Journal of Pragmatics 43 (2011) 3345–3361

Contents lists available at ScienceDirect

Journal of Pragmatics

journal homepage: www.elsevier.com/locate/pragma

Relevance theory and citations

Howard D. White *

College of Information Science and Technology, Drexel University, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Received 18 September 2010
Received in revised form 11 June 2011
Accepted 7 July 2011
Available online 8 September 2011

Keywords:
Authors
Bibliometrics
Cognitive effects
Least-effort behavior
Maximal and optimal relevance
Term frequencies

ABSTRACT

Relevance theory (RT) holds that the relevance of communications is determined by their cognitive effects and the effort needed to process them. The evidence is usually drawn from dialogues between speakers and hearers. Self-communing scholars and scientists afford a new source of real-world evidence as they cite various works over time. Authors cite works with the intention of strengthening their claims in context – a cognitive effect for them as well as for readers – and the works they choose to cite most frequently are found through least-effort behavior. Indicators of least effort include heavy reliance on self-citation, re-citation of a limited number of acquaintances and orienting figures known through reading, and the use of closely related vocabulary across texts. Such practices produce the power-law distributions of citations and terms that are ubiquitous in bibliometrics. These distributions accord well with the claimed universality of RT's Cognitive Principle on maximization of relevance. Authors maximize the relevance of citations for themselves, then optimize those citations for readers under the Communicative Principle. Examples are drawn from a set of course readings, the citation records of three authors, and word-association data. Major tenets of RT have considerable power in explaining various findings from citation research.

1. Introduction

1.1. Toward a new synthesis

In writing a new document, an author's utterances may include formal citations to other documents. This article makes use of relevance theory (Sperber and Wilson, 1995) to examine the behavior of scholars and scientists in their authorial capacity as citers of precedent work. It also devotes some attention to their readers.

S&W define relevance as a property of inputs that individuals hear or see or read or call to mind. Many of these inputs will have non-negligible cognitive effects and will require varying degrees of effort to process. As stated in Wilson (2007:4)[1]:

Relevance (of an input, in a context)

The greater the cognitive effects, the greater the relevance.

The smaller the processing effort required to derive these effects, the greater the relevance.

* Tel.: +1 215 895 2481.
E-mail addresses: whitehd@drexel.edu, whitehoward@comcast.net.

[1] This refers to Pragmatic Theory, Deirdre Wilson's online course that was publicly available for several years on the Web. It is occasionally quoted here for its terse and recent formulations of RT ideas expressed at greater length in her other publications. The lectures are unpaginated and referred to by number.

0378-2166/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.pragma.2011.07.005

The two components interact simultaneously as a ratio: *relevance = cognitive effects/processing effort* (S&W, 1996; Goatly, 1997:139). Spoken news, for instance, will be very relevant to the hearer whom it instantly and greatly affects.

Relevance theory (RT) has drawn much of its data from imaginary dialogues between persons who are jointly present, like Peter and Mary in S&W (1995). In their analyses of dialogues, relevance theorists have always focused on the reasoning of hearers, so as to explain how speakers' meanings are derived from what they actually say. However, since RT's principles are supposed to apply to communication in general, it is natural to try extending them to persons who are not jointly present, such as authors and readers. These form an obvious parallel to the basic speaker–hearer division in RT (cf. MacKenzie, 2002; Owtram, 2010).

Writing differs from most talk, of course, in that the utterances are durably stored—for example, in the oeuvres of scholars and scientists. The analysis of such oeuvres can provide something that unrecorded talk does not: evidence of factors that shape authors' attempts to be relevant. In one of the public-to-private-to-public chains of representation that Sperber writes about (2006:436–439), authors take published writings into their mental libraries, then make them explicit again when they cite. Although these are utterances of a very restricted kind, they nevertheless record communicative acts from which authorial thought processes may be inferred.

Scholars and scientists can generally say why they cite individual items (White, 2004a; Harwood, 2009), but it is doubtful that they heed their patterns of citation over time. Yet their thought processes are revealed less by what they cite in a single work than by what they cite across multiple works—across the papers, articles, chapters, lectures, and books that constitute their oeuvres. As oeuvres grow, the citation choices of authors accumulate. Those of many thousands of authors are now recorded in electronic form, notably in online indexes such as the Web of Science, Scopus, and Google Scholar. These indexes allow software to reveal patterns in what authors choose to cite, especially the works they cite repeatedly (White, 2001). The patterns relate directly to the second pillar of Sperber and Wilson's definition of relevance, processing effort.

Although S&W have persuaded many that relevance goes up when processing effort goes down, their own examples are not meant to be high in verisimilitude. They evoke differences in effort by juxtaposing plausible with extremely implausible made-up sentences. For example, as answers to a dinner guest's question about what is on the menu, they contrast "We are serving chicken" with "Either we are serving chicken or $(7^2 - 3)$ is not 46" (W&S, 2004:609). These sentences are logically equivalent, but the latter induces a Pinteresque pause while one works it out, making it a decidedly less relevant reply. In contrast, citations provide relevance theory with naturally occurring as opposed to deliberately artificial data.

When citation features are analyzed, citers appear to economize on effort. In RT, least-effort behavior is usually associated with hearers (or readers) as they interpret utterances. However, in the citation records that authors compile over time, preferences for certain kinds of effort reduction consistently appear. Citing authors may also want to reduce readers' effort, but they first reduce their own in highly patterned ways, as will be shown.

Aggregated citation data are thus a new body of potential evidence for RT. They have long been analyzed in bibliometrics–the subdiscipline of information science devoted to the quantitative study of literatures–but they are scarcely touched in linguistic pragmatics. Aggregated citations do represent a sharp break with RT's traditional data–inferences from individual sentences–by moving analysis to noun phrases and the statistical level of language. But why remain exclusively with the traditional data? In recent years, claims from RT have been statistically tested in the emerging field of experimental pragmatics; those studies involved face-to-face trials with human subjects (e.g., Van der Henst and Sperber, 2004; Gibbs and Bryant, 2008). The present paper signals another turn—to statistical data gathered unobtrusively from online records. After all, the major principles of RT should be open to corroboration on more than one analytical level.

Information scientists routinely assume citations to be "relevance related" (Saracevic, 1975:330). In their work on document retrieval, including citation-based retrieval, they take the relevance of documents to users' queries as their central concern. A synthesis of citation research and RT is thus worth attempting. It follows three large-scale introductions of RT into information science – Harter (1992) and White (2007a,b) – and a fair number of smaller ones.

Admittedly, the present continuation of that line is atypical of the RT literature in many ways. It involves a shift from speech to writing, from hearing to self-communion, from sentences to noun phrases, from ordinary to specialized discourse, and from processes in the moment to processes that take months or years. Nevertheless, relevance theory as used here can parsimoniously explain a type of behavior outside its traditional purview. The basic tenets of RT, it turns out, are highly applicable to certain findings about authors and their citations.

1.2. Readers, citers, and cognitive effects

The author's text and the reader's situation are both parts of context, defined in Wilson (2007:1) as "the set of mentally-represented assumptions actually used in interpreting a given utterance." A cognitive effect occurs when an input of newly presented information interacts with a context of existing assumptions in one of three ways: by *strengthening* an assumption; by *contradicting and eliminating* it; or by *combining* with it to *yield a new conclusion*–one derived from neither the new information alone, nor the context alone, but from the new information and the context combined (Wilson, 2007:3).

Citations can affect readers either in their own right, as packets of bibliographic data, or in combination with the sentences around them, such as a passage that attacks a cited book (Hyland, 2000:20–40). For these effects to be within the proper domain of pragmatics, according to Wilson (2007:1), they must exemplify what she calls *overt* information transmission. (Her other two categories of information transmission, covert and accidental, will be taken up shortly.) Overt information transmission must be intended by the communicator (here, citing authors) and the intent must be recognized by

the audience (here, readers). That readers can do so, as RT predicts, is shown by research on citers' motives–specifically, studies in which analysts infer motives from the prose in which citations are embedded (White, 2004a: 107–111). For example, a partial list of citers' motives from Garfield (1977:85) includes "paying homage to pioneers; giving credit for related work; identifying methods, equipment, etc.; providing background reading; correcting one's own work; correcting the work of others; criticizing previous work; substantiating claims; alerting researchers to forthcoming work." Citers seldom state their motives explicitly, and so labels such as Garfield's, though tenable, are clearly readers' inferences. They combine new information (the citations and surrounding prose) with readers' existing assumptions about motives to yield new conclusions.

While motives and effects like these vary across citations, there is at least one effect that all citers overtly intend, and that is to strengthen readers' assumptions of adequate documentation. Citers invite readers' lookups, and readers understand this intent even if few ever accept the invitation.[2] The practice of tying present to past texts, which scientists and scholars share with journalists, is a straightforward technique for building authority and trust. Citers use it at given textual points to strengthen *their own* assumptions by revealing their basis in other documents. This effect is especially plain when citers follow scholarly convention and attribute the source of a direct quote:

> If I quote something from another document, I must cite my source.
> I have quoted something from another document.
> I must cite my source.

But citers feel the same strengthening effect in all other cases as well. They themselves control documentation; passages from precedent works – e.g., a theory, a finding, a method – are texts they can point to for reinforcement if their claims are challenged. They therefore experience the strengthening effect every time they cite or re-cite a work in a new context, such that omitting the work in that context would be felt as a mild to severe documentary loss. Moreover, they expect readers to feel the same effect they do. Hence the ubiquity of citation in learned writing.

If citers intend further manipulative effects on the side, they exemplify Wilson's second kind of information transmission, which she calls *covert*. Examples would include self-promotion (citing oneself), promotion of one's in-group (citing close acquaintances), currying favor (citing the powerful), and appropriating prestige (citing the famous). A few critics have argued that covert intentions like these pervade learned writing (cf. White, 2004b). Citers, however, have never admitted them on any large scale. The one intended effect they would all admit to – and hence the one that brings citation analysis most clearly under RT – is documentation, which is universally under their control.

Passages containing citations may also affect readers in ways not controlled by citers. Because these effects are not intentional, they qualify as Wilson's third kind of information transmission, called *accidental.* For instance, authors cannot know what effects their cited items will have through inferential combinations in readers' minds. An author could bring to light a paper with strong implications for the reader's own research—a pleasant outcome unless the paper scoops or contradicts something the reader wants to publish. An author might be found to have mistranslated a cited passage (MacKenzie, 2002:131–146). An author could fail to cite works that a reader thinks should have been included (MacRoberts and MacRoberts, 1986). And so on. Outcomes like these, authors are not likely to foresee.

1.3. Citers and maximal relevance

Long before readers encounter the citations in any work, the author's take on other documents comes into play. Authors are always members of prior readerships, and in order to cite, they must consider not only the passages they are composing but also the contents of their own mental libraries, which may include works they themselves have written. Such processing falls under the First or Cognitive Principle of RT: "Human cognition tends to be geared to the maximisation of relevance." The First Principle applies to cognition in general and not just communication. Individuals, that is, will seek the most relevant inputs not only from other individuals but also from what might be called contexts of self-communion, which include their own perceptions, recollections, and inferences. In such contexts, MacKenzie writes (2002:31), "Achieving maximum relevance involves selecting the context that enables one to achieve the best possible balance of effort against effect." He goes on to quote S&W (1995:147):

> Each thought process leaves the mind in a state characterised by an initially given context and possible extensions. If we are right in assuming that the train of human thoughts is steered by the search for maximal relevance, then the mind should try to pick out, from whatever sources it has available, including its own internal resources, the information that has the greatest relevance in the initial context: that is, which has the greatest contextual effects and requires the smallest processing effort. Such information is to be sought in accessible extensions of the context, whether they involve encyclopaedic memory, the short-term memory store, or the environment.

[2] The full interpretation of citations is a two-step process. As Langham puts it (1995:361): "The function of any citation-signaller is to alert the reader to some kind of association between the citing text and the cited text. Citation-signallers may additionally, by using page-references or chapter numbers, single out a particular part of the cited text as especially relevant. This additional information can be very useful if it is genuinely the intention of the citing author that the cited text be read by his or her own readership. We have to assume, I think, that this is always any citing author's principal reason for citing, that any citation-signaller is intended as a kind of command to the reader to consult the cited text, or at least the specified part thereof; and that only by reading that cited text or cited text element will the reader stand any chance maximally of understanding the citing text."

Table 1
Two key formulations of relevance theory.

Communicative Principle (for speakers/writers)
Every utterance communicates a presumption of its own optimal relevance. An utterance is optimally relevant if: (a) it is relevant enough to be worth the audience's processing effort, and (b) it is the most relevant utterance compatible with the communicator's abilities and preferences.
Comprehension procedure (for hearers/readers)
Follow a path of least effort in deriving cognitive effects [of an utterance in a context]: consider interpretations in order of accessibility; stop when expectations of relevance are satisfied.

Examples of self-communion occur in RT (e.g., in Goatly, 1997:138; W&S, 2004:608–609), but they involve reasoning about stimuli external to the individual, such as the arrival of a train. In authorial self-communion, the stimuli originate *within* the individual and are externalized as prose on the page. Citations in a sense are responses to stimuli that authors themselves have created.[3] It is as if the speaker–hearer pairs typical of RT are no longer separate individuals (like Peter and Mary), but are lodged within each author. An internal speaker proposes possible utterances by recording them; an internal hearer acts like an editor or critic who accepts some utterances as maximally relevant while rejecting others. Such deliberations presumably account for the redraftings of documents, including the citations in them, that depend on authorial reflection rather than reader input or feedback.

My exhibits will show several ways in which authors seek "the best possible balance of effort against effect" for themselves when they cite. Audience-conscious authors will also seek this balance for readers, sensing with Wilson and Wharton (2006:1569) that "the extra effort a writer puts into redrafting a text may save the reader some effort in comprehension." But readers are multiple, absent, and dispersed; they vary in their assumptions, and many will be strangers. Authors cannot make their citations maximally relevant to them all; there is no way to discover what would best suit each individual.[4] Because of the human differences involved, the only feasible course for citing authors is to model readers' needs on what the citers themselves find maximally relevant.

Citers tend to reduce their own effort in at least three ways: (1) by favoring works they have already read – or written – over those that must be newly read; (2) by favoring works whose technical vocabularies are close to their own; and (3) by reusing a selection of these works across their oeuvres. The frequencies of reuse show what is most cognitively accessible to them over time. Their choices turn out to be egocentric as well—not in the sense of being egotistical, but in the sense of reflecting their own intellectual histories and biographical details (such as with whom they have studied or worked). Repetitive egocentrism is a powerful means of economizing on effort and thus maximizing personal relevance. The process, while not impossible to override, is the norm for innumerable citers. Except for occasional critics, readers take the results for granted.

1.4. Citers and optimal relevance

In communicating with readers, citers are guided by RT's Second or Communicative Principle; readers, by the relevance-theoretic comprehension procedure. These are seen in Table 1.

Higashimori and Wilson (1996:2) gloss *maximal relevance* as "the greatest possible effects for the smallest possible effort," and *optimal relevance* as "adequate effects for no unjustifiable effort." The argument here is that authors choose citations whose relevance is maximal for them and presumably optimal for their readers.

Authors optimize relevance in part by citing works in sufficient detail to identify them uniquely. A marker such as "Richard (1993)" in body text points to a fuller reference in footnotes or endnotes, such as "Richard, M. 1993. Attitudes in context. *Linguistics and Philosophy* 16:123–148." This string of routine phrases has implicatures that all academic readers understand. It optimizes its own value as documentation, since it implies that a particular work can be retrieved at an address in a particular source. It also implies the documentary value of the work referred to, conveying something like "You will find this item relevant, in terms of effects and effort, to the context in which it was cited." With few exceptions, such as parodies or faked scholarship, authors invite any such implicature to be checked: what else explains the convention of citing specific *editions* of works, which makes it easier for readers to find and check them? Moreover, when authors give the exact pagination of *part* of a larger work instead of citing the work as a whole, as Langham (1995) recommends, it makes checking easier still.

Suboptimal citations are undesirable because "...a stimulus is worth processing only if it is more relevant than any alternative input available at the time..." (W&S, 2004:612). Even so, the presumption of optimal relevance has an escape clause. The Communicative Principle adds in clause (b) that what seems optimally relevant to authors will be conditioned by their abilities and preferences. Given the contexts in which citations are made, readers such as editors or referees might infer, for example, that an author's choices are *not* better than any alternative. This could be a failure of the author's abilities (as

[3] These stimuli are listed in writings on citers' motives, such as Brooks (2010). Citers typically have motives beyond mere documentation—e.g., giving credit for the use of intellectual property while also making an argument more persuasive.
[4] Conversations, in which speakers and hearers can alter each other's understandings in real time, are plainly different.

when the omitted works are unfamiliar to him), but it could also result from misunderstandings or biases on the part of readers. RT explicitly recognizes that communications are not always successful (S&W, 1995:158–159).

When S&W rewrote the 1986 version of the Communicative Principle and mentioned "preferences" as well as "abilities," they did so to accommodate the speaker's preferred level of effort (1995:268): "Clause (b) of the [superseded] presumption of optimal relevance should in any case have allowed for the speaker's right to be lazy or prudish, i.e., to have her own preferences and take them into account. In later publications or oral presentation, we amended this effort clause to say that no *unjustified* or *gratuitous* effort was to be demanded" (italics theirs). The test of speakers' effort in RT is thus whether it makes utterances reasonably easy for *hearers* to process. Likewise for authors' effort. If it saddles readers with gratuitous burdens, it is insufficient effort.

What and how well authors cite are matters of preference, because the amount of work they are willing to do is not fixed. Some might obsess for days over bibliographic minutiae or spend weeks tracking down a paper they will cite only once. As a rule, however, they will prefer items that spring to mind, are quickly obtained, and whose bibliographic details are not tiresome. They may also try to eliminate gratuitous burdens for readers, such as by not citing works that are too mathematical, or in foreign languages, or redundant, or published in obscure sources. But they will tend to see their own easiest choices as sufficiently reader-oriented. Being too scrupulous on this score slows down composition intolerably.

1.5. Preferences and predictions

S&W say that speakers might prefer to be lazy. In serious cases, this could cause attempts at optimal relevance to fail. It should therefore be noted that citers *en masse* have been criticized for what amounts to lazy scholarship. For example, Latour writes (1987:33–34): "First, many references may be misquoted or wrong; second, many of the articles alluded to might have no bearing whatsoever on the claim and might be there just for display; third, other citations might be present but only because they are always present in the author's articles, whatever his claim, to mark affiliation and show with which group of scientists he identifies. . ." So understood, citations are not genuinely worthwhile (not "positive") in their effects and only *seem* relevant (S&W, 1995:264).

Latour continues (1987:40): "A given paper may be cited by others for completely different reasons in a manner far from its own interests. It may be cited without being read. . .or to support a claim which is exactly the opposite of what its author intended; or for technical details so minute that they escaped their author's attention; or because of intentions attributed to the authors but not explicitly stated in the text. . ." To get away with preferences like these, citers must assume that equally lazy readers will never check anything; if they did, the result would be "disastrous" (Latour, 1987:33).

There are indeed cases in which citers demonstrably *should* do more on readers' behalf (cf. Garfield, 1990; Wright and Armstrong, 2008). However, the economy of effort seen in the exhibits below has nothing to do with laziness. As a deep-seated preference or habit of mind, it reveals itself not only in citations that, if pursued, will cause readers gratuitous difficulties, but also in the far more numerous citations they will find perfectly acceptable.

Relevance theory and the critical view thus make different predictions. RT would predict that authors intend their citations to be optimally relevant in given contexts, and that, following the comprehension procedure, readers will usually find them so. Readers, that is, understand citations to guarantee potentially relevant lookups, and the lookups, if performed, bear out expectations. RT can thus explain why scholars and scientists maintain citation as a system. But RT also recognizes that the guarantee may fail (S&W, 1995:158–159). Impediments may lie in authorial abilities or preferences, but also in the unpredictable assumptions that readers bring to interpretation.

In contrast, the critical view predicts only that the relevance of citations will be undercut by what citers covertly or accidentally reveal. When they cite themselves or their acquaintances, for example, critics may say they do it to inflate citation counts (Meho, 2007:32). Latour, as noted, accuses citers of covertly padding their reference lists with spurious entries and distorting cited works for their own ends. In his account they also err by accident, such as when they inadvertently misrepresent works or make bibliographic mistakes. (Bad mistakes can render items unfindable.) And, according to Langham (1995), citers are accidentally ambiguous when they fail to specify *what part* of a cited work relates to the citing sentence and what the relationship is. He explores a tradeoff that many published citations produce: less effort by citer → greater effort for reader.[5]

This brings us to citers' actual behaviors over time. As stated earlier, these are marked by apparent effort reduction, in the sense that more demanding behaviors are not hard to imagine (and will be, below). The relevance-theoretic view is that citers economize on effort. The critical view is that they *over*-economize on it. The charges of critics, by no means always justified, cannot be pursued here. But, given the prominence of processing effort in RT, it is significant that the critics' major complaints relate to it in various ways.

[5] To illustrate, Langham (1995:362) contrasts two citation-bearing sentences:

(1) The planet Mars has been claimed to consist largely of low-grade pork luncheon meat (Smith, 1977).

(2) Saturn is widely believed to be composed almost entirely of bratwurst (e.g., by Schmidt, 1983:120–124).

In (1), it is not clear whether Smith (1977) should be read in whole or in part. Nor is it clear what component of (1) Smith (1977) actually relates to. One assumes it is the overall claim, but it could be the planet Mars, luncheon meat, low-grade pork luncheon meat, etc. Finally, it is not clear what Smith's relation to the claim actually is—originator, proponent, critic? In (2), all three matters are settled in the sentence itself, making less work for the reader if more for the author.

1.6. Behavior and bibliograms

Citation choices are not determined by grammar or the canons of good prose. They may be affected by, e.g., disciplinary conventions, social ties, or the injunctions of referees, but such forces are not irresistible. So, exercising their preferences over time, prolific authors could decide that they would never cite the same work twice (embracing novelty), or that they would find contexts for citing large numbers of their influences in every work of their oeuvres (embracing comprehensiveness). They could forgo citing themselves (embracing modesty) or any member of their in-group (embracing detachment). They could resolve to cite only works likely to strike readers as fresh and unforeseeable in the new context (embracing creativity). And so on.

Do they do any of these things? Generally, no. They do not refrain from citing themselves or their in-groups, nor do they routinely make "fresh and unforeseeable" choices. Moreover, regardless of verbal content, their ranked citation counts develop markedly similar structures over time. They do not plot as a roughly normal distribution. Nor do they plot as a uniform distribution, in which every writing is cited once or cited in more or less equal measure with others. Instead, in a pattern ubiquitous among citers, a small "core" of items at the top receive disproportionately heavy citation; below them, a longer "scatter" of decreasingly cited items ends in a tail of many items cited only once. This is, broadly speaking, a power-law distribution, sometimes plotted as a "reverse J," as in Fig. 1. Such skewed distributions will be interpreted here as indicators of the relevance of citations to citers (not readers), with the most frequently cited items being most relevant.

In White (2005a) I introduced the term "bibliogram" to make it easier to discuss bibliometric power-law rankings as linguistic objects. The coinage is needed because so many bibliometricians ignore verbal content and discuss only the statistics and mathematics of the frequency counts. Bibliograms co-emphasize the words being counted and open them to relevance-theoretic interpretations. They also concentrate data that would otherwise be highly diffuse.

A bibliogram is a verbal construct made when noun phrases from extended stretches of text are ranked high to low by their frequency of co-occurrence with one or more user-supplied seed terms. Each bibliogram has three components: (1) a seed term that sets a context. (2) Words that co-occur with the seed across some set of records. (3) Counts (frequencies) by which co-occurring words can be ordered high to low (White, 2005b).

Among the seed terms in the present paper are the names of citing and cited authors in pragmatics and physics, as found online in ProQuest's Dialog information service. Deirdre Wilson's online course in pragmatics is also used as a seed to rank items from her reading lists.

Dialog software counts *the different works* in which an author has cited a document (not references within works). My claim is that each work represents a new context for the document and hence a separate cognitive effect on the citer. A document cited in three of the author's works, for example, has had more effects than a document cited in one. S&W (1995:130–131) have always held that perceptions of relevance cannot be exactly measured; they are at most a matter of degree. Accordingly, citation counts are not put forward here as exact measures of effects. They nevertheless lend themselves to ordinal-level statements: the claim, for instance, that the more contexts in which a document appears, the greater its relevance to the citer's arguments.

This claim, however, presupposes ranked citation counts for a list of documents—that is, bibliograms. When citations are considered one at a time, no document is more relevant to the citer than any other. It is as if, for any document actually cited, the ratio *cognitive effects/processing effort = relevance* translates to 1/1 = 1. In the numerator, an "effect weight" of 1 stands for a

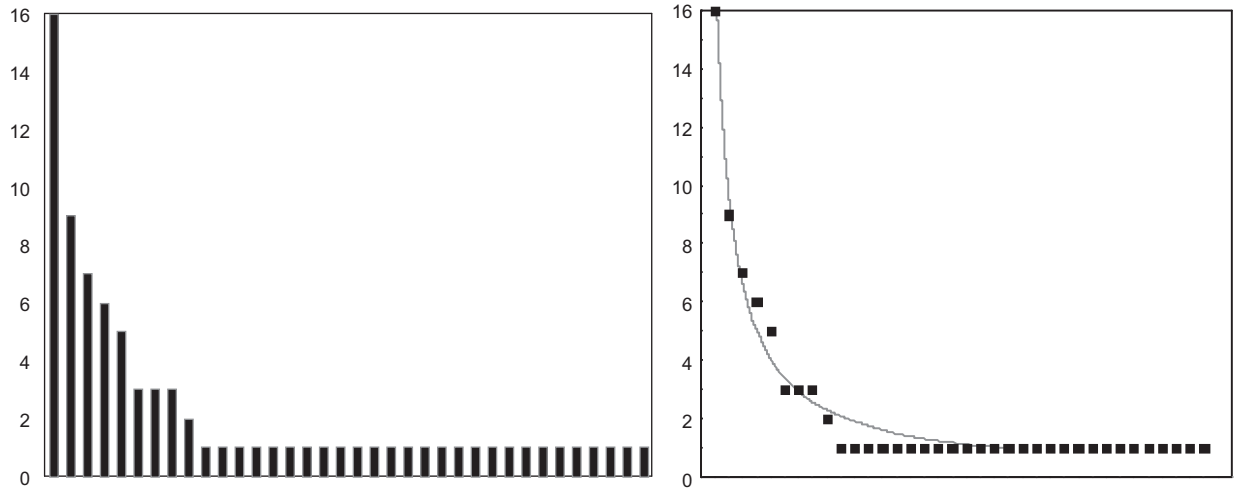

Fig. 1. Plots of citation frequencies for works by 36 individual authors or author-teams. The vertical axes show frequency of citation. The horizontal axes are ranks, from 1 at left to 36 at right (numbers omitted). The two top-ranked bylines were cited 16 and 9 times, while those in the long tail were cited once. At right, the power-law curve takes the "reverse-J" shape. Power-law distributions are highly skewed: they involve relatively few cases with high values and relatively many cases with low values. Such "few-high, many-low" distributions are found throughout bibliometrics.

constant: the effect of strengthening the *citing* work through documentation. In the denominator, an "effort weight" of 1 stands for the exact level of effort at which a particular *cited* work is acceptable to the citer. When relevance itself equals 1, the citation is made.

If a citer concludes that a possible citation is not maximally relevant (because, e.g., its bibliographic details are problematical), it is as if she senses that the effort weight in the denominator *exceeds* 1, causing relevance to drop *below* 1, the all-or-nothing threshold (for example, 1/2 = 0.5). The citation is then rejected.

Bibliograms aggregate all final 1's into citation counts for specific works, and the counts can be compared. Appreciable differences indicate that some works are more relevant to the citer than others. Focusing on effects, we can say they were cited in more contexts of assumption—more *citing* works by the citer. Focusing on effort, we can say that a particular *cited* work came most readily to mind in all those contexts. Thus, if 1 citing work/1 cited work = 1 citation, then 14 citing works/1 cited work = 14 citations to the same item. Scoring documents in this way is like gauging the strength of a habit by noting how frequently it manifests itself.

To explain why the counts take their characteristic core-and-scatter shape, we need an explanation of why the noun phrases associated with the counts occur as they do. Power-law distributions are now explained probabilistically. Newman (2005:341) uses urban population sizes and citation counts as two examples, since power-law distributions occur across a wide range of phenomena:

> Thus the probability of a city gaining a new member is proportional to the number already there; the probability of a paper getting a new citation is proportional to the number it already has. In many cases this seems like a natural process. For example, a paper that already has many citations is more likely to be discovered during a literature search and hence more likely to be cited again. Simon (1955) dubbed this type of "rich-get-richer" process the *Gibrat principle*. Elsewhere it also goes by the names of the *Matthew effect* (Merton, 1968), *cumulative advantage* (Price, 1976), or *preferential attachment* (Barabási and Albert, 1999). [Reference style adapted for this journal.]

The relevance-theoretic account of cognitive operations seen here is compatible with a probabilistic explanation of citation frequencies. That is, the more an individual has cited something for its worthwhile effects and acceptable effort, the more likely he or she will cite it again.

1.7. Re-citation of authors and works

Higher probabilities of re-citation are associated with citers' preferences for authors and works with certain characteristics. Frequent re-citation of such items builds core, while items cited markedly less become scatter. Core items appear to be easiest for citers to process over time, while scatter items are progressively less so.

At or near the top of the core are self-citations. Recall that, in Zipfian rankings of word frequencies, the top-ranked words are *short*, which Zipf (1965:63) ascribed to a human preference for words that cost relatively little effort to produce. Self-citations are not that. Rather, they reflect works that are, in every sense, *most accessible to citers*: instantaneous to recall, gratifying to use, remembered in some detail, easy to consult in full, and easy to describe bibliographically (often by recycling soft copy already on hand). We rarely learn exactly why authors self-cite, but the record shows that, as they consider what is relevant in given contexts, their preferences frequently lead them to choose their own works.

Cores also comprise a relatively small set of other authors whom citers see as having topical interests close to their own. Most citers orient themselves, through years of study, to such "coalitions in the mind" (Collins, 1998:19–53). The frequent invocation of these authors allows citers to reuse reading already done, thereby avoiding new demands on their time. (For these works, too, they may have recyclable soft copy.) Significantly, the citer may also know many members of the coalition personally (in varying degrees, with various sentiments). I hasten to add that authors cite primarily on intellectual, not social, grounds; for example, they make countless references to persons they have not or could not have met, some of whom may well be major influences on their work. Textual knowledge is thus more basic than social familiarity. But if their citees are also salient to them as acquaintances, it almost surely reduces the effort of reading them and of later calling their works to mind; otherwise, the names of such citees would not turn up so regularly in the top ranks of bibliograms (White, 2001:98–99).

Following Newman on probabilities, we can predict that the more someone is cited – acquaintance or not – the easier it is to cite that person again in new works. It would buttress this claim, however, if citers *in general* tended to agree on who goes with whom in the upper ranks of bibliograms. We can test this possibility by answering two questions and comparing the results. First, whom does an author frequently cite? Second, with whom is the same author frequently cited by others? If the two sets of names are quite similar, that indicates convergence on what might be called the easily accessible or obvious associations. A lack of overlap, on the other hand, would reveal that the citer's top choices are idiosyncratic and not comparably obvious to others. Bibliograms for this test are quickly obtainable from Dialog and will be seen in examples below.

The same test can also be put to any reader who is familiar with a specialty. In the present article, readers versed in pragmatics will be able to decide whether top-ranked items in some bibliograms are easier *for them* to process than lower or bottom-ranked items. This makes the interpretability of associations a test within the RT framework; greater ease signals their greater cognitive accessibility and relevance in the context set by a seed term.

The main feature that eases the association of core *works*, as opposed to the names of their authors, is similarities of vocabulary. Marks of similarity between citing and cited works include exact or near-exact matches in prominent terms

(e.g., in titles and abstracts), matches of terms within body texts, and clear instances of semantic relatedness even if terms do not match. Further similarities may be found in the names of journals or monographic series that publish the works, and the names of specialties and disciplines in which they are created.

If an author has discussed a work at some length *within* a text, it may have a higher probability of re-citation *across* that author's writings. An author may also routinely re-cite a work because it stands for a concept (such as "scalar implicature") that then need not be elaborated. When many authors cite the same work for this reason, Small (1978) calls the work a "concept symbol."

All of these factors lessen authors' processing effort in choosing what to cite. The result is relatively small "personal anthologies" of frequently re-cited authors and works. The point resembles one that Wilson (2007:3) takes from psycholinguistic research on frequency of use: "The more often a word, a concept, a sound, a syntactic construction or a contextual assumption is used, the less effort is required to process it." The relative frequencies with which works and authors have been cited in individual oeuvres are long-term frequency-of-use data.

In exhibits to follow, citers' tendencies to cite themselves and their acquaintances will be demonstrated with bibliograms of authors' names. (Such names can designate either oeuvres or persons: "He has a collection of Sperber"; "Sperber is in Paris.") Similarities of topic will be demonstrated, shallowly, with bibliograms of titles or of words drawn from titles. Comparisons of core items (e.g., 20 top-ranked words) with scatter items (e.g., 20 bottom-ranked words) are intended to maximize contrasts in limited space.

Notions of core and scatter naturally lead to the question of where core stops and scatter begins. Many bibliometricians divide them mathematically, segmenting ranked counts into various zones. A meaningful but non-mathematical way of doing this, seen in White (2001), is to divide a citation bibliogram into items cited at least twice (re-citations) and items cited only once (unicitations). Here, rather than attempting a strict division of zones, I will merely compare perceptible relevance in the highest and lowest ranks of some bibliograms.

2. Exhibits

2.1. Wilson's course

Table 2 has the nine documents that Wilson (2007) cited in *at least two* lectures of her 10-lecture course Pragmatic Theory. More than 40 other documents she cited in only one. What is being counted, in other words, are occurrences of works in the reading lists at the end of lectures (not mentions of works within them). The counts serve to rank the documents that co-occur with the course name, which is here the seed of the bibliogram.

Distinctions not preserved in the table are that (1) some readings are assigned, while others are merely recommended and (2) books are sometimes cited in full, but more often in part (e.g., a specific chapter). But assuming that the items in the table all had about the same cognitive effects for Wilson herself (in strengthening her assumptions of adequate documentation), the difference in ranks shows that, as she composed contexts in her lectures, she accepted certain documents for use more frequently than others. That implies they were easier for her to process, whereas little or no difference in ranks would imply that differences in effort were negligible. In the present case, however, contextual assumptions did not become easier by recurring in conversations. They recurred in self-communion. We therefore need another reason for why they keep cropping up.

It can be found in straightforward documentary characteristics. Take Wilson's top-ranked title: it is her own and Sperber's well-known book. Why it occurred to her anyone can understand; in fact, it is hard to imagine her doing the course any other way. Further items on the list by her and Sperber reinforce the notion of easy cognitive accessibility. Also reinforcing are the major book on RT by Robyn Carston, Wilson's colleague and former doctoral student at the University of London; two works on "literal meaning" (relevant to Wilson's discussion of metaphor) by François Recanati, Sperber's colleague at the Jean Nicod Institute in Paris; the essays by Paul Grice, from whose work RT derives as a counterforce; and Stephen Levinson's standard Gricean textbook.

Again, however, none of the items in Table 2 is there by iron law. Suppose its top title were Karl von Frisch's *The Dance Language and Orientation of Bees*, a book Wilson cites once. Instead of being obvious, that choice would be startling,

Table 2
Titles cited as readings for at least two lectures in Wilson's online course in pragmatics.

Authors	Titles	Counts
Sperber, D.; Wilson, D.	*Relevance: Communication and Cognition*	9
Carston, R.	*Thoughts and Utterances: The Pragmatics of Explicit Communication*	6
Wilson, D.; Sperber, D.	Relevance Theory	5
Grice, H. P.	Logic and Conversation	3
Grice, H. P.	*Studies in the Way of Words*	3
Wilson, D.; Sperber, D.	Truthfulness and Relevance	3
Levinson, S.	*Pragmatics*	2
Recanati, F.	*Literal Meaning*	2
Sperber, D.	Pragmatics, Modularity and Mindreading	2

Table 3
Frequencies of selected words and word-stems in titles of readings for Wilson's online course in pragmatics.

Relevance	28	Children	1
Communicat–	19	Codes	1
Pragmatics	15	Competence	1
Theory	14	Disambiguation	1
Irony	10	Echo	1
Cognition	9	Homophonic	1
Conversation–	8	Idioms	1
Understanding	7	Maxim	1
Thoughts	6	Mindblindness	1
Explicit	6	Nouns	1
Utterances	6	Philosophy	1
Implicat–	5	Pie	1
Truthfulness	4	Poetics	1
Logic	4	Presumptions	1
Lexical	4	Psycholinguistics	1
Meaning–	4	Rhetoric	1
Context	4	Sarcastic	1
Language	4	Schizophrenia	1
Verbal	4	Surface	1
Pretense	4	Universe	1

and people would wonder by what mental gymnastics – what effort – Wilson could make bees relevant to the entire course.[6] Or consider a more plausible example of citations meant to be "fresh and unforeseeable." Since Wilson is herself a novelist (and Sperber the son of one), why not structure the whole course around snippets from fiction that are analyzed to demonstrate points from RT? Wilson has used Jane Austen in this manner without needing to cite her formally. But would it not make an interesting reading list to see scores of formal citations to fiction writers and playwrights—authors we could then rank by frequency of use? Interesting, yes, but a huge amount of new processing effort for Wilson, who already has or can make up whatever snatches of dialogue she needs. Similarly, while there are now many introductions to RT that she *could* choose to cite (perhaps to spread credit among her exponents), they are essentially summarizing work by S&W. So why make the extra effort of weaving them into her text when she is their source and can call up what she needs from memory?

Consider, too, the vocabulary seen in Table 2. Wilson's course introduces pragmatics by focusing on RT, and, just in that small table, "pragmatics" and "relevance" each appear three times. Elaborating, Table 3 shows the frequencies of substantive words in the titles of *all* her course readings each time the titles appear. (For example, since Carston's book is cited at the end of six lectures, "thoughts" and "utterances" have counts of six.) The display is limited to words and stems that occur at least four times, and to an arbitrary 20 of the 80 or so words that occur once. The left-hand column reveals at a glance how Wilson's citations thematize the course, even on the level of single title words. Many experts in pragmatics could presumably guess the implied subject matter simply by seeing it. The miscellany at right implies nothing so clear; nor does any other set of singleton terms. Moreover, since bibliograms can be created from writings in their entirety, a ranking of lexical words from Wilson's lectures *in full* would elevate exact or partial matches with the words at left over matches with the words at right. This is merely to claim that authors tend to make their citations obviously relevant to contexts through verbal repetitions and reinforcements. "Obvious" translates here as "easy to process, highly accessible" for both writer and reader.

The evidence thus far demonstrates that this particular course is *Wilson's*; we infer from the ranked documents and vocabulary her topical priorities. If other academics (such as Geoffrey Leech, Stephen Levinson, or Lawrence Horn) taught an introduction to pragmatic theory with numerous readings, they would surely prioritize different texts, choosing most frequently the items *they* found easiest to process. What is not likely to differ, however, is the distribution of the associated counts, which would still have a power-law shape.

[6] This possibility, though remote, has an analogue in Wilson's (1991) novel, *Slave of the Passions*. Her heroine, Grace, enrolls in "a two-year MA course in human behaviour" at what is called the London Institute of Behavioural Science. She is put in a group whose tutor, a twit named Bastable, assigns the first tutorial essay on molluscs. The ensuing comedy of frustration includes this exchange between Grace and a fellow group member in a pub. He speaks first (p. 93):

'What do you think about molluscs, then?'

'Basically, I'd say they were irrelevant.'

'Ah, to you and me. It takes a mollusc to appreciate another mollusc.'

'Dr. Bastable seems to like them.'

'There you are, then.'

'Yes, but what are we supposed to *write*? I thought this was a course on human behaviour.'

'Well, it's all the same, isn't it? Molluscs, squids, people.'

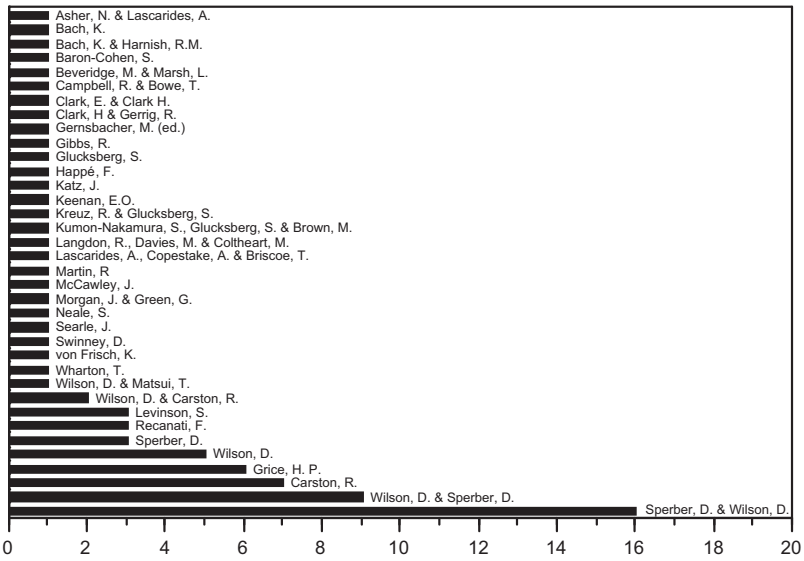

Fig. 2. This is Fig. 1 with the axes reversed for easier labeling of the bars. The frequency counts reflect total citations to works by authors or author-teams in the readings for Wilson's online course in pragmatics. The larger counts may involve more than one work by the authors named (in byline order). Here, a few bylines have many citations on the scale (higher repetition, higher cognitive accessibility, higher ease of processing, higher relevance over time). Many bylines have few citations on the scale (lower repetition, lower cognitive accessibility, lower ease of processing, lower relevance over time).

I noted that bibliometricians who simply analyze distributions of counts without looking at the full bibliogram treat this shape simply as a matter of probabilities—a kind of "natural process," in Newman's phrase, in which "the probability of a paper getting a new citation is proportional to the number it already has." It is a natural process, all right, but one driven by citers' preferred levels of effort. For many reasons, it is simply easier for Wilson to structure her course around *Relevance: Communication and Cognition* than around, say, Levinson's *Pragmatics* (not to mention von Frisch on bees), which explains their relative counts. At the same time, her book in this context is optimally relevant for her students. A content-blind analysis misses these causal explanations from RT.

Fig. 2 completes the picture by revealing how frequently Wilson cited the authors and co-authors of all 54 of her readings. Out of 81 citations to the readings, 33 are to works by her and Sperber, singly or in combination; three more are to works she co-authored with Robyn Carston or Tomoko Matsui—intellectual ties that are also social ties. Together, these counts make up 44% of her total, which suggests even more strongly than Table 2 how ease of processing is structured.

This is, again, not to imply egocentricity in a bad sense. Aside from the fact that Wilson is introducing a specialty she co-founded, legions of other academics are prolific self-citers. Linking present to past items in one's oeuvre is no more egocentric than publishing with a byline (White, 2001:102–103). And citing acquaintances is often inevitable, given the small worlds typical of research specialties.

2.2. Identities and images for Carston and Recanati

Different kinds of bibliograms can be created, depending on the seed chosen and the types of terms that co-occur with it. Two kinds I have named and studied are *citation identities* and *citation images* (White, 2001, 2004a). Both are found in Table 4, in which the seed authors are Robyn Carston and François Recanati, whose patterns are typical of innumerable other authors. The table has the top 20 ranks of their much longer bibliograms.

In citation identities, the seed is a *citing* author, and the co-occurring terms are the authors he or she has cited over time. The counts reflect how frequently an author has cited any given citee. (More precisely, they reflect the number of journal publications in which, e.g., Carston has cited at least one work by, e.g., Diane Blakemore.)

In citation images, the seed is a *cited* author, and the co-occurring terms are the authors with whom the seed has been cited over time by citers in general. On top of the image is the seed author's own citation count. The subsequent counts reflect the number of times the seed and any other author have been cited *jointly*. (More precisely, they reflect the number of journal publications in which at least one work by, e.g., Recanati has been cited with at least one work by, e.g., John Searle.) This joint citation of pairs is known as *author co-citation* (White, 2004a).[7]

Identities. The shape of identities has a least-effort interpretation regardless of verbal content. Prolific citers, that is, will seldom read and think their way to deep knowledge of everyone they cite; they do this only for certain orienting figures. But verbal content provides additional evidence of their conservation of effort.

[7] See appendix for details.

Table 4
Top 20 names, by frequency, in the citation identities and images of two seed authors in pragmatics. In each author's identity, matches with the image are italicized. Further matches exist in ranks not shown.

Carston identity		Carston image		Recanati identity		Recanati image	
Sperber D	14	Carston R	354	*Recanati F*	26	Recanati F	679
Carston R	13	Sperber D	244	*Kaplan D*	17	Bach K	202
Wilson D	12	Grice HP	190	*Sperber D*	12	Sperber D	182
Grice HP	10	Recanati F	139	*Bach K*	10	Grice HP	177
Levinson SC	8	Wilson D	139	*Searle JR*	10	Searle JR	158
Blakemore D	7	Levinson SC	129	*Grice HP*	9	Kaplan D	148
Horn LR	7	Bach K	115	Nunberg G	8	Carston R	139
Recanati F	6	Horn LR	97	Clark HH	7	Lewis D	114
Burton-Roberts N	4	Blakemore D	85	*Evans G*	7	Austin JL	107
Fodor JA	4	Chomsky N	61	Fillmore CJ	7	Perry J	104
Gazdar G	4	Stanley J	60	Fodor JA	7	Schiffer S	93
Kempson R	4	Searle JR	76	*Perry J*	7	Evans G	88
Bach K	3	Lewis D	52	*Austin JL*	6	Wilson D	88
Cappelen H	3	Gazdar G	46	Putnam H	6	Kripke S	86
Chomsky N	3	Kempson R	45	Quine WVO	6	Davidson D	84
Cormack A	3	Cappelen H	44	Ducrot O	5	Stanley J	84
Smith NV	3	Bezuidenhout A	41	*Kripke SA*	5	Frege G	79
Atlas JD	2	Lakoff G	41	*Lewis D*	5	Neale S	78
Barlev Z	2	Gibbs RW	39	Partee BH	5	Salmon N	77
Borg E	2	Fodor JA	37	*Salmon NU*	5	Soames S	76

As expected, then, we see self-citation placing the two seed authors in Table 4 at or near the top of their identities.[8] Recanati heads his. Carston's near-tie with Sperber occurs because she so often relates her own work to *Relevance: Communication and Cognition*, for which he gets first-author credit. (The Wilson course data had a similar pattern.) The different ranks suggest that, in Carston's case, it is somewhat easier for her to devise contexts in which she cites Sperber, Wilson, Grice, and herself than, say, Noam Chomsky (let alone, say, Walter Kintsch, whom she cites once). In that sense, her own oeuvre and those of her other top three choices are more relevant to her as a writer. The same interpretation holds for Recanati's ranked citees.[9]

Earlier, I guessed that experts in pragmatics could quickly infer the nature of Wilson's course from recurrent words in the titles of her readings (Table 3). Carston's and Recanati's identities permit a similar guess—that the same experts would be able to make good sense of the names as ranked in Table 4. Knowing that these are the top 20 citees for Carston and Recanati, they would presumably find the overall patterns unsurprising and could volunteer quite specific facts as to why so-and-so appears where. But, as with the low-ranked vocabulary in Table 3, they would be hard-pressed to make comparable inferences about the names in Table 5, a selection of those Carston and Recanati cite once. If present readers find they can relate Table 4 names to Carston and Recanati more easily than Table 5 names, that is consistent with the claim that some citees are more relevant to their work than others.

Images. The consistency argument can be taken further. When an image and an identity have the same seed author, they can be directly compared for content. The images in Table 4 can thus serve as a check on claims about what is going on in the identities. The key question was anticipated earlier: Are the top ranks of both filled with many of the same names? If so, the image confirms that those names occur easily to authors who also cite Carston or Recanati (or both). The data come, furthermore, from many authors, since citation images are aggregated (by Dialog software) from many identities.[10] Conversely, if the top names in the image bear little resemblance to the top names in the identity – which is possible – such confirmation is lacking.

As it turns out, we see quite strong confirmation in Table 4. Carston's image matches her identity on the 14 italicized names. Of the top 10 names in her identity, nine are duplicated. Recanati has 12 matching names in his top ranks overall, with seven duplicates in his top 10. For both authors, many further identity-image matches may be found in the ranks below the top 20.

[8] Seed authors will not *necessarily* rank first (or tied for first) in their identities; for various reasons, their self-citation count may not be high enough. But they will always rank first (or tied for first) in their images, because the writings that cite them must be at least as numerous as the writings that cite any co-occurring name.

[9] There are authors whose citations do not form identities as described here. For example, some authors take up wholly different subjects in everything they write, and so their documentation, unlike that of typical scholars and scientists, does not repeat cross-textually. It may, however, repeat *within* their works, in "op. cit." and "ibid." style (cf. Herlach, 1978). This once more shows the preferential nature of citation effort.

[10] When an image is created – say, Recanati's – Dialog automatically counts all the times he has been co-cited with others by citers in general, *including himself*. That is, any time he cites himself in an article, he necessarily co-cites himself with everyone else in his list of references. Unless this is corrected, his identity would here be compared with an image containing that very identity. However, a brief Boolean command in Dialog makes it easy to remove the set of works by Recanati from the set of works by everyone else from which the image is created. That has been done with all seed authors here, so that, e.g., Recanati's identity reflects citations by him, while his image reflects co-citations of him and other authors by everyone else.

Table 5
Selected authors cited once by two seed authors in pragmatics.

Carston identity		Recanati identity	
Asch S	1	Almog J	1
Bar-Hillel Y	1	Benbaji Y	1
Crystal D	1	Campbell R	1
Dennett D	1	Dominicy M	1
Elton M	1	Field H	1
Fauconnier G	1	George A	1
Gineste MD	1	Hintzman DL	1
Hudson R	1	Kanger S	1
Johannessen JB	1	Lakoff G	1
Kintsch W	1	Marr D	1
Leslie AM	1	Noveck I	1
Matsui T	1	Ortony A	1
Newmeyer F	1	Panaccio C	1
Pietroski P	1	Rawls J	1
Pollard C	1	Rorty R	1
Putnam H	1	Saka P	1
Reiter R	1	Schlick M	1
Strawson P	1	Tennant N	1
Tsohatzidis SL	1	Vermazen B	1
Wedgwood D	1	Wierzbicka A	1

Table 6
Top 20 names, by frequency, in the citation identity and image for a seed author in physics. In the identity, matches with the image are italicized. Further matches exist in ranks not shown.

Newman identity		Newman image	
Newman MEJ	91	Newman MEJ	5192
Watts DJ	45	Albert R	2408
Albert R	43	Barabasi AL	2178
Dorogovtsev SN	38	Watts DJ	2115
Barabasi AL	31	Dorogovtsev SN	1469
Strogatz SH	28	Pastor-Satorras R	952
Molloy M	26	Strogatz SH	742
Bollobas B	25	Erdos P	714
Faloutsos M	23	Jeong H	713
Girvan M	22	Bollobas B	613
Amaral LAN	19	Barrat A	557
Krapivsky PL	19	Guimera R	533
Wasserman S	19	Amaral LAN	509
Guimera R	17	Cohen R	498
Holme P	17	Goh KI	456
Moore C	17	Girvan M	447
Redner S	17	Holme P	443
Broder A	16	Faloutsos M	404
Cohen R	16	Krapivsky PL	382
Erdos P	16	Barthelemy M	352

2.3. Bibliograms for Newman

The previous results are not peculiar to an in-group of relevance theorists. Table 6 displays the identity and image for Mark Newman, the above-quoted physicist who specializes in research on complex networks.[11] It reveals that he has 15 duplicate names in his top 20 ranks and nine matches in his top 10—a pattern strongly resembling the one seen for the pair of authors in the far-distant field of pragmatics.

Mathematical network scientists tend to be impersonal in explaining power-law distributions, in part because such distributions are often found in phenomena not affected by human agency. But also these scientists are not given to cognitive explanations of human behavior, especially that of named individuals. For them, citation patterns are best explained by abstract probabilistic forces, such as the "cumulative advantage" of scientific papers or the "preferential attachment" of Web

[11] Newman's data were gathered in 2010 from Scisearch (the 1990 to the present file) on Dialog. His identity is based on 109 journal publications in various specialties, including many in which he was not first author. I would add that my style of analysis is best suited to publications by sole authors or very small teams. Although most of Newman's works fit this description, very large teams of authors have become common in physics and other sciences. Scores or hundreds of bylines on a paper (sometimes called hyperauthorship) do not fix responsibility for citation, and inferences about cognition are harder to draw.

Table 7
Top 20 works cited by M.E.J. Newman in his specialty. Abbreviations of journal and book titles are from Thomson Reuters. Italicized titles are books.

First author, year, journal or book	Title	Count
Watts D.J., 1998, Nature	Collective dynamics of 'small-world' networks	40
Barabasi A.L., 1999, Science	Emergence of scaling in random networks	31
Albert R., 2002, Rev Mod Phys	Statistical mechanics of complex networks	29
Newman M.E.J., 2003, SIAM Rev	The structure and function of complex networks	29
Strogatz S.H., 2001, Nature	Exploring complex networks	26
Molloy M., 1995, *Random Struct Algor*	A critical point for random graphs with a given degree sequence [chapter]	25
Faloutsos M., 1999, Comp Comm R	On power-law relationships of the Internet topology	23
Albert R., 1999, Nature	Diameter of the world-wide web	22
Dorogovtsev S.N., 2002, Adv Phys	Evolution of networks	21
Amaral L.A.N., 2000, P Natl Acad Sci US	Classes of small-world networks	19
Newman M.E.J., 2001, Phys Rev E 2	Random graphs with arbitrary degree distributions and their applications	19
Redner S., 1998, Eur Phys J B	How popular is your paper? An empirical study of the citation distribution	17
Girvan M., 2002, P Natl Acad Sci US	Community structure in social and biological networks	16
Wasserman S., 1994	*Social Network Analysis*	16
Newman M.E.J., 2001, P Natl Acad Sci US	The structure of scientific collaboration networks	15
Broder A., 2000, Comput Netw	Graph structure in the web	14
Cohen R., 2000, Phys Rev Lett	Resilience of the Internet to random breakdowns	14
Krapivsky P.L., 2000, Phys Rev Lett	Connectivity of growing random networks	14
Molloy M., 1998, *Comb Probab Comput*	The size of the giant component of a random graph with a given degree sequence [chapter]	14
Bollobas B., 1985	*Random Graphs*	12

pages. As Newman wrote above: "For example, a paper that already has many citations is more likely to be discovered during a literature search and hence more likely to be cited again." True, but Newman's example, based on the actions of citers in general, diverts attention from the intimacies of cognition implied by identity content. The probability of a work's being cited is best explained by how easily it occurs to *individual* citers. That is what produces the result for *aggregated* citers in Newman's example.

Consider the works he himself has cited most frequently in his articles, as seen in Table 7. Well-cited though these works may be, it is unlikely he found them through literature searches. Some he wrote himself. Others are by co-authors or colleagues whose papers he is in a position to read early, in their pre-citation phase. Acquaintances he cites, to name only his co-authors from Table 6, include Barabàsi, Girvan, Holme, Moore, Strogatz, and Watts. Nor is co-authorship likely to be the only personal tie in the bibliograms of Newman and the pragmatics authors above.

If, furthermore, Newman or anyone else cited publications primarily because they were already well-cited, their bibliograms would contain few or no works with low citation counts, and that is not the case (cf. White, 2004b). Even if Newman did find and cite something because it was already well-cited, that would not explain why he cites it again and again, as seen in Table 7. His favored re-citations there include a limited set of works (1) by himself or acquaintances, (2) in a few specialties of physics, mathematics, and computing, and (3) in elite journals. (*Nature* and the *Proceedings of the National Academy of Sciences* [U.S.] appear three times; *Physical Review Letters*, twice.) These preferences continue in documents below the top 20 shown in the table. The data indicate that Newman is winnowing his mental library down to a personal anthology—and indeed many of these same authors and some of the same works appear in an actual anthology he co-edited (Newman et al., 2006). The simplest explanation for his pattern is that, like the authors in pragmatics, he habitually saves on effort by using the same congenial items in context after context.

S&W (1995:77) explain such behavior as an outcome of repeated processing: "...as a result of some kind of habituation, the more a representation is processed, the more accessible it becomes. Hence, the greater the amount of processing involved in the formation of an assumption, and the more often it is accessed thereafter, the greater its accessibility." Wilson's own oeuvre and personal anthology very likely comprise items that, as representations or assumptions, she has processed longer and more intensively than the remainder of what she cites. Whenever she can use them in new works, she does so, thereby reducing her own effort. In this, she is no different from Newman and many thousands of other authors. Certain documents engage their minds over long periods. The more they have preferred an item for its relevance, the more it becomes second nature to cite it again. This gives a psychological basis for Newman's extra-cognitive claim: "...the probability of a paper getting a new citation is proportional to the number it already has."

The probabilities of citation, based on counts, can thus be clarified by applying RT to bibliograms. As soon as citers have published a few works, differences in the frequencies with which authors are cited – and hence in their probabilities over the total identity – are very likely to appear.

Table 7 also shows how the titles of Newman's most cited works convey his main specialty, which is the mathematical analysis of scientific and Internet-based networks. (Appearing four times: "random graphs"; three times: "complex networks"; twice: "degree sequence," "structure," "random networks," "small-world networks," "social network," "Internet," and "Web.") In one's own oeuvre, familiar similarities of topic make it easy to relate a present writing to earlier works. The same goes for the personal anthology of works by others. These cognitive factors prompt the reuse of works in

new contexts. They do so even when the works reused have *not* been well-cited by citers in general or published in elite journals; the latter quality filters are thus less basic in creating relevance.

Tables 6 and 7 both support inferences of least-effort processing by citers, but they show that it is achieved in more than one way. While Newman is his own most cited *author*, he did not write the *work* he cites most heavily. Authors vary in how their citations to specific works add up to counts for entire oeuvres. The self-citations of prolific authors themselves form bibliograms, in which some of their works appear repeatedly and others only once or not at all. This fact disappears when the self-citations are aggregated in their identities.

2.4. Bibliograms, associagrams, and RT

When the citation records of *multiple* authors are aggregated, as in the images for Carston, Recanati, and Newman, the patterns take the same power-law shape as those of individual authors. This latter process seems related to the one that produces the response frequencies in word association tests with multiple respondents. Clark (1970:273) writes of such tests: "The common associations – i.e., the responses other people are most likely to give – are produced more quickly than the uncommon ones. This suggests that we can attach greater importance to the fastest, more frequent associations, for hypothetically they are the product of the basic association mechanism." My chain of inference here for both individual and aggregated citers is: quicker association → less processing effort → more frequent use.

Of course, image seed names such as "Carston" and "Recanati" in Table 4 are not literally stimuli in word association tests, but they can serve as such in a thought exercise. Imagine 100 experts in pragmatics being asked what other author (or authors) they think of when they hear "Robyn Carston," then "François Recanati," and so on, down a list. I leave open how many names would be elicited per respondent, but it seems a good bet that the more allowed, the more the ranked responses will come to resemble the images in Table 4. In contexts defined by "Carston" or "Recanati," the experts will converge on certain authors as most obviously relevant, while those perceived as less so will receive progressively fewer mentions.

Actual word association tests have the same pattern as bibliograms. As I wrote in White (2005a:445): "The most compelling reason for relating bibliograms to word association norms is that the counts of the words or phrases in the two kinds of lists both have core-and-scatter distributions. It is tempting, in fact, to conjecture that bibliograms (from literatures) and lists of word association norms (from people) are internally coherent for broadly similar reasons. Accordingly, I will here call tabulated word association norms 'associagrams' to stress the parallel with 'bibliograms.'" Granted, there are also differences. For example, associagrams are created in minutes; bibliograms accrete over years. Associagrams tap general vocabularies; bibliograms, specialized vocabularies and authors' names. But the comparability of the two forms remains strong.

In the associagrams in Table 8, the data for the word "Sardine" are from 102 students (Marshall and Cofer, 1970). For "Library" and "Marriage" they are from 100 students (Edinburgh Word Association Thesaurus, n.d.). The responses for "Sardine" and especially "Library" are highly concentrated on certain words, while for "Marriage" they are more diffuse. My guess is that the responses "Sperber" and "Wilson" would be similarly highly concentrated for "Carston" and that the names associated with "Recanati" would be somewhat more diffusely distributed. Be that as it may, some responses in Table 8 occurred to many respondents; others occurred to relatively few. The top-ranked terms were judged most relevant to the stimulus term. Once the full list is seen, the reasoning that put certain terms on top is plain to everyone; less obvious

Table 8
Words associated with three stimulus nouns by two groups of *ca.* 100 students. Additional words associated a single time with Sardine and Marriage are not shown.

Sardine		Library		Marriage	
Fish	39	Books	48	Love	9
Can	24	Book	32	Sex	7
Oil	5	Work	3	Divorce	5
Sandwich	5	Quiet	2	Guidance	5
Oily	3	Silence	2	Wedding	5
Subway	3	Coffee	1	Lines	4
Crowded	2	Go	1	Church	3
Good	2	Laughter	1	Death	3
Smell	2	Lavatory	1	Licence	3
Tight	2	Must	1	Bed	2
Anchovy	1	Papers	1	Bride	2
Bones	1	Read	1	Husband	2
Close	1	Reading	1	Ring	2
Cramped	1	Service	1	Union	2
Crowd	1	Study	1	Wed	2
Horrible	1	Table	1	Wife	2
Onions	1	Throbbing	1	Men	1
Salmon	1	Working	1	Partnership	1
Salty	1			People	1

associations appear only in lower ranks. Given "sardine," for instance, people whose least-effort reply was "fish" far outnumbered those who came up with the more creative "subway."

RT thus bears on the premise of a popular television game show. *Family Feud* awards cash prizes to contestants whose responses to a stimulus phrase match any of the top responses, hidden from view, that people gave in a prior survey. Prizes are awarded, in other words, for convergent rather than divergent thinking. Suppose the phrase is "Things that go with marriage," and one contestant says "love" while another says "partnership." Even before what "Survey says..." is revealed, television viewers would be eccentric not to know that "love" is likelier to win. Just so, in a (very improbable) *Family Feud* for pragmatics experts, the phrase "Authors most frequently cited with Carston" seems likelier to elicit "Grice" than "Lakoff." The stimulus "Recanati" seems likelier to elicit "Bach" than "Frege." The "prior survey" here would be the images in Table 4, and the associations would be instantaneous relevance judgments based on experts' mental encyclopedias.[12]

Commenting on *lexical* relevance as revealed by word association tests, Miller (1996:157) writes: "Can the pooled data for a large group of people be taken as representative of the mental lexicon for any single individual? The answer is yes. The psychological validity of these associations can be tested with the priming technique..." After giving an example, he concludes: "The statistical distribution of responses to a word association test, therefore, reflects the distribution of response strengths within each individual." Although knowledge of related writings in a learned specialty is not merely lexical, Miller's generalization may hold for citation data as well. That is, the statistical distributions of response strengths in citation images, which aggregate data from many authors, may be quite similar to the distributions of strengths for individual authors in citation identities.

As a non-experimental test of this, I correlated the counts of the top 20 names in Carston's, Recanati's, and Newman's identities with the counts for the same names in their images. Many of the latter names and counts appear in Tables 4 and 6. When they do not, I was always able to find matches in the lower ranks of the images retrieved from Dialog. The resulting Pearson *r*'s are: Carston 0.86, Recanati 0.89, and Newman 0.96. All three *r*'s are highly unlikely to have occurred by chance; by one-tailed test ($n = 20$), they are statistically significant at $p < 0.000$. (A p value states the probability that an r value was produced by chance rather than by a genuine relationship of the variables.[13]) The distributions of counts in the images – Miller's "response strengths" – are thus strongly related to those uppermost in the identities. Identities tend to resemble images in both verbal content and numeric structure.

The significance of this finding, if upheld, lies in what it says about the activation of one kind of encyclopedic knowledge during long periods of time. The effort of processing inputs in this case is indicated by frequencies of responses as authors compose new contexts in their oeuvres. If we look at the most frequent responses in their citation identities, there are non-circular reasons – the qualitative characteristics of identities given in section 1.7 – for believing these responses were relatively easy to make. This outcome would be notable even if every seed author's identity were idiosyncratic. It becomes more so when, in the seed author's image, many other authors find it easy to make roughly the same responses, in roughly comparable strengths. It appears that the same cognitive processes operate throughout literature-based specialties; the easy associations for individual citers are also the easy ones for their fellow scholars and scientists. Some citations are more relevant than others because they cost those who make them less effort. The striking regularities in bibliograms thus accord well with the claimed universality of RT's Cognitive Principle.

3. Conclusion

Bibliograms are distantly related to the ranked word lists that led George Zipf to formulate his well-known "principle of least effort" (1965), and this needs a brief comment. The higher ranking words on Zipf's lists are short, easily said, and easily understood. Carston (2005:276–277) uses these facts in discussing "least effort" for speakers and hearers from a relevance-theoretic standpoint. The term for hearers, seen in the comprehension procedure in Table 1, is all-important in RT. It connotes stopping at the most plausible interpretation of an utterance in context and accepting it as the *only* interpretation of what the speaker meant. By contrast, the term for speakers connotes minimizing the syllables to be pronounced, à la Zipf—for RT, a minor consideration. Speakers may sometimes prefer "the most economical form of words," but other contexts will bring other possibilities to the fore, such as a preference for roundabout expressions or learned vocabulary. Furthermore, given the speaker's goal of being understood, effort-saving shorter words may not be fully optimal. For instance, Carston contrasts "Do you have any siblings?" with "Do you have any brothers and sisters?" "Siblings" costs the speaker less breath, but as a relatively uncommon word may be harder for the hearer to process. "Brothers and sisters" is just the opposite. It should be preferred to accommodate the hearer, unless the speaker wants "siblings" for, e.g., its greater formality.

[12] Asked what other author pops into their heads when "Strindberg" is mentioned, most literary people would probably say "Ibsen." And indeed in Strindberg's citation image in Thomson Reuters's Arts & Humanities Search, Ibsen is the author most frequently co-cited with him. Tied for second are Friedrich Nietzsche and Michael Meyer, a biographer and translator of both Ibsen and Strindberg (White, 2009).

[13] Named for the statistician Karl Pearson, *r* is a summary measure of linear (i.e., straight-line) relationship between the values of two numeric variables. It ranges between −1 and +1, with the midpoint 0 indicating no relationship. The decimal fractions seen above, which are all close to +1, indicate that identity counts tend to be high when image counts are high, and low when they are low. If the seed authors are removed from the data (because they may inflate correlations), the Pearson *r*'s are Carston 0.87, Recanati 0.62, and Newman 0.87. These latter results remain highly significant by one-tailed test ($n = 19$); for Carston and Newman, $p < 0.000$; for Recanati, $p = 0.002$. At least over some range of the higher ranks, image counts are good predictors of identity counts, which tap individual cognition.

The bibliograms displayed here have nothing to do with economy of articulation. They reveal least-effort behavior that is deeper and more personal. The authors' equivalent of least-effort speech would be primer-level writing. Bibliograms reveal instead a pervasive preference for learned vocabulary, chosen for its relevance in context rather than on the basis of length.[14] A speaker in Carston's example uses such language to impress hearers. By and large, academic authors use it simply to meet the intellectual and rhetorical standards of their fields. Within disciplinary contexts, it makes for efficient communication.

For information scientists, the language associated with seed terms in bibliograms serves to *index* the seed terms. When individual authors (rather than professional indexers) supply that language, they are in effect indexing themselves; their citation identities and ranked vocabulary lists are constructs in which they stereotype their distinctive concerns. This self-indexing can be used to retrieve their works, which is why it interests information scientists. But to RT, bibliograms are of *psychological* interest: they show the names, words, and phrases most cognitively accessible to authors. By thus focusing attention on producers of utterances, bibliograms can be related to maximal relevance and the Cognitive Principle, even before recipients – here, readers – and the Communicative Principle are taken into account.

In both shape and verbal content, bibliograms support a conjecture of Sperber's in his book *Explaining Culture* (1996:114). After introducing "the effect–effort balance" from RT in this anthropological context, he writes: "It is plausible that individuals should be equipped so as to tend to optimize the effect–effort balance not just on the input side, but also on the output side. Public productions, from bodily movements, to speech, to buildings, even when they are modeled on some previous productions, are likely to move towards forms where the intended effect can be achieved at minimal cost." Scientific and scholarly writings, from which bibliograms derive, are a prime example of output in Sperber's sense. Relevance theory bears on the authors of these writings no less than on their readers. It bears, moreover, on authors not merely as they perceive but as they actively create.

Appendix

Some technical background on the identity and image data. They were obtained in 2010 from the Thomson Reuters (TR) citation databases on Dialog—specifically from Social Scisearch combined with Arts & Humanities Search, both of which cover journals in linguistics. (When the same journal is covered in both, I have removed duplicate records.) Although the Web of Science, TR's own product, offers the same databases, the bibliograms seen here can be made only with Dialog software and its Rank command. Dialog can create most citation identities and images very quickly. Instructions for doing so are given in White (2001).

Dialog extracts cited author names from strings that TR calls "cited references," which, to minimize data-entry and storage costs, are full of abbreviations and truncations. For instance, the string "Sperber D, 1986, Relevance Communicat," taken directly from Dialog output, stands for the 1986 edition of Sperber and Wilson's *Relevance: Communication and Cognition*. The strings "Bach K, 1994, V9, P124, Mind Lang" and "Wilson D, 2002, V111, P583, Mind" give just enough data to identify articles by Kent Bach and Deirdre Wilson. Scores of strings like these would be unwieldy and often cryptic in the present article. Ranked authors' names are far more convenient to tabulate and compare.

The citations in the TR databases can be *to* documents of any kind (articles, books, book chapters, working papers, conference proceedings, and so forth). But these citations are taken only *from* the journal literature—that is, from items published in the thousands of journals covered by TR. The authors that Recanati cites in Table 4, for example, come from 46 of his journal pieces but not from any of his books.

What Dialog counts in Table 4 are the *articles* (or other journal contributions) in which citations occur and not the citations to an author *within* those articles. Some hypothetical examples: If a Carston article cites three different works by Horn, her count of citations to him increases by only one, not three. It also increases by one if her article references a single work by Horn more than once. If a Recanati article cites both the Bach and Wilson articles above, the count for Bach and Wilson as a co-cited author pair increases by one.

With different software, citations *within* works could be counted, and this would yield somewhat different results from counts of citing works. But the Dialog output serves to illustrate my present points.

The databases on Dialog allow only sole or first authors to be counted in the rankings. Thus, when Carston cites the Sperber-Wilson book above, only Sperber gets credit in the counts.

The names of thousands of authors in the citation databases appear in more than one way (such as "Grice P" and "Grice HP"), each with its separate count. I have attempted to combine the fragmented counts in the table. The Thomson Reuters databases were originated by the Institute for Scientific Information, and TR has retained ISI's practice of identifying authors only by their surnames and initials. This conflates the counts of authors with the same names—a problem especially severe with common names like "Wilson D." But the Boolean AND logic built into the Rank command serves as a disambiguator. For example, when "Wilson D" co-occurs with "Recanati F," the Deirdre Wilson of relevance theory emerges.

Changes in these conventions (such as incorporation of data from authors' books) would alter counts and rankings, but not, I think, in a way that would radically affect the arguments made here. The high-ranking names in present citation identities and images would very likely remain in the upper ranks, even if their relative standings changed a bit.

[14] If one must cite something by, say, Brigitte Endres-Niggemeyer, one does not ordinarily seek a work by an author with a shorter surname.

References

Barabási, Albert-László, Albert, Réka, 1999. Emergence of scaling in random networks. Science 286, 509–512.
Brooks, Terrence, 2010. Citer motivations [ELIS classic], Encyclopedia of Library and Information Sciences, 3rd ed., vol. 1, pp. 1038–1044.
Carston, Robyn, 2005. Relevance Theory. Grice and the neo-Griceans: A response to Laurence Horn's 'Current issues in neo-Gricean pragmatics.' Intercultural Pragmatics 2, 303–319.
Clark, Herbert H., 1970. Word associations and linguistic theory. In: Lyon, John (Ed.), New Horizons in Linguistics. Penguin Books, Harmondsworth, Middlesex, UK, pp. 271–286.
Collins, Randall, 1998. The Sociology of Philosophies: A Global Theory of Intellectual Change. Belknap Press of Harvard University Press, Cambridge, MA.
Edinburgh Word Association Thesaurus, n.d. http://www.eat.rl.ac.uk/.
Garfield, Eugene, 1977. Can citation indexing be automated? In: Essays of an Information Scientist, vol. 1, pp. 84–90. http://garfield.library.upenn.edu/volume1.html.
Garfield, Eugene, 1990. Journal editors awaken to the impact of citation errors. How we control them at ISI. In: Essays of an Information Scientist, vol. 13, pp. 367–375. http://garfield.library.upenn.edu/volume13.html.
Gibbs, Raymond W., Bryant, Gregory A., 2008. Striving for optimal relevance when answering questions. Cognition 106, 345–369.
Goatly, Andrew, 1997. The Language of Metaphors. Routledge, London and New York.
Harter, Stephen P., 1992. Psychological relevance and information science. Journal of the American Society for Information Science 43, 602–615.
Harwood, Nigel, 2009. An interview-based study of the functions of citations in academic writing across two disciplines. Journal of Pragmatics 41, 497–518.
Herlach, Gertrud, 1978. Can retrieval of information from citation indexes be simplified? Multiple mention of a reference as a characteristic of the link between citing and cited article. Journal of the American Society for Information Science 29, 308–310.
Higashimori, Isao, Wilson, Deirdre, 1996. Questions on Relevance. UCL Working Papers in Linguistics, vol. 8. http://www.phon.ucl.ac.uk/home/PUB/WPL/uclwpl8.html.
Hyland, Ken, 2000. Disciplinary Discourses: Social Interactions in Academic Writing. Longman, Harlow, England.
Langham, Thomas, 1995. Consistency in referencing. Journal of Documentation 51, 360–369.
Latour, Bruno, 1987. Science in Action: How to Follow Scientists and Engineers Through Society. Harvard University Press, Cambridge, MA.
MacKenzie, Ian, 2002. Paradigms of Reading: Relevance Theory and Deconstruction. Palgrave Macmillan, New York.
MacRoberts, Michael H., MacRoberts, Barbara R., 1986. Quantitative measure of communication in science: A study of the formal level. Social Studies of Science 16, 151–172.
Marshall, G.R., Cofer, C.N., 1970. Single-word free-association norms for 328 responses from the Connecticut cultural norms for verbal items in categories. In: Postman, Leo J., Keppel, Geoffrey (Eds.), Norms of Word Association. Academic Press, New York, pp. 321–360.
Meho, Lokmin I., 2007. The rise and rise of citation analysis. Physics World (January), 32–36.
Merton, Robert K., 1968. The Matthew effect in science. Science 159, 56–63.
Miller, George A., 1996. The Science of Words. Scientific American Library, New York.
Newman, M.E.J., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46, 323–351.
Newman, Mark, Barabási, Albert-László, Watts, Duncan J. (Eds.), 2006. The Structure and Dynamics of Networks. Princeton University Press, Princeton and Oxford.
Owtram, Nicola T., 2010. The Pragmatics of Academic Writing; A Relevance Approach to the Analysis of Research Article Introductions. Peter Lang, Bern, Switzerland.
Price, Derek J. de Solla, 1976. A general theory of bibliometric and other cumulative advantage processes. Journal of the American Society for Information Science 27, 292–306.
Saracevic, Tefko, 1975. Relevance: A review of and a framework for the thinking on the notion in information science. Journal of the American Society for Information Science 26, 321–343.
Simon, Herbert A., 1955. On a class of skew distribution functions. Biometrika 42, 425–440.
Small, Henry G., 1978. Cited documents as concept symbols. Social Studies of Science 8, 327–340.
Sperber, Dan, 1996. Explaining Culture; A Naturalistic Approach. Blackwell, Oxford.
Sperber, Dan, 2006. Why a deep understanding of cultural evolution is incompatible with shallow psychology. In: Enfield, N.J., Levinson, Stephen C. (Eds.), Roots of Human Sociality: Culture, Cognition, and Interaction. Berg, Oxford, England, pp. 431–449.
Sperber, Dan, Wilson, Deirdre, 1995. Relevance: Communication and Cognition, 2nd ed. Blackwell, Oxford.
Sperber, Dan, Wilson, Deirdre, 1996. Fodor's frame problem and relevance theory (reply to Chiappe & Kukla). Behavioral and Brain Sciences 19, 530–532.
Van der Henst, Jean-Baptiste, Sperber, Dan, 2004. Testing the cognitive and communicative principles of relevance. In: Noveck, I.A., Sperber, D. (Eds.), Experimental Pragmatics. Palgrave Macmillan, Basingstoke, Hampshire, England; New York, pp. 141–171.
White, Howard D., 2001. Authors as citers over time. Journal of the American Society for Information Science and Technology 52, 87–108.
White, Howard D., 2004a. Citation analysis and discourse analysis revisited. Applied Linguistics 25, 89–116.
White, Howard D., 2004b. Reward, persuasion, and the Sokal Hoax: A study in citation identities. Scientometrics 60, 93–120.
White, Howard D., 2005a. On extending informetrics: An opinion paper. In: Proceedings of ISSI 2005, the 10th International Conference of the International Society for Scientometrics and Informetrics. Karolinska University Press, Stockholm, vol. 2. pp. 442–449.
White, Howard D., 2005b. Bibliogram. Wikipedia. http://en.wikipedia.org/wiki/Bibliogram.
White, Howard D., 2007a. Combining bibliometrics, information retrieval, and relevance theory, part 1: First examples of a synthesis. Journal of the American Society for Information Science and Technology 58, 536–559.
White, Howard D., 2007b. Combining bibliometrics, information retrieval, and relevance theory, part 2: Some implications for information science. Journal of the American Society for Information Science and Technology 58, 583–605.
White, Howard D., 2009. Pennants for Strindberg and Persson. In: Celebrating Scholarly Communication Studies: A Festschrift for Olle Persson at His 60th Birthday. E-Newsletter of the International Society for Scientometrics and Informetrics, vol. 5-S, pp. 71–83. http://www.issi-society.info/ollepersson60/.
Wilson, Deirdre, 1991. Slave of the Passions. Picador, London.
Wilson, Deirdre, 2007. Pragmatic Theory [PLIN 2002], a course of the Department of Phonetics and Linguistics, University College London. Lectures cited: 1. The nature and goals of pragmatics. 3. Relevance: The cognitive principle. 4. Relevance: The communicative principle.
Wilson, Deirdre, Sperber, Dan, 2004. Relevance theory. In: Ward, Gregory L., Horn, Laurence R. (Eds.), Handbook of Pragmatics. Blackwell, Oxford, pp. 607–632.
Wilson, Deirdre, Wharton, Tim, 2006. Relevance and prosody. Journal of Pragmatics 38, 1559–1579.
Wright, Malcolm, Armstrong, J. Scott, 2008. The ombudsman: Verification of citations: Fawlty towers of knowledge? Interfaces 38, 125–139 (includes comments by others).
Zipf, George K., 1965. Human Behavior and the Principle of Least Effort; An Introduction to Human Ecology. Hafner, New York (Facsimile of 1949 ed.).

Howard D. White After taking his PhD in librarianship in 1974 at the University of California, Berkeley, Howard D. White joined Drexel University's College of Information Science and Technology, where he is now professor emeritus. He co-authored *For Information Specialists: Interpretations of Reference and Bibliographic Work* (1992); a later book is *Brief Tests of Collection Strength* (1995). In 1993 he won the Research Award of the American Society for Information Science and Technology (ASIST), and in 1998 he and Katherine McCain received the society's annual best paper award. In 2004 he was given ASIST's highest honor for career achievement, the Award of Merit. The International Society for Scientometrics and Informetrics honored him in 2005 with the biennial Derek de Solla Price Memorial Medal for contributions to the quantitative study of science.