THE NATURE OF NEAR-SYNONYMIC RELATIONS

Chrysanne DiMarco
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
cdimarco@logos.uwaterloo.ca

1 Introduction

The words *gawp*, *gaze*, and *stare* all denote a kind of prolonged look: they are *near-synonyms*, or *plesionyms* [Cruse 1986]. However, as we learn from their individual entries in the *Oxford advanced learner's dictionary* (*OALD*; fourth edition, 1989), to gaze is to look long and steadily; to stare is to do this with the eyes wide open; and gawping has the additional requirement that the act be impolite or stupid. In recent work [DiMarco, Hirst, and Stede 1993; DiMarco and Hirst 1993], we address the problem of representing the lexical features that distinguish groups of near-synonyms.

Our lexical features for differentiation are not intended to be any kind of primitives for decompositional semantics: they are not being used to represent whole meanings, but rather to represent *differences* between meanings. These differences between plesionyms can be shades of denotation or connotation, or emphases on different components of the meaning.

Our eventual goal is a representation for a lexicon in which semantic and stylistic distinctions can be made between synonyms and plesionyms, both within and across languages, for the purpose of lexical choice in natural language generation and machine translation. The nature of these distinctions suggests that they can be viewed as *relations* between near-synonyms. In this paper, we undertake a study of the characteristics of near-synonymic relations as a step towards a knowledge representation for lexical discrimination.

2 Previous research

As a first step, which we described in [DiMarco, Hirst, and Stede 1993], we carried out a study of dictionary usage notes in order to compile a list of the kinds of dimensions that are used frequently as denotative or connotative differentiae. We produced a preliminary list of 26 denotational dimensions and 12 connotative dimensions (including a few that we added from the discussion on lexical aspects by Vinay and Darbelnet [1958]). (This set is not yet complete or definitive, of course, but we have managed to include a fairly comprehensive selection of the most common differences between near-synonyms.) Some of the dimensions are simple binary choices; others are continuous. We show a representative sample in Table 1. Each line of the table shows a dimension of differentiation followed by example sentences in which two plesionyms vary along that dimension.

3 Chaffin and Herrmann

3.1 Basic theory

Chaffin and Herrmann [1988] have provided a theoretical approach for accounting for semantic relations that we will apply to *near-synonymic* relations. They describe a systematic study of the nature of semantic relations, beginning with a catalogue of the relation properties that an adequate theory of semantic relations should explain, followed by a list of sample relations. These relations (*e.g.*, *synonymity*, *pseudoantonym*) are then broken down into relation elements (*e.g*, *symmetrical position*, *locative inclusion*), which are formally defined. Chaffin and Herrmann's study culminates with an explanation of how this relation-element approach can be used to account for each of the relation properties. We will undertake a similar kind of study in proposing a theoretical account of near-synonymic relations. However, unlike Chaffin and Herrmann, who began with readily recognizable semantic relations and then defined relation elements, we find that in our study of near-synonyms, it is more appropriate to begin with the relation elements, which are more easily identified, and then move on to the construction of the relations, which are more difficult to define.

We will begin by examining four properties that Chaffin and Herrmann believe any theory of semantic relations should account for and we will show that these properties are also relevant to any theory of near-synonymic relations.[1]

[1] Chaffin and Herrmann include *relation discrimination*, but as our whole study is of lexical differentiae, *all* our relation properties have something to do with discrimination. They also include *relation verification*, but a demonstration of this property would involve psychological testing, which we have not yet undertaken.

Denotational dimensions	Connotative dimensions
Intentional/accidental: She {*stared at* | *glimpsed*} him through the window.	**Formal/informal:** He was {*inebriated* | *drunk*}.
Continuous/intermittent: Wine {*seeped* | *dripped*} from the barrel.	**Abstract/concrete:** The {*error* | *blunder*} cost him dearly.
Immediate/iterative: She {*struck* | *beat*} the drum.	**Pejorative/favorable:** That suit makes you look {*skinny* | *slim*}.
Sudden/gradual: The boy {*shot* | *edged*} across the road.	**Forceful/weak:** The building was completely {*destroyed* | *ruined*} by the bomb.
Degree: We often have {*mist* | *fog*} along the coast.	

Table 1: Examples of features that dictionary usage notes adduce in word differentiation (adapted from [DiMarco and Hirst 1993]).

Relation comparison. The primary property is *relation comparison*: pairs of near-synonyms can be compared and judged as more, or less, similar to each other than others. For example, there is something similar in the relationship between *stingy/frugal* and between *fat/plump*. In each case, the first word (*stingy*, *fat*) is pejorative while the second (*frugal*, *plump*) has a nuance of being admirable or attractive. This relationship would not be maintained if, for example, we replaced *fat/plump* by *rotund/plump*.

Relation expressions. The second relation property is *relation expressions*, which refers to people's ability to use common words and phrases to express near-synonymic relations. For example, *mistake* and *error* both refer to something done incorrectly or improperly, but *mistake* is *more general* than *error*, according to the usage note in the *OALD*.

Relation complexity.[2] The property of *relation complexity* refers to the need to represent different relations between the same pair of near-synonyms, on more than one level of complexity; we need to be able to include nuances that are relevant to a given situation and ignore others.

Relation creativity. Chaffin and Herrmann observe that "the production and recognition of relations is a creative ability", so that the relation between two words "can be readily identified although the reader may never have considered the relation of these particular terms before" [p. 292]. We will show that *relation creativity* is equally necessary to a theory of near-synonymic relations. For example, the relation of *arrange/organize*[3] can be recognized as one that contrasts correctness with functionality, and we might then detect this same relationship for other pairs of near-synonyms (*e.g.*, *trim/shave*).

In summing up the importance of these relation properties to a theory of semantic relations, Chaffin and Herrmann state that "these diverse phenomena must be explained by theories of relations" and "we will find that in order to explain relations it is necessary to assume that relations are normally composed of more primitive elements that account for their characteristics and for people's abilities to make judgments about them" [p. 292]. We believe these observations are equally true of theories of plesionymic relations and we will show that a relation-element theory of near-synonymy will account for these relation properties.

3.2 Theoretical assumptions

In developing their theory of semantic relations, Chaffin and Herrmann make the following representational assumptions [paraphrased from pp. 293–294]:

- A relation R between two concepts x and y is composed of a set of dyadic relation elements $(E_a, \ldots, E_n)$:

$${}_xR_y \longrightarrow (E_a,\ \ldots,\ E_n)^4$$

- Relation elements may be hierarchically organized so that the presence of one element depends on the presence of another, or elements may be independent of one another. In the following representation, independent

[2] Chaffin and Herrmann [1988] use the somewhat misleading term *relation ambiguity*, but we believe it is more accurate and less confusing to use the term *relation complexity*.

[3] "**Arrange** is to put in a pleasing or correct order ... **Organize** is to put into a working system" (from the usage note in the *OALD*).

[4] This notation should be read as "the relation R decomposes to the relation elements ...".

elements are separated by commas and dependent elements appear in parentheses following the element that they depend on:

$$_xR_y \longrightarrow (E_1(E_2(E_3)), E_4, \ldots, E_n)$$

- Relations may share one or more elements. The greater the proportion of elements two relations have in common, the more similar they are.

Two examples of Chaffin and Herrmann's semantic relations are *synonymity*[5] and *pseudoantonym*, which they define in terms of the following sets of relation elements:[6]

synonymity: intersection (inclusion (bilateral))

pseudoantonym: dimension (bipolar, connotative)

We will adapt these representational assumptions[7] to our study of plesionymy and use them in constructing near-synonymic relations from the relation elements to be defined below.

4 The relation elements

Chaffin and Herrmann define a set of *relation elements* of which semantic relations are composed. These relation elements are described as "elements that the relations had in common and elements that distinguished the relations from each other" [p. 301]. We observe that, for our purposes, a relation element is a denotational or connotative feature that is part (or all) of a description of a near-synonymic relation; and hence near-synonymic relations can be differentiated by these various elements. Thus, given this observation, we can consider our features of differentiation, as illustrated in Table 1, to be examples of the relation elements that compose, and distinguish, near-synonymic relations.

We stipulate that our near-synonymic relation elements are unitary, that is, they represent lexical relationships that need not be decomposed any further. While we expect that relation elements will be language-independent, the degree to which relation elements need to be decomposed might differ from language to language: we will refine the element to exactly the level of distinction necessary for the plesionyms of the language, and no further.

[5] Note that Chaffin and Herrmann treat synonymy as a single semantic relation, while we are interested in the many different near-synonymic relationships that can exist.

[6] Lack of space precludes a full explanation of these relation-element structures, but it is not necessary for understanding the work we will present.

[7] Chaffin and Herrmann also make processing assumptions, including one that relates to psychological verification of their relations; we do not use these assumptions in this paper.

5 The relations

We will work through several examples, showing how near-synonymic relations can be constructed from relation elements. All the word descriptions in the examples below will be taken from usage notes in the *OALD*. Our first example is the representation of the distinction between *ask* and *beseech*:

> [**Ask**] is the most usual and informal word ... **beseech** [is] stronger and more formal than **beg**.

From this usage note, and our own native-speaker knowledge, we identify the relation elements that distinguish each word:

ask: general; informal

beseech: formal; forceful

This notation can be read as "*ask* is more general and more informal than *beseech*; *beseech* is more formal and more rhetorically forceful than *ask*." We construct the relation between *ask* and *beseech* by taking the complement of their respective relation-element structures and then indicating dependencies between the resulting relation elements:

ask/beseech: (general (formal$_{ji}$, forceful$_{ji}$))

This relation states that *ask* is less formal and less forceful than *beseech*. (The ji subscript on *formal* and *forceful* should be read as indicating the direction of the relationship between word j and word i.) This notation also shows that the relation elements *formal* and *forceful* are both dependent on the dominant element *general*.

Chaffin and Herrmann show that for semantic relations, the same relation can hold between more than one pair of words. As this next example for *thin* and *emaciated* shows, the same *plesionymic* relation (in this case, the relation between *ask/beseech*) can hold between more than one pair of words. The *OALD* usage note for *thin* and *emaciated* describes them as follows:

> **Thin** is the most general word. It may be negative, suggesting weakness or lack of health ... **Emaciated** indicates a serious condition resulting from starvation.

We identify the relation elements that distinguish *thin* and *emaciated*:

thin: general

emaciated: formal; forceful

We take the complement of these relation elements and indicate the dependencies:

thin/emaciated: (general (formal$_{ji}$, forceful$_{ji}$))

A more complex example is the relation between *quarrel* and *row*. The *OALD* usage note describes their distinguishing features as follows:

> A **quarrel** is a sharp, often angry, exchange of words between people ... A **row** is angry and may involve shouting, usually for a short time ... A **row** can also take place between public figures or organizations.

There are two ways we can construct the relation between *quarrel* and *row*, depending on whether the argument is between people or inanimate organizations:

quarrel/row:

(forceful, formal_{ji}, $\text{emotional}_{ji}(\text{vectorial}_{ji})$)

quarrel/row:

(forceful, $\text{inanimate}_{ji}(\text{formal}_{ji})$)

The first relation states that *row* is more formal and more emotional; *quarrel* is more forceful. It also indicates that the greater emotion of a *row* is linked to a difference in scale, the *vectorial* element, which in this case refers to the different lengths of time of a *quarrel* and a *row*. The second relation notes that a *row* can involve inanimate entities but, if it does, then the effect is more formal. Thus, we can have different relations between plesionyms, depending on the different usages of the words.

By following the same kind of approach, we can construct relations for some other pairs of near-synonyms:

frown/grimace:

(general (formal_{ji}, forceful_{ji})

mistake/blunder:

(general (formal_{ji}, forceful_{ji}, careless_{ji}))

fat/plump:

(general (forceful ($\text{polite}_{ji}(\text{attractive}_{ji})$)))

We observe that the same or similar relation can hold between different pairs of near-synonyms, for example, *ask/beseech*, *thin/emaciated*, and *frown/grimace*. This is analogous to the case of semantic relations, which, as Chaffin and Herrmann note, are readily recognizable and nameable. Near-synonymic relations cannot be so easily labelled, but we can still see that some basic set of relations might be defined and could be used to construct new relations. For example, we showed that the relation between *ask* and *beseech* could be represented by the following structure:

ask/beseech:

(general (formal_{ji}, forceful_{ji}))

We saw how this basic relation could also apply to *thin/emaciated* and *frown/grimace*; this suggests that, for lexical-choice processing, we will want to keep a catalogue of existing relations from which new relations could be built. Another pair of near-synonyms, *mistake* and *blunder*, share the same distinctions, except that *blunder* is often the result of carelessness (*OALD*). So we add to the existing specification to obtain the following relation:

mistake/blunder:

(general (formal_{ji}, forceful_{ji}, careless_{ji}))

Lastly, dependencies can lead to quite complicated relations, as in the case of *fat/plump*, where the distinction of politeness (impoliteness) is related to different dependencies for each near-synonym: the nuances of force and impoliteness are interdependent, as are those of politeness and attractiveness.

6 The relation properties

In Section 3.1, we set out a list of relation properties that any theory of near-synonymic relations should be able to account for. In this section, we discuss how a relation-element approach addresses these issues.

Relation comparison. By breaking down the relations between plesionyms into relation elements, we can obtain a finer degree of discrimination between similar words for the task of lexical choice in generation. As we discuss in [DiMarco, Hirst, and Stede 1993], many of the semantic distinctions between plesionyms do not lend themselves to neat, taxonomic differentiation; rather, they are fuzzy, with plesionyms often having an area of overlap. For example, the boundary between *forest* and *wood* is vague, and there are some situations in which either word might be equally appropriate. The problem is compounded when we are dealing with more than one language, for the 'breakpoint' between small and large tracts of trees is different for different languages. For multilingual generation, we can compare plesionyms in different languages in terms of their different element structures, so that it should be easier to choose the particular word in a particular language that fits a given situation.

Relation expressions. We have seen that often the distinctions between near-synonyms need to be expressed using common words and phrases. But we have shown that there are ways of expressing relations using fairly common vocabulary to represent these distinctions. The ease of relation identification may contribute towards relation verification: we can anticipate that psychological tests, of the sort Chaffin and Herrmann carried out for semantic relations, could be used to verify our relations and relation elements, as we can meaningfully and precisely represent the subjects' intuitions about the distinctions between near-synonyms.

Relation complexity. Relations may need to be described at more than one level of complexity, so that the distinctions between two words may be identified in more than one way. We have shown how a relation-element approach allows us to define different relation structures for the same pair of near synonyms (*e.g.*, *quarrel/row*).

Relation creativity. We have noted in previous work [DiMarco, Hirst, and Stede 1993] that the representation of the distinctions between near-synonyms would seem to require a constrained, but not finite, vocabulary. With a relation-element approach, we have seen how a basic set of relations might be constructed; new relation elements may be added, but we may be able to incorporate them into existing relations, so that the catalogue of relations need not grow uncontrollably. Thus, we can produce new relations by elaborating on existing, well-known relations or by concatenating existing relations [p. 322].

7 Implementing near-synonymic relations

We are currently investigating different systems for implementing a relational theory of near-synonymy. The first system that we are looking at is WordNet [Miller et al 1990], which seems particularly relevant as words are organized both by semantic relations and by "synsets" (synonym sets).

WordNet contains definitions of nouns, verbs, and adjectives; for now, we are concentrating on the representation of adjectival near-synonyms. In keeping with the philosophy of WordNet, we envisage the use of a pointer for each type of near-synonymic relation in our catalogue, so that we might represent the relations between plesionyms as follows:

r1 : (general (favourable$_{ji}$))
r2 : (general (formal$_{ji}$, forceful$_{ji}$))
r3 : (general (favourable, forceful$_{ji}$))

Currently, the coding of a synset of adjectives would look as follows in WordNet:

{ thin, slender, emaciated, thin1, & }

where "thin1, &" indicates that members of this synset are related to the 'concept' *thin1* by the similarity relation.

We can imagine imposing additional structure on a synset and making use of a catalogue of near-synonymic relations to obtain the following coding:

{ [thin, slender, r1], [thin, emaciated, r2], [slender, emaciated, r3], thin1, & }[8]

[8] In WordNet, square brackets are used to indicate a lexical relation between words, rather than a semantic relation between concepts.

While such a representation of near-synonymic relations would be very easy and natural in WordNet, it relies on the solution of a number of challenging problems, specifically, how to generate a complete set of near-synonymic relation elements, and how to define a constrained and reusable catalogue of near-synonymic relations.

8 Conclusion

Our eventual goal is a knowledge representation for the discrimination of near-synonyms. We have taken a step towards such a representation by investigating the nature of near-synonymic relations. We have set out four properties that any theory of near-synonymic relations should address, and we have shown how relation-element theory gives us a way of representing distinctions between near-synonyms that accounts for these properties. We are currently investigating the representation of our relational approach to near-synonymy in the WordNet system.

Acknowledgements

We thank Graeme Hirst for reading and giving incisive comments on earlier drafts of the paper. We also thank the anonymous referees for suggesting an extension of this work. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada.

References

DiMarco, Chrysanne and Hirst, Graeme. "Usage notes as the basis for a representation of near-synonymy for lexical choice." *Proceedings, Ninth Annual Conference of the University of Waterloo Centre for the New Oxford English Dictionary and Text Research*, Oxford, September 1993, 33–43.

DiMarco, Chrysanne; Hirst, Graeme; and Stede, Manfred. "The semantic and stylistic differentiation of synonyms and near-synonyms." *Proceedings, AAAI Spring Symposium on Building Lexicons for Machine Translation*, Stanford, March 1993.

Chaffin, Roger and Herrmann, Douglas J. "The nature of semantic relations: A comparison of two approaches." In: *Relational models of the lexicon*, Martha W. Evens (editor), Cambridge University Press, 1988, 288–334.

Cruse, D.A. *Lexical semantics.* Cambridge University Press, 1986.

Miller, George A.; Beckwith, Richard; Fellbaum, Christiane; Gross, Derek; and Miller, Katherine. "Five papers on Wordnet." CSL Report 43, July 1990, Cognitive Science Laboratory, Princeton University.

Oxford advanced learner's dictionary of current English, fourth edition. Oxford University Press, 1989.

Vinay, J.P. and Darbelnet, J. *Stylistique comparée du français et de l'anglais.* Beauchemin ltée, 1958.