Evolution of concept networks and implications for knowledge representation

Jung-ran Park

College of Information Science and Technology, Drexel University, Philadelphia, Pennsylvania, USA

Abstract

Purpose – The purpose of this paper is to present descriptive characteristics of the historical development of concept networks. The linguistic principles, mechanisms and motivations behind the evolution of concept networks are discussed. Implications emanating from the idea of the historical development of concept networks are discussed in relation to knowledge representation and organization schemes.

Design/methodology/approach – Natural language data including both speech and text are analyzed by examining discourse contexts in which a linguistic element such as a polysemy or homonym occurs. Linguistic literature on the historical development of concept networks is reviewed and analyzed.

Findings – Semantic sense relations in concept networks can be captured in a systematic and regular manner. The mechanism and impetus behind the process of concept network development suggest that semantic senses in concept networks are closely intertwined with pragmatic contexts and discourse structure. The interrelation and permeability of the semantic senses of concept networks are captured on a continuum scale based on three linguistic parameters: concrete shared semantic sense; discourse and text structure; and contextualized pragmatic information.

Research limitations/implications – Research findings signify the critical need for linking discourse structure and contextualized pragmatic information to knowledge representation and organization schemes.

Originality/value – The idea of linguistic characteristics, principles, motivation and mechanisms underlying the evolution of concept networks provides theoretical ground for developing a model for integrating knowledge representation and organization schemes with discourse structure and contextualized pragmatic information.

Keywords Semantics, Structures, Pragmatism

Paper type Conceptual paper

Introduction

Language is arguably the salient characteristic that defines humanity (Park and Park, 2005). Givon (1979, p. 352) posits language as "a system of representation of knowledge, acquisition of new knowledge, remodeling-change of knowledge and the communication of new knowledge". An examination of the historical development of semantic senses yields pertinent implications for knowledge representation studies. For instance, the phenomena of synonymy (i.e. related semantic senses across terms/words) and polysemy (i.e. multiple semantic senses of a term/word), both of

which add dynamic complexity and creativity to natural language use, at the same time engender great hindrances in information retrieval owing to inherent lexical ambiguities. The key principles and characteristics underlying the formation of concept networks represented in polysemy and synonymy can be applicable and exploitable to the design of knowledge representation and information retrieval schemes.

In this paper, I will follow Brugman and Lakoff's sense of concept network. Brugman and Lakoff (1988, p. 480) state that concept networks exhibit a "radial" structure:

> Categories may contain a great deal of internal structure – for instance, that one member of a category should be more exemplary of that category than some other member; that the boundaries of the category are not always clear-cut. The category structure utilized here is called a "radial" structure, with a central member and a network of links to other members.

In this sense concept networks concern semantic categories that share core common concepts and semantic sense relations. Polysemy, which involves multiple related semantic senses, is an example of a concept network.

The mechanisms behind the evolution of concept networks manifest a close linkage with discourse-pragmatic contexts and accordingly with socio-cognitive factors. Context-dependent meaning changes and new meaning creation, as well as key principles governing the process of development of concept networks, are manifest across languages (Traugott and Dasher, 2002; Yap, 1999; Park, 2003). This characteristic is applicable to designing knowledge representation schemes geared to multilingual and multicultural resources.

This study aims to present descriptive characteristics of the historical development of concept networks. The principles, mechanisms and motivations behind the evolution of concept networks, as well as implications and potential applications of such principles and characteristics in relation to knowledge representation and information retrieval, will also be discussed.

For this, natural language data including both speech and text are analyzed by examining discourse (both speech and text) contexts in which a linguistic element such as a polysemic word or homonym occurs. Use of speech based language data is derived from the fact that one of the critical conditions of language evolution underlies frequent verbal use of certain linguistic element. For example, Hawaiian pidgin, which lacks the principal grammatical elements of standard language, originated from a multicultural environment owing to waves of immigration from different countries and ethnically heterogeneous plantation life. In this environment, communication was facilitated by employing verbal use of only the core elements of English lacking any structured grammatical elements such as defined word order or morphemic rules. Over time, by frequent verbal use a pidgin may evolve into a creole, which employs a fully-fledged grammatical text and is structured virtually to the same extent as an established language. In this sense, analysis of speech-based data as well as text is critical to the study of information science and documentation.

Finally, existent knowledge representation schemes such as AACR2 (Anglo-American Cataloging Rules) and LCSH (Library of Congress Subject Headings) are examined, focusing on the treatment of homonymy and polysemy by applying linguistic principles underlying the evolution of concept networks. The rapid proliferation of digitization projects by libraries and other organizations calls for

serious research on semantic interoperability in relation to metadata quality evaluation. The critical roadblock to achieving semantic interoperability based on metadata quality control is posed by the lack of a mediation mechanism that cataloging professionals can readily refer to during the process of metadata creation and mapping (see also Heery, 2004; Barton *et al.*, 2003). Owing to the flexibility and complex structure of natural language, which allows for the representation of a concept in various ways, even utilization of a single metadata standard such as Dublin Core (DC) presents challenges in achieving the goal of maximizing semantic interoperability in the process of metadata creation and mapping. Thus, this paper will touch on implications and potential applicability of principles behind the process of concept network development to the design of a mediation mechanism for knowledge organization aimed at enhancing semantic interoperability and resource discovery across digital collections.

Characteristics and principles of the evolution of concept networks

The historical development of a linguistic element encompasses form and structure together with meaning. In this section, key characteristics and principles of the historical development of a linguistic element will be discussed, focusing on the semantic-pragmatic level that deals with meaning in language use.

Characteristics of source concepts

The starting point for the historical development of concept networks can be found in concrete lexemes which serve as source concepts for the development of concept networks. Thus, a concept network is formed from a source concept comprising a concrete, referential and propositional lexical sense. Concrete lexemes denote core entities of human experiences. Heine *et al.* (1991, p. 33) argue that the source concepts used for the historical development of a concept network refer to some of the most elementary human experiences. They are typically derived from the physical state, behavior or immediate environment of man and are frequently referred to in human thought and communication. Bybee *et al.* (1994, p. 10) maintain that it is the reference plane of the irreducible notions in human experience – whether concerning existence or movement in space or psychological or social states, perspectives, and events – that serve as the basis for the grammatical functions and semantic meanings in human languages.

In this sense, concrete lexemes which function as source concepts are employed to solve the problems of expressing abstract concepts such as moral values, perspectives, attitudes and beliefs. Thus, the target concept derived from the source concept denotes subjective, epistemic and abstract concepts. This process is accomplished principally through the employment of metaphoric transference between two conceptual domains, i.e. source and target domains, to be discussed in the following sections. In this view, the human faculty for creativity is fully engaged throughout the process of historical semantic sense development of a lexical element.

For instance, perceptual and sensory verbs such as "look/see" are good examples of this representation of the most elementary human experiences. In this sense, they provide fertile ground for the development of abstract concept which will be illustrated in the following sections.

Historical processes of synonymy and polysemy development
There are several processes that may act on a source concept which is derived from concrete lexemes in the process of its development into a target concept (Hopper and Traugott, 1993; Bybee *et al.*, 1994, pp. 4-9); Diessel, 1999, pp. 116-17). On the semantic and pragmatic level, the concrete lexeme which serve as a source concept may acquire a new function and as a result generate synonymy and polysemy (i.e. related multiple semantic senses) that contribute significantly to the complexity of natural language use. On the other hand, this historical development of synonymy and polysemy greatly hinders knowledge representation and information retrieval (Park, 2002, 2005a).

The dynamic interconnections between different linguistic levels which concern sound, word, semantic meaning, sentence, text and pragmatic language use in socio-cultural contexts, as well as non-linguistic influences such as socio-cognitive factors, can be observed in the process of the development of concept networks. Put another way, there are several processes that may act upon the source concept in the process of its development into the target form (Hopper and Traugott, 1993; Bybee *et al.*, 1994, pp. 4-9; Diessel, 1999).

Hopper (1991, p. 23) presents five principles of the evolution of a source concept. Among these, three principles are particularly related to the development of synonymy as well as polysemy.

The first principle is termed layering: "Within a broad functional domain, new layers are continually emerging . . . the older layers are not necessarily discarded, but may remain to coexist with and interact with the newer layers." For example, there are several functional and grammatical elements for denoting the same English future tense, as in the following italicized illustrations:

- *be going* + *to*: I *am going to* travel to Canada next month.
- *be* + *-ing*: She *is coming* home late from work.
- *be* + *to*: She *is to* give a birth in the near future.

These historically layered elements denote the same grammatical function. Thus, when a new grammatical marker evolves it may co-exist with the old one. Synonym development is in part derived from this layering principle.

Second, there is the divergence principle: The source concept and a target item that has undergone historical change from this source concept diverge and autonomously undergo linguistic change. Thus, polysemic senses derived from the same etymology frequently appear. This can be illustrated by the two following examples:

(1)   Hey, *look* at the picture.

(2)   *Look*, let me think about it.

In illustration (1), the bold-typed lexical item *look*, which has not undergone the historical process, is used as a visual perception verb with the sense of "physical sight". However, in (2) the same lexeme *look*, having undergone semantic sense development, does not contain the sense of "physical sight"; instead it is utilized as an attention-getter from the speech participant. In other words, by qualifying the speaker's utterance through use of the attention-getter *look*, the speaker draws mental attention from the speech participant in a given speech event. Thus, the abstract

semantic sense (i.e. mental attention) of *look* in (2) is derived from the concrete source concept of *look* in (1) concerning physical vision and physical attention.

As Traugott (1995) notes concerning the semantic relevance between source and target items, target elements can be traced to their original source lexical item. Hopper (1991) identifies this phenomenon as the principle of persistence, which is the third principle: linguistic elements that have undergone historical changes are to some degree semantically transparent and remain relevant to their original source item. It is this characteristic that allows one to trace the source of the target concept.

In the above illustrations, the semantic sense of the polysemic attention-getter *look* as in (2) is related to its original source concept in (1) which denotes physical sight. Owing to the principle of persistence, the semantic sense relations in concept networks are logical, non-arbitrary and systematic.

Motivation and mechanisms of semantic sense development

What are the driving forces behind the development of new semantic senses? How are the subjective and epistemic senses and abstract concepts developed? In what manner does such sense development occur?

Epistemicity and subjectification

In this paper, usage of the terms subjectivity and subjectification is based on studies of analysis of linguistic discourse (Traugott, 1982, 1989, 1995; Traugott and Dasher, 2002; Langacker, 1990; Finegan, 1995; Lyons, 1982 among others); accordingly, references to other traditions of subjectivity are not included. According to Traugott (1989, 1995), a concrete propositional sense that serves as a source concept evolves in discourse situations to meet the purposes of creating text and expressing the speaker's attitudes. In other words, the concrete and lexical form having propositional meaning evolves to serve in creating text and in "expressing self and representing the speaker's subjective belief state/attitude toward the proposition" (Traugott, 1995, p. 31). Finegan (1995, p. 1) defines subjectivity as "expression of self" together with the speaker's attitudes and beliefs (Lyons, 1982; Traugott, 1995). Finegan (1995, p. 6) points out that the notion of subjectivity is central to human interaction and "an emerging view of discourse as an instrument not solely, perhaps not centrally, designed for communicating ready-made content, but as an expression of self and in part, its creation". Along with this notion of subjectivity, Traugott (1995, p. 46) argues that subjectification, which can be defined as "structures and strategies that languages evolve in the linguistic realization of subjectivity (Finegan, 1995, p. 1)", is characteristic of development of linguistic elements:

> The reason for the apparent ubiquity of subjectification presumably lies in the speaker's attempts to communicate the relevance of what is being uttered to the communicative event, which includes hearers as well as the speaker. For the speaker's communicative purposes to be achieved, forms are constantly being recruited from lexical domains expressing concrete, objective meanings, and are construed in terms of the perspective of the speaker, the speech event, and the discourse context.

Traugott (1982, p. 257) proposed the following path of semantic sense development:

$$\text{Propositional} > \text{Textual} > \text{Expressive}$$

In this developmental path, concrete referential meanings (i.e. propositional) develop toward increasing discourse coherence (i.e. textual) and more abstract and subjective meanings (i.e. expressive). However, meanings do not necessarily follow precisely the above developmental path. That is, concrete referential meaning can develop toward discourse coherence or toward subjective meaning. Such text and speaker-based semantic sense development conforms to the principle of regularity; i.e. from less to more subjective, but not vice-versa.

Finegan (1995, pp. 4-5) categorizes three main areas of study involving subjectification:

(1) speaker's perspectives as shaping linguistic expressions;

(2) speaker's expression of affect towards the propositions; and

(3) speaker's expression of the modality or epistemic status of the propositions.

For instance, adverbs are one of the grammatical categories that contribute to the expression of the speaker's epistemic status, the third category of Finegan, as shown below:

(3)   The presenter *surely* failed in giving an interesting workshop.

The bold-typed modal adverb *surely* qualifies the speaker's point of view of the truth-value of the proposition. Thus, it serves to modulate the epistemic status of the proposition.

The impetus for semantic sense development is seen as based on socio-cognitive grounds: expressing the speaker's perspective or attitude in social grounds; and increasing discourse coherence by drawing attention from the hearer to upcoming speech events and discourse contexts in cognitive ground. As noted by Traugott (1995) and Finegan (1995), the speaker's cognitive need for increasing the information content through discourse coherence as well as the social need for expressing interpersonal attitudes and beliefs underlies the semantic sense development. That is, in the need for signifying the relevance of an utterance to a given communicative event in order to build discourse coherence as well as the speaker's desire for expressing self, interpersonal attitudes and subjective beliefs underlie semantic sense extensions from lexical elements denoting concrete and objective meanings.

Metaphor and metonymy: devices for the evolution of concept networks
As mentioned in an earlier section, a concept network is formed from a source concept comprising a concrete, referential and propositional lexical sense. The target concept derived from the source concept denotes subjective, epistemic and abstract concepts. As discussed above, such conceptual derivation is motivated by the socio-cognitive needs of language users.

These socio-cognitive factors influence the process of semantic sense development through metaphoric transference between the two conceptual domains, i.e. a shift from the concrete to the abstract domain. Abstract and functional entities such as pragmatic elements denoting interpersonal politeness, solidarity and distance and other functional elements denoting mood, aspect and are derived from concrete and basic lexemes. This process is accomplished principally through the employment of metaphoric constructions.

The attention-getter *look* in (2) and the epistemic modal marker *surely* in (3) can be seen as semantic extensions from the original sources. Put another way, the historically

developed semantic senses have been derived through metaphoric transference from the original meaning. Thus, it can be seen that metaphoric transference involves the connection of two domains. One domain is physical space as the origin and source of the connection, as with the perceptual visual verb *look*, which is in the physical domain as illustrated in (1). The other domain is the mental space serving as a target of the connection. However, this connection is conditioned by discourse and pragmatic forces since transference between the two domains occurs in context-bound situations. As Heine *et al.* (1991) suggest, this process is a denotation of the human creative capacity for solving the problem of expressing an abstract idea.

The following table from Rohrer (2005) illustrates in a very general way this transference between the two domains in context-bound situation. In this case, the context concerns cyberspace (see Table I).

This metaphoric transference between source and concept domains can be seen with a perceptual visual verb such as *see*. Traugott and Konig (1991, pp. 391, 396) notes the characteristics of the source element of the perceptual visual verb in the following way:

Most speech act verb categories include members that have been derived from terms related to the field of vision, specifically terms for *seeing*, *showing* ...

Advise < Lat. Advisare "*watch, observe*" < videre "see"

Regard < OF regarder "*look at*" (< "guard, watch") ...

Thus, the verbs *advise* and *observe* located in the mental domain were historically derived from the verb *see* located in the physical domain and grounded in concrete human experience. As well, the verb *regard*, also in the mental domain, was historically extended from the verb *look at* from the concrete physical domain:

Verbs of seeing (and also hearing) in general tend to be extended to the field of intelligence. Thus, *I see = I understand*. That verbs related to visual perception are a source for SAVs (Speech Act Verbs) suggests that there is a very widespread belief, however unconscious, that seeing is believing and understanding.

As well, Sweetser (1990, p. 21) clearly suggests a metaphorical connection between the perceptual visual verb and abstract concepts related to mental activities such as

Highway (source)	Internet (target)
Highway	Transmission pathways (cables, etc.)
Space	Cyberspace
Vehicles	Computers (telephones, TV, etc.)
Goods transported	Information
Fuel	Electricity
Drivers	Users
Destinations	Information supply sites
Journey	Downloading (or uploading) information
Marketplace	Commercial information suppliers
Impediments to motion (roadblocks, bumps, mechanical trouble, etc.)	Technological difficulties

Source: Rohrer (2005)

Table I.
Mapping of Internet as a
Highway (Cyberspace)

understanding, knowledge and ascertaining of epistemic certainty, in the following way:

> *Physical sight* → *knowledge, intellection*. The basis for this metaphor is to be found in the primacy of the sense of vision as a source of information about the world around us. Not only does English have expressions such as "I saw it with my own eyes" to indicate certainty, but studies of evidentials in many languages show that direct visual information is considered to be the basis of the most certain kind of knowledge.

> *Physical vision* → *mental vision*. This metaphor is probably based on the strong connection between sight and knowledge, and also on the shared structural properties of the visual and intellectual domains – our ability to focus our mental and visual attention and to monitor stimuli mentally and visually.

It is noteworthy that Sweetser (1990) points out that the metaphorical connection between physical activities and mental activities does not occur arbitrarily but rather with motivational ground; that is, "shared structural properties" between the two domains. In the case of the connection between the perceptual visual verb and concepts in the mental domain, it is "our ability to focus our mental and visual attentions."

This semantic shift and transference also shows regularity; that is, developmental direction is regular and unidirectional in two domains from basic and concrete concepts such as the visual perception verb *see* in the physical domain to abstract and functional concepts such as the epistemic modal marker as in the illustrations (2) and (3) in the mental domain.

In addition to metaphor, there is another mechanism that underlies semantic sense development: metonymy. Hopper and Traugott (1993, p. 81) describe conceptual metonymy as "meaning changes arising out of contiguity in linguistic (including) pragmatic contexts". Semantic transfer through metaphor is "analogical and iconic"; on the other hand, semantic transfer through contiguity is "indexical" based on reanalysis (Anttila, 1989, pp. 141-2: cited in Hopper and Traugott (1993, p. 82). They conceptualize that "metonymy points to ('indexes') relations in context" (Hopper and Traugott, 1993, p. 82). In this paper, the notion of "analogical and iconic" is drawn from studies in pragmatics, cognitive linguistics and discourse analysis centering on the issue of semantic change. In relation to the semiotic account on iconic and indexical notions, works by Peirce (1955), de Saussure (1966) and Seboek (1994) are the most pertinent. Mai (2001) introduces Peircean semiotics while applying this to the subject indexing process.

These two mechanisms, metaphor and metonymy, underlie semantic sense development. Traugott and Konig (1991, p. 213) point out the interconnection of the two mechanisms in the process of semantic change in the following way:

> Metaphor is largely correlated with shifts from meanings situated in the external described situation to meanings situated in the internal evaluative, perceptual, cognitive situation, and in the textual situation. Metonymy is largely correlated with shifts in meanings situated in the subjective belief-state or attitude toward situation, including the linguistic one...while metaphor is correlated primarily with solving the problem of representation, metonymy and conventionalizing of conversational meanings are associated with solving the problems of expressing speaker attitudes.

As mentioned earlier, the transference between the two conceptual domains occurs in a regular and unidirectional manner in that the direction of transference occurs from the

concrete physical domain to the abstract mental domain, but not the reverse. Such regularity in semantic change is widely recognized across languages (Traugott and Dasher, 2002). For instance, Park (2003) examines cross-linguistic phenomena stemming from the development of polysemic pragmatic markers derived from basic concepts such as perception verb "look/see," existential element "there is," and the deictic elements "this, that, here, there". By examining the historical development from these source concepts to diverse pragmatic and functional markers, Park (2003) attests cross-linguistic generalization in the development of these source concepts. Yap (1999) also presents the manner in which polysemy derived from the verb *give* is formed across languages. Such cross-linguistic phenomena can also be applicable to the design of knowledge representation schemes geared to multilingual and multicultural resources.

Discourse structure and pragmatic knowledge in semantic sense relations of concept networks

In previous sections, I presented descriptive characteristics of the historical development of concept networks in natural language: principles, motivation and mechanisms. Key principles such as regularity and motivation for semantic change, i.e. subjectivity, as well as other principles such as layering, divergence and persistence which account for the development of synonym and polysemy underlie the evolution of concept networks. The principle of regularity allows for the semantic relations of a concept network to be non-arbitrary and systematic. Also, the metaphoric and metonymic mechanisms behind the evolution of the concept network are closely intertwined with discourse-pragmatic contexts and accordingly socio-cognitive force.

This section will briefly introduce semantic sense relations of multiple polysemic senses and will touch on the treatment of homonymy and polysemy, which greatly hinder knowledge representation and access owing to their lexical ambiguities in existent knowledge representation schemes such as AACR (Anglo-American Cataloging Rules) and LCSH (Library of Congress Subject Headings). Implications and potential applicability of principles behind the process of semantic sense development to the design of knowledge representation scheme geared toward enhancing semantic interoperability across digital collections will also be discussed.

Semantic sense relations on a continuum scale

Let me first illustrate the contextual attributes of multiple polysemic senses by employing the classification of pragmatic markers. The lack of a unified nomenclature and categorization for pragmatic markers in linguistic interpretation is indicative of the elusive and context-sensitive characteristics of these markers. The diverse terminology used to refer to these elements (e.g. pragmatic markers – Brinton, 1996; discourse markers – Schiffrin, 1987; pragmatic expressions – Erman, 1987; pragmatic operators – Ariel, 1999; interjections – James, 1973; hedges – Brown and Levinson, 1987; discourse particles – Schourup, 1990; pragmatic particles – Ostman, 1995) points to this multifunctional nature. It can be seen, then, that contextual attributes significantly affect the interpretation of sense relations of pragmatic markers.

Following are some definitional excerpts from the discourse coherence research perspective:

Pragmatic markers can have either a local or a global orientation in the discourse, expressing a local (between two adjacent utterances) or global (between discourse segments further apart) connection for the hearer. They are thus vitally important for the establishment of an understanding of coherence in conversation (Lenk, 1998, p. 52).

They signal a relationship between the segment they introduce, S2, and the prior segment, S1 (Fraser, 1999, p. 950).

The above definitions underscore the fact that the functional aspects of pragmatic markers are principally oriented around discourse coherence and sequential coordination.

To illustrate:

(4)   Interviewer: But?
       Interviewee: *Well* there is no but.
       (Schiffrin, 1987, p. 119).

The above conversation occurs in a television talk show on controversial issues; the interviewer was not successful in eliciting a contentious viewpoint from the interviewee. Thus, by employing *but?* the interviewer directly attempts to elicit a controversial perspective from the interviewee. To this, the interviewee employs *well*, which marks his inability (i.e. *there is no but)* to answer the interviewer's question. In the above illustration, the pragmatic marker *well* functions for discourse coherence and conversational continuity by marking an insufficient response (Lakoff, 1973) and insufficient relevance for the interpretation of the impending utterance (Jucker, 1993).

By contrast, below are definitional excerpts from the interpersonally-oriented approach:

. . . pragmatic particles . . . as their primary function to implicitly anchor utterances vis-à-vis the communicative restraints of a culture and society, the demands of aspects of interactive politeness, and the prevalent norms of affect and involvement (Ostman, 1995, p. 100).

. . . vehicles for the establishment and maintenance of interpersonal relations between interlocutors (James, 1983, p. 193).

To illustrate:

(5)   a. His idea sounds ridiculous.
       b. *Well*, I do not think so.

In the example above, the primary function of the pragmatic marker *well* can be found in its interpersonal relations to the speech participant. Specifically, *well* indexes the speaker's subjective stance to mitigate the illocutionary force of an upcoming utterance – disagreement to the hearer (e.g. as in [5a.] His idea sounds ridiculous). In other words, by qualifying her utterance by using the pragmatic marker *well*, the speaker softens the disagreeing effect of the utterance.

All these multiple senses are interrelated and permeable to a certain extent even though one sense is foregrounded depending on the context in which it occurs. This permeability and fuzziness is best illustrated as a continuum scale. Ariel (1999) envisions pragmatic markers as being situated at a point on a continuum scale based on the semantic transparency between the pragmatic markers and the original source

concepts. The historical development of multiple senses of polysemic pragmatic markers (PM) is seen on such a scale in Figure 1.
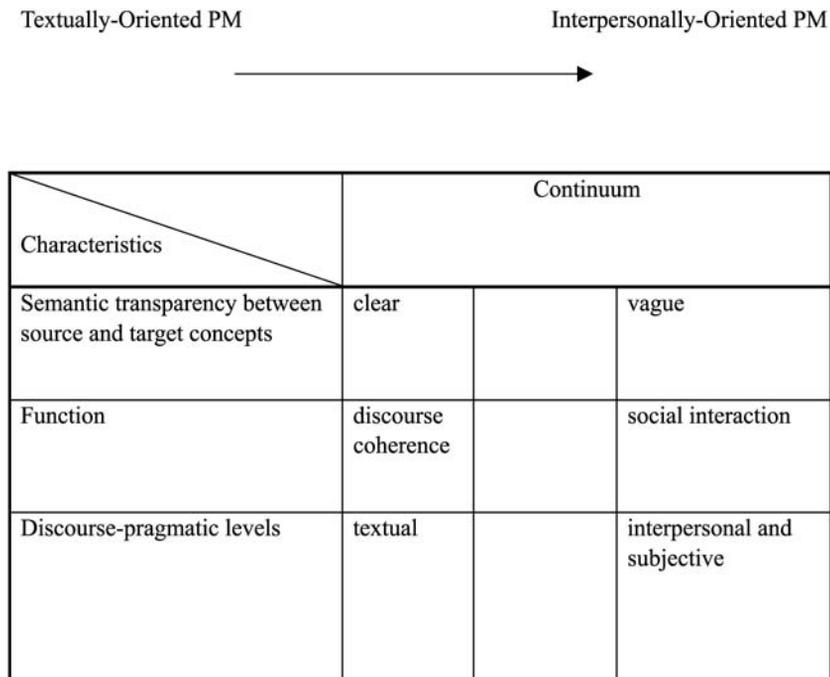
In Figure 1, textually-oriented pragmatic markers which maintain referential meaning from their original source concepts tend to function for discourse and text coherence on the discourse-pragmatic levels. These markers are likely to evolve to interpersonally-oriented pragmatic markers which retain a vague semantic transparency from their source concepts. At the same time, these markers evolve to functioning social interaction on the discourse-pragmatic levels.

The interrelationship and permeability of multiple senses are captured in the above continuum scale by three principal parameters:

(1) concrete semantic sense;

(2) discourse and text structure; and

(3) pragmatic contexts represented in socio-cognitive dimension.

This suggests that the critical factor in the design of knowledge representation and information retrieval schemes is the linkage of discourse and pragmatic information to current knowledge representation schemes.

Dolan *et al.* (2000) point out the critical importance of pragmatic information for word sense disambiguation (WSD) in the following way: "Since pragmatic information is often a valuable cue for WSD, this is a serious concern. Yet the idea that dictionaries somehow isolate lexical from pragmatic knowledge, failing utterly to represent world knowledge, is incorrect." As well, Lascarides and Asher (1996) argue for the

Textually-Oriented PM   Interpersonally-Oriented PM

Characteristics	Continuum		
Semantic transparency between source and target concepts	clear		vague
Function	discourse coherence		social interaction
Discourse-pragmatic levels	textual		interpersonal and subjective

Figure 1.
Evolution of polysemic senses on a continuum scale

inextricable interconnection between pragmatics, discourse and semantics in natural language interpretation:

> In order to place a theory of lexical disambiguation into a discourse context, we require three ingredients. First, we require theory of discourse structure, which stipulates how rhetorical relations affect the structure of discourse, the constraints they impose on coherent discourse, and the semantic effect they have on the constituents they relate. Second, we require an accompanying theory of discourse attachment, which models how pragmatic knowledge resources are used to infer which rhetorical relations hold between two given discourse constituents. Finally, we need a formal language for representing lexical information; ... These three ingredients must be mixed together in a unified account of NL interpretation, which makes precise the above accounts of how information flows between words and discourse.

In order to reflect the linguistic nature of the development of concept networks, it is essential that knowledge representation schemes capture contextualized pragmatic information and discourse structure (see also Blair, 1990; Mai, 2001, 2004; Hjorland and Albrechtsen, 1995; Hjorland, 2003; Smiraglia, 2002; Svenonius, 2000).

Homonymy treatment in AACR2
Homonyms have the same lexical form (e.g. bank [building] *v.* bank [river]; Hyde Park (London, England) *v.* Hyde Park (Tampa, Florida); John Smith (firm) *v.* John Smith (person)) but manifest unrelated meanings that are arbitrarily developed. On the other hand, in the case of polysemy, as discussed in the above sections, semantic sense development presents systematic and predictable sense relations of concept networks comprising polysemic senses. Thus, the distinction between homonymy and polysemy is based on the representation of semantic relations; in other words, in contrast to polysemy, in the case of homonymy, semantic relations between the two concepts are arbitrary and non-systematic. In this sense, homonymy creates contrastive ambiguity (e.g. bank [building] *v.* bank [river]).

Concerning lexical disambiguation techniques, *AACR2* chapters 22-26 present pragmatically constrained disambiguation methods for the names of persons, corporate bodies and places by differentiating contexts. The integration of contextual information to the name authority file delivers successful lexical disambiguation for homonyms. For example, to disambiguate identical names, birth and/or death dates are added (e.g. John Q. Smith (1904-1972). John Q. Smith (1905-); in the case of ambiguous corporate body names, a qualifier is added (e.g. John Smith (firm) *v.* John Smith (person); to disambiguate identical geographic names, contextual contrast is employed: Hyde Park (London, England) *v.* Hyde Park (Tampa, Florida). Thus, all of the above mentioned disambiguation techniques incorporate contextual information into a knowledge representation system that delivers effective information retrieval.

Polysemy in knowledge representation
In the case of polysemy, which manifests semantic extensions based on a concrete basic concept, lexical semantic relations are not contrastive as opposed to those of homonyms. Polysemic semantic relations are systematic and non-arbitrary (logical) and these multiple senses share a basic core sense. Semantic senses in this case are permeable, overlapping and fuzzy to a certain extent owing to this shared core sense. For example, the following pair illustrates such a logical and non-arbitrary relation

between the two polysemic senses of *saw* (6a – physical vision) and *see* (6b – mental understanding). These multiple senses are interrelated and permeable to a certain extent, even though each different sense (i.e. 6a – physical vision), (6b – mental understanding) is foregrounded dependent on the context in which it occurs as in below:

(6) a. I *saw* that John entered into the building.
     b. I *see* what you mean.

Let me illustrate this using a written text. Figure 2 is from the *Oxford English Dictionary (OED)*, under the entry of *crosswalk*.

One of the valuable features of the *OED* from an evolutionary perspective is its presentation of chronological citations of language uses. Language usage illustrated in the *OED* tends to be mostly from written sources, such as novels, magazines and newspapers, rather than from spoken language uses. As shown in Figure 2, the core meaning of *crosswalk*, from a corpus example appearing in 1744, is *a concrete and instrumental object which assists in crossing a garden*. This meaning can be seen to evolve with civilization environmental change, as in the usage cited from 1979. Along

---

Crosswalk. Also cross walk and with hyphen. [CROSS- 4 a.]

 A. path or walk that crosses another, esp. in a garden.

1744 F. Moore *Voyage to Georgia* 30 The Garden is laid out with Cross-walks planted with

Orange-trees. **1813** Jane Austen *Pride & prej*. II. V. 53 Leading the way through every walk and

cross walk.

B. N. Amer. And *W. Austral.* A pedestrian crossing. Also *fig.*

1853 *Children's Aid Society* (First Public Announcement) Mar., The girls … are the cross-walk

sweepers, the little apple-peddlers, and candy-sellers of our city. **1904** *N.Y. Even. Post* 16 May 7

The government service answers very well as a cross-walk in getting over a trying period in a

young man's life. **1915** *Amer. City* Sept. 184/2 It was decided to recommend … that the

following be adopted for general use …Education of the public to use crosswalks at

intersections. **1962** *Coast to Coast 1961-62* 173 He went towards the crosswalk. He had to wait

for the traffic. **1964** M. McLuhan *Understanding Media* xxii. 224 Witness the portent of the

crosswalk, where the small child had power to stop a cement truck. **1979** P. Theroux *Old*

Patagonian Express (1980) iii. 48 No traffic waited at the red lights, no pedestrians at the

crosswalks. **1983** *Austral. Women's Weekly* Aug. 20/3 Pedestrian crossings? Not in Perth. There,

a pedestrian crossing is a crosswalk.

Figure 2.
An example of polysemy,
crosswalk

with the extension of concrete core meaning, as shown in the 1979 citation (*No traffic waited at the red lights, no pedestrians at the crosswalks*), we can observe evolving figurative and abstract uses of the word, as those from 1853 (*The girls … are the cross-walk sweepers, the little apple-peddlers, and candy-sellers of our city*) and 1904 (*The government service answers very well as a cross-walk in getting over a trying period in a young man's life.*)

In the context of semantic mapping of metadata schemes, the extended meaning and usage of *crosswalk* is readily observed in the literature, as in *Crosswalks is a program … from the ALCTS/CCS/SAC/Subcommittee Report (1999). This usage would seem to be well positioned as an entry for future revised *OED* editions. I will illustrate this below, extending Figure 2 (see Figure 3).

As illustrated in an earlier section, semantic extensions through the metaphoric mechanism results in polysemic senses. The example of *crosswalk* above also demonstrates that language usage of a lexical item is closely intertwined with its context.

The issue of polysemy calls for attention from the knowledge representation research community (see also Green and Fraser, 2004). Park (2002) delves into the limitations regarding the manner in which polysemy is treated in knowledge representation schemes such as the Library of Congress Subject Headings (LCSH). LCSH can be seen as analogous to a monomorphic language in which lexical items are given a single meaning and lexical ambiguities are dealt with by scope notes (see Figure 4) and other notations such as *see* and *see also* references.

As a monomorphic language, the power of semantic expressiveness in LCSH is limited, inasmuch as dynamic sense relations are not sufficiently captured. As well, semantic relations among synonymous terms rather than that of multiple polysemic senses are captured by a syndetic structure as shown in Figure 5.

The power of semantic expressiveness in natural language is realized through contextual attributions of multiple senses. Permeability and overlapping among multiple senses are realized through contextual attributes.

Crosswalk. Also cross walk and with hyphen. [CROSS- 4 a.]

1999 ALCTS/CCS/SAC/Subcommittee *Report on metadata and subject analysis*. A crosswalk

is a program or algorithm to map elements in different metadata schemes.

Computational linguistics (May Subd Geog)

Here are entered works on the applications of computers in processing and analyzing language.

Works on the computer processing of natural language for the purpose of enabling humans to

interact with computers in natural language are entered under Natural language processing

(Computer science).

Figure 5.
An example of a LCSH
syndetic structure

Doerr (2001) also points out that "problems arising from complementary polysemy for thesaurus design have not yet been studied". Also stated is that polysemy presents the most serious hindrances in interoperability faced in the process of mapping thesauri, especially considering the fact that the relationship among terms is not simply monohierarchical but also polyhierarchical.

As illustrated by the processes of evolution of concept networks, the motivation for linguistic evolution can be found in socio-cultural and cognitive factors. That is, language evolves to reflect the socio-cultural communicative needs of language users. Polysemy, then, is caused and affected by the socio-cultural context in which written and spoken discourse appears. Research into polysemic disambiguation, therefore, would be most productive if concentrated along contextual lines.

Concept networks as a mediation mechanism for knowledge organization
The rapid proliferation of digitization projects by libraries and other organizations calls for serious research on metadata quality evaluation. As evinced through information-sharing for non-networked traditional bibliographic collections through authority control, successful resource discovery and exchange across ever-growing distributed digital collections demands semantic interoperability based on accurate and consistent resource description. Successful resource sharing and access across ever-growing distributed digital collections demand semantic interoperability based on accurate, consistent and complete resource description. The critical roadblock in achieving semantic interoperability based on metadata quality control is posed by the lack of a mediation mechanism that cataloging professionals can readily refer to during the process of metadata creation and mapping (see also Heery, 2004; Barton *et al.*, 2003). Owing to the flexibility and complex structure of natural language, which allows for the representation of a concept in various ways, even utilization of a single metadata standard such as Dublin Core (DC) presents challenges in achieving the goal of maximizing semantic interoperability in the process of metadata creation and mapping.

Lynne Howarth (2001, 2003) presents challenges on matching between common namespace, so called, set of labeled categories totaling 17 labels, which is based on analysis of mapping of nine metadata schemes (i.e. EAD; DC, GILS; TEI; VRA; CIMI; CSDGM; ONIX; MARC) and their corresponding definitions by conducting focus group methods. According to the findings of the study, only three category names (i.e.

language, physical format, date and time period) are most readily understood; others ($n = 14$) require clarification of definition and additional definitions. Howarth (2003, p. 183) underscores the critical nature of context in which a term and the corresponding definition occurs: "the provision of brief, simple definitions devoid of topical or other perspective was insufficient to clearly and unequivocally identify all elements. In short, how the category 'fit' within a specific or a broader context was seen as being essential to uniquely understanding each term."

Park (2005a, b) examine semantic interoperability across digital collections based on content analysis of 659 unqualified Dublin Core metadata item records created by catalogers. The analysis of 659 metadata item records shows evidence of frequent inaccurate, incomplete and inconsistent metadata element uses that will hinder semantic interoperability. Some examples: the "physical description" field is either inaccurately used as DC "format" or "description"; there is great confusion in employing the DC elements "type" and "format" and they are interchangeably used; the DC elements "source" and "relation" are inconsistently used; the DC element "relation" is interchangeably used with cataloger-defined field names such as "digital collection". Some of the most frequently identified locally added field names concern provenance information such as "contact information", "ordering information", and "acquisition".

The results of the studies (Park, 2005a, b) suggest the critical need for mediation mechanisms that provide contextual relations among metadata elements and their corresponding definitions and usage to facilitate metadata creation and mapping process by lessening semantic ambiguity. Current utilization of unqualified Dublin Core metadata, especially non-cataloging professionals, further increases the necessity for such a mediation mechanism. As an example, a significant number of data providers for Open Archive Communities such as OLAC (Open Language Archives Community) are non-cataloging professionals who do not have any education and practicum related to information organization and access (Park, 2004).

As illustrated in both Howarth (2003) and Park (2005a, b), hindrance and problems in metadata mapping results from the lack of the provision of surrounding context in which a term and its usage (i.e. definition) occur. This lack of context creates semantic ambiguity and consequently produces inaccurate, inconsistent and incomplete metadata item records. Knowing and locating where a vocabulary is visually placed in a concept network is a critical part of acquiring meaning of the term (Miller *et al.*, 1990). Knowing the meaning of a term in a concept network has great potential to improve usage of metadata elements and consequently improve the metadata creation and mapping process between DC metadata elements and author or cataloger generated vocabularies. In this sense, concept networks can be utilized as a mediation mechanism which enhances the metadata creation and mapping process by disambiguating semantic ambiguities caused by isolation of a term and its corresponding definition from the relevant context.

The structure of a concept network can be designed to comprise conceptual categories that share a core semantic property. To illustrate: the concept of "author", "creator", or "performing body" share the core semantic property of "intellectual responsibility of a work" and this can be categorized into the same conceptual category under "name". Conceptual categories can be organized into a hierarchical structure, i.e. conceptual taxonomy. A concept network may also consist of a concept description that represents conceptual relations among terms and conceptual categories. Instances

(i.e. examples), brief definition, and a scope note if necessary for further disambiguation of a concept can also be part of the structure of concept networks. Conceptual relations are expressed by a variety of semantic features such as thing (i.e. object), people (i.e. agent, actor), event (i.e. process), time/aspect, place (i.e. location), and instrument. Concept networks can be visually expressed by employing a small number of notations and symbols. For instance, a rectangular box may represent a conceptual category. Conceptual relations between concepts and conceptual categories can be represented by nodes (points) and conceptual relations between concepts can be expressed by links. Instances may also be represented by a vertical line.

The concept networks can be modified and enhanced through an iterative process of analyzing conceptual structure. In other words, addition or deletion of the instantiation of a concept can affect the structure of concept networks in aspects such as conceptual taxonomy, conceptual relations, definition and scope note. Concept networks have good potential to facilitate proper interpretation of metadata concepts and accurate and consistent usage of the data elements during the metadata creation and mapping process, for both non-cataloging professionals as well as cataloging professionals.

Conclusion and future studies

The characteristics, principles, motivation and mechanisms behind the evolution of concept networks provide theoretical ground for improving semantic sense relations in knowledge representation schemes. The characteristic of regularity shows good potential for knowledge representation schemes in defining systematic and predictable semantic sense relations of concept networks.

The metaphoric transference between two conceptual domains is conditioned by discourse and pragmatic forces, inasmuch as transference between the two domains occurs in context-bound situations. The impetus for semantic sense development can be found in the cognitive need of speech participants for increasing information content through discourse coherence as well as the social need for expressing subjective interpersonal cohesion and an epistemic stance.

In consequence, pragmatic contexts represented in the socio-cognitive dimension and discourse structure are closely intertwined with semantic sense relations, as illustrated in Figure 1. This suggests that it is critical to replicate discourse structure and pragmatic knowledge in the knowledge representation scheme. Thus, the path for future studies lies in development of a model for integrating pragmatic and discourse information into knowledge representation schemes. Developing a well-defined interface for linking both modules (i.e. discourse structure and contextualized pragmatic information) to knowledge representation schemes should also be a primary goal for future studies.

It is important to note that the key principles, mechanisms and impetus behind the evolution of concept networks are widely recognized across languages. This cross-linguistic generalization has been attested in linguistic studies by natural language data derived from cross-linguistic and cross-cultural contexts. It seems certain to bear approaches for impending research issues concerning knowledge representation studies in global contexts (Beghtol, 2004; Williamson and Beghtol, 2003).

References

ALCTS/CCS/SAC Subcommittee on metadata and classification (1999), "Final report", available at: www.ala.org/alcts/organization/ccs/sac/metaclassfinal.pdf

Ariel, M. (1999), "Pragmatic operators", *The Encyclopedia of Languages and Linguistics*, Pergamon Press, Oxford.

Asher, N. and Lascarides, A. (1996), "Lexical disambiguation in a discourse context", *Lexical Semantics: The Problem of Polysemy*, Clarendon Press, Oxford.

Barton, J., Currier, S. and Hey, J.M.N. (2003), "Building quality assurance into metadata creation: an analysis based on the learning objects and e-prints communities of practice", *Dublin Core Conference, Washington, September 28-October 2, 2003*, available at: http://purl.oclc. org/dc2003/03barton.pdf

Beghtol, C. (2004), "Naïve classification systems and the global information society", paper presented at the 8th International ISKO Conference, London, 13-16 July, 2004.

Blair, D. (1990), *Language and Representation in Information Retrieval*, Elsevier Science Publishers, New York, NY.

Brinton, L.J. (1996), *Pragmatic Markers in English: Grammaticalization and Discourse Functions*, Mouton de Gruyter, Berlin.

Brown, P. and Levinson, S. (1987), *Politeness: Some Universals in Language Usage*, Cambridge University Press, Cambridge.

Brugman, C. and Lakoff, G. (1988), "Cognitive topology and lexical networks", in Small, S., Cottrell, G. and Tanenhaus, M. (Eds), *Lexical Ambiguity Resolution: Perspective from Psycholinguistics, Neuropsychology and Artificial Intelligence*, Morgan Kaufmann, San Francisco, CA.

Bybee, J.L., Pagliuca, W. and Perkins, R. (1994), *The Evolution of Grammar: Tense, Aspect, Modality in the Language of the World*, University of Chicago Press, Chicago, IL.

de Saussure, F. (1966), *Course in General Linguistics*, McGraw-Hill, New York, NY.

Diessel, H. (1999), *Demonstratives: Form, Function, and Grammaticalization*, John Benjamins Publishers, Amsterdam.

Doerr, M. (2001), "Semantic problems of thesaurus mapping", *Journal of Digital Information*, Vol. 1. No. 8, available at: http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Doerr/

Dolan, W., Vanderwende, L. and Richardson, S. (2000), "Polysemy in a broad-coverage natural language processing system", in Ravin, Y. and Leacock, C. (Eds), *Polysemy: Theoretical and Computational Approaches*, Oxford University Press, Oxford.

Erman, B. (1987), *Pragmatic Expressions in English: A Study of 'You Know,' 'You See,' and 'I Mean' in Face-to Face-Conversation*, Almqvist & Wiksell, Stockholm.

Finegan, E. (1995), "Subjectivity and subjectivisation: an introduction", in Stein, D. and Wright, S. (Eds), *Subjectivity and Subjectivisation: Linguistic Perspectives*, Cambridge University Press, Cambridge.

Fraser, B. (1999), "What are discourse markers?", *Journal of Pragmatics*, Vol. 31, pp. 931-52.

Givon, T. (1979), "Language and ontology: on constructing a universe", in Givon, T., *On Understanding Grammar*, Academic Press, New York, NY.

Green, R. and Fraser, L. (2004), "Patterns in verbal polysemy", paper presented at the 8th International ISKO Conference, London, 13-16 July 2004.

Heery, R. (2004), "Metadata future: steps towards semantic interoperability", in Hillman, D.I. and Westbrooks, E.L. (Eds), *Metadata in Practice*, American Library Association, Chicago. IL.

Heine, B., Claudi, U. and Hunnemeyer, F. (1991), *Grammaticalization: A Conceptual Framework*, University of Chicago Press, Chicago, IL.

Hjorland, B. (2003), "Fundamentals of knowledge representation", *Knowledge Representation*, Vol. 30 No. 2, pp. 87-111.

Hjorland, B. and Albrechtsen, H. (1995), "Toward a new horizon in information science: domain analysis", *Journal of the American Society for Information Science*, Vol. 46 No. 6, pp. 400-25.

Hopper, P.J. (1991), "On some principles of grammaticalization", in Traugott, E.C. and Heine, B. (Eds), *Approaches to Grammaticalization*, John Benjamins Publishers, Amsterdam, pp. 17-35.

Hopper, P.J. and Traugott, E.C. (1993), *Grammaticalization*, Cambridge University Press, Cambridge.

Howarth, L.C. (2001), "Designing a metadata-enabled namespace for enhancing resource discovery in knowledge bases", in Guerrini, M., Gambari, S. and Sardo, L. (Eds), *Proceedings of the International Conference Electronic Resources: Definition, Selection and Cataloguing*, available at: http://w3.uniroma1.it/ssab/er/relazioni/howarth_eng.pdf

Howarth, L.C. (2003), "Designing a common namespace for searching metadata-enabled knowledge repositories: an international perspective", *Cataloging & Classification Quarterly*, Vol. 37 Nos 1/2, pp. 173-85.

James, A. (1983), "Compromisers in English: a cross-disciplinary approach to their interpersonal significance", *Journal of Pragmatics*, Vol. 7, pp. 191-206.

James, D. (1973), "Another look at, say, some grammatical constraints, on, oh, interjections and hesitations", in Corum, C., Smith-Stark, T.C. and Weiser, A. (Eds), *Papers from the 9th Regional Meeting of the Chicago Linguistics Society*, Chicago Linguistics Society, Chicago, IL.

Jucker, A.H. (1993), "The discourse marker well: a relevance-theoretic account", *Journal of Pragmatics*, Vol. 19, pp. 435-52.

Lakoff, R. (1973), "Questionable answers and answerable questions", *Issues in Linguistics*, University of Illinois Press, Chicago, IL.

Langacker, R.W. (1990), "Subjectification", *Cognitive Linguistics*, Vol. 1, pp. 5-38.

Lenk, U. (1998), *Marking Discourse Coherence: Functions of Discourse Markers in Spoken English*, G. Narr, Tubingen.

Lyons, J. (1982), "Deixis and subjectivity: loquor, ergo sum?", in Jarvella, R.J. and Klein, W. (Eds), *Speech, Place, and Action: Studies in Deixis and Related Topics*, John Wiley, Chichester and New York, NY, pp. 101-24.

Mai, J-E. (2001), "Semiotics and indexing: an analysis of the subject indexing process", *Journal of Documentation*, Vol. 57 No. 5, pp. 591-622.

Mai, J-E. (2004), "Classification in context: relativity, reality, and representation", *Knowledge Representation*, Vol. 31 No. 1, pp. 39-48.

Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D. and Miller, K.J. (1990), "Introduction to WordNet: an on-line lexical database", *International Journal of Lexicography*, Vol. 3 No. 4, pp. 235-44.

Ostman, J-o. (1995), "Pragmatic particles twenty years after", in Warvik, B. (Ed.), *Organization in Discourse*, University of Turku, Turku.

Park, J-r. (2002), "Hindrances in semantic mapping among metadata schemes: a linguistic perspective", *Journal of Internet Cataloging*, Vol. 5 No. 3, pp. 59-79.

Park, J-r. (2003), "A study of selected Korean pragmatic markers: synchronic and diachronic perspectives", PhD dissertation, Department of Linguistics, University of Hawaii at Manoa, Manoa, HI.

Park, J-r. (2004), "Language-related open archives: impact on scholarly communities and academic librarianship", *E-JASL: The Electronic Journal of Academic and Special Librarianship*, Vol. 5 Nos 2-3, available at: http://southernlibrarianship.icaap.org/content/v05n02/park_j01.htm

Park, J.-r. (2005a), "Semantic interoperability across digital image collections: a pilot study on metadata mapping", in Liwen V. (Ed.), *CAIS/ACSI 2005 Data, Information, and Knowledge in a Networked World, Proceedings of the 2005 Annual Conference of the Canadian Association for Information Science held with the Congress of the Social Sciences and Humanities of Canada at the University of Western Ontario, London, Ontario, June 2-4, 2005*, available at: www.cais-acsi.ca/proceedings/2005/park_J_2005.pdf

Park, J-r. (2005b), "Human language: resources from linguistics and beyond", *College & Research Libraries News*, Vol. 66 No. 3, available at: www.ala.org/ala/acrl/acrlpubs/crlnews/backissues2005/march05/humanlanguage.htm

Park, J-r. and Park, S-j. (2005), "Digital collection management software employed by libraries and museums: evaluation of metadata semantic mapping functionality", paper presented at the poster session at ALISE (Association/Library and Information Science Education) Annual Conference, 11-14 January 2005, Boston, MA.

Peirce, C.S. (1955), *Philosophical Writings of Peirce*, Dover Publications, New York, NY.

Rohrer, T. (2005), "Conceptual blending on the information highway: how metaphorical inferences work", available at: http://philosophy.uoregon.edu/metaphor/iclacnf4.htm

Schiffrin, D. (1987), *Discourse Markers*, Cambridge University Press, Cambridge.

Schourup, L. (1990), *Common Discourse Particles in English Conversations: 'Like', 'Well', 'Y'know*, Garland, New York, NY.

Seboek, T.A. (1994), *Signs: An Introduction to Semiotics*, University of Toronto Press, Toronto.

Smiraglia, R. (2002), "The progress of theory in knowledge representation", *Library Trends*, Vol. 50 No. 3, pp. 330-51.

Svenonius, E. (2000), *The Intellectual Foundation of Knowledge Representation*, MIT Press, Cambridge, MA.

Sweetser, E. (1990), *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*, Cambridge University Press, Cambridge.

Traugott, E.C. (1982), "From prepositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization", *Perspectives on Historical Linguistics*, John Benjamins Publishers, Amsterdam.

Traugott, E.C. (1989), "On the rise of epistemic meanings in English: an example of subjectification in semantic change", *Language*, Vol. 65, pp. 31-55.

Traugott, E.C. (1995), "Subjectification in grammaticalisation", in Stein, D. and Wright, S. (Eds), *Subjectivity and Subjectivisation*, Cambridge University Press, Cambridge, pp. 31-54.

Traugott, E. and Dasher, R.B. (2002), *Regularity in Semantic Change*, Cambridge University Press, Cambridge.

Traugott, E.C. and Konig, E. (1991), "The semantics-pragmatics of grammaticalization revisited", in Traugott, E.C. and Heine, B.H. (Eds), *Approaches to Grammaticalization*, Vol. 1, John Benjamins Publishers, Amsterdam, pp. 189-218.

Williamson, N.J. and Beghtol, C. (Eds) (2003), *Knowledge Representation and Classification in International Information Retrieval*, Haworth Information, Binghamton, NY.

Yap, F-H. (1999), "Crosslinguistic implications of the grammaticalization of *bagi* and other 'give' morphemes: the polysemy of 'give' constructions in Malay and other languages", dissertation, UCLA, Los Angeles, CA.

Further reading

ALCTS/CCS/SAC Subcommittee on metadata and subject analysis (1999), "Subject data in the
metadata record recommendations and rationale: a report from the ALCTS/CCS/SAC
Subcommittee on metadata and subject analysis", available at: www.ala.org/alcts/
organization/ccs/sac/metarept2.html

Duranti, A. (1994), "Language in context and language as context: the Samoan respect
vocabulary", in Goodwin, C. and Duranti, A. (Eds), *Rethinking Context: Language as an
Interactive Phenomenon*, Cambridge University Press, Cambridge.

Hjorland, B. (2003), "Afterword: ontological, epistemological and sociological dimensions of
domains", *Knowledge Representation*, Vol. 30 Nos 3/4, pp. 239-45.

Holmes, J. (1986), "Functions of 'you know' in women's and men's speech", *Language in Society*,
Vol. 15, pp. 1-22.

Ostman, J-o. (1981), *You Know: A Discourse-Functional Approach*, John Benjamins Publishers,
Amsterdam.

Svartvik, J. (1980), "'Well' in conversation", in Greenbaum, S. (Ed.), *Studies in English Linguistics
for Randolph Quirk*, Longman, London, pp. 167-77.

Corresponding author
Jung-ran Park can be contacted at: jung-ran.park@ischool.drexel.edu