

A Touch of Color: A Linguistic Analysis of the Use of Color Terms in the Language of Tourism

Viviana Gaballo

Abstract

In contemporary linguistics, color is an important area of investigation since it provides insights into human cognition and categorization of reality. Drawing on Berlin & Kay (1969) and the theories developed after their seminal work (MacLaury 1995, Wierzbicka 1996), this study aims at analyzing the evolution of color naming approaches across time and assessing the impact of color terms in their context of use, namely in the language of tourism, while bringing to the forefront the methodological difficulties behind a linguistic study of color terms based on corpus studies.

The focus of the study is on describing language in use, in a contextual perspective and by means of a corpus-based analysis of such parameters as collocation, colligation, semantic preference and semantic prosody (Sinclair, 1991 and 1996). It also evidences cases of diachronical changes in semantic preference and suggests potential applications in the area of language learning and translation studies.

Introduction

The literature on color terms in the English language is so extensively represented that it would hardly require any further study. After the publication of Berlin and Kay's (1969) seminal study on Basic Color Terms (BCT), it appears that more than 3000 works have been written that have color terms as their major focus. However, most of these studies are essentially context-free studies with an anthropological perspective. Of the few linguistic studies available, many rely on introspection as their source¹. This work aims to offer a more empirical approach and a broadened scope encompassing languages other than English. Its main aim is the evaluation of the behavior of color terms in a contextual perspective and by means of a corpus-based analysis of language in its context of use, namely in the language of tourism. In particular, conducting an empirical analysis of such parameters as collocation, colligation, semantic preference and semantic prosody (Sinclair: 2003), we will attempt to demonstrate the strong correlation between color terms and the language of tourism. At the same time, we will address some methodological issues related to core aspects of the research.

¹ Notable exceptions are Lucy (1997) and Dedrick (1998).

Although our primary goal is not to investigate color terms from the cognitive and/or anthropological perspective, we will acknowledge the interdisciplinary character of color term research and move freely among these fields while maintaining a focus on linguistic aspects.

As part of a multilingual project investigating the contextual use of color terms in the language of tourism in English, Italian and Japanese² based on corpus linguistics methods, i.e. on both the corpus-based and the corpus-driven approaches as described in Tognini-Bonelli (2001), this study presents preliminary results from two monolingual corpora which have been built to investigate each language under scrutiny here (i.e. British English and Japanese).

Besides shedding light on the behavior of color terms in the context of use of the language of tourism, the preliminary results of this study may also contribute to the areas of Language Learning and Translation Studies as teachers may use this empirical evidence and the intercultural background that emerges to train language learners and translation trainees to enrich the texts they produce to present local tourism attractions or services by adding a “touch of color”. Considering that the play on and contrast of color terms may have an impact on the choice of a tourism product, colors can therefore be thought of as a subliminal³ strategy in the promotion of tourism packages.

Theoretical background

Color is the object of many scientific studies, from natural sciences to human sciences. By this term, we understand a complex system of physical, neurophysiological, psychological and linguistic-cultural phenomena that are produced by a physical stimulus and depend on human perception and cognition. In physical terms, color is a phenomenon of light caused by how our eyes detect differing qualities of projected or reflected light. As we experience it, it is best understood as having a three-dimensional structure, where it can vary on any of the dimensions of *hue*, *saturation* or *brightness* (Thompson, 1995). *Hue* (jap. 色相 (*shikisō*)) is a circular dimension, in which the colors vary from red to orange, yellow, green, blue, purple and finally back to red again. *Brightness* (jap. 明度 (*meido*)) is orthogonal to the hue plane, and simply corresponds to how light or dark a color is, while *Saturation* (jap. 彩度 (*saido*)) corresponds to the degree to

² The study includes some of the research results presented with Antonio Manieri (Daito-Bunka University, Tokyo) at the 18th Pragmatics and Language Learning conference in Kobe, Japan in 2010.

³ Although no psychological theories or data support such a system of recognition, marketers intentionally embed subliminal stimuli—usually sexual objects, symbols, or words—in advertisements claiming these hidden, subliminal stimuli affect us in powerful ways of which we are unaware.

which a color is free from dilution by white, so that very vivid colors are high in saturation, while black, white and gray have a zero degree of saturation.

Color naming transforms the continuous color space into distinct color categories, to which we assign a word. The question at stake is how human perception succeeds in partitioning the perceptual continuum into discrete, linguistically defined categories and whether, when a language does develop a color lexicon, there is a universal tendency for languages to lexically label similar areas of the color space in similar ways (Moore et al. 2002).

Research on color naming has cultivated three different main approaches: firstly, the approach of cross-cultural comparison that emerged from the anthropological tradition; secondly, the psychophysical approach that developed out of vision research; thirdly, the developmental approach that has supplemented color naming research with insights concerning how color names and color vision are interrelated in child development.

One anthropological tradition of research has been the empirical verification of the so-called Sapir-Whorf Hypothesis, based on the assumptions of linguistic relativity – the world is differently experienced and conceived in different linguistic communities – and linguistic determination – language influences perception and thinking. Whorf's work was harshly criticized in the 1960s and 1970s, especially after the publication of Berlin and Kay's (1969) work on color terminology, in which they posit that lexical labels for basic color terms are not arbitrary but follow universal principles.

The contribution of the Berlin and Kay (1969) work in this regard is that all languages tend to acquire terms for various colors in a similar order [(black, white) > (red) > (yellow, green) > (blue) > (gray, orange, brown, pink, purple)] and that the color terms in different languages refer to the same specific areas of the color space, i.e., that there are universal constraints (or regularities) on color terminologies (Kay & Maffi, 1999). In their scheme, in languages that have the full set of eleven basic color terms (BCT)⁴, these terms would label similar areas of the color space.

⁴ The Russian term синий, 'dark blue', is considered a basic color term (BCT) as well as голубой, 'light blue'. In Hungarian, instead, two are the terms denoting RED: piros, 'red', and vörös, 'dark red'. In these languages, therefore, the full set of BCTs would account to twelve. However, as both голубой and vörös have a more limited context of use than their co-hyponyms, their belonging to the BCTs is arguable (see Paramei 2005 for a discussion on context restrictedness as the legitimate argument for contending the basic status of a salient color term). On the other, there are also languages that merge the category of 'blue' and 'green' to one Basic Color Term, a phenomenon that has been circumscribed as "grue" (Berlin & Berlin 1975; MacLaury 1997).

The universalist perspective introduced by Berlin and Kay's paradigm was followed by MacLaury's *Vantage Theory* and Wierzbicka's *Natural Semantic Metalanguage*. The former (MacLaury, 1992:141) is described as a

model of category dynamics according to which people create, maintain and change a color category by analogy to the manner in which they keep track of their own position in physical space. People form the analogy specifically at the level of fixed and mobile co-ordinates, not merely between a space and a category as general entities.

In MacLaury's theory, the introduction of similarity and distinctiveness as fixed or mobile coordinates (cf. the quotation above) determines the size of the category (1997:93). One interesting feature of MacLaury's theory is that its identification of a number of different semantic relations in the domain of color (1997:150-152). Of these relations, some are quite familiar, e.g. *near synonymy* and *complementation*. This is also true of *inclusion*, which corresponds to the term *hyponymy* in linguistics. *Coextension*, however, had not formerly been described in linguistics. Coextension in the domain of color is characterized as a relation in which the ranges of two color terms encompass each other's focus (See footnote 4 for an example).

According to Wierzbicka's *Natural Semantic Metalanguage*, all languages share about 60 semantic primitives, i.e. elementary units of meaning that are building blocks of complex and culturally specific meanings. Color, however, is not considered a universal concept – the term 'color' does not even show in many languages – and should be investigated less in terms of color universals than in terms of universals of a semantics of 'seeing'. Wierzbicka does not question the original results of Berlin and Kay's study, but she claims that landmarks are experientially grounded and that the universal pattern originates in universal landmarks, for instance the sky (in the case of 'blue'). In Wierzbicka's view (1996: 330), the meaning of the most basic color terms can be seen as quotations incorporating "fundamental and visually salient features of human environment: the sky, the sun, vegetation, fire, the sea, the naked earth, the earth covered with snow."; color terms should therefore be linked to universals of human experience (Wierzbicka, 1990:145):

[I] have explicated a number of color concepts via "environmental" concepts such as those encapsulated in the English words *fire*, *sun*, *sky*, *clouds*, *water*, and *ground*. These "environmental" concepts [...] are regarded as constructs built by human beings on the basis of their experience of life on earth.

One of Wierzbicka's most notable contributions is the introduction of the visual descriptor (Wierzbicka 2005: 218), a new perspective which opposes the Anglo-centric vision that dominated previous studies on color terms. This new perspec-

tive, by complementing the universal elements existing in all languages with culturally and linguistically relative elements⁵, creates a possible bridge between both sides of this never-ending debate.

However, more recent studies have renewed support for some of Whorf's ideas (Lucy 1992, Robertson et al. 2000, Saunders & van Brakel 2002), and even the universality of basic color terminology has been questioned (e.g., Levinson 2000). Some linguists have challenged the general findings of Berlin and Kay (1969) and Kay and Maffi (1999), suggesting that color term systems do not conform to predictable rules, at least not to the extent that they had been claimed to. Levinson (2001) suggested that the idea of color as a domain for linguistic categorization, in which we would expect to find a set of co-hyponyms each denoting specific ranges of color, might not be applicable to all languages⁶. He noted – thus confirming similar findings and considerations provided by Conklin (1955), Snow (1971) and Kuschel & Monberg (1974) – that some languages conflate other properties, such as texture or variegation, with color⁷. This contributes to difficulties in determining which color terms in particular languages are basic, because it makes it more difficult to decide which words are color terms at all.

Methodological issues

The perspective and research methodology this study is grounded on are based on empirical evidence, which strives to avoid the flaws of ethnocentrism. The analytical claims made here are derived from the methodological approach used, which provides the foundations for soundness of the arguments. Complementary methods such as descriptive studies, introspection or previous psycholinguistic and anthropological research are used to add further weight to a particular piece of evidence.

⁵ The Dani (New Guinea) terms *mili* and *mola* would then be justified and no longer be regarded as doubtful cases of basic color terms, but defined as *global visual descriptors* (Wierzbicka 2005: 218). The same applies to the terms of the *Maa* language that Payne (2003) calls *color-plus-design terms* which codify elements beyond the three classic color properties. In addition, dignity would be returned to all color terms created by metonymy, which in most languages are more common than abstract terms (Wierzbicka 2005: 239-240).

⁶ A color term in any one language denotes a specific range of color, which can be denoted differently in any other language; e.g. the Japanese color term 青い (*aoi*) conceivably corresponds to three different color terms in Italian: *blu*, *azzurro* and *celeste*. Sometimes difference is substantiated diachronically: e.g., the Italian *nero* and the French *noir* have lost the brightness feature which distinguished *niger* (more bright) from *ater* (less bright) in Latin.

⁷ In the Maasai tribe's language, *Maa*, for instance, there are about twenty terms which denote a combination of color and design, such as the color-plus-design term *ɲabóli*, which means 'brown with large dark brown stripes' (Payne 2003: 177-178).

The purpose of this study is the analysis of color terms in English and Japanese as related to the language of tourism, but not with the intent to establish the evolutionary hierarchy of color terms (Berlin & Kay, 1969), nor to determine the correspondence between color terms and the color spectrum (MacLaury (1995, 1997), or to highlight similarities and differences in the perception of colors (Hansen et al., 2007). Rather, this study concerns the description of language use within the framework of a contextual analysis of language conducted with the help of text corpora. Here lies the main difference with psycholinguistic studies (See former section). What matters is language in use, i.e. instances of actual performance – attested and genuine – and not the search for linguistic universals perceived as such based on conceived and isolated sentences produced in order to validate hypotheses. The description of the behavior of color terms therefore requires the study of their context of use of the language as it is understood by native speakers, in naturally occurring conditions. The scholar intervenes a posteriori, because the object of study, i.e. the “text”, has not been produced to be studied but for real communication purposes. A corpus-based analysis offers the opportunity to investigate authentic texts of language in use, which are representative of the language according to criteria that the linguist clarifies in advance. In a corpus-based study, texts are subjected to computational analysis, which allows a considerable amount of texts to be analyzed with minimal effort on the analyst’s part.

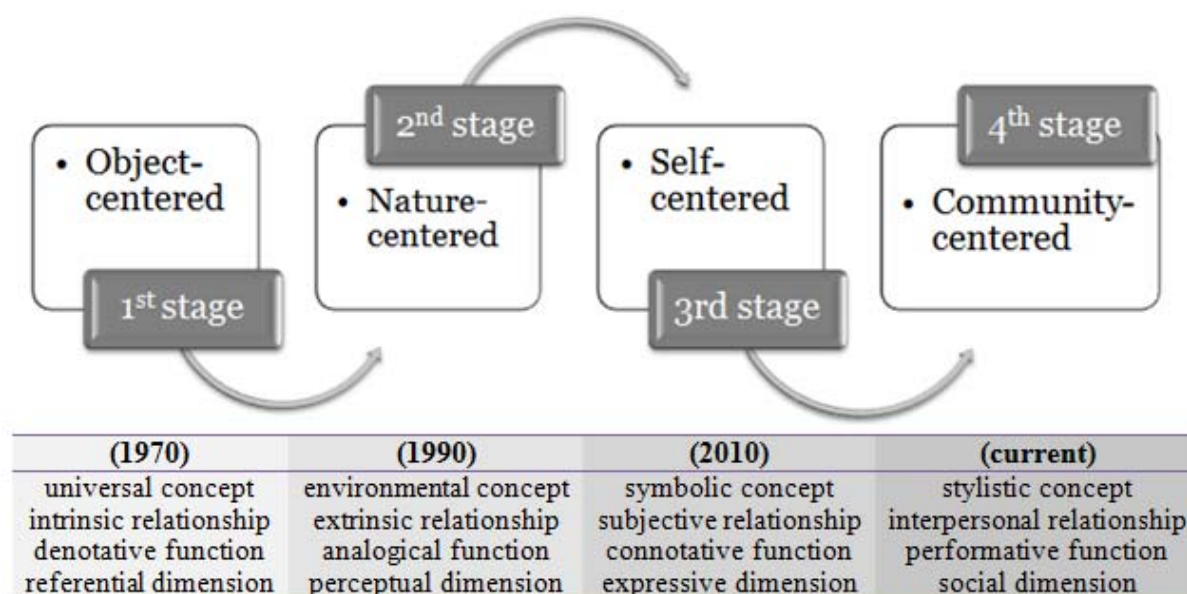
With a corpus-based study, moreover, information is easily obtained as to the frequency of words and their associations in their contexts of use, thus ensuring the empirical validity of the quantitative analysis, without imposing a linguist’s interpretation. If the attention is therefore paid to the repeated (linguistic) event, we may think that this repetition is the reflection of social practice; therefore, the corpus describes a general tendency of speakers, rather than individual acts of communication (Tognini-Bonelli, 2001: 3-4).

An additional feature of a corpus-based analysis is the attention given not only to word form or meaning, according to the parameters of traditional semantics and grammar, but also the interest in the word as the inseparable unity of form and meaning, by virtue of the interdependence between vocabulary and grammar (*lexico-grammar*), so that each meaning is associated with a form, based on the principle of co-selection.

It can be assumed, therefore, that the use of a corpus can facilitate research on the meaning of color terms, on their frequencies in relation to other words, on their lexical-grammatical constraints. Nevertheless, when corpus analysis is applied to the investigation of color terms, methodological issues arise as to how the color data contained in a corpus should be retrieved. Tognini-Bonelli (2001: 65-98) offers two different approaches: a corpus-based and a

corpus-driven approach. The former mainly involves in assessing the level of coverage and the adaptation of a particular model to the data⁸. The latter is more likely to lead to generalizations based on hypotheses that have been derived directly from the data observed⁹. Priority is given to the latter approach in this study.

Further methodological issues are connected with the evolution of technology in the past decade. Color naming techniques have been developed and shown to track closely the subjects' chromatic response functions (Hardin & Maffi 1997). However, current literature on color terms is still neglecting a new scenario that is developing along with the newest technologies globally available nowadays to communities of speakers. Web 2.0 technologies have extended the level of interactivity between actors and their resources; as a consequence, a new type of relativism has been introduced which has affected many sectors, including color naming, as illustrated in Table 1. The stages in the table illustrate the evolution of the concept of color terms since Berlin and Kay's seminal work in 1969; however, rather than self-excluding, they should be seen as an accumulative process, as every next step in the timeline also incorporates the preceding one(s) [visualized in the table by the darkening shades of gray]¹⁰.



⁸ In the corpus-based approach, reliance on data is not very rigid or systematic; for example, the frequency or absence of a given structure can be ignored in the formulation of a theoretical statement concerning the language system (Tognini Bonelli, 2001: 81).

⁹ In the corpus-driven approach, data is analyzed in a comprehensive and systematic way, avoiding any preconceived selection (Tognini Bonelli, 2001:84).

¹⁰ More on the notion of "semantic accumulation" as applied to the historical development of colour-symbolic meanings can be found in Ohtsuki's pioneering study (2000).

Table 1 – Timeline of the approaches to colors and color naming

As in vision research in general, the question about the relationship between color perception and naming is crucial¹¹. From Berlin & Kay's object-centered (1969) and Wierzbicka's nature-centered (1990) approaches color terms seem to have shifted to a self-centered approach (van Leeuwen, 2011) which is now turning to or giving more prominence to a community-centered approach¹².

Traditional color theory¹³, starting from an initial range of colors that was significantly smaller than the range of colors that appear in the physical world, tried to get around this limitation by exploring the perceptual and expressive dimensions of color, eventually opening up to a social dimension¹⁴. An example of the mentioned evolution¹⁵ is offered by the color terms *pink* (universal concept), *peach* (environmental concept), *peachpuff* (symbolic concept) and *Love Is In The Air* (stylistic concept). The denotative function of *pink* finds a new referent in the natural analogy with *peach*, which is then enriched with personal connotations in *peachpuff* and, finally, almost given sublimity in a term, *Love Is In The Air*, which combines both symbolism and lifestyle (see note 14).

¹¹ The development of emerging color terms was also analyzed by Rakhilina and Paramei (2011:130), who underlined the significance of the cognitive boundary between natural and artifact color: "Along with bestowing manufactured objects with the chromatic distinctness needed for effective communication, dyes, pigments and lights apparently evoke new qualities of perceived colour, not seen in nature, that call out to be reflected linguistically." What we argue here, however, is that the mentioned taxonomic boundary (natural objects vs. artifacts) is moving towards a more complex, socially charged system of designation which conflates artifacts, personal and interpersonal experience.

¹² Also in van Leeuwen's social semiotic analysis of color in contemporary life (2011: 85-99) the idea of coexisting trends is made clear: "Semiotically, colour has always been able to convey ideational, interpersonal and textual meanings, but at certain times and in certain periods one or other of these functions has dominated over the others, in the practice of using colour as well as in colour discourses." (van Leeuwen, 2011: 97).

¹³ We must not forget Johann Wolfgang von Goethe's contribution to color theory: his 1,400-page treatise on color published in 1810, which reformulated Newton's view of color in an entirely new way. Goethe anticipated Hering's "opponent-color" theory, which is one basis of our understanding of color vision today.

¹⁴ Van Leeuwen (2001: 85-7), in his social semantic investigation of color in contemporary life, discussed three types of identity in which color can play a role: social identity, individual identity and lifestyle identity. All three types of identity are reflected in the community-centered approach to color naming discussed in this article, with a greater emphasis on the last one, i.e. lifestyle identity. As van Leeuwen specifies (2011: 86): "The signifiers of lifestyle rest predominantly on connotation, on signs that are already loaded with cultural meaning but not yet subject to prescription or tradition. More specifically, they rest on *composites of connotation*."

¹⁵ Even if referred to variants (in *hue*, *saturation* or *lightness*) of the BCT *pink*.

Zooming in on the turn between the last two stages – i.e. the passage from a self-centered, introspective approach to color naming, towards a more open, shared concept of color which reflects not only personal, but also community lifestyles – will allow us to reconnect to the initial idea of the increased level of interactivity that the new paradigm of Web 2.0 technologies offers, and look – from the investigator’s viewpoint – into the increasing phenomenon of user-generated designations¹⁶.

Some examples of user-generated designations for the color *pink* will prove how current color naming trends extend the units of meaning (Sinclair, 1996) to encompass not only patterns formed by either the Adj/N alone or Adj+N in a variety of combinations, e.g. *Poink, pomegranate, intensity, Bikini, Flexeril, Lucious, Olivia, Punch, Watermillion, oh my blush, oh Piggy, Strawberry Lipgloss, miami hotel, romantic boy, Gold Chupa, Blushing Bride, Spike of Interest, Ready for the Floor, easy on the eyes, Funny Like You, innocent love, twin pinks, Lover’s Kiss, victoria’s secret, ♥ to ♥* or even by Adj+Adv, e.g. *deep within*, but also the full sentence structure S+V or S+V+C, e.g. *energize, memories remain, imagine roses, Be Pink, Bless Your Hearts!, Velveteen Sucks, I breathed pink, I heart you, We are the Love, my heart is yours, You Really Got Me, Love Is In The Air*¹⁷.

It is beyond the purpose of this study to further analyze the structure and meaning of these community-centered varieties of color terms. However, one distinctive feature is the impressive variety¹⁸ and number (close to infinity) of the possible designations generated by users, which makes it a daunting task for the corpus analyst to identify the occurrences of “colors” in a corpus where, theoretically, every single adjective, noun or verb (or combinations thereof) can represent or symbolize a color.

We will therefore adopt two different approaches in our corpus analysis in order to reduce dispersion of data and excessively time-consuming search,

¹⁶ A Color Name Survey was conducted by *xkcd* (<http://aram.xkcd.com/color>) in March 2010 among bloggers, who were shown the RGB color palette and asked to type a name (word or phrase) they might use for each color. An interesting result of the survey is the sheer amount of people who contributed to naming colors (over five million colors named across 222,500 user sessions) and the gender-based analysis of the outcome, which proved women’s more discreet sensitivity to colors and more flowery ways of denoting them (e.g. dusty teal, blush pink, dusty lavender, butter yellow, dusky rose) against men’s less fanciful, more prosaic attitude to color naming (e.g. penis, gay, WTF, dunno, baige [sic]).

¹⁷ Color names were crowd-sourced by using *Colors*, software for Android-based smart phones and tablets.

¹⁸ Variety which goes beyond the “bewildering range of ‘colorimetric’ methods” mentioned by van Leeuwen (2011:50) but is “as much motivated by the dominant ‘interests’ of modern urban life as the colour names of the Hanuhóo or the Ovahero were by their interests in plants and cattle.” (van Leeuwen, 2011:51).

which would spoil the benefit of working with electronic corpora. In the case of BCTs, as the color terms to be searched are known, they will be investigated by directly analyzing the relevant concordance lines. Non-BCTs, instead, shall first have to be identified in the corpus; therefore, a suitable method should be used, considering that perusing a simple wordlist would most probably be as time-consuming as working without a corpus tool. A more suitable method to elicit data from a corpus, especially if a specialized corpus, is to produce a keyword list¹⁹. A keyword list is likely to be more useful in suggesting lexical items that could warrant further investigation, as it gives a measure of saliency, while a simple wordlist only provides frequency. Sample applications of these tools are discussed in the following sections.

A final point to be made concerns potential diachronic issues that arise due to volatility of web contents. On the one hand, one cannot be sure that linguistic data collected from a web resource at a certain point of time will remain unchanged. On the other hand, changes in the texts of web resources offer us the opportunity of investigating diachronic linguistic variation. Some of the findings are presented in the next section.

As part of an ongoing project aiming at describing the occurrence of color terms in the specialized field of tourism in languages including but not limited to English, we compiled two corpora²⁰ by collecting English (E) and Japanese (J) web pages introducing tourism attractions. In order to build the corpora, we chose to gather texts that were representative of three main sectors of tourism: 1) cultural tourism, including Japanese traditional attractions, such as traditional castles, festivals, Buddhist temples, Shintoist shrines; 2) ecotourism, including natural parks, seaside and mountain tourism, spas; 3) hospitality industry, such as hotels, B&Bs, and *ryokan* (a traditional Japanese inn). The web sub-corpora assembled by the time this paper was prepared consisted of: (E) 832 written texts, with a total of 545,000 words; and (J) 268 written texts, with a total of 275,478 characters.

Our first interest in this study lies in assessing the function of color terms in the specialized language of tourism. Therefore, through concordance analysis (conducted using AntConc 3.2.1²¹), we have sketched a collocational profile of some frequent color terms according to the well-known parameters elaborated

¹⁹ A keyword list finds unusually frequent words in a specialized target corpus compared with a general reference corpus, and also displays unusually infrequent keywords (or negative-keywords) (Graeme 1998).

²⁰ The project also includes Italian – not analyzed in this study – and is open for expansion to additional languages.

²¹ This corpus analysis software is particularly suitable for investigation of idiographic languages <http://www.antlab.sci.waseda.ac.jp/software.html>.

by Sinclair (1991, 1996), namely *collocation*, *colligation*, *semantic preferences* and *semantic prosodies*. In relation to the aims of this study, the latter ones were particularly emphasized, since as “the functional choice which links meaning to purpose” (Sinclair, 1996: 88), they give us insights into the pragmatic level of written texts on tourism. Exploring the repeated linguistic events and revealing linguistic patterns “can enable us to make sense of the ways that language is used in construction of discourses (or ways of constructing reality)” (Baker, 2006: 1).

Observation of the concordance results revealed some noteworthy features related to synchronic (Japanese corpus) and diachronic (English corpus) aspects, which are reported in the next sections.

A synchronic perspective

We conducted a collocational analysis of some color terms in order to reveal salient lexical patterns. We decided to focus on basic color terms (BCTs) considering that no previous corpus analysis of Japanese color terms had been conducted before²². Since some color terms in Japanese can belong to different word classes (e.g. *aka* is a noun, *akai* is an adjective) and because of the complicated orthographic variation of Japanese language, we decided to consider terms written as common Chinese characters²³ belonging to any word class occurrence. Preliminary analysis revealed five basic color terms in the corpus with a frequency of more than ten occurrences: *midori* 'green' (64 occ.), *shiro* 'white' (40), *aka* 'red' (32), *ki* 'yellow' (27), and *ao* 'blue/green' (19).

²² In Japan, the linguistic tradition known as *kokugogaku* ('national language studies') has been dealing with color terms since 1955, when Satake Akihiro published his essay on color names in Old Japanese, demonstrating that they do not denote certain ranges on the spectrum, but rather brightness and saturation.

²³ The search in fact consisted not only of single characters, but also of groups of characters, both in the form of *kanji* only (corresponding more or less to a noun, e.g. 黒, 白, 赤, 青, 緑, 紫), and in the form of adjectives with the adjectival ending (*okurigana*) ~い, (e.g. 黒い, 白い, 赤い, 青い). As a matter of fact, a text written in modern Japanese may contain a mix of *kanji* and *kana*, with the former being used to express the semantic function of a word and the latter being used to express its grammatical function (although this is just a trend, not a rule). As a consequence, the same word can be written in various ways: e.g. the word *momiji*, 'fall leaves', appears in the analyzed corpus in three distinct forms: 179 times in *kanji* 紅葉, 20 times in *hiragana* もみじ, 17 times in *katakana* モミジ. This variation becomes more complex as also the *okurigana* may change, and the complexity is even greater in the case of compound nouns. When entering our search word in the web search engine and in the concordancer, we needed to consider all the various written forms.

緑 *Midori* / 'green'

We will start our survey with the most frequent color term: *midori*, 'green'. Looking at the left co-text, we noticed that it frequently collocates with names of trees and 'nature', 'garden', 'outside'. The right co-text often includes the agentive particle *ni* and in 9 cases *kakomareta*, the past passive form of the verb *kakomu*, 'surround/encircle', in 2 cases *oowareta*, the past passive form of the verb *oou*, 'surround', in 1 case *ahureta*, from the verb *ahureru* 'overflow', and in another case *kakusareta*, from the verb *kakusu* 'hide'. This clearly reveals a certain semantic preference for verbs referring to 'covering', to the extreme case of 'hide' (*kakusu*).

The verb *kakomareta* (or the synonym *oowareta*) is often in an appositive phrase, so we can highlight the common sequence:

- (1) *midori-ni kakomareta* 'name'
green-BY cover-PASS-PST 'name'

A similar preference is revealed in some more cases where verbs meaning 'extension in the space', e.g. *hirogaru* 'spread out', occur.

However, other interesting non-immediate collocates of *midori* are *kirei* 'fine', *utukusii* 'beautiful' or *subarasii* 'wonderful', *seisei* 'relieved', *tyoowa* 'harmony', *gensoo* 'fantastic', a set of words which are not strictly synonyms, but all express some sort of pleasant, esthetics-related feeling. Another collocation contributes to this kind of semantic prosody of *midori* in the corpus: *yutaka* 'rich/opulent' (12 occ.).

白 *Siro* / 'white'

Turning to *siro* 'white', we noticed that it has a certain tendency to collocate with conjunctions such as *ya* 'and' (6 occ.), *to* 'and' (2 occ.) or *matawa* 'or' (1 occ.), with the coordinate element being one or more color terms. This is reiterated in two other occurrences where *siro* collocates with color terms, but instead of being coordinated by particles, color terms are juxtaposed and separated by a *ten*, the Japanese comma. Other color terms occur in five concordance lines, even if not in the immediate co-text. To sum up, we found 16 cases of strong attraction between *siro* 'white' and another color term.

Looking at other collocations, we find terms related to plants, specifically names of various trees and *hana* 'flower'. Actually, in 6 out of 16 occurrences, *hana* is part of the expression *yu no hana* (or the compound form *yubana*), lit. 'flowers of hot water', which are mineral deposits left by hot springs. Verbs such as *saku* 'bloom', *tutumu* 'wrap', *tukeru* 'put' appear in the 10 occurrences related to plants.

Noteworthy, the (loanword) color term *pinku* 'pink' has a similar collocation. This lexical association of *siro* 'white' (as well as *pinku* 'pink'), with *hana* 'flowers', and *saku* 'bloom' immediately recalls the well-known Japanese tradition of *hanami*, lit. "flower viewing", the custom of enjoying the beauty of flowers, especially cherry or *ume* (*prunus mume*) blossoms, from mid-January to early May.

赤 *Aka* / 'red'

Aka shows a tendency to co-occur with other color terms, e.g. *ki* 'yellow'. As for verbs near the node *aka*, *someru* 'to dye' and its intransitive counterpart *somaru* are found very often, usually in a colligational sequence like:

- (2) *aka-ni someru/somaru*
red-ADV dye-TR/INTR

Other verb collocates include *irozuku* 'change color' and *irodoruru* 'color', both related to the semantics of changing color and both in the same lexical-grammatical relation as example 2. Another concordance line is also close in meaning:

- (3)
シラカバやダケカンバ、ナナカマドなど、黄色や赤の鮮やかなドレスを身にまといます。

Here the node *aka* is associated with *kiiro* 'yellow' and with the expression *doresu o mi ni matoimasu* 'be dressed in', i.e. another (metaphorical) way of referring to the fact that trees change color in the fall. This semantic preference for 'changing color' is expressed also by using the verb *naru* 'to become', preceded by *akaku*, the adverbial form of the adjective *akai* 'red'.

Twenty concordance lines show strong collocation with names of trees, as we have already seen in example 3, and as is illustrated in the following examples:

- (4a)
溪流には赤や黄に染まった木々が兩岸からせり出し、水の流りに映えます。

- (4b)
イタヤカエデの赤が浮かび上がり、睡蓮沼では山々を背景に紅葉が広がり、記念写真を撮る観光客でにぎわいを見せます。

- (4c)
赤や黄色に染まった樹木と滝の白色のコントラストが絶景です。

In example 4b we also find *momizi* 'fall leaves' (lit. 'red leaves', read also *kooyoo*), which is particularly interesting since it is written in Chinese characters meaning 'crimson leaves', more generally 'colorful fall leaves'. *Momizi* are to the Japanese fall what cherry blossoms are to spring; the viewing of fall leaves has

been a popular activity in Japan for centuries, and it still draws large numbers of travelers to famous *kooyoo* spots. In addition to this, the collocation between *aka* and tree names is so strong that *aka* is often a member of a compound with a tree name, as in *akamatu* 'red pine'.

The presence of red in these examples is really impressive and looking at a more distant co-text, we often find words related to the semantic field of 'beautiful' (*utukusii* 'beautiful', *gensoo* 'fantastic'). Therefore, as with *midori* 'green', we can again elicit semantic prosody that connotes the promoted experience as being an enjoyable and pleasing one.

黄色 *Kiiro* / 'yellow'

Ki or *kiiro*, the color term for yellow shows distinct tendencies. In 15 out of 27 occurrences, it collocates with other color terms, especially (11 times) with red-area terms. In 3 cases, instead, *midori* 'green' appears. In the right co-text there is a frequent occurrence of constructions with the verb *someru* 'dye', followed by *irozuku* 'change color' and by *naru* 'become' construction. In the left co-text, collocations with plant names (in particular, hypernyms and hyponyms of tree) prevail. It seems that also *kiiro* 'yellow' displays the same contextual behavior as *aka* 'red', although affective semantic prosody distinguishes the two terms: no word from the semantic area of 'beautiful' is found in the vicinity of *kiiro*.

青 *Ao* / 'blue/green' ('grue')

Ao 'blue/green' is not very frequent in the corpus: just 19 occurrences, often with *umi* 'sea' or *sora* 'sky'. What has to be stressed about *ao* is its semantic prosody of 'beautiful', as expressed by collocates such as *utukusii* 'beautiful', *kirei* 'nice', *migotona* 'admirable', or of 'harmony', as expressed by *tyoowa* 'harmony', *odayakana* 'quiet', and *yoi kanzi* 'good feeling'.

The table below provides an overview of the collocational profiles of frequent color terms in the Japanese sub-corpus. The semantic preference of the colors to collocate with expressions related to a natural setting is quite clear.

In addition, expressions of beauty and harmony occur with a certain frequency in the more distant co-text, which means that the colors attract words indicating pleasant and enjoyable experiences, which contributes to their establishing a definitely positive semantic prosody.

Table 1: Collocational profile of frequent color terms in the Japanese sub-corpus

Color term		青 'blue'	緑 'green'	白 'white'	赤 'red'	黄色 'yellow'
Immediate co-text	N.	'sky' 'sea'	trees 'nature' 'park' etc.	trees other colors	trees other colors	plants other colors
	V.	No verb	'surround' 'cover' 'hide'	'bloom' 'wrap'	'dye' 'change color'	'dye' 'change color'
SEMANTIC PREFERENCE			NATURAL SETTING			
Non-immediate co-text	adjectives in the semantic area of 'beauty' adjectives and nouns in the semantic area of 'harmony'					No adjective
SEMANTIC PROSODY		PLEASANT/ENJOYABLE EXPERIENCE				NEUTRAL

A diachronic perspective

Besides conducting the customary collocational analysis of some of the color terms selected from the keyword list, we also had the opportunity to compare a sub-set of the British English corpus, the UK-Hotels sub-sub-corpus²⁴, with the corresponding corpus collected five years before²⁵. The objective was to reveal any salient lexical change that may have taken place in the past five years in the use of colors to promote hotel accommodation in the UK. The findings are worth discussing.

After purging the results of all unchanged or irrelevant instances (nearly half of the texts), we focused on the remaining diachronic data (1997 texts are referred to as T1; 2012 texts as T2) and were able to evidence three phenomena that we called²⁶: 'de-coloring', 'over-coloring', and 'clustering'.

²⁴ The UK-Hotels sub-sub-corpus contains texts downloaded in 2012 from the web sites of 325 hotels in the UK.

²⁵ This corpus contains texts downloaded in 1997 from the web sites of the same hotels as the ones downloaded in 2012.

²⁶ As corpus compilation is still ongoing, the preliminary findings presented here shall eventually have to be corroborated after completion of the corpus. However, considering that the relevant data refer to nearly half of the UK-Hotels files, the preliminary findings already outline well-defined trends.

De-coloring

By de-coloring we mean the tendency to minimize or totally leave out color terms in T2 as compared to T1. This tendency is very pronounced in the corpus: color terms are either totally omitted, as in the example below²⁷, or downgraded in their vividness or purity (e.g. *subdued hues*), or at most replaced by implicit references (e.g. *Irish linens*).

T1	T2
<p>The bedrooms feature a <i>subtle color</i> scheme of taupe, brown and cream in <i>leather</i>, <i>walnut</i> and <i>suede</i>, alongside <i>stainless steel</i> and <i>sandblasted glass</i> flat screen TV's, DVD's, Hi-Fi's and oversize beds are fitted as standard in deluxe rooms, while the sexy bathrooms feature <u>Philippe Starck</u> fittings against <i>limestone</i> mosaics.</p>	<p>A haven for the weary traveler! Wrap yourself in a waffle bathrobe, flick on the TV and make yourself a hot drink, or order something special from room service. Snuggly duvets with hand made mattresses together with non allergenic pillows and luxurious bed linen offer one of the best night's sleep you can have anywhere. If you have to work during your stay, we'll make it stress-free with wireless internet throughout the hotel and a large work desk in every room.</p>

The example above clearly shows how the vivid description of bedrooms in T1 has given way to a dreary, even aseptic scene (e.g. *non allergenic pillows*) in T2, whose complete lack of any direct or indirect mention to colors testifies of the shift from the referential function of the colorful language describing the physical properties of the rooms (T1) to the conative function of the colorless language, which engages the addressee directly, and is best illustrated by the use of vocatives and imperatives (T2).

Over-coloring

The presence of both 'de-coloring' and 'over-coloring' in the same corpus might seem quite improbable; however, the data in the *UK-Hotels* corpus confirm that the latter is almost totally confined to a particular category of hotels: "boutique" hotels. Boutique hotels are small, trendy properties marketed on bringing unique, location-specific experiences to travelers. They are also called lifestyle or design hotels²⁸. The overabundant color terms used in the description of the boutique hotel experience are in line with the extreme attention paid to details and commitment to luxury quality sustained by the underlying philosophy: "In

²⁷ Italics, underlined and bold typefaces in the examples are the author's choice. Color terms are in bold, color-related words are italicized (e.g. materials, etc.), while co-branding (see note 28) products are underlined.

²⁸ The term itself is a little fuzzy, particularly as major hotel brands have adopted it as a marketing term, and the hotel industry does not recognize a clear-cut definition. Most boutique hotels, however, do share some common attributes in their design, style and appeal.

the business of making dreams a reality, unable to accept the ordinary”. An example of ‘over-coloring’ is provided below:

T1	T2
<p>The rooms at XXX all have a unique and individual design, ranging in both styles and colours. Rooms feature sumptuous beds and furnishings, high ceilings, and antique prints; providing the guest the utmost in luxury.</p> <p>The guest rooms provide all the features you would expect from 5-star accommodations, including a mini-bar, wireless Internet access, and pay-per-view films. In the gorgeous bathrooms you will also find an array of deluxe toiletries.</p>	<p>The <i>colour</i> schemes are daring, stunning and dramatic, all designed with a vibrant spirit, the passion for travel and life, as it should be, romantic and thought provoking. Black and mustard, rich cardinal reds, lavender, vanilla, washes of tea rose and a room that is white on white on white, offering style and elegance to the discerning traveller.</p>

From the standard, denotative description of apparently colorless bedrooms in T1, the text in T2 is unexpectedly expressive, supercharged as it is with a frenzy of color terms, which appeal to the traveler’s unconscious and emotions as well as to their reason. Unlike texts depicting standard hotels where BCTs are rarely used and ‘elaborate color terms’ (e.g. *taupe*, *cream*, *limestone*) are more likely to appear, texts depicting luxury, “boutique” hotels make stunning use of the very basic color terms (e.g. *black*, *red*, *white*).

Clustering

Lexical clusters²⁹ – i.e., collections of related words that possess statistically similar linkage structures with other words – emerge in both T1 and T2, as shown in the example below. However, discursive clusters, i.e., ‘clouds’ of color-related keywords, seem to have become less dispersed in T2, where colors conflate with color-related items, thus reducing dispersion and conceptual distance. An example of ‘clustering’ is provided below:

T1	T2
<p>The <i>colours</i> used throughout the bedrooms and reception areas - striking combinations of Chinese blue, mustard, plum and stone - are derived from the original 19th Century floor tiles which you will see in the entrance hall.</p> <p>[...]</p> <p><i>Curtains</i> and <i>bed hangings</i> are <i>natural linen</i>,</p>	<p>All with <u>Beaudesert beds</u> featuring hand made mattresses along with an antique writing desk and separate table to enjoy breakfast without disturbing your work. The <i>colours</i> used throughout the bedrooms are mustard, plum and stone, derived from the original 19th Century floor tiles of the entrance hall. A <i>limestone</i> bathroom fea-</p>

²⁹ The idea of *lexical clusters* is based on Deniz Yuret’s lexical attraction model and Steven Larson’s clustering of statistically similar data. For an application of the idea to corpus analysis, see Raymond Hickey’s “Corpus Presenter: Software for Language Analysis” (2003).

whilst the <i>headboards</i> and <i>valances</i> are hand embroidered. The <i>beds</i> are dressed with <i>cotton sheets</i> , large square <i>feather pillows</i> , and traditional <i>pure woollen blankets</i> .	tures luxury <u>Penhaligon toiletries</u> .
---	---

A fourth phenomenon, *co-branding*³⁰, is only indirectly linked to, but equally important in the use of color terms in the language of tourism, and hotel accommodation in particular. Co-branding is the practice of using multiple brand names together on a single product or service. In hotel and accommodation management, it reflects the market-driven need to strategically exhibit the brand of specific products used in hotel rooms in order to captivate prospective guests (e.g., Frette linen, Bennison fabrics, Penhaligon toiletries, Molton Brown bath products, Bang&Olufsen TV). All co-brands are purposefully selected to deliver the perfect balance with the “house” colors of the hotel

Conclusions

The purpose of this study was to investigate the use of color terms in the specialized language of tourism with an empirical approach and from a contextual perspective. The study presented preliminary results from an ongoing multilingual project investigating the contextual use of color terms in the language of tourism in English and Japanese based on corpus linguistics methods, i.e. on both the corpus-based and the corpus-driven approaches as described in Tognini-Bonelli (2001).

After reviewing the milestones of research on color terminology across languages – from the seminal work by Berlin and Kay, *Basic Color Terms* (1969), through MacLaury’s *Vantage Theory* (1995) and Wierzbicka’s *Natural Semantic Metalanguage* (1996) to Levinson’s (2001) unpredictability of color terminology – we introduced a new type of relativism in color naming, induced by the increasing phenomenon of user-generated color names made possible by the latest Web 2.0 technologies. The current development in approaches to color naming was also outlined to include the new, present-day notion of symbolic and stylistic concepts, which highlight expressive and social dimensions in the conveyed meaning of a color term along with a subjective, as well as interpersonal relationship with the denoted field.

Among several tendencies of color terms, this paper focused on the pragmatic relevance of colors in tourist texts, and investigated it from both a synchronic and a diachronic perspective.

³⁰ Co-branding, also called brand partnership, is a marketing strategy whereby two or more companies form an alliance to work together, creating marketing synergy.

Over the past few years, improvements in computer technology have made it possible to handle huge amounts of linguistic data available in the Net, on a short timescale; however, the inherent volatility of web contents remains a key dilemma in the management of web resources. On the other hand, changes in the web texts collected across years offer us the opportunity to investigate linguistic variation. We therefore compared a corpus of UK hotels' web pages collected in 1997 with the corresponding corpus of web pages (same hotels) collected in 2012 to investigate any variation in the language used to describe the accommodation and services offered. Three main phenomena were identified: *de-coloring* (the tendency to use a 'colorless' language in current texts), *over-coloring* (the tendency of "boutique" hotels to abound in colorful descriptions), and *clustering* (the tendency for colors and color-related items to co-occur and to be present in higher frequencies in particular sections of text). A fourth phenomenon was also observed that is applicable, although only indirectly related to colors: *co-branding* (the tendency to use multiple brand names together on a single product or service).

The study also proved that the strong correlation of tourism, seasons and nature in Japanese. Each period of the year seems to be dominated by a certain hue: from the white (*siro*) and pink (*pinku*) of spring *hanami*'s sceneries, through the lush green (*midori*) of nature in the summer, to the red (*aka*), with its nuances and gradations, of the dying nature in the fall. In other words, a change in season brings about a change in the dominant color in tourist text descriptions, through a process that include all the main basic color terms except black (*kuro*). The presence of adjectives related to the semantic field of beauty, wonder and peace is a key to interpret the positive disposition created by the semantic prosody that colors acquire in tourism texts. We therefore reiterate that the Japanese tend to recreate the emotive experience of visiting the promoted tourist location through colors and that color terms, with their semantic prosody, play a central role in attracting potential visitors to the locations advertised.

This study contributes to our understanding of how color terms can affect the general economy of tourist texts and their emotional impact on readers. Hence, writing for tourism with the intent to introduce tourist attractions and promote recreational services should take into account a "touch of color", considering that the play on and contrast of color terms may have an impact on the choice of a tourist product. A thought through choice of color terms can therefore pursue a subliminal strategy in the promotion of tourist packages.

References

- Berlin, Brent and Kay, Paul, *Basic Color Terms. Their Universality and Evolution*, Berkeley, University of California Press, 1969.
- Bornstein, M. H., Color vision and color naming: a psychophysiological hypothesis of cultural difference. *Psychological Bulletin*, 80.4, 1973.
- Bowerman, Melissa and Levinson, Stephen C., *Language acquisition and conceptual development*, Cambridge, Cambridge University Press, 2001.
- Conklin Harold C., Hanunóo Color Categories, *Southwestern Journal of Anthropology*, 11, 1955.
- Dann, Graham M.S., *The Language of Tourism: A Sociolinguistic Perspective*, Wallingford, CAB International, 1996.
- Dedrick, Don, *Naming the Rainbow: colour language, colour science, and culture*. Dordrecht, Kluwer, 1998.
- Eco, Umberto, *Trattato di semiotica generale*, Milano, Bompiani, 1975.
- Hansen, T., Walter, S., & Gegenfurtner, K. R., Effects of spatial and temporal context on color categories and color constancy, *Journal of Vision*, 7(4), 2007.
- Hardin C.I. and Maffi, Luisa (eds.), *Color Categories in Thought and Language*, Cambridge, Cambridge University Press, 1997.
- Jafari, Jafar (ed.), *Encyclopedia of Tourism*, London, Routledge, 2000.
- Levinson, Stephen C., Yéli Dnye and the Theory of Basic Color Terms. *Journal of Linguistic Anthropology*, 10.1, 2001.
- Kay, Paul, Synchronic variability and diachronic change in Basic Color Terms, *Language in Society*, 4.3, 1975.
- Kay, Paul, and Maffi, Luisa, Color Appearance and the Emergence and Evolution of Basic Color Lexicons. *American Anthropologist*, 101, 1999.
- Kay, Paul, Berlin, Brent, Maffi, Luisa, and Merrifield, William R., Color Naming Across Languages. In C. L. Hardin & L. Maffi (eds.) *Color Categories in Thought and Language*, Cambridge, Cambridge University Press, 1997.
- Kay, Paul, Berlin, Brent, and Merrifield, William R., Biocultural implications of systems of color naming. *Journal of Linguistic Anthropology*, 1, 1991.
- Kay, Paul and Regier, Terry, Language, thought and color, Recent developments, *Trends in Cognitive Sciences*, 10.2, 2006.
- Kennedy Graeme, *An Introduction to Corpus Linguistics*, London and New York, Longman, 1998.
- Kuschel, Rolf and Monberg, Torben, 'We don't talk much about colour here': A study of colour semantics on Bellona Island, *Man*, 9, 1974.
- Lucy, John A., *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*, Cambridge, Cambridge University Press, 1992.
- Lucy, John A., The linguistics of "color", in C.L. Hardin & L. Maffi, eds., *Color categories in thought and language*, Cambridge, Cambridge University Press, 1997.
- MacLaury Robert E., Vantage Theory, in J.R. Taylor and R.E. MacLaury (eds.), *Language and the Cognitive Construal of the World*, Berlin and New York, Mouton de Gruyter, 1995.
- MacLaury, Robert E., *Color and Cognition in Mesoamerica: Construing Categories as Vantages*, Austin, University of Texas Press, 1997a.
- MacLaury, R. E., Ethnographic evidence of unique hues and elemental colors, Commentary on Saunders and van Brakel, *Behavioral and Brain Sciences*, 20.2, 1997b.

- Moore, Carmella, Romney, A. Kimball, and Hsia, Ti-Lien, Cultural, gender, and individual differences in perceptual and semantic structures of basic colors in English and Chinese, *Journal of Cognition and Culture*, 2.1, 2002.
- Ohtsuki, M., *A cognitive Linguistic Study of colour Symbolism*, Tokyo, Institute for the Research and Education of Language, Daito-Bunka University, 2000.
- Payne L. D., Maa Color Terms and Their Use as Human Descriptors, *Anthropological Linguistics*, 45.2, 2003.
- Paramei, Galina, Singing the Russian blues, An argument for culturally Basic Color Terms, *Cross-Cultural Research*, 39.1, 2005.
- Rakhilina, Ekaterina V. and Paramei, Galina V., Colour terms, Evolution via expansion of taxonomic constraints, Amsterdam, Benjamins, 2011.
- Roberson Debi, Color Categories Are Culturally Diverse in Cognition as Well as in Language, *Cross-Cultural Research*, 39.1, 2005.
- Roberson, Debi, Davies, Ian R.L. and Davidoff, Jules, Color Categories are Not Universal, Replications and New Evidence from a Stone-Age Culture, *Journal of Experimental Psychology, General*, 129.3, 2000.
- Rosch, Eleanor H., Natural Categories. *Cognitive Psychology*, 4, 1973.
- Saunders, Barbara A. C., *The Invention of Basic Color Terms*, Utrecht, ISOR, 1992.
- Saunders, Barbara A.C. and van Brakel, Jaap, The trajectory of color, *Perspective on Science*, 10.3, 2002
- Snow, David L., Samoan color terminology: A note on the universality and evolutionary ordering of color terms, *Anthropological Linguistics*, 13, 1971.
- Tognini-Bonelli, Elena, *Corpus linguistics at work*. Philadelphia/Amsterdam: Benjamins, 2001.
- Satake Akihiro, Kodai nihongo ni okeru shikimei no seikaku, *Kokugo kokubun*, 24.6, 1955.
- Sinclair, John M., *Corpus, Concordance, Collocation*, Oxford, Oxford University Press, 1991.
- Sinclair, John M., The Search for Units of Meaning, *Textus IX.1*, 1996.
- Sinclair, John M., *Reading concordances*, Harlow, Longman, 2003.
- Thompson, Evan, *Colour vision: a study in cognitive science and the philosophy of perception*, New York, Routledge, 1995.
- Van Leeuwen, Theo, *The Language of Colour, An Introduction*, New York, Routledge, 2011.
- Wierzbicka, Anna, The Meaning of Color Terms, Semantics, Culture, and Cognition, *Cognitive Linguistics*, 1.1, 1990.
- Wierzbicka, Anna, *Semantics, Primes and Universals*, Oxford, Oxford University Press, 1996.
- Wierzbicka, Anna, There are no 'color universals', but there are universals of visual semantics. *Anthropological Linguistics*, 47.2, 2005