

The dissertation lays the foundations for a naturalistic account of word meaning capable of addressing the conflicting intuitions that word meanings are both “out there”, world-involving and objective (the “objectivist” intuition) and in the heads of speakers, i.e., cognitive and perspectival (the “mentalist” intuition). The strong naturalization constraint endorsed in this project has it that the sought-after syncretic notion of word meaning must be nonmysterious and constitute a potential object for the natural sciences. The objectivist intuition is discussed within the framework of semantic externalism and the theory of semantic deference. Whereas the importance of the phenomenon of semantic deference (i.e., the fact that speakers defer to semantic standards for the fixation of the meaning of their words) is recognized, it is shown that taking the normativity of meaning evidenced by semantic deference at face value leads to embracing a form of meaning objectivism that is incompatible with naturalism. On the one hand, the objectivist/externalist commitment to independent meaning-determining realities could be stronger than the commitments actually undertaken by the natural sciences themselves. On the other hand, the degree of idealization inherent in the objectivist account makes it oddly disconnected from and ultimately irrelevant to actual linguistic practice. However, usage-based accounts – which have meanings determined by the way words are actually used rather than determined “from outside” – notoriously struggle to provide a satisfactory account of the normativity of meaning. The proposed move consists in biting the bullet and treating the inherent normativity of meaning as a form of cognitive illusion, albeit an unavoidable illusion and one which must be taken seriously in order to explain the properties of linguistic understanding. A strictly usage-based account is shown to be viable and even to be able to account for the objectivist explanandum, once it is coupled with biological functionalism. Word meanings “out there” turn out to be viable natural objects, yet quite unlike the apparent objects of our pre-theoretical intuitions. “Complete”, world-involving word meanings are complex functional kinds (like organs or artifacts) constituted (rather than determined) by speakers’ actual dispositions and relevant environmental factors. As such, complete meanings – whether at the communal level (conventional meanings) or at the level of the individual speaker (idiosyncratic patterns of use) – are essentially opaque to speakers and can only be identified from a theoretical point of view on the basis of functional considerations. Moreover, the environmental factors intuitively corresponding to the traditional notion of objective reference or extension cannot be considered independently of the other internal and relational meaning-constitutive factors. The view of meaning defended is thus supportive of a certain form of anti-realism, where reference and truth are relativized to evolved interests, yet it is not supportive of any global form of anti-realism, for the presuppositions of biological normativity still provide

a realist anchor to natural-language meanings. From this theoretical perspective, the mentalist intuition is taken to concern the internal, cognitive sub-components of complete meanings. Internal meanings are the cognitive kinds associated with word types (lexical meanings) or word tokens (ways in which words are understood/interpreted on an occasion of use). It is argued that internal meanings – whether stable or occasion-specific – have an irreducible abstract dimension for which no naturalistically plausible worldly counterpart is to be found. The experience of aboutness of the concepts intuitively encoded and expressed by words is again to be treated as a cognitive illusion, on a par with the illusion of the inherent normativity of word meaning. However, the abstract nature of internal meanings explains some of the key properties of linguistic understanding – aboutness, compositionality, co-reference – without which productive thought and linguistic communication would be impossible. The proposed account thus makes room for compositional-extensional semantics and shared understanding, as long as these are fully internalized. The connection with the external components of complete meanings is indirect, mediated by procedures whose workings are to a large extent opaque to users. The main consequence of the proposed framework is the incommensurability of internal meaning and complete meaning, and therefore a rejection of the possibility of an articulation of internal meaning and complete meaning compatible with the commonsense view from which traditional accounts of semantic deference and semantic externalism are built.

Cette thèse jette les bases d'une théorie naturaliste de la signification des mots à même de rendre compte de deux intuitions en apparence conflictuelles : d'une part, l'intuition selon laquelle les significations des mots ont une existence extérieure objective et impliquent le monde (l'intuition « objectiviste ») ; d'autre part, l'intuition selon laquelle les significations sont dans la tête des locuteurs, c'est-à-dire correspondent à des réalités cognitives et perspectivales (l'intuition « mentaliste »). La contrainte naturaliste assumée dans ce projet veut que la notion synchrétique de signification que l'on cherche à développer puisse constituer un objet potentiel d'investigation pour les sciences naturelles, c'est-à-dire qu'elle soit, au moins en principe, localisable dans le monde naturel. L'intuition objectiviste est débattue dans le cadre de l'externalisme sémantique et de la théorie de la déférence sémantique. Bien que l'importance du phénomène de la déférence sémantique (le fait que les locuteurs défèrent à des standards sémantiques pour la fixation de la signification des mots qu'ils emploient) soit pleinement reconnue, l'argument poursuivi mène à la conclusion que la normativité de la signification que semble imposer la déférence sémantique ne doit pas être prise pour argent comptant, sous peine

d'épouser une forme d'objectivisme de la signification incompatible avec le projet de naturalisation stricte. D'une part, l'engagement ontologique objectiviste/externaliste vis-à-vis de réalités indépendantes déterminant les significations pourrait être plus fort que les engagements ontologiques implicites des sciences naturelles elles-mêmes. D'autre part, le degré d'idéalisation propre au point de vue objectiviste le rend étrangement détaché de la pratique linguistique effective, et en définitive sans pertinence pour rendre compte de celle-ci. Cela étant dit, les théories fondées sur l'usage – pour lesquelles les significations sont déterminées par la façon dont les mots sont effectivement employés plutôt que déterminées « de l'extérieur » – sont en général critiquées pour leur incapacité à rendre compte de la normativité de la signification. La proposition que fait la thèse consiste à assumer cette conséquence d'une théorie fondée sur l'usage et à considérer la normativité intrinsèque de la signification comme une forme d'illusion cognitive, bien qu'une illusion inévitable et devant être prise au sérieux s'il s'agit d'expliquer les propriétés remarquables de la compréhension linguistique. Une théorie strictement fondée sur l'usage est viable et même capable de rendre compte de l'intuition objectiviste, une fois que cette théorie est couplée avec un fonctionnalisme biologique. Les significations « extérieures » des mots sont bien des objets naturalisables, quoique fort différents des objets apparents de nos intuitions pré-théoriques. Les significations « complètes », c'est-à-dire impliquant le monde, correspondent à des espèces fonctionnelles complexes (à la manière des organes ou des artefacts) qui sont constituées (plutôt que déterminées) par les dispositions effectives des locuteurs et les facteurs environnementaux pertinents. En tant que telles, les significations complètes – que ce soit au niveau de la communauté linguistique (significations conventionnelles) ou au niveau du locuteur individuel (usages idiosyncrasiques) – sont fondamentalement opaques pour les locuteurs et ne peuvent être identifiées qu'à partir d'un point de vue théorique externe et sur base de considérations fonctionnelles. En outre, les facteurs environnementaux correspondant intuitivement à la notion traditionnelle de référence ou d'extension objective ne peuvent être considérés indépendamment des autres facteurs internes et relationnels constitutifs de la signification. La théorie de la signification défendue suggère donc une certaine forme d'anti-réalisme, dans lequel la référence et la vérité sont relativisées à des intérêts spécifiques produits par l'évolution naturelle. Cette théorie ne sert pour autant guère d'appui à un quelconque anti-réalisme global, car les présupposés de la normativité biologique continuent à fournir un ancrage réaliste aux significations linguistiques. Une fois cette perspective théorique sur les significations impliquant le monde adoptée, on fait la supposition que l'intuition mentaliste concerne les sous-composantes internes et cognitives des significations complètes. Les significations internes sont

les espèces cognitives associées avec les types lexicaux (significations lexicales) et avec les tokens lexicaux (façons dont les mots sont compris/interprétés lorsqu'ils sont employés). Il est avancé que les significations internes – qu'elles soient stables ou propres à une occasion d'usage – ont une composante abstraite irréductible à laquelle ne correspond aucune contrepartie mondaine acceptable d'un point de vue naturaliste. L'expérience de l'« être-à-propos » (*aboutness*) des concepts intuitivement encodés et exprimés par les mots doit encore une fois être considérée comme une illusion cognitive, à l'instar de l'illusion de la normativité intrinsèque de la signification. Cependant, la nature abstraite des significations internes explique certaines des propriétés centrales de la compréhension linguistique – être-à-propos, compositionnalité, co-référence – sans lesquelles la pensée productive et la communication proprement linguistique seraient impossibles. La théorie proposée fait donc une place à la sémantique compositionnelle-extensionnelle et à la compréhension partagée, pour autant que celles-ci soient complètement internalisées. La connexion avec les composantes externes des significations complètes est indirecte, médiée par des procédures dont le fonctionnement est en grande partie opaque aux utilisateurs du langage. La conséquence principale du cadre proposé est l'incommensurabilité de la signification interne et de la signification complète et, partant, le rejet de la possibilité d'une articulation entre les deux types de signification qui soit compatible avec le point de vue de sens commun à partir duquel sont construites les théories traditionnelles de la déférence sémantique et de l'externalisme sémantique.