Knowing opposites and formalising antonymy

Keith Begley – PhD, Member of the Trinity Plato Centre. Trinity College Dublin, the University of Dublin. College Green Dublin 2, D02 PN40, Ireland; e-mail: begleyk@tcd.ie

This paper discusses knowledge of opposites. In particular, attention is given to the linguistic notion of antonymy and how it represents oppositional relations that are commonly found in perception. The paper draws upon the long history of work on the formalisation of antonymy in linguistics and formal semantics, and also upon work on the perception of opposites in psychology, and an assessment is made of the main approaches. Treatments of these phenomena in linguistics and psychology posit that the principles of minimal difference and invariance are centrally important. It will be suggested that the standard approach employing meaning postulates fails to capture the relevant notion of antonymy, in part because it is not informed by these principles, and in part due to a number of other problems with this kind of approach, many of which may be overcome by building in the central principles from the beginning. The paper also discusses the issue of whether we can know that opposites necessarily exclude each other and, if so, how. This issue is intertwined with what is known as the colour incompatibility problem that Wittgenstein wrangled with at various times during his life. The paper assesses various solutions to these problems including an approach that was first put forward by Jerrold J. Katz. The relation between this approach and the theory of determinables and determinates is also examined. A further development upon this approach is proposed and then applied to the case of the formalisation of antonymy. It is argued that this approach avoids the problems suffered by the main approaches discussed earlier in the paper.

Keywords: Opposites, Antonymy, Semantics, Perception, Meaning Postulates, Jerrold Katz

Знание о противоположностях и формализация антонимии

Кит Бигли – доктор философии. Тринити-колледж, Университет Дублина. College Green Дублин 2, D02 PN40, Ирландия; e-mail: begleyk@tcd.ie

В этой статье обсуждается знание о противоположностях. В частности, рассматривается лингвистическое понятие антонимии и то, как оно представляет отношения противоположности, обычно встречающиеся в восприятии. Исследование опирается на длительную историю формализации антонимии в лингвистике и формальной семантике, а также на то, как восприятия противоположностей изучаются в психологии. Дается оценка основных подходов к проблеме. Лингвистический и психологический анализ восприятия противоположностей помещают в центр принципы минимального различия и инвариантности. В статье делается предположение о том, что использующий постулаты значения стандартный подход не улавливает релевантное понятие антонимии отчасти потому, что он не основан на этих принципах, отчасти же из-за ряда иных затруднений, многие из которых можно преодолеть, если

основываться на принципах минимального различия и инвариантности изначально. В статье также обсуждается вопрос о том, можем ли мы знать, что противоположности с необходимостью исключают друг друга, и если да, то как. Эта проблема тесно связана с так называемой проблемой несовместимости цветов, к которой на протяжении своей жизни не раз обращался Витгенштейн. В статье характеризуются различные решения этих проблем, включая подход, предложенный Джеральдом Дж. Катцем. Также исследуется связь между этим подходом и теорией определяемых и определяющих. Предлагается некоторое развитие идеи Катца, которое затем применяется к формализации явления антонимии. Приводятся аргументы в пользу того, что наш подход позволяет избежать затруднений, которые свойственны обсуждавшимся в статье стандартным подходам.

Ключевые слова: противоположности, антонимия, семантика, восприятие, постулаты значения, Джеральд Катц

1. Minimal Difference and Invariance of Opposites

'Antonymy' is the linguists' name for a paradigmatic sense relation of opposition between two lexemes. That is, it is the relation between the meanings of signs that are substitutable for each other in a syntactic structure, and which are considered to be opposites. According to Jones [2002, p. 9], the word 'antonymy' was originally coined in 1867 by C.J. Smith in the preface to his thesaurus of *Synonyms and Antonyms*.[1] Smith characterises it using an analogy to the word 'synonym' or 'synonymy'

---

[1] Contrary to what Jones reports, the word 'antonymy' does not appear in Smith's book, rather, the word 'antonym' does. Although this may have been the first avowedly technical usage in English, in the context of lexicography, a simple *Google Ngram* search reveals earlier usages of the word 'antonym' at least as early as 1860 in a manual for the instruction of fifth grade students in Milwaukee. Here we find 'antonym' used in a rather apt piece of pedagogical advice: "The force of a word may frequently be better understood by reference to its opposite or antonym than by the use of a synonym" [By-laws of the Board of Commissioners, Rules for the Government of the Schools, and Laws Relative to Common Schools in the City of Milwaukee, 1860, p. 37].

   The error and/or inaccuracy is now commonplace, and repeated often; the most recent example that I am aware of is from 2014 [Gao & Zheng, 2014, p. 234]. The source of the error seems likely to be the introduction to *Webster's New Dictionary of Synonyms* 1968, authored by Rose F. Egan [cf. Egan, 1984, p. 15a]. That is, despite Egan being careful to point out later [1984, p. 26a] that Smith, in his own words, "ventured, not to coin, but to reissue" [Smith, 1867, p. v]. Smith does not mention any prior usage of the term, rather, his qualificatory remark seems to be due to a possible usage of the word 'antonym' in grammar to refer to a pronoun, and Smith sees little harm in repurposing this "Greek word" given that the Latin terms are already current in grammatical contexts [1867, p. v–vi]. There are earlier uses of similar terms in other languages, such as the French 'antonymie' or 'contraire', and the German 'gegensatz' or 'gegenbigriff'.

(i.e., identity, or similarity between the meanings of words), and as a technical term for the sense relation between words that are 'opposite' to each other:

> Words which agree in expressing one or more characteristic ideas in common he [the author] has regarded as Synonyms, those which negative one or more such ideas he has called Antonyms [Smith, 1867, p. v, Preface].

This was a very loose description of the relation for lexicographical purposes, which is somewhat inaccurate and naive in the sense that it implies that pairs of antonyms are maximally opposed to, or separated from each other. Indeed, this is often how we conceive of opposites pre-theoretically, as being maximally opposed.

However, if we consider any pair of antonyms, for example, 'hot' and 'cold', it is clear that they are both kinds of temperature, and share everything about their meaning in common except along a single dimension, a scale on which they occupy opposite poles. This property of antonyms has been recognised, by linguists such as Cruse, to hold a certain fascination for us, given that they are both semantically close to, and distant from, each other. Cruse called it the 'unique fascination' of opposites:

> Opposites possess a unique fascination, and exhibit properties which may appear paradoxical. Take, for instance, the simultaneous closeness, and distance from one another, of opposites. The meanings of a pair of opposites are felt intuitively to be maximally separated. Indeed, there is a widespread idea that the power of uniting or reconciling opposites is a magical one, an attribute of the Deity, or a property of states of mind brought about by profound meditation, and so on. The closeness of opposites, on the other hand, manifests itself, for instance, in the fact that the members of a pair have almost identical distributions, that is to say, very similar possibilities of normal and abnormal occurrence. It is also reflected in the frequency of speech errors in which the intended word is substituted by its opposite [Cruse, 1986, p. 197].

Jones notes two examples of language acquisition that may also provide evidence of what Cruse calls our 'unique fascination' with antonymy. Firstly, he says that: "It has been widely documented that children tend to grasp the concept of oppositeness at a very early age, often learning antonyms in pairs rather than as single items." Secondly, he says that: "we are drawn to 'opposites' when learning a new language and feel more comfortable with the precise meaning of a word in our native tongue if we are familiar with its corresponding antonym" [Jones, 2002, p. 3]. Further, Cruse also points to the importance of opposites throughout the history of philosophy, in the continuation of the above quotation regarding the 'unique fascination' of opposites:

> Philosophers and others from Heraclitus to Jung have noted the tendency of things to slip into their opposite states; and many have remarked

on the thin dividing line between love and hate, genius and madness, etc. The paradox of simultaneous difference and similarity is partly resolved by the fact that opposites typically differ along only one dimension of meaning: in respect of all other features they are identical, hence their semantic closeness; along the dimension of difference, they occupy opposing poles, hence the feeling of difference [Cruse, 1986, p. 197].

In the linguistic treatment of antonymy there is a principle called the *minimal difference* between antonyms. This principle states that the 'best' antonyms, the ones that are recognised as such by most competent speakers of the language, differ with regard to only a single dimension of sense, while they are similar in respect of having all or most other senses in common. In the linguistic literature, this principle is apparent at least as early as 1970 [Clark, 1970, p. 275], and is commonly relied upon in most contemporary literature on the subject.

This property of semantic closeness or minimal difference is important not least because it allows for the regularity and, consequently, the generalisation of antonymous sense relations. If antonymy were instead some kind of relation of maximal difference, or unrelatedness, it would be non-recurring in a lexical system. Antonymy can be seen as a kind of relation that obtains between many pairs or sets of terms, but there are many other relations between terms that are not amenable to being generalised across the entire language. That is, they are non-recurring and for this reason are usually not named. In order to emphasise this point, Cruse names the relation between 'dog' and 'banana' (somewhat ironically) "dogbananonmy" [Cruse, 2000, p. 145]. This is not a relation that would hold between any other lexemes. It does not recur like antonymy or the other canonical sense relations do.

Consider, the question '*Hot* is to *Cold*, as *Alive* is to *what*?'; most competent speakers of the language will give the answer 'Dead', because they recognise the relation between 'Hot' and 'Cold' as being a recurring one and, pre-theoretically, as an opposition. Now consider a second question of the form '*Dog* is to *Banana*, as *Rose* is to *what*?'[2] There is no principled way to answer this question, because the sense relation of *dogbananonmy* is not a recurring one. That is, it does not hold between any other lexemes. An important consequence of these considerations is that we cannot describe antonymy simply in terms of logical negation, because this would not be enough on its own to individuate this relation. For example, purely denotatively, 'not-hot' includes 'cold' as well as 'dog', 'banana', 'rose', 'pig', 'green', and every other term besides 'hot'. This, of course, is not how natural language functions.

---

[2] Lyons used the example of 'rose' and 'pig' to make a similar point to that which was made by Cruse, but he did not take the further intentionally absurd step of coining a name for this sense relation [Lyons, 1977, p. 288].

An interesting analogue of the principle of *minimal difference* has also appeared as one of the results of recent psychological research into the visual perception of opposition. The researchers involved consider their own work to complement the investigation into the 'unique fascination' of opposites [Bianchi & Savardi, 2008, p. 11]. They investigated the perception of what they call 'contraries', a technical term they employ so as not to be confused with the linguistic study of antonyms [Ibid., p. 13]. They concluded that the perception of opposition between objects relies on there being invariance between the objects. That is, the objects must be similar enough with regard to most of their properties, in order to be perceived as opposites of each other.

The other side of this finding is that the recognition of global opposition, that is, opposition directly perceived without recourse to analysis, does not increase when more of the local properties of the perceived objects are different from, or opposed to, each other. In fact, this is more likely to be perceived as 'diversity' or 'difference' rather than opposition [Bianchi & Savardi, 2008, p. 149; Branchini et al., 2021, pp. 2–3]. This provides a perceptual analogue of Cruse's *dogbananonmy*, or non-recurring, non-canonical sense relations. If all we perceive between two objects is a collection of relations between dissimilar properties, then there will be no global opposition apparent. In contrast, the simple relation presented by, for example, two isosceles triangles pointing in opposite directions, immediately strikes one as an oppositional arrangement, with directional differences such as these being the most salient examples of this effect. So, there is a close parallel between linguistic and psychological phenomena corresponding to the *minimal difference* and invariance between oppositions in both of these domains, and the authors are keen to stress that "the role of opposites in cognition should not be relegated to formal logical matters such as those traditionally exemplified by 'the square of opposition'" [Branchini et al., 2021, p. 2][3]. As we shall see, this principle of minimal difference is entirely absent from attempts to formalise antonymy relations using predicate calculus.

---

[3] The authors cite the Béziau & Basti edited volume (2017), in which the topic of antonymy is briefly discussed by Bobenrieth [2017, p. 162f] (Note that Bobenrieth's name appears in reverse order, as 'B.M. Andrés'). He partially quotes an explanation by Katz of the relationship between contradiction between sentences or propositions, i.e., truth-bearers, and antonymy between non-truth-bearing constituent expressions or 'terms'. By restricting antonymy to a "domain of discourse", Bobenrieth is effectively also adopting a form of the principle of minimal difference, but the precise approach is not specified. (Note that Bobenrieth provides an incorrect citation for the work he is quoting from as Katz' *Philosophy of Language* (1966), when he is actually quoting from an excerpt of Chapter 2 of Katz' *Semantic Theory* (1972), printed in Margolis & Laurence [1999, p. 144ff].)

2. Formalising Antonymy
Using Predicate Calculus

In this section, I will discuss the problem of defining antonymy (semantic opposition) in terms of logical opposition. This is an issue that has been continually noted by a number of linguists. For example, Mettinger (1994) recognised that the problem, which has arisen in 20th century philosophy of language and linguistics, can be traced to its origins in ancient Greek philosophy:

> The question whether semantic opposition is a manifestation or function of logical opposition has given rise to a number of discussions […] with as many arguments in favour of a logical treatment of semantic opposition as against it. The problem seems to be founded in the dichotomy of "language" vs. "thought" and consequently, of "meaning (=Bedeutung)" vs. "concept (=Begriff)" dating back to Greek philosophy [Mettinger, 1994, p. 15].

Mettinger himself is on the side of those who are against defining semantic opposition in terms of logical opposition, and points to John Lyons' shift to the use of the gradable/ungradable distinction as a semantic basis for the distinction between kinds of antonymy [Ibid., pp. 16–17]. Jones et al. [2012] suggested that mere logical incompatibility is not enough to pick out only antonyms, because it picks out other incompatible lexemes also:

> One could define oppositeness in terms of logical incompatibility – that is, if a thing can be described by one of the members of an antonym pair, it is impossible for it to be described by the other. So, if a person is a *man,* he is not also a *woman.* If a piece of string is *long* with reference to some contextual standard, it cannot also be *short* with reference to the same standard. But logical incompatibility is an insufficient criterion for defining oppositeness, since many pairs of lexemes are semantically or logically incompatible, but this does not lead to their use as antonyms. So, while it is unlikely for something to be both a *limerick* and a *pencil*, this is not reason enough to think of *limerick* and *pencil* as opposites [Ibid., p. 3].

Later, and more directly, they state that:

> We question the validity of truth-conditional entailment (i.e. given the truth of p, [not p] is false) as an adequate method of confirming how people interpret antonymous meanings with negation in natural language [Ibid., p. 90].

According to this, logical opposition is, at least, insufficient for defining antonymy. Nevertheless, this practice has persisted in formal semantics, and it is commonplace to state schemata for the various kinds of

opposites in terms of what are known as *meaning postulates*.[4] These are universally quantified conditional or biconditional expressions given in a predicate calculus and used to state axioms for the extra-logical expressions of a language. In what follows, I will briefly discuss one approach that can be taken as a representative example of the formal methods of defining antonymy in terms of meaning postulates, which was presented by Cann (1993).[5]

Cann's definitions for each of the kinds of antonymy are given in terms of meaning postulate schemata. These schemata each present the form that a meaning postulate for a specific pair of predicates must take in order to be counted as an instance of the kind of antonymy being defined. Cann employs the symbols X' and Y' for the parameters of these schemata,[6] which would be replaced by the predicate constants representing antonyms in particular instances of these meaning postulates in a language.

Cann's most general definition, 'Opposites', names what is really just a kind of incompatibility relation between sets of things that are the extensions of expressions given in a predicate calculus. This is a very broad use of the term 'opposite' because it would also include relations between words that are merely incompatible with each other.[7] His definition is as follows:

> **Opposites**: X is an opposite of Y iff there is a meaning postulate relating X' and Y' of the form: $\forall x \, [X'(x) \rightarrow \sim Y'(x)]$ (the extension of X is distinct from that of Y) [Ibid., p. 220].

This broad definition spells out a necessary condition for most kinds of antonym. In each instance of antonymy, involving monadic predicates, the terms involved must be related in this way, if not in some other ways also. We will forgo an exposition of the polyadic case, but it is analogous. Further, as Cann says, "the different types may be distinguished by

---

4   This is a method that was introduced by Rudolf Carnap, but another important figure in this regard is the Polish philosopher Kazimierz Ajdukiewicz who employed *axiomatic meaning directives,* in a similar way to that in which Carnap employed meaning postulates [Woleński, 2016].

5   Cann's presentation is useful for our purposes primarily because he provides us with clear definitions of most of the kinds of antonymy.

6   The 'superscript prime' notation employed by Cann, which is the apostrophe: ', "indicates that the symbol is not a word in the object language, e.g. English, but an expression in the translation language […] For example, the English word forms *gives, give, gave, given,* are all associated with the citation form *give* and the translation of any of these word forms into [a logical language] is thus *give'*" [Cann, 1993, p. 35].

7   It is quite odd that he uses this terminology only two pages after also calling 'tulip' and 'rose' "opposites", in the sense of being "incompatible co-hyponyms" of 'flower' [Cann, 1993, pp. 218, 220], which is, of course, a much more restricted notion than that of mere incompatibility.

imposing further restrictions" [Cann, 1993, p. 220]. He defines complementary antonyms as follows:

> **Complementaries**: X is the complementary of Y iff there is a meaning postulate relating X' and Y' of the form: $\forall$x [(X'(x) → ~Y'(x)) & (~X'(x) → Y'(x))] (the extensions of X and Y are distinct and the complement of the extension of X is equal to the extension of Y) [Ibid.].

He then defines what he calls 'antonyms' (in the narrow sense of gradable opposites) as being "opposites restricted to a particular domain", as follows:

> **Antonyms**: X and Y are antonyms iff there is a meaning postulate relating X' and Y' of the form $\forall$x [$\forall$P[(X'(P))(x) → ~(Y'(P))(x)]] (the extensions of X and Y are distinct for some given domain) [Ibid., p. 221].

'*P*' here stands for some property of the object *x*. Cann uses the example of being an elephant. So, what is called 'big' cannot be called 'small' in the same respect, i.e., that of being an elephant. However, even a small elephant could be called 'big' in respect of being an animal. It should be fairly clear from my reconstruction above that '*P*' must at least name a property of the object. However, it is not clear that Cann's system can provide a mechanism for determining what this property should be in each case in which the postulate would be relied upon.[8] Cann also provides definitions for other kinds of oppositions including converses, however, we will not let these detain us here as they do not materially affect the focus of our discussion.

3. Katz' Arguments against Meaning Postulates

In 1977, Katz made the following sociological observation regarding the entrenchment of Carnap's use of predicate calculus as means of representing the semantics of natural languages:

> Carnap's theory is not only the current orthodoxy for philosophers who do not subscribe to Quinian skepticism about meaning but it is becoming the hottest bandwagon for linguists of all persuasions. To say, at the present time, that it is widely believed that semantic representations of sentences from natural languages are formulas in some predicate calculus would be extreme understatement. The achievements of modern logic in the study of implication in artificial languages are impressive, and philosophers and linguists, impressed with them, have based their ideas about semantic representation on such artificial languages. Thus, the assumption that semantic

---

[8] We will leave a discussion of this problem aside for present purposes. For a system that can provide such a determination, see Katz [1972, p. 255ff] who happens to discuss the very same example.

representations of sentences in natural languages ought to be written as for-
mulas of some predicate calculus is now a truism, serving theorists of other-
wise rival persuasions as a common means of characterizing the problem of
semantic description for natural languages in its most general form [Katz,
1977, p. 381].

This 'truism' is connected to the one that is still encountered to this
day when one is asked to consider the logical form of an argument in-
dependently of the meanings of the ('extra-logical') words contained
in the sentences of the argument. The situation is one in which, in addi-
tion to this, the semantic properties of, and relations between these words
are also represented solely in terms of the logical relations contained
in the axioms in which these words feature.

The method of defining semantic relations in terms of logical rela-
tions by way of meaning postulates has continued long after Katz' obser-
vation that this method is not entirely adequate for its purpose. This has
been noted by the linguist Lynne Murphy (2003):

While meaning postulates are regularly employed in model-theoretic se-
mantics, it has been repeatedly pointed out that the postulates explain
nothing about relations among meanings [e.g., Katz, 1972; Lakoff, 1972].
Since the only "meanings" involved are extensional sets, meaning postu-
lates essentially express relations among things (i.e., the denoted things in
the words' extensions) rather than among words or word senses. Further-
more, they simply assert those relations – they do not explain why certain
relations (and not others) hold between certain expressions [Murphy,
2003, p. 64].

Indeed, while it still appears to be the method favoured by linguists
working in formal semantics, and is used in textbooks, linguists working
outside of the formal setting complain of a difficulty in defining seman-
tic relations such as antonymy, and often merely repeat the logical dis-
tinctions between the different sub-categories of such relations without
the aid of a logical calculus. In the previous section, we discussed one ex-
ample of defining antonymy in terms of meaning postulates, which was
presented by Cann (1993). Let us now consider some decisive criticisms
proposed by Katz.

If we wish to represent a semantic entailment that is based on
the antonymy between the meanings of (extra-logical) words contained
in two sentences, this will not be possible in an unmodified predicate cal-
culus given that the representation could only take an invalid form. Con-
sider the example, 'The cat is dead' entails 'The cat is not alive', which
has the invalid form: 'Dc' entails '$\neg Ac$', where 'D' stands for 'Dead',
'A' for 'Alive', and 'c' for a particular cat. This can be represented as
a valid inference in the predicate calculus if we stipulate, as an axiom of
the language, the meaning postulate that represents the complementarity
between the meanings of 'Dead' and 'Alive', i.e., '$\forall x[(D(x) \rightarrow \neg A(x))$ &

$(\neg D(x) \rightarrow A(x))]$'. Given this meaning postulate, it is then possible to represent the entailment as an argument of the following form: 'Dc', '$\forall x(D(x) \rightarrow \neg A(x))$' entails '$\neg Ac$'. However, this meaning postulate approach does not explain *why* the meanings of 'Dead' and 'Alive' are antonymous and thereby lead to the validity of this representation as such, it merely stipulates that they are so related [Katz, 1972, p. 185]. It views them, as it were, from the outside via their external relations. Analogously, we might just as well give a numeral in place of each proposition above (*à la* propositional calculus) and argue that 'P_1' entails 'P_2' just as long as we postulate that '$P_1 \rightarrow P_2$' [Katz, 1977, p. 397; Katz, 1986, pp. 67–68], thereby conducting predicate logic without need of employing a predicate calculus, which is absurd. It is clear, extrapolating from the above example, that similar instances can be constructed for Cann's other definitions that rely on meaning postulates, and that they would be vulnerable to analogous argumentation.

Katz (1992) presents an argument that is independent to the one given above. He first provides the following two lists of antonyms: (i) "blind/having sight, orphaned/having a living parent, amorphous/having definite form"; (ii) "red/blue, happy/sad, odd/even". Katz points out that the antonyms in (i) are privative, that is, the sense of the first term is that of a lack of a feature included in the sense of the second term of each pair, whereas the antonyms in (ii) are not privative. Given that his original argument is quite clear and succinct I shall reproduce it in full below:

> [...] in so far as the meaning postulate approach accounts for the antonym pairs in both [i] and [ii] on the basis of postulates of the form "$(\forall x)(F(x) \rightarrow \neg G(x))$", it cannot explicate the fact that the pairs in [i] are privation/possession relations while those in [ii] are not. Given nothing more than an assignment of the same extensional structure to the antonym pairs in both cases, there is no means of accounting for the fact that the pairs in [i], but not those in [ii], are asymmetrical with respect to negation – one member of the former pairs, e.g., "blind", but not "having sight", being a privation term, is inherently negative but otherwise the same in content as the other term. To distinguish the privation/possession antonyms like [i] from antonyms like [ii] and to predict which member of the pairs in [i] is a privation term, it is necessary to refer to decompositional sense structure.
>
> The inexpressibility of the asymmetry is particularly clear when it is necessary, as it is in cases like the last pairs in [i] and [ii], to represent the antonyms with a biconditional. The equivalence of "$(\forall x)(\neg F(x) \leftrightarrow G(x))$" and "$(\forall x)(F(x) \leftrightarrow \neg G(x))$" thus leads to the false claim that antonyms of both kinds are symmetrical with respect to negation [Katz, 1992, p. 704].

The argument makes two decisive criticisms of meaning postulates: First, that a distinction between those kinds of antonymy relations that are privative and those that are not, cannot be made in terms of the usual meaning postulates, given that these relations must both be represented

by the same kind of postulate and, therefore, no difference can be represented in this regard. Second, in the case of the privative antonyms, a theory given in terms of meaning postulates will be unable to predict which of the two terms is negative (privative) and which is positive (possessive). Only a theory that can give an account of the decompositional sense structure of these terms and is able to reveal an underlying negative element will be able to do this.

There is another important difference between the meaning postulate approach to antonymy and the decompositional approach, which should be noted, but which seems to have gone largely unnoticed in the literature regarding the general discussion and comparison of these systems.[9] This is a difference that comes to the fore especially when antonymy is the kind of relation being considered. Although meaning postulates provide definitions of incompatibility relations between predicates in the language, these definitions do not themselves make reference to the predicates of the language that are superordinate to the opposed predicates. That is, subordination and its structure are treated as being extraneous to the definition of antonymy by the meaning postulate approach. They have nothing to contribute to the definitions of antonymy provided. However, we began our analysis by pointing out that such minimal difference or invariance between opposites, both in language and perception, has been identified as a central principle both by linguists and psychologists.

The linguists' intuitive conception of antonymy, as being a relation between 'incompatible co-hyponyms', is not captured by the meaning postulate approach as being an inherent feature of antonymy, but rather as a mere coincidence of relations based on postulates. Thus, there is a marked difference here between a meaning postulate approach and a decompositional approach such as the one employed by Katz, in which the sense structure that underlies antonymy is inherently bound-up with the sense structure that underlies hyponymy. That is, meaning postulates are only capable of capturing external relations between the extensions of their predicates, whereas a decompositional approach, in Katz' sense, aims to capture internal relations between the concepts themselves by positing underlying structures. It is clear that Cann later recognised this consequence of theories utilising meaning postulates, but considers it an advantage:

> A theory that utilises meaning postulates treats the meaning of words as atomic with their semantic relations specified directly. So, although traditional sense relations, both paradigmatic and syntagmatic, can easily be reconstructed in the system (see [Cann, 1993] for an attempt at this) they do not follow from the semantics of the words themselves [Cann, 2011, p. 476].

9　An exception can be found in the discussion of superordination in [Pitt, 1999, p. 146]. Pitt is a former student of Katz.

Cann argues that atomic theories are more parsimonious than decompositional theories, because the latter must postulate atoms in any event. However, this would miss the force of Katz' earlier argument regarding the choice of the level at which atoms are postulated, that is, the point at which bedrock is reached [cf. Begley, 2021, pp. 321–322]. Even if one was to put up a spirited defence to the above criticisms on behalf of the meaning postulate approach, we would still be left with the basic issue, which is that meaning postulates do not relate meanings beyond relating the extensions of words, that is, sets of things, individuals in the world.

4. Colour Incompatibility and Antonymy

A problem that was addressed by Katz especially over the period 1990–2002, is a generalisation of the problem that Wittgenstein came to realise was left unresolved by the *Tractatus*, known as *the colour incompatibility problem*. Katz begins his discussion of the problem with the following quotation from the *Tractatus*:

> It is clear that the logical product of two elementary propositions can neither be a tautology nor a contradiction. The statement that a point in the visual field has two different colours at the same time is a contradiction [Wittgenstein, 1974, p. 86: 6.3751].

When we consider this quotation together with a proposition such as "The spot is red and blue", we come up against one of the problems that eventually led Wittgenstein to reject his early philosophy as presented in the *Tractatus*. The proposition is a conjunction of two elementary propositions, namely, "The spot is red" and "The spot is blue". As such, and by the first sentence of Wittgenstein's statement above, it cannot be a logical contradiction. However, by the second sentence of the statement above it is recognised intuitively that the conjunction is nevertheless contradictory.

In a review of the *Tractatus*, from 1923, Frank Ramsey pointed out that Wittgenstein's attempt at handling this problem by appealing to a reduction in terms of physical theory does not succeed because "[…] Wittgenstein is only reducing the difficulty to that of the *necessary* properties of space, time, and matter or the ether" [Ramsey, 1923, p. 478]. The reason that this cannot work is that the propositions that describe these physical necessities and impossibilities will themselves be of the same form as the original proposition, and will again fail to take the form of a logical contradiction, '*p* and *not p*'. Wittgenstein became dissatisfied with his original treatment of this problem in the *Tractatus,* and after various attempts at resolving it through, for example, the *ad hoc* modifications of the truth table system, which he introduced in the paper 'Some Remarks on Logical Form' from 1929, and others [cf. Monk, 2014,

pp. 326–327], he abandoned any hope of finding a formal method for dealing with the problem. Wittgenstein eventually arrived at some notion of a use theory of meaning, that is, broadly speaking, a pragmatic approach as opposed to a formal semantic one. Carnap's meaning postulates are another way out of the colour incompatibility problem but, as we saw earlier, it is also an *ad hoc* method that is confounded by many other difficulties.

The main issue for a decompositional semantic approach like Katz' is that while the surface form of the sentences and their component expressions do not contain a negative element, the decompositional analysis of them must include such an element in order for a contradiction (or analyticity) to arise. This, however, is difficult given that we cannot simply say that the sense of 'red' contains the sense 'not blue' and *vice versa,* because the sense of 'red' would have to contain the senses of all the other colour terms also, in the form of, e.g., 'not green', 'not yellow', etc., in order to account for the sense properties of sentences involving these other colour terms [Katz, 1998, p. 564]. However, this implies that the sense of 'red' would have to contain the senses of all the other colour terms and, by parity of reasoning, this would lead to a vicious regress, for these senses themselves would also have to contain the senses of all the other colours; furthermore, this regress of senses would include multiple instances of senses of the form 'not red'; so, the original sense of 'red' would itself contain 'not red', which would make the system untenable. Problems such as these have led other philosophers to think that colour incompatibility statements such as "This spot is red and blue" and analytic statements like "Red is not blue" are instead synthetic. At the time of writing, Katz described this as the majority view [Ibid.].

To find a way around this difficulty, Katz draws upon W.E. Johnson's distinction between determinables (e.g., colour) and determinates (e.g., red, blue, etc.), in which "The special mode of difference that unites the determinates under a determinable is that no two of them can simultaneously characterize the same thing" [Ibid., p. 562]. Further, following Johnson, he adopts the position that colour senses are primitive. However, he does not treat them as being simple, rather, as a complex of markers that are not definable in terms of other markers. In the case of colour primitives, one of these markers will be 'colour', which dominates its particular determinate.

We will not go further into the minutia of Katz' formalism here.[10] The basis of his solution to the problem is that the formal devices that are already available to him in his semantic theory, allow him to incorporate Johnson's insights and make a distinction between the meaning of a word *in isolation* and the meaning of the word *in a sentence*. As, Katz explains:

---

[10]   For further some discussion, see [Begley, 2021, p. 322ff].

> […] the sense of a word in a sentence can have sense components which are not part of its sense as a lexical item but come into the latter from the senses of other constituents in the sentence. Compositionality makes the difference. Failing to note this, philosophers assimilate the case of words in a sentence to the case of words in isolation, thereby taking a genuine impossibility intuition about the latter to apply to the former, where it does not apply at all [Katz, 1998, pp. 572–573].

In light of this, Katz shows that the relation between primitive colour terms, and other determinates of determinables, is one of necessary exclusion, that can be known on the basis of underlying structure. Katz' semantic theory thus has a degree of uniformity in how it treats these kinds of incompatibility and standard oppositions involving subordination. Katz defines antonymy generally as follows:

> Two constituents C_i and C_j are antonymous (on a sense) if and only if they are not full sentences and they have, respectively, readings R_i and R_j such that R_i is identical to R_j except that R_i contains a semantic marker (M_i) and R_j contains a semantic marker (M_j) and the semantic markers (M_i) and (M_j) are distinct members of the same antonymous n-tuple of semantic markers [Katz, 1972, p. 52].

An *antonymous n-tuple* is simply a set of markers for senses that are grouped under a superordinate sense. We should notice the stark contrast here with the meaning postulate approach. Here the readings, R_i and R_j, which are each made up of a tree structure of semantic markers, contain the same or similar relations of subordination all the way up the tree. That is, the linguistic and psychological principles of minimal difference and invariance that we mentioned earlier as being central are built into the account.

The first distinction that Katz' semantic theory allows us to draw is the classical one between complementary antonymy and contrary antonymy. These are two of the three main kinds of antonymy that Katz explicitly recognises in his main technical work, *Semantic Theory*, from 1972. Katz' calls them 'Contradictories' and 'Contraries', respectively. The third kind that he mentions is that of 'Converses', which we will leave aside for present purposes and for want of space [Ibid., p. 159].

The distinction between these first two kinds of antonymy is one that arises, in a rather elegant and natural way. Put simply, when an antonymous n-tuple contains only two senses, the antonymy relation between those senses will be a complementary one. On the other hand, when an antonymous n-tuple contains more than two senses the antonymy relation between any two of those senses will be a contrary one. Hence, there is no need for direct logical postulation of the differences between these two kinds of antonymy. Instead, their differing logical properties arise from differences in their underlying sense structure.

Thus, the most basic distinction between these two kinds of antonymy is based simply on a count of the senses in an antonymous n-tuple. If there are two senses (represented by semantic markers) in an n-tuple, the application of the antonymy operator to one of them will return the other. If there are more than two senses in an n-tuple, then the application of the antonymy operator to one of them will return a disjunction of the other senses in the n-tuple. These two facts explain the logical properties of the two kinds of antonymy that were expressed in terms of meaning postulates by "$\forall x [(X'(x) \rightarrow {\sim}Y'(x)) \& ({\sim}X'(x) \rightarrow Y'(x))]$" and "$\forall x [X'(x) \rightarrow {\sim}Y'(x)]$", respectively [Cann, 1993, p. 220].

For the purposes of illustration, it is instructive to consider in particular the reason why the second conjunct of the first postulate, i.e., '${\sim}X'(x) \rightarrow Y'(x)$', does not hold for contraries. The result of the application of the antonymy operator on a particular member, e.g., 'X', of an antonymous n-tuple containing the underlying senses corresponding to 'X', 'Y', 'Z', …, returns a disjunction of the senses corresponding to 'Y' and, at least a third sense, corresponding to 'Z'. Thus, the implication '$\forall x [{\sim}X'(x) \rightarrow Y'(x)]$' is not valid for instances of contrary antonymy. However, there is no need for this to be stipulated, because it is a logical property that arises from the grouping of senses into antonymous n-tuples where $n > 2$. The other logical properties of the two kinds of antonymy arise in an analogous manner.

Conclusion

We began by attending to treatments of the phenomena of opposition in language and perception. We saw that the principles of minimal difference and invariance are centrally important, but that a similar principle did not inform the theoretical definitions of antonymy provided in formal semantics employing meaning postulates. That is, such definitions do not accord with the principle of minimal difference.

Further, we saw that meaning postulates merely assert relationships between sets of individuals in the world and do not provide an explanation of why semantic relations such as antonymy hold. We saw that Katz objected to the use of meaning postulates on the basis that they were merely *ad hoc* and effectively did not employ a fine-grained enough representation to provide an explanation of relations such as antonymy. That is, although these sense relations can be represented in this way, they are not treated as arising from the senses of the words involved. Katz also argued that antonymy defined in terms of meaning postulates is incapable of capturing differences between privative and possessive, and positive and negative antonyms.

We noted that these issues may be overcome in part by building in the central principle of minimal difference as part of the system of

representation. We then saw that there is indeed a decompositional system that does this, which was put forward by Katz and developed over several decades. The mature version of this system, incorporating insights from Johnson's theory of determinables and determinates, is even able to tackle what is known as the *colour incompatibility problem*, which had troubled the early Wittgenstein and others. It does this by treating the senses of colour terms as primitives of the system that are nonetheless complex in that they contain markers representing their determinable, that is, invariant components of sense structure.

Finally, we concluded by showing that Katz' decompositional system is able to provide definitions of the kinds of antonymy. We witnessed the elegant way in which the system distinguishes between contrary and complementary antonyms, effectively on the basis of a count of the senses in the relevant antonymous n-tuple, which amounts to a set of minimally different senses with invariant structure in common. That is, in contrast to the meaning postulate approach, sense relations such as antonymy arise from the sense structure of the words involved rather than being postulated to apply from without.

References

By-laws of the Board of Commissioners, Rules for the Government of the Schools, and Laws Relative to Common Schools in the City of Milwaukee. Milwaukee: Burdick's Ben Franklin Printing Office, 1860.

Begley, 2021 – Begley, K. "Atomism and Semantics in the Philosophy of Jerrold Katz", in: U. Zilioli (ed.). *Atomism in Philosophy: A History from Antiquity to the Present*. London: Bloomsbury Academic, 2021, pp. 312–330.

Béziau, Basti, 2017 – Béziau, J.Y., Basti, G. (eds). *The Square of Opposition: A Cornerstone of Thought*. Basel: Birkhäuser, 2017.

Bianchi, Savardi, 2008 – Bianchi, I., Savardi, U. *The Perception of Contraries*. Roma: Aracne, 2008.

Bobenrieth, 2017 – Bobenrieth, A. "The Many Faces of Inconsistency", in: J.Y. Béziau, G. Basti (eds). *The Square of Opposition: A Cornerstone of Thought*. Basel: Birkhäuser, 2017, pp. 145–168.

Branchini, Capitani, Burro, Savardi, Bianchi, 2021 – Branchini, E., Capitani, E., Burro, R., Savardi, U., Bianchi, I. "Opposites in Reasoning Processes: Do We Use Them More Than We Think, but Less Than We Could?", *Front. Psychol.*, 2021, 12:715696, pp. 1–13.

Cann, 1993 – Cann, R. *Formal Semantics: An Introduction*. Cambridge: Cambridge University Press, 1993.

Cann, 2011 – Cann, R. "Sense Relations", in: C. Maienborn, K. von Heusinger, P. Portner (eds). *Semantics: An International Handbook of Natural Language Meaning*. Berlin: de Gruyter, 2011, pp. 456–478.

Clark, 1970 – Clark, H.H. "Word Associations and Linguistic Theory", in: J. Lyons (ed.). *New Horizons in Linguistics*. Baltimore: Penguin, 1970, pp. 271–286.

Cruse, 1986 – Cruse, D.A. *Lexical Semantics*. Cambridge: Cambridge University Press, 1986.

Cruse, 2000 – Cruse, D.A. *Meaning in Language: An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press, 2000.

Egan, 1984 – Egan, R.F. "Introductory Matter", in: *Webster's New Dictionary of Synonyms*. Springfield: Merriam-Webster, 1984.

Gao, Zheng, 2014 – Gao, C., Zheng, Q. "A Linguistic Study of Antonymy in English Texts", *Journal of Language Teaching and Research*, 2014, vol. 5, no. 1, pp. 234–238.

Jones, 2002 – Jones, S. *Antonymy: A Corpus-based Perspective*. London: Routledge, 2002.

Jones, Murphy, Paradis, Willners, 2012 – Jones, S., Murphy, M.L., Paradis, C., Willners, C. *Antonyms in English: Construals, Constructions and Canonicity*. Cambridge: Cambridge University Press, 2012.

Katz, 1966 – Katz, J.J. *The Philosophy of Language*. New York: Harper & Row, 1966.

Katz, 1972 – Katz, J.J. *Semantic Theory*. New York: Harper & Row, 1972.

Katz, 1977 – Katz, J.J. "The Advantage of Semantic Theory over Predicate Calculus in the Representation of Logical Form in Natural Language", *The Monist*, 1977, vol. 60, no. 3, pp. 380–405.

Katz, 1986 – Katz, J.J. *Cogitations*. New York: Oxford University Press, 1986.

Katz, 1992 – Katz, J.J. "The New Intensionalism", *Mind*, 1992, New Series, vol. 101, no. 404, pp. 689–719.

Katz, 1998 – Katz, J.J. "The Problem in Twentieth-Century Philosophy", *The Journal of Philosophy*, 1998, vol. 95, no. 11, pp. 547–575.

Lyons, 1977 – Lyons, J. *Semantics*. Vol. 1. Cambridge: Cambridge University Press, 1977.

Margolis, Laurence, 1999 – Margolis, E., Laurence, S. (eds.). *Concepts: Core Readings*. Cambridge, MA: MIT press, 1999.

Mettinger, 1994 – Mettinger, A. *Aspects of Semantic Opposition in English*. Oxford: Oxford University Press, 1994.

Monk, 2014 – Monk, R. "The Temptations of Phenomenology: Wittgenstein, the Synthetic a Priori and the 'Analytic a Posteriori'", *International Journal of Philosophical Studies*, 2014, vol. 22, no. 3, pp. 312–340.

Murphy, 2003 – Murphy, M.L. *Semantic Relations and the Lexicon: Antonymy, Synonymy, and Other Paradigms*. Cambridge: Cambridge University Press, 2003.

Pitt, 1999 – Pitt, D. "In Defense of Definitions", *Philosophical Psychology*, 1999, vol. 12, no. 2, pp. 139–156.

Ramsey, 1923 – Ramsey, F. "Review: Tractatus Logico-Philosophicus", *Mind*, 1923, New Series, vol. 32, no. 128, pp. 465–478.

Smith, 1867 – Smith, C.J. *Synonyms and Antonyms*. London: Bell & Daldy, 1867.

Wittgenstein, 1929 – Wittgenstein, L. "Some Remarks on Logical Form", *Proceedings of the Aristotelian Society*, 1929, Supplementary Volumes, vol. 9, pp. 162–171.

Wittgenstein, 1974 – Wittgenstein, L. *Tractatus Logico-Philosophicus*. Trans. by D.F. Pears, B.F. McGuinness. London: Routledge & Kegan Paul, 1974.

Woleński 2016 – Woleński, J. "Ajdukiewicz on Analyticity", *Studies in East European Thought*, 2016, vol. 68, no. 1, pp. 5–10.