hypothesized that Dante plagiarized and killed the true author of *The Divine Comedy*, Adriano, and deceived everyone, so in fact, the proposition “The Creator of *The Divine Comedy*” cannot be used to describe or refer to “Dante.” Of course, this is possible and contingent; that is, a world may exist where these conditions are true. Using descriptive methods, referring to proper names depends

on whether the referent can be selected uniquely, and there are different possible situations in different worlds, so such descriptions cannot exist. Gary Kemp integrated David Kaplan's viewpoint and proposed that proper names and proper names are directly referring expressions, "they do not refer using some kind of conceptual representation of the object such as a Fregean sense (a mode of presentation)[3]". Similarly, like proper nouns, natural kind terms are not referred to through some conceptual representation of the object. In the Twin Earth experiment, Putnam indicated that the microstructure of H₂O is the essence that distinguishes water on Earth from water on the twin Earth. Searle's refutation using the index method seems to have not touched upon the essence. Because under the experimental system of Twin Earth experiment, the natural kind term "water" has determined that its reference is based on its microscopic essence (because H₂O and XYZ are not different in other aspects), in a world people can say that "water" is an example of "something flowing in a river (the river is in this world, too.)" in language use, which does not conflict much with Searle's explanatory path of language use. However, it cannot be assumed that "water" and "something flowing in a river [3]" are synonymous. If Earthlings use "things flowing in rivers" to describe "water" because they rely on the external world to make decisive judgments, where what flows in rivers is actually H₂O, then within the scope of this Earth world, "water" can be used to describe H₂O. In the twin Earth world, according to the above logic, "water" will be used to describe XYZ due to external influences. So, according to the explanation of Putnam's natural class words, Searle's rebuttal is not enough to prove the failure of externalism in this thought experiment.

2.2 Indexing and Semantic Externalism are not Incompatible

The previous text has proven that based on the definition of "natural kind terms", we can distinguish them from the descriptive indexing method. In this section, this article will continue to prove that even the indexing method is essentially not incompatible with the semantic externalism position demonstrated by the Putnam Twin Earth experiment.

After refuting Putnam's Twin Earth experiment in an indexing sense, Searle created a scenario with semantic referential specificity similar to Putnam's Twin Earth experiment based on the idea of Twin Earth: assuming there is a couple of Jones and Sally on Earth, and there is also a pair of Jones and Sally on Twin Earth who are exactly the same as the Jones couple on Earth, they can be called Twin Jones and Twin Sally. When Jones said "Sally," he referred to Sally instead of the twin Sally, which is related

to his previous and current external environmental factors. Assuming that Sally and the twin Sally both work for the space agencies of their respective planets, after completing a space mission, Sally comes to the twin Earth due to some mistake. The twin Sally comes to Earth, but since the Earth and the twin Earth are the same, they have not found any anomalies. In this case, Jones inevitably referred to the twin Sally as "Sally," when he said "Sally," he meant Earth Sally. This can also be explained by the index method. Jones undoubtedly made a mistake because he had an index connotation about Sally in his mind, but he made a perceptual mistake when using indexes to judge real objects [6]. However, it can also be concluded that meaning is determined internally, and meaning cannot be regarded as the result of external factors because changes in external factors have not changed the semantics of "Sally". However, in fact, in Searle's case, the importance of the external world's formation of the connotation of "Sally" in Jones's mind seemed to have been overlooked before the exchange between the two Sally.

The intentional content generated before language expression is determined by index labels, obtained by the subject from the external world through external causal effects using specific perspectives and viewpoints as tools. The key is that subjectivity is reflected in the personalized shaping of the external world in the brain (where "shaping" is only used as a descriptive term), rather than being reflected as the only material that constitutes personalized labels. Horowitz also has some objections to Searle's refutation of internalism. He indicates that Searle simplified the concept of "intension determines extension"[7]. The uncertainty of the microstructure of two types of water under the Twin Earth experimental system leads to the proposition that the connotation of "identical in structure with this stuff" is not enough to truly determine the extension. He indicates, "What completes that determination in the TE case are extra-mental facts [8]". Frege believes that people do not always have complete and comprehensive knowledge about referents, and not all inner sensations have external counterparts as their referents [9]. This also demonstrates that meaning is not simply derived from the mind. The explanatory path of Dretske's information semantics may also help. He proposed that semantic information should be transformed through structural information, and the formation of structural information cannot be separated from the subject's acceptance of the influence of the external world and the cognitive activities of learning from the external world [10]. Similarly, let's assume that we stand from a third-party perspective (neither an Earthling perspective nor a twin Earthling perspective) that clearly recognizes the difference between H₂O and XYZ, independently referred to as "water" by two worlds.

We can see that due to a lack of understanding of microstructure, the conditions set by Earthlings for their “water,” also known as index conditions (which does not include an understanding of the microstructure of water), are no different from the index conditions set by twin Earthlings for their “water.” However, their minds do not rely on the same substance H₂O and XYZ, when determining index conditions and meanings. This determines that when we stand from a clear and omniscient perspective on this matter, when people from both worlds say the word “water”, although there is a principle of relying on the mind to achieve the correspondence between external objects and index conditions, after all, the objective facts and the practical basis formed by index conditions indicate that the meaning of “water” in the two worlds is not the same. Horowitz said, “descriptivism does not entail internalism”[8], so even acknowledging the existence of indexing does not necessarily point to internalism and exclude externalism. Through layer-by-layer analysis, it can be seen that both externalist and internalist elements determine the indexing conditions in the indexing method determine the indexing conditions in the indexing method. Therefore, it can be considered that the refutation of Searle’s semantic internalism cannot effectively undermine the idea that “meaning just ain’t in the head” in the Twin Earth experiment.

3. Summary

Above all, Putnam’s definition of natural kind terms is sufficient to counter Searle’s refutation of the Twin Earth experiment, and it can be seen from this that the influence of externalist elements on semantics is indeed an open and comprehensive perspective. Continuing Searle’s explanation of indexing methods, this article argues that indexing methods fundamentally cannot threaten Putnam’s semantic externalism perspective. By analyzing and refuting, it can be seen that behind the close connection between the demonstration path of Searle’s internalism and the construction of the mind, the influence of the external world on semantics is crucial and cannot be ignored. This article does not believe that based on externalism necessarily leads to “the same environment, the same psychological

content”, but equally acknowledges the personalized role played by the mind and recognizes the role of psychological connotation in shaping meaning (just as described in the previous text). Based on this article, there is a possibility of continuing to infer the accompanying relationship between semantics and externalist elements, and it is possible to clarify further the impact of changes in external factors on semantics, which is indeed present and not completely hidden in the shadow of semantic internalist elements. The exploration of semantic externalism has important value and significance for studying language philosophy. In the future, further exploration and extension can be made to clarify and deeply explain the relationship between semantic externalism factors and semantic internalism factors.

References

- [1]Hilary Putnam.The meaning of ‘meaning’, in *Mind,Language and Reality:Philosophical papers*, Volume 2.Cambridge, UK: Cambridge University Press,1975,pp.215-271
- [2]He Dan, Chen Qun. Meaning in the Mind - A Brief Discussion on Putnam’s Semantic Externalism. *Research on Dialectics of Nature*. 2009 (25). 4-8
- [3]Gary Kemp.What is this thing called Philosophy of Language. Routledge .2024
- [4]John Searle. *Mind: A Brief Introduction*.Oxford University Press.2004
- [5]John Searle. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press, Cambridge.1983
- [6]Chen Gang. On Searle’s Intrinsic Language Philosophy. *Philosophical Research*.2003,(10):77-83
- [7]Camap R. *Meaning and Necessity*. Chicago University Press, Chicago.1947
- [8]Horowitz. Putnam, Searle, and Externalism. *Philosophical Studies*.1996,(81):27-69.
- [9]Frege, Gottlob.(1892/1970).“On Sense and Reference”,translated by Max Black. Geach and Black ,1970,pp.57-58
- [10]Wang Shipeng. Interpretation and Debate on the Information Semantics of the Twin Earth experiment. *Research on Dialectics of Nature*.2011,(8):7-11