Implicatures and Naturalness

Igor Douven

Abstract Pragmatics postulates a rich typology of implicatures to explain how true assertions can nevertheless be misleading. This typology has been mainly defended on the basis of a priori considerations. We consider the question of whether the typology corresponds to an independent reality, specifically whether the various types of implicatures constitute natural concepts. To answer this question, we rely on the conceptual spaces framework, which represents concepts geometrically, and which provides a formally precise criterion for naturalness. Using data from a previous study, a space for the representation of implicatures is constructed. Examination of the properties of various types of implicatures as represented in that space then gives some reason to believe that most or even all types of implicatures do correspond to natural concepts.

Keywords Conceptual spaces · Implicatures · Multi-dimensional scaling · Natural concepts · Pragmatics

Linguists and other language researchers customarily distinguish between syntax, semantics, and pragmatics, where (roughly) the first pertains to the ways words can and cannot be combined into sentences, the second to word and sentence meaning, and the third to language use. This paper is concerned with a question central to pragmatics, specifically with the scientific status of so-called implicatures, which play a key explanatory role in this field. More specific still, we are interested in the question of whether all types of implicatures that the current literature distinguishes between are *natural* concepts, where the notion of a natural concept will be understood as defined by researchers working on psychological spaces. The question is important insofar as only natural concepts deserve a place in mature scientific theories (Lewis 1983; Boyd 1991).

To address this question, we use data from a study reported elsewhere (Douven and Krzyżanowska 2019) to construct a psychological space for the representation of implicatures. In that space, we examine the properties of various types of impli-

I. Douven (✉)
SND/CNRS/Sorbonne University, Paris, France
e-mail: igor.douven@paris-sorbonne.fr

© The Author(s) 2021
S. Löbner et al. (eds.), *Concepts, Frames and Cascades in Semantics, Cognition and Ontology*, Language, Cognition, and Mind 7,
https://doi.org/10.1007/978-3-030-50200-3_7

catures, with a special interest in seeing whether they satisfy an important criterion for naturalness (convexity—see below) as proposed in the psychological spaces literature. The outcome will be seen to provide some support for holding that most or even all types of implicatures do correspond to natural concepts.

1 Theoretical Background

The basic insight at the root of pragmatics is that we can mislead our audience not only by telling lies, but also by telling nothing but the truth. Suppose someone asserts,

(1) President Obama has one daughter.

The assertion is true yet misleading, given that it suggests that Obama has *exactly* one daughter—which is false. What is suggested is not asserted, but it is nonetheless conveyed due to a normally warranted presumption of a kind of cooperativeness that goes beyond merely telling the truth. In the present example, we may suppose that the speaker was in a position to assert, and could with just as much effort have asserted, that Obama has *two* daughters, which would have been true as well but would in addition have been more informative. Precisely because we expect each other to be cooperative in this kind of way—to try to make our contributions to a conversation not only true but also relevant, clear, and informative—a person unaware of how many daughters Obama has would be justified to infer from an assertion of (1) that he has exactly one daughter. That Obama has *exactly* one daughter is said to be an *implicature* of (1), whose *semantic* content is only that Obama has one daughter, possible among many more.

There exist a number of different typologies of implicatures, which are partly independent of each other. One broad division is that between *conventional* and *conversational* implicatures, where the former are said to arise due to the meaning of specific words, and the latter due to the context in which an assertion is made. For instance, the word “although” in

(2) Although Obama won a second term as president, dolphins are mammals.

Suggests the existence of a contrast between the two conjuncts in this sentence (which strikes us as wrong, given that the conjuncts appear unrelated). On the other hand, there is no single word in (1) that might lead a hearer to think that Obama has *exactly* one daughter. *That* suggestion can arise for the reason mentioned above: because we would normally assume that (1) is the strongest statement the speaker can make regarding the number of daughters Obama has. Indeed, there are conversational contexts where this assumption would not be warranted. For instance, if it has just been asserted that anyone who has at least one daughter qualifies for a certain special government program, we would not interpret an assertion of (1) as suggesting that Obama has exactly one daughter. Rather, we would take the speaker’s point to be that Obama meets the requirement for the government program.

This brings us to second distinction. We just said that although an assertion of (1) would, in normal circumstances, implicate that Obama has exactly one daughter, there are circumstances in which this implicature would not arise. Grice (1989, p. 37 f) calls implicatures of this type “generalized conversational implicatures.” He differentiates them from what he calls “particularized conversational implicatures,” which arise only in specific conversational contexts. For instance, if we are at a party and you ask me what time it is, you may interpret my assertion of

(3) The guests are leaving.

As indicating that it is already late, even if asserting (1) normally does not engender this suggestion.

It is fair to say, though, that most attention in the literature has gone to a sub-typology of conversational implicatures which is based on the various types of expectations—each brought about by the overarching expectation of cooperativeness—that the implicatures exploit. For instance, the aforementioned implicature of (1) is said to be of a scalar type, because we can represent numbers (e.g., numbers of children) on a scale, and the expectation of informativeness then requires that we go as far out on that scale as is warranted by our evidence. So someone’s asserting (1) implicates that she knows, or has good evidence for believing, that Obama has exactly one daughter. By contrast, someone asserting that

(4) Kate Middleton gave birth to a son and she married Prince William.

Is offending the expectation that we report events in an orderly fashion, which in this instance means: in the order in which they occurred. Thus, the obviously wrong implicature generated by an assertion of (1)—that the event mentioned first also happened first—is said to be of an order type.

Scalar implicatures have given rise to a further sub-typology, this one being based on the different scales that can underly the production of these implicatures. The main subtypes are the *quantificational implicatures*, which involve a scale of quantifiers (e.g., some–many–most–all); the *gradable adjective implicatures*, which exploit some scale of adjectives that can apply to differing degrees (e.g., soft–audible–loud–blaring); the *ranked ordering implicatures*, which involve orderings (like beginner–intermediate–advanced); and the *cardinal number implicatures*, which involve some cardinal number scale, as in our example (1).

This paper will focus on the typology which starts by branching off the conversational and conventional implicatures and which then has the further branches for the conversational implicatures described in the previous two paragraphs. This typology has been mainly defended on the basis of a priori considerations, more specifically on what are sometimes called “linguistic intuitions.” However, such intuitions are known to be not always reliable. Indeed, while the said typology is still part of mainstream pragmatics, parts of it have been contested. For instance, some authors deny that sentences like (1) carry the “exactly *n*” reading as a matter of implicature, claiming that, rather, the “exactly” reading is part of the semantics of numerals (see Scharten 1997 and Breheny 2008). And Bach 1999 has argued that the belief in the existence of conventional implicature rests upon a myth.

Bach's arguments have in turn been challenged (e.g., Potts 2005) and in any event my aim is not to question the reality of any part of the aforementioned typology. Rather, I am interested in the metaphysical status of the various types that occur in it. It has often been said that we do not just want scientific theories to be predictively accurate, but also want them to inform us about what, deep down, underlies the phenomena (e.g., Psillos 1999). And that requirement can be satisfied only if these theories "carve nature at its joints," that is, only if their core concepts are *natural* ones (Lewis 1983). Against this background, the question I am asking is whether the above typology latches on to some independent, fundamental reality. Do, for instance, so-called order implicatures constitute a *natural* class of implicatures? More generally, are *all* types of implicatures natural? Or better perhaps, if Lewis (1983) is right that naturalness permits of degree, are they all *equally* natural?

To address these questions, we need some understanding of what it takes for a concept to count as natural. It has been argued that a concept is natural if it figures in one or more laws of nature (e.g., Putnam 1983). But this is problematic, given that it is hard to say what makes a regularity a law of nature (or otherwise) without making reference to natural concepts (Douven and van Brakel 1998). To characterize naturalness of concepts, it is actually more helpful to turn to recent work on conceptual spaces, in which a criterion for distinguishing natural from nonnatural concepts has been proposed that is backed by a considerable amount of experimental evidence.

We will construct a conceptual space later on, and will then go into details. For now, it suffices to say that a conceptual space is a one- or multidimensional metric space, where the dimensions represent fundamental qualities that items can have to varying degrees and with respect to which they can be compared to each other. Distances in such spaces are supposed to be inversely related to similarities: the greater the distance between (the representations of) two items in a given space, the more dissimilar the items are in the respect represented by the space. For example, CIELAB space is a three-dimensional Euclidean color space, and distances in the space are meant—and have been shown—to predict accurately how similar people will judge different shades to be: the closer two shades are in CIELAB space, the more similar they tend to appear to human observers (Fairchild 2013). Many other conceptual spaces are known in the literature, and although the best-known ones all pertain to perceptual concepts (next to color spaces, such as CIELAB, there are vowel spaces, odor spaces, taste spaces, etc.), more recently conceptual spaces have been developed for more abstract concepts, including moral, epistemic, and scientific concepts.

What makes conceptual spaces especially valuable is that they allow us to represent concepts geometrically, as regions in some given space. Thereby, the study of concepts becomes both formally rigorous and empirically testable. For instance, the concept of redness can be thought of as a region in CIELAB space, which means we can carry out all sorts of mathematical operations on it—like measuring its volume—and at the same time use it for conducting all sorts of experimental work (e.g., concerning the nature of vagueness: see Douven et al. 2013; Decock and Douven 2014; Douven and Decock 2017; Douven et al. 2017; Douven 2018).

If concepts are regions in conceptual spaces, is any region in any conceptual space a concept? “Concept” is, to a high degree, a term of art, and so we are free to answer this question in the positive. However, the more worthwhile question is whether any region represents or could represent a *natural* concept. And it takes little imagination to appreciate that now the answer is definitely negative. In color space, there are infinitely many regions that contain all the colors in the rainbow. Surely such regions represent gerrymandered rather than natural concepts.

Now that we can think of concepts formally, can we also distinguish formally between those regions that represent or can represent natural concepts and those that can *not*? Gärdenfors (2000, p. 71) proposes a topological criterion, which he calls

Criterion P: A *natural concept* is a convex region of a conceptual space,

where a region $\mathcal{R}$ is convex if and only if, for any pair of points $x, y \in \mathcal{R}$, if $z \in \overline{xy}$ then $z \in \mathcal{R}$. As Gärdenfors (2000, p. 70) explains, Criterion P can be thought of as a principle of cognitive economy, given that “handling convex sets puts less strain on learning, on your memory, and on your processing capacities than working with arbitrarily shaped regions.” He also cites important empirical work on color naming which shows that color concepts like BLUE, RED, GREEN, and so on, which we tend to regard as *natural* color concepts, all form convex regions in CIELAB space (see also Jraissati and Douven 2018). Douven (2016a) presents further empirical evidence for Criterion P, showing that the concepts BOWL and VASE come out as convex in the appropriate shape space.

Whereas Criterion P is a plausible *necessary* condition for natural concepts, it is debatable whether it is also *sufficient*.¹ Gärdenfors (2000, p. 70) already expressed doubts on this point, and Douven and Gärdenfors (2018) argue explicitly that further conditions are needed to single out the natural concepts. However, in addressing the question of whether all types of implicatures are equally natural concepts, we will content ourselves with considering whether the various types of implicatures, when represented in a conceptual space we are about the construct, satisfy Criterion P. If some fail to do so, that is an indication that they are not natural concepts. And if some or all do satisfy the criterion, that is at least some evidence for holding that they *are* natural concepts.

To build the requisite conceptual space for representing types of implicatures, we need input data. The data we are going to use are taken from a study reported in Douven and Krzyżanowska (2019). We briefly describe the data in the next section, and then go on to construct a conceptual space in Sect. 3.

¹There has been some discussion about whether Criterion P is even necessary. See Gärdenfors (2018) and references given there.

2 Input Data

Douven and Krzyżanowska (2019) were interested in three questions, all related to the semantics–pragmatics interface. First, they sought to investigate empirically whether ordinary speakers’ responses to true but supposedly pragmatically infelicitous sentences—true sentences that generate a false implicature—are in line with linguists’ and philosophers’ ideas about how semantic and pragmatic aspects of language are to be sorted. Specifically, they were interested in whether people reliably distinguish between the *truth* and the *assertability* of sentences in a way that accords with mainstream thinking in linguistics and philosophy.

Second, Douven and Krzyżanowska (2019) were interested in possible differences in responses brought about by the various types of implicatures. For instance, might people systematically deem true sentences generating false *conventional* implicatures *more* unassertable than true sentences generating false *conversational* implicatures? Might the different types of conversational implicatures be evaluated differently in this respect?

And third, they were interested in individual differences among participants. Previous research (Spychalska, Kontinen, and Werning 2016) had suggested that some people are more inclined to judge the truth values of sentences purely on the basis of what according to theorists are the semantic contents of those sentences, whereas other people might base their truth judgments also, at least to some extent, on the sentences’ pragmatic aspects, so that they might be more inclined to judge a true sentence with a false implicature as *false*.

To investigate these questions, Douven and Krzyżanowska used materials consisting of the 24 items listed in Table 1 together with a great variety of filler items which were meant to conceal from the participants the purpose of the study. The test items were meant to generate six types of false implicatures, where each type was instantiated by four different sentences: quantificational implicatures (items 1–4); gradable adjective implicatures (items 5–8); ranked ordering implicatures (items 9–12); cardinal number implicatures (items 13–16); temporal order implicatures (items 17–20); and conventional implicatures (items 21–24).

In both studies reported in Douven and Krzyżanowska (2019), the participants were divided into three groups, where participants in one group were asked about the items’ *truth*, participants in a second group were asked about the items’ *assertability*, and participants in the remaining group were asked about the items’ *believability* (the questions about believability were related to a secondary research goal, which we leave aside here; see Douven 2010, 2016b, and Douven and Krzyżanowska 2019). The difference between the two studies was that participants in the first were always asked to give yes/no answers, whereas participants in the second study were asked to indicate on a 7-point Likert scale the extent to which they agreed that an item was true/assertable/believable.

As for the first research question, neither study revealed any significant differences among the responses from the three groups (nor were there significant differences between the two studies). Figure 1 presents the proportions of positive responses from

Table 1 Items used in the studies reported in Douven and Krzyżanowska (2019)

1.	Some patches are blue. ^a
2.	Some roses are flowers
3.	Most patches are red. ^b
4.	Most laptops are computers
5.	The tiger finds the boy's cereal moderately sweet. ^c
6.	The female basketball player Margo Dydek (7 ft 2 in/2.18 m) was tall for a woman
7.	Bill Gates is relatively rich
8.	On the North Pole, winter temperatures are somewhat cold
9.	In the UK, people over the age of 85 have the right to retire
10.	In principle, all American citizens over the age of 25 have the right to vote in federal election
11.	In the UK and the US, children under the age of 15 are prohibited from buying hard drugs
12.	In the US, people who earn more than \$200,000 a year are obliged to pay taxes
13.	Alfred Hitchcock made two movies
14.	President Obama has one daughter
15.	In the last Olympic games, the US won four medals
16.	At the height of its power, Great Britain owned 12 ships
17.	The tiger looks for the bread in the toaster and the boy puts a piece of bread into the toaster. ^d
18.	Princess Diana died in a car accident and she divorced Prince Charles
19.	The man comes up with a bogus answer and the boy asks how the load limit on bridges is determined. ^e
20.	Kate Middleton gave birth to a son and she married Prince William
21.	Although Prince William had fallen in love with Kate Middleton, the 2014 Winter Olympics will be in Russi
22.	<i>Harry Potter and the Sorcerer's Stone</i> was a box office hit, therefore Obama is the president of the US
23.	Although Obama won a second term as president, dolphins are mammals
24.	Mitt Romney lost the 2012 presidential election, therefore U2 is a rock band

^aShown with a series of only blue patches. ^bShown with a series of only red patches. ^cShown with a comic strip in which a tiger is seen finding a boy's cereal extremely sweet. ^dShown with a comic strip in which a boy first puts bread in a toaster and then a tiger looks into the toaster. ^eShown with a comic strip in which a boy first asks the question and then the man answers it

the first study, which shows how close the responses from the three groups were to each other. The graphs of the mean responses from the second study, not shown here, are virtually indistinguishable from those shown here; see Douven and Krzyżanowska (2019). So, as far as these results go, it hardly appears to matter whether we ask people to judge the truth, believability, or assertability of a sentence that is true according to standard semantics but that generates a false implicature. More generally, Douven and Krzyżanowska (2019) found no evidence that the semantics–pragmatics divide,

Fig. 1 Proportions of positive responses per item from the first study in Douven and Krzyżanowska (2019); labels refer to the numbering of items in Table 1

however useful from a theoretical perspective perhaps, is reflected in how ordinary speakers tend to evaluate sentences like those in Table 1.

As stated above, Douven and Krzyżanowska (2019) were also interested in possible differences in responses due to the various types of implicatures generated by their materials. Just eye-balling the results in Fig. 1, it appears that proportions of positive responses tend to be in the same range for each type separately, but not so much across types. In line with this, Douven and Krzyżanowska’s analysis revealed a significant effect of type of implicature on the responses. They again obtained the same result for the responses from their second study. Hence, the answer to their second question was positive.

For the third question—whether participants can be split into logical responders and pragmatic responders—they looked at the correlations between the responses for any pair of items. If a division between logical and pragmatic responders exists, then at a minimum one would expect these correlations to be rather high: some participants—the supposedly logical responders—would then tend to judge all items in Table 1 to be true, while others—the supposedly pragmatic responders—would tend to judge all those items to be false. But that turned out not to be the case. Figure 2 is reproduced from Douven and Krzyżanowska (2019) and shows the correlations among the “truth” responses from the first study; the correlations from the second study were essentially the same. It is clearly visible that, whereas both the responses to the quantificational items and the responses to the conventional items correlate amongst themselves, they do not even moderately correlate with most of the other

Fig. 2 Correlations among “truth” responses from the first study in Douven and Krzyżanowska (2019); labels refer to the numbering of items in Table 1

items, nor do the responses to those other items tend to correlate even moderately among themselves.

Given that in no interesting respect were there significant differences between the two studies reported by Douven and Krzyżanowska, we in the following consider only the data from the first study.

3 Building an Implicature Space

In Sect. 1, we mentioned that, whereas most conceptual spaces to be found in the literature are for *perceptual* concepts, there is nothing that prevents us from constructing spaces for other types of concepts, as is witnessed by some recent proposals for modeling abstract concepts spatially. Here, I am going to make a further such proposal, to wit, a proposal for constructing an implicature space. I am not aware of any previous attempts to create such a space, but the idea of a conceptual space for the representation of implicatures certainly makes sense.

At least, the idea makes sense *prima facie*—there is a *concept* of conventional implicature, a *concept* of order implicature, and so on—but one must always reckon with the fact that trying to construct a conceptual space leads nowhere. To see how this may happen, it is first to be noted that conceptual spaces are typically constructed by means of a dimensionality-reduction technique, the one most commonly used being multidimensional scaling (MDS). In an MDS procedure, we construct a spatial representation of a set of items, taking as input similarity judgments, or confusion probabilities, or correlation coefficients, pertaining to those items. There is no guarantee, however, that the resulting representation will be any good. Specifically, what we aim at in an MDS procedure is a space which (i) is *low-dimensional*, ideally, with no more than three dimensions; (ii) has good fit, which in this context is expressed in terms of *stress*, where lower stress values indicate more faithful representations of the similarities/confusion probabilities/correlations related with the items we are trying to represent; and (iii) has *interpretable dimensions*, in that we can associate each dimension with some fundamental attribute the items can be said to have to some degree. An outcome of an MDS procedure may fail to satisfy some or all of these criteria.

The items we are going to use to construct an implicature space are the ones given in Table 1, and the specific input data are the correlations among the responses to those items reported in Douven and Krzyżanowska (2019) and briefly described and depicted in the previous section.

To start building our space, we must first turn those correlations into distances. There are many options for measuring such distances, but the most common ones are all instances of the so-called Minkowski metric, which is defined thus:

$$\delta_k(p, q) = \left(\sum_{i=1}^n |x_i - y_i|^k \right)^{1/k}$$

with $p = \langle x_1, \dots, x_n \rangle$ and $q = \langle y_1, \dots, y_n \rangle$. For $k = 1$, this yields the so-called city-block or Manhattan metric, and for $k = 2$, the more familiar Euclidean metric.

It is generally held that the Euclidean metric is appropriate for measuring distances between similarity ratings (confusion probabilities, correlations) when the “dimensions” underlying those ratings are *integral* in the sense that they cannot be experienced independently of each other (for instance, one cannot separately experience the hue and the saturation of a shade). If, by contrast, the relevant dimensions are *separable* (i.e., not integral), then the city-block metric is generally considered to be the right choice (see, e.g., Torgerson 1958; Garner 1962; Shepard 1964; and Nosofsky 1986).

In the present case, it is not immediately clear which, or how many, dimensions are going to be necessary to faithfully represent our items, supposing we can obtain a faithful representation at all. Thus, in particular, it is not clear whether we should expect the dimensions to be integral or separable. For that reason, we derive distances from the correlation coefficients both via the Euclidean metric and via the city-block metric, and then carry out MDS procedures for each separately.

Fig. 3 Stress values for MDS solutions with 1 to 10 dimensions, both for the city-block and for the Euclidean distances

Once distances are derived—in the present case done via the `dist` function that is part of the base R language (R Core Team 2017)—one faces a further choice, to wit, whether to apply metric or nonmetric multidimensional scaling. The former tries to represent objects geometrically in a way which preserves as faithfully as possible the *distances* between those objects in the distance matrix that is given as input. By contrast, the latter tries to represent objects geometrically in a way which preserves as faithfully as possible the *ordering* of the distances between those objects according to the distance matrix; so, the smaller the distance between objects according to the matrix, the closer they are in the geometric representation, though no linear mapping of matrix distances onto distances in geometric space is aimed for. When distances derive from subjective assessments, nonmetric multidimensional scaling is generally recommended (Bartholomew et al. 2008, pp. 56–62). Given that, in our case, the distances do come from subjective assessments—people’s responses to the items in Table 1—nonmetric multidimensional scaling will be used in the following.

Specifically, we conduct the MDS procedures using the function `metaMDS` that is included in the `vegan` package for R. All configurations are centered and rotated to a principal axes orientation (see Borg and Groenen 2010, Sect. 7.10). MDS procedures are conducted for 1–10 dimensions and their stress levels are compared. The various stress values for the outcomes are shown in Fig. 3. We see immediately that we can obtain better solutions for the city-block distances than for the Euclidean distances. According to Johnson (2008, p. 205), in MDS we look for stress values less than 20. This criterion is met already by the two-dimensional solutions.

There is a second type of plot commonly used to assess the goodness-of-fit of an MDS solution, the so-called Shepard plot, in which input and output distances are plotted against each other. Figure 4 shows such plots for the best two- and three-dimensional MDS solutions, so plotting the city-block distances among the correlations (the observed dissimilarities) against the city-block distances in the solutions. We see that, in both cases, the fit is excellent, with an R^2 value of .98 for the two-dimensional solution and of .99 for the three-dimensional one. Especially in the latter

Fig. 4 Plot of distances among correlations against distances in best MDS solution

Fig. 5 Two-dimensional MDS solution for the city-block distances; different categories of items are differently colored

case, the plotted points are grouped very tightly around the monotonically increasing line corresponding to perfect fit (for the nonmetric case). The actual solutions are displayed in Figs. 5 and 6.

So far, the best solutions satisfy two out of the three criteria (i)–(iii) mentioned above: they are low-dimensional, and they have excellent fit. How about the third criterion, that of having interpretable dimensions? While coming up with an interpretation of the dimensions of an MDS solution is often challenging (see Douven 2016a), it seems doable in the present case, at least for the first two dimensions (the only two, if we are happy to go with the two-dimensional solution).

From much of the pragmatics literature one comes away with the impression that utterances either are or are not infelicitous, depending on whether they generate a false implicature, as if that were a categorical matter. That seems as wrong, however, as the suggestion, also encountered in some of the same literature, that an utterance

Fig. 6 Different viewpoints on the three-dimensional MDS solution for the city-block distances

either does or does not generate an implicature. The two wrong suggestions may well be related: failure to observe that utterances can be more or less felicitous may stem from a failure to observe that implicatures can be stronger or weaker.

Consider, for instance, an example from Douven (2012). In the example, a graduate student tells her supervisor,

(5) You have published some papers that I really like.

The supervisor can see two different possible explanations of why the student uttered this sentence. One is that the student wanted to convey that she read some of his papers and liked all of those; the other is that she read some or all of his papers and liked some of those she read and some not so much. The supervisor may think the first explanation tops the second and therefore infer that the student did not read all of his papers. However, the point the example is meant to illustrate is that because of the presence of an alternative explanation of why the student uttered (3), and an alternative that is close in explanation quality to the first explanation, the inference can only be guarded, so that, as a result, the implicature is only a weak one. Put differently, if it should turn out that the student read all of her supervisor's papers, an utterance of (3) would at most be minimally infelicitous.

Once this is observed, it is not too speculative to think that the first dimension represents something like degree of felicitousness (or conversely, degree of potentiality to mislead one's audience). Consider the four items most to the right in the two-dimensional space (6, 10, 12, 18), and compare them with the quantifier items (1–4) and the conventional items (21–24): All eight of the last items strike one as being much more infelicitous than the first four items. And all of the cardinal number items (13–16) do strike us as being more infelicitous than, for instance, item 6, but not quite as infelicitous as the quantifier or conventional items. More generally, that felicitousness is a matter of degree should be uncontroversial and is directly related to the claim made in Douven (2012) that implicatures can vary in strength. The latter claim was defended in terms of explanation quality—an implicature can

be part of the best explanation of why the speaker said what she said in the context in which she said it, but the extent to which the best explanation stands out as being the best can vary, and can have a significant impact on people's willingness to infer the truth of that explanation, as has recently been verified experimentally in Douven and Mirabile (2019).

Some support for this suggestion also comes from considering that item 18, which mentions Princess Diana's death first and her divorce second, carries basically *no* risk of misleading anyone about the order of the events, given that a divorce requires a person to be alive. Here, semantics (the meanings of "divorce" and "death") and world knowledge simply prohibit the implicature of the "wrong" temporal order to arise from an utterance of item 18. This is different for temporal order items 17 and 19: both suggest a temporal order of the events that is perfectly possible given the meanings of the terms involved and general knowledge about the world but that happens to be contradicted by the comic strips the sentences pertained to. Perhaps temporal order item 20 does not fit this interpretation quite as well, given that, in the context of the British royal family, it seems rather improbable, *a priori*, that the wife of a successor to the throne becomes a mother, or even becomes pregnant, while being unmarried. On the other hand, as the saying goes, the times they are a-changin'.

The strict split between conventional and conversational implicatures, mentioned in Sect. 1, may in fact be due to another false dichotomy. Against the widespread assumption that an implicature arises either due to the conventional meaning of some term or due to context plus the assumption of speaker cooperativeness, some authors have pointed out that there can be differences in the frequencies with which contexts occur that give rise to this or that implicature, and these differences may have an effect on the degree to which an implicature comes to be felt as being part of the meaning of a given expression. Hopper and Traugott (2003, Sect. 4.3) refer to this process as "semanticization," citing the following characterization of it:

[I]f some condition happens to be fulfilled frequently when a certain category is used, a stronger association may develop between the condition and the category in such a way that the condition comes to be understood as an integral part of the meaning of the category. (Dahl 1985, p. 11)

Given that the frequency with which the condition may be fulfilled in contexts in which an expression is used may vary, one would suppose that the situation that the condition is understood as part of the meaning of the expression is a limiting case, and that the strength of the association between condition and expression can vary.

To make this more concrete, compare, for instance, items 1, 8, and 22. It is difficult to imagine a context in which use of the word "therefore" does *not* suggest an inferential relationship between the clauses it connects. Helping us indicate the presence of such a relationship seems to be the *only* use we have for the word. So, it is felt as being part of the *meaning* of "therefore" that there is an inferential relationship between the connected clauses, even if for theoretical reasons it may still be better to attribute this suggestion to pragmatics—specifically, "therefore" generating a conventional implicature—than to semantics.

At the other extreme, looking at item 1, it is *very easy* to conceive of contexts in which we do not at all intend “some” to have a “not all” reading. Suppose I utter,

- (6) John is going to organize a party, and knowing him, he’s going to play loud music. Some people in the neighborhood will be annoyed.

I may utter these sentences without having any evidence, and without meaning to imply, that not all people in the neighborhood are going to be annoyed by the loud music at John’s party. What I know for sure is that *some* people are going to be annoyed, but while I am not in the stronger epistemic position to assert that *all* people are going to be annoyed, I do not wish to suggest that that is not an open possibility. And my audience, reasonably supposing that I have not surveyed all people in the neighborhood on this matter, also will not likely take me to be suggesting as much, and so will not likely be misled.

Finally, consider item 8. In virtually all contexts, we will take “somewhat cold” simply to *mean* “not extremely cold.” On the other hand, on our best current theoretical analyses of gradable adjectives (such as “cold”), these implicitly refer to standards, and such standards are known to be sensitive to contextual variation. Consider a discussion in which a group of adventurers are planning an expedition, where it is already decided that the expedition is going to be to some extremely cold place. Then the modifier “somewhat” in an utterance of item 8 might be appropriate in the context of their conversation if they had just been considering places to go where it is even colder than at the North Pole in the winter. (I am assuming, for the sake of the example, that such places exist, which I have not verified.) Even in that context, we may presume, none of the adventurers would want to deny that winter temperatures at the North Pole are extremely cold.

Perhaps similar considerations apply to the cardinal number items (13–16). Recall the context, from Sect. 1, where it would be entirely appropriate to assert that Obama has one daughter. Or consider this exchange:

Quizmaster: “Name one country that won at least four medals in the last Olympic games.”

Candidate: “France won four medals.”

Such contexts may not be very common, but they are also not extremely rare. (As for the item about Hitchcock, that may not have been well chosen, given that especially a younger generation may have little familiarity with Hitchcock or his movies.)

Based on the above considerations, and given that the conventional items are all near the bottom of the scale constituted by the second dimension, the quantifier items all at the top of that scale, and the degree modifier items as well as the cardinal number items are in between, my best guess concerning the second dimension is that it represents something like context-sensitivity or degree of semanticization.

In short, the proposed interpretations of the first two dimensions are degree of felicitousness (or degree of misleadingness) and degree of semanticization, respectively. It appears harder to come up with an interpretation of the additional dimension for the three-dimensional solution and we leave this as an open issue here. It is to be emphasized that because the MDS procedures were conducted on the basis of

relatively sparse data, any interpretation of the dimensions is at best an exploratory hypothesis, to be confirmed in follow-up research, ideally involving a richer set of materials.

4 Naturalness

We finally come to the question concerning naturalness: Are the concepts associated with the various types of implicatures *natural* ones? We did much of the necessary stage-setting in the previous section, due to which we now have available an implicature space (or two, if we like), which will make answering the aforementioned question much easier. After all, as was remarked in Sect. 1, in the conceptual spaces framework the notion of naturalness has a precise meaning, or at least the framework provides a precise criterion for naturalness, viz., convexity. (It will be recalled that a region is convex if and only if, for any pair of points lying in the region, the line segment connecting them lies in its entirety in the region as well.) As mentioned, there is a wealth of evidence supporting this criterion; for instance, in color space, we find only shades of red between any pair of shades of red, and not also (say) shades of blue or green or orange. Does a similar conclusion hold for the various types of implicatures as represented in our implicature space(s)?

We start by considering again the two-dimensional solution shown in Fig. 5. We observe that, in this solution, the quantifier items (1–4) are tightly grouped together, as are the cardinal number items (13–16) and the conventional items (21–24). The same is true for three of the four gradable adjective items (5, 7, 8), the outlier being 6. One reason why this may not be very surprising is that the first three items all concern so-called degree modifier phrases (“*X* is relatively/moderately/somewhat *Y*”), whereas the outlier involves a comparison class phrase (“*X* is *Y* for a *Z*”). In the pragmatics literature, these are commonly distinguished, and so it might have been better if Douven and Krzyżanowska had kept them separate in their work; they might for instance have included four items of each subtype among their materials. In any case, the types seem to trigger somewhat different pragmatic inferential mechanisms: degree modifier phrases implicate that the utterance would be false, or at least further from the truth, were the modifier omitted, while comparison class phrases implicate that the utterance would be false, or at least further from the truth, if the comparison class were not mentioned or were replaced by the normally implicit default comparison class (“Trump is rich for an American president” implicates that he is not rich *tout court*, or not rich for an American, generally speaking).

There may be an even simpler explanation for the outlier. The assertion that Margo Dydek was tall for a woman will normally generate the implicature that she is *not* tall for a person (when the men are included in the comparison class), which is false in the present case. But while thereby an assertion of item 6 would *normally* generate a false implicature, and so would *normally* be misleading, Douven and Krzyżanowska could only assume their participants to see the falsity of the implicature by adding, in parentheses, the height of Margo Dydek (everybody knows Bill Gates, and knows

Fig. 7 Two-dimensional MDS solution with convex hulls added

that he is rich, but not so many will have heard of Margo Dydek). However, with the basketball player’s height being explicitly mentioned in the sentence, even if only parenthetically, the risk of generating a false implicature is automatically reduced to zero: the sentence, while somewhat awkwardly formulated perhaps, will have no tendency to mislead anyone into thinking that Margo Dydek was not tall for a person (being over 2 m, as the sentence asserts her height is, counts as tall by any reasonable standard). In retrospect, then, this was probably a poorly chosen item in Douven and Krzyżanowska’s materials.

At first blush, the picture appears to be more troubling for the ranked ordering items (9–12) and the temporal order items (17–20). In neither group do the items seem to hang together very tightly. More importantly still, they do not appear to form convex regions in the space. Whereas, as just mentioned, we do not find shades of blue or green among the shades of red in color space, from Fig. 5 it looks as though the ranked ordering items and the temporal order items *are* interspersed. (The fact that both types refer to some kind of *ordering* could lead one to believe that maybe these items form actually only one type of implicature, which might then be represented by a convex region. But that would be a mistake: the orderings have nothing essentially in common, ranked ordering implicatures implicitly referring to some scale, and temporal ordering implicatures explicitly referring to different points in time, even if the points in time can remain unspecified.) This becomes easier to see still when we add, as is done in Fig. 7, the convex hulls for the different types of implicatures to the MDS solution. (The convex hull of a set of points is the smallest convex set encompassing all points in the set.)

The three-dimensional MDS solution scored better on stress than the two-dimensional one, and it might be that all types of implicatures do form convex regions in three-dimensional space. This is *almost* the case, but here, too, the ranked ordering items and the temporal order items have partly overlapping convex hulls (all

Fig. 8 Three-dimensional MDS solution with convex hulls added

other convex hulls are cleanly separated from each other). This can be seen *somewhat* from Fig. 8, although it is only really clear if one rotates the figure in *Mathematica*, the software that was used to produce the plots.

So, we might be inclined to conclude that either ranked ordering implicatures or temporal order implicatures (or both) fail to constitute a natural concept, or at any rate not one as natural as the other types of implicatures. I doubt, however, whether that conclusion would be warranted. Specifically, I doubt whether we should assume that all alleged ranked ordering items and all temporal order items in Douven and Krzyżanowska’s materials generate the implicatures they were supposed to generate.

When considering an interpretation of the first dimension of the implicature space (or spaces), we already noted that some conjunctions that relate events in the wrong temporal order will nonetheless not lead hearers to make any false inferences about that order. That is simply because some events can only occur in a given order, for logical reasons, or probably more often for reasons of how the world is organized, whether physically, biologically, legally, socially, or in some other respect. In particular, item 18, about Princess Diana, will not have led anyone to believe, even if only for a moment, that she first died in a car accident and then had a divorce. And rerunning the whole MDS procedures described in the previous section but now leaving item 18 out does produce a space in which all types of implicatures form convex concepts.

This is not necessarily to say we should put all the blame on item 18. Some of the ranked ordering items may not have been as happily chosen either. For instance, it is conceivable that item 12, about Americans earning over \$200,000 a year having to pay taxes, may for some of Douven and Krzyżanowska’s participants not even have generated a weak implicature to the effect that Americans earning less are exempt from paying taxes. That is because the item is easily interpretable as making an assertion about a specific income group with no intention to suggest anything about any other income group, or so it seems.

More generally, at this point it is probably best not to make too much of the apparent clash of the temporal order and ranked ordering implicatures in the two-

and three-dimensional spaces, and rather to take the finding as motivating further research, with a richer set of materials, which is at the same time better geared to the specific purpose of constructing a conceptual space.

5 Concluding Remarks

The main question addressed in this paper was whether the various types of implicatures postulated by modern-day pragmatics constitute natural concepts. The question is an important one insofar as serious scientific theories are supposed to feature precisely such concepts. To answer this question, some preparatory work had to be done, mainly in the form of constructing an implicature space. We followed a common procedure for constructing such spaces, noting however that there was no guarantee that the procedure would work. But we were lucky and ended up with a two-dimensional implicature space that met all criteria by which conceptual spaces are commonly judged. We also obtained a three-dimensional space that appeared to fit the input data even better, although here we had some difficulty interpreting all three dimensions (a problem that might be overcome by gathering further data and rerunning the analysis).

Examination of where in our space (or spaces) the items that had served as input were located showed a tight within-type clustering of most of those items. More importantly still, items belonging to the same type tended to span convex regions in that items belonging to one type lay mostly not between items belonging to some other type. While this is not *proof* that the various types of implicatures correspond to natural concepts—given that convexity is only a necessary criterion—it is at least some first evidence that they do correspond to such concepts indeed.

Admittedly, there were some violations of the convexity criterion. The results might in fact lead one to speculate that temporal order implicatures do not constitute a natural class, or not a highly natural one (if naturalness comes in degrees). One might even be able to back this speculation up theoretically, by pointing out that there may not be a one-to-one relation between respecting temporal order in a sentence and risk of misleading one's audience by uttering that sentence, given that the latter may be prevented by world knowledge even if the sentence relates events in the wrong order. But, as noted in the previous section, this speculation is probably best not taken too seriously at the moment, given that our results were based on relatively sparse materials, which on top of that were not chosen with an MDS-kind of analysis in mind.

What we have, then, is a proof of principle that implicatures can be represented in a conceptual space, and that this can help answering an important theoretical question about them. That is good news for researchers interested in experimental pragmatics, as conceptual spaces make it easy to generate empirical predictions about which factors will determine the classification of whichever items are representable in them. And it is equally good news for advocates of the conceptual spaces framework, who are constantly looking for ways to generalize their framework to domains beyond

those of perceptual concepts. But to see exactly how much research on implicatures can benefit from the current approach, more empirical work is called for, along the lines hinted at at various junctures in this paper.²

References

- Bach, K. (1999). The myth of conventional implicature. *Linguistics and Philosophy*, 22, 327–366.
- Bartholomew, D. J., Steele, F., Moustaki, I., & Galbraith, J. I. (2008). *Analysis of multivariate social science data* (2nd ed.). Boca Raton FL: CRC Press.
- Borg, I. & Groenen, P. (2010). *Modern multidimensional scaling* (2nd ed.). New York: Springer.
- Boyd, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, 61, 127–148.
- Breheny, R. (2008). A new look at the semantics and pragmatics of numerical quantified noun phrases. *Journal of Semantics*, 25, 93–140.
- Dahl, Ö. (1985). *Tense and aspect systems*. Oxford: Blackwell.
- Decock, L., & Douven, I. (2014). What is graded membership? *Noûs*, 48, 653–682.
- Douven, I. (2010). The pragmatics of belief. *Journal of Pragmatics*, 42, 35–47.
- Douven, I. (2012). The lottery paradox and the pragmatics of belief. *Dialectica*, 66, 351–373.
- Douven, I. (2016a). Vagueness, graded membership, and conceptual spaces. *Cognition*, 151, 80–95.
- Douven, I. (2016b). *The epistemology of indicative conditionals*. Cambridge: Cambridge University Press.
- Douven, I. (2018). The rationality of vagueness. In R. Dietz (Ed.), *Vagueness and rationality*. New York: Springer, in press.
- Douven, I., & Decock, L. (2017). What verities may be. *Mind*, 126, 386–428.
- Douven, I., Decock, L., Dietz, R., & Égré, P. (2013). Vagueness: A conceptual spaces approach. *Journal of Philosophical Logic*, 42, 137–160.
- Douven, I., & Gärdenfors, P. (2018). What are natural concepts? A design perspective, manuscript.
- Douven, I., & Krzyżanowska, K. H. (2019). The semantics-pragmatics interface: An empirical investigation. In A. Capone, M. Carapezza, and F. Lo Piparo (Eds.), *Further advances in pragmatics and philosophy: Theories and Applications* (pp. 81–100) (Vol. II). New York: Springer, in press.
- Douven, I., & Mirabile, P. (2018). Best, second-best, and good-enough explanations: How they matter to reasoning. *Journal of Experimental Psychology: Memory, Learning, and Cognition*, in press.
- Douven, I., & van Brakel, J. (1998). Can the world help us in fixing the reference of natural kind terms? *Journal for General Philosophy of Science*, 29, 59–70.
- Douven, I., Wenmackers, S., Jraissati, Y., & Decock, L. (2017). Measuring graded membership: The case of color. *Cognitive Science*, 41, 686–722.
- Fairchild, M. D. (2013). *Color appearance models*. Hoboken NJ: Wiley.
- Gärdenfors, P. (2000). *Conceptual spaces*. Cambridge MA: MIT Press.
- Gärdenfors, P. (2018). Convexity is a testable criterion in the theory of conceptual spaces. In M. Kaipainen, F. Zenker, A. Hautamäki, & P. Gärdenfors (Eds.), *Conceptual spaces: Elaborations and applications*. New York: Springer, in press.
- Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.
- Grice, H. P. (1989). Logic and conversation. In M. A. Cambridge (Ed.), *His studies in the ways of words* (pp. 22–40). Harvard University Press.
- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (2nd ed.), Cambridge: Cambridge University Press.

²I am greatly indebted to two anonymous referees for helpful comments on a previous version.

- Jraissati, Y., & Douven, I. (2018). Delving deeper into color space, manuscript.
- Johnson, K. (2008). *Quantitative methods in linguistics*. Oxford: Blackwell.
- Lewis, D. K. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61, 343–377.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115, 39–57.
- Potts, C. (2005). *The logic of conventional implicatures*. Oxford: Oxford University Press.
- Psillos, S. (1999). *Scientific realism*. London: Routledge.
- Putnam, H. (1983). *Realism and reason*. Cambridge: Cambridge University Press.
- R Core Team (2017) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, <http://www.R-project.org/>.
- Scharten, R. (1997). *Exhaustive interpretation: A discourse-semantic account*, Ph.D. dissertation, University of Nijmegen.
- Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1, 54–87.
- Spychalska, M., Kontinen, J., & Werning, M. (2016). Investigating scalar implicatures in a truth-value judgement task: Evidence from event-related brain potentials. *Language, Cognition and Neuroscience*, 31, 817–840.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: Wiley.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.