

Paradigmatic relations and semantic information in MTT: some questions

M Lynne Murphy

University of Sussex
Brighton BN1 9QH UK
lynnem@cogs.susx.ac.uk

Résumé – Abstract

Cet article soulève et développe trois questions sur la structure de la théorie sens-texte, à partir de considérations sur le rôle que les fonctions lexicales paradigmatiques jouent dans son modèle du lexique: la distinction paradigmatique/syntagmatique a-t-elle un sens? Les fonctions lexicales paradigmatiques représentent-elles de l'information sémantique redondante (et si oui, la redondance est-elle justifiée)? Et enfin, les définitions de lexèmes dans la zone sémantique peuvent-elles être remplacées par une conception du sens lexical en tant que propriété émergente, sur le modèle des vues "structuralistes-associationnistes" du lexique? Bien qu'aucune réponse à ces questions ne soit proposée, cet article offre une panoplie de données pour contribuer au débat sur ces problèmes.

This paper raises and discusses three questions about the structure of Meaning-Text Theory, based on consideration of the role that paradigmatic Lexical Functions play in its model of the lexicon: Is the paradigmatic-syntagmatic distinction meaningful? Do paradigmatic LFs represent redundant semantic information (and if so, is the redundancy justified)? And finally, might definitions of lexemes in the Semantic Zone be replaced by an understanding of lexical meaning as an emergent property—on the model of structuralist-associationist views of the lexicon? While no answers to these questions are provided, a range of data are offered to contribute to the debate on these issues.

Keywords – Mots Clés

Lexicologie, fonctions lexicales paradigmatiques, Théorie Sens-Texte, antonymes, hyponymes, meronymes, synonymes

Lexicology, paradigmatic relations, Meaning-Text Theory, antonymy, hyponymy, meronymy, synonymy

1 Introduction

Theories of lexical relations are often compared as to whether their authors are ‘lumpers’ or ‘splitters’—in other words, whether they distinguish many or few types of lexical relation. The proponents of Meaning-Text Theory fall generally into the category of splitters, with Mel’čuk (1996) identifying 64 basic relations (25 of them paradigmatic), and others (e.g., Percova, 1996) adding more relations as they see fit. Furthermore, the Lexical Functions in MTT provide the raw materials for the creation of additional, combined Lexical Functions, which describe even more specific lexical relations. Thus there seems to be no inherent limit to the number of possible relations described by MTT.

Faced with the prospect of an indefinite number of lexical relations, my own work on paradigmatic relations has reached quite the opposite conclusion, that there are no definite boundaries among lexical relations and therefore the exercise of defining subtypes of lexical relations misses the point that all these relations have basic properties in common. In past work (Murphy, 1995, 2000), I have argued that all paradigmatic relations can be derived from two general relational principles, and in forthcoming work (Murphy, at press), I reduce that number of principles to one, which defines paradigmatic relations on the basis of minimal relevant difference in context. Thus, one might say that I am the utmost ‘lumper’. It must be granted that MT theorists and I have come to our respective splitting and lumping positions by means of very different goals and assumptions. My position is that of a psycholinguistically-oriented theoretical lexicologist with a bent toward the pragmatic (or some might say ‘cognitive’) resolution of linguistic problems. MTT’s roots and applications in NLP (and its focus on the function of language in and as text) led its developers to be particularly concerned with providing an accessible and explicit description of a language’s lexicon. Nevertheless, MTT has been proposed as a linguistic theory as well as a computational approach, and so can be critiqued from both vantage points. Below I discuss three questions of theoretical importance that are raised when a lumper like me considers the structure of MTT and the semantic metalanguage and lexical functions within it. Each of these questions is a corollary of a more basic question, specifically: if Lexical Functions describe lexical co-occurrence, then why are paradigmatic relations included among the lexical functions? This question has been raised before by, for example, Heylen (1995) and Alonso Ramos and Tutin (1996). Below I consider some different sorts of data and assumptions that might affect how one approaches these questions.

2 Is the paradigmatic/syntagmatic distinction meaningful?

From its beginnings, most proponents of MTT have observed a division between syntagmatic and paradigmatic lexical functions, although the terminology and definitions for these categories have varied over the years. The most basic definition (e.g., Mel’čuk, Žolkovskij, 1970) holds that syntagmatic LFs describe the tendency of words to co-occur, while paradigmatic LFs describe words that can substitute one for the other in text. The problem here is that these allegedly dichotomous categories have overlapping membership, as has been noted by Grimes (1990). So, for example, items related by the **Gener** LF frequently co-occur in sentential contexts, as in (1):

- (1) *a mango **fruit** compote*
*Oranges are not the only **fruit***
***Fruits**, like mangoes and kiwis, are good for you.*

Antonyms also co-occur in sentences at great rates, and thus Jones (2002) and Murphy (2003) raise the question of whether antonymy can instead be classed as a type of syntagmatic relation. Indeed Žolkovskij and Mel'čuk (1970) classed **Anti** as a syntagmatic (in their terms, *a parameter*) LF, although Alonso Ramos and Tutin object that antonyms do not co-occur in contexts like (2).

- (2) **He is small big.*

In a sense, this is an unfair criticism in that the example is ungrammatical, rather than semantically prevented. In contexts in which it is grammatical to concatenate words from the same grammatical class, we can find instances of antonymic co-occurrence, as in (3), and co-occurrence of other types of opposites, such as those in (4).

- (3) *Little Big Man* (film title)
She's been playing same-differents (i.e., a game—from CHILDES database)
both warm and cool colors
Her feelings went from hot to cold.
- (4) *a parent-child conference*
syntagmatic-paradigmatic divide
I love to hate you (song title)

Similarly in Chinese opposites co-occur in lexicalised compounds that indicate the dimension or generic category in which the opposites exist, as in (5).

- | | | | |
|-----|-----------------|---------------------------|----------------------|
| (5) | <i>hǎo-huài</i> | 'good-bad' = quality | |
| | <i>dà-xiǎo</i> | 'big-small' = size | |
| | <i>kuài-màn</i> | 'fast-slow' = speed | |
| | <i>shuǐ-tǔ</i> | 'water-earth' = climate | |
| | <i>fù-mǔ</i> | 'father-mother' = parents | (Li, Thompson, 1981) |

Such co-occurrences are counter-examples to the old-fashioned claim that paradigmatic relations involve words that do not co-occur, while syntagmatic relations involve co-occurrence. Alonso Ramos and Tutin (1996) offer definitions that avoid this problem: A paradigmatic LF associates a keyword with another lexeme (its value) such that the two lexemes share a "non-trivial semantic component" (p. 151), whereas a syntagmatic LF associates a keyword with another lexeme such that the two words have *restricted* lexical co-occurrence in phrases (Mel'čuk, Polguère, 1987). In cases where paradigmatically related words co-occur, the co-occurrence is semantically motivated and unrestricted. On this understanding, antonyms co-occur because people have reasons to use opposite meanings in text, not because of any arbitrary association between one and the other.

But because these definitions apply differently to different instantiations of the same LF, we still cannot reliably distinguish between paradigmatic and syntagmatic LFs. There are both syntagmatically related expressions that share a semantic component and paradigmatically related expressions that have been claimed to be in restricted co-occurrence.

In the first case, we have syntagmatically related pairs that seem to share significant semantic contents.¹ For example, **Pred** (Predicator) is categorized as a syntagmatic LF (Alonso, Tutin 1996), and exemplified by the relation between *debtor* and *owe*. Since a debtor is one who owes, the amount of semantic content shared between the two is certainly not trivial, and so this example also rides the boundary between the syntagmatic and the paradigmatic.

In the second case, many, including Gross et al. (1989) and Fellbaum (1995) have claimed that co-occurrence is restricted in some cases of lexical opposition. For a particular word, like *large*, many words are semantically possible antonyms, such as *little*, *small*, *minute*, *slight*, *tiny*, *short*, *peewee*, and so forth. But one of these words, *small*, co-occurs with *large* at far greater rates than any of the others, indicating (they argue) that the co-occurrence is more than semantically motivated, it is a privileged (or restricted) relation. Similarly, Cruse (1986) has argued that cross-linguistic evidence proves that antonym pairings are arbitrary, since, for example, the antonyms of translational equivalents across languages need not be translational equivalents themselves. Using markedness phenomena as evidence, Cruse shows that the German antonym pairing *gut/schlecht* has more in common with English *good/poor* than with the normal English pairing *good/bad*. Thus, he argues, these antonym pairings must be specified in the lexicon, since if they were semantically predictable, they would be consistent cross-linguistically. If these theoreticians are correct in arguing that some values of paradigmatic relations are in restricted co-occurrence with their headwords, then the boundary between the paradigmatic and the syntagmatic is blurred.

Contra Cruse and the WordNet authors, I've argued that such cases of lexical opposition are predictable (Murphy, 1995, 2000, at press). However, it's not clear that MTT can represent the predictability of these relations, since in order to predict them, both fine-grained semantic descriptions and non-semantic information must be available to the relational mechanism (see Murphy, at press). Such fine-grained semantic distinctions may not be available in the Semantic Zone of MTT, since the restricted vocabulary used in definitions might not allow for the types of slight differences needed to distinguish words like *big* and *large*. It is also not clear to me whether non-semantic factors, such as register and morphological complexity, are accessible to relational decision-making in MTT. It's entirely predictable on registral grounds that *dead* is a better opposite for *alive* than *croaked* or *expired*, and so it seems that to specify the *dead/alive* contrast through a Lexical Function is to include non-arbitrary information in the lexicon.

It would be easy to just ignore the alleged distinction between the syntagmatic and the paradigmatic were it not for the fact that Alonso Ramos and Tutin find that the division has explanatory value within MTT. Specifically, categorizing an LF as syntagmatic or paradigmatic allows for generalizations to be made about what types of other LFs that lexical function can combine with and what type of combined LF will result. These generalizations include: (a) when two simple LFs are combined to make a combined LF, only one of the pair can be paradigmatic, and (b) if the first LF in a combined LF is paradigmatic, then the LF will be of the compound (rather than complex) type. This sounds like an interesting generalization, but both of these generalizations about paradigmatic LFs do not apply to the paradigmatic LF **Anti** (as noted by Alonso Ramos and Tutin). The fact that these generalizations have exceptions again calls into question the validity of the paradigmatic and syntagmatic categories.

¹ Since Alonso Ramos and Tutin (1996) do not define what they mean by "non-trivial", it is difficult to test their claim that paradigmatic relations involve the sharing of "non-trivial semantic components".

Thus, we've seen both that instantiations of LFs in particular lexical entries (*large*, *debtor*) flout the definitions of 'paradigmatic' and 'syntagmatic' and that one LF (**Anti**) flouts the generalizations based on those definitions. What is clear is that the values of LFs differ in the extent to which they are motivated by restricted co-occurrence or semantic relatedness. The next question is whether lexical functions should be used to describe relations that are based on the sharing of semantic components.

3 Are paradigmatic Lexical Functions redundant?

Models of lexical knowledge can be contrasted in terms of two metaphors: the dictionary metaphor and the thesaurus metaphor. Models that follow the dictionary metaphor posit that lexical knowledge includes definitions of lexemes; these include the componential approaches to lexical meaning. Models that follow the thesaurus metaphor hold that lexical knowledge includes links among related lexemes; this is the structuralist, or associationist, approach to meaning. MTT falls in the middle of the continuum between dictionary- and thesaurus-style models, since it holds that lexical entries include both a definition (in the Semantic Zone) and lexical functions (in the Lexical Co-occurrence Zone).

Since paradigmatically related words share semantic information, their relation can usually be derived from the words' definitions. But as we've seen for antonymy, there is evidence that some relations are not just semantically based but also lexically based. LFs for such relations are well-motivated, then, as they demonstrate restricted lexical co-occurrence. This is consonant with the common (and influential) position that the lexicon stores only the arbitrary and idiosyncratic facts of a language (e.g., Bloomfield 1933, Chomsky 1965).

In MTT, all of a word's LF specifications need not be arbitrary—they can be derived from the word's definition. However any LF must at least sometimes have arbitrary specifications in order to justify its position in the model—otherwise the LF is simply redundant codification of semantic information. Mel'čuk and Wanner (1994, 1996) argue that the values of LFs can be partly motivated by the semantic properties of words, since words with LF values in common (or words that are common values of LFs for a set of words) often have semantic characteristics in common. For example, *plea*, *proposal*, *advice*, and *offer* are all nouns that can be the object of *reject*, and each of these nouns describes a verbal suggestion of some sort. So membership in the 'suggestion' class of nouns makes it likely that the word will have *reject* as one of the values for **Oper**. In order to efficiently represent the information that entire classes of words have particular LF values, information can be inherited from the class. The problem for doing this is that full semantic classes rarely act as one. So, instead of instituting a general hierarchy with automatic inheritance of properties from superordinate to subordinate, particular LFs in a subordinate word's lexical entry are marked for inheritance from the superordinate word's LF specifications.² Interestingly, however, this information is inherited from the word's definition (its semantic zone) rather than from another LF. So, while the LF **Gener** should identify the superordinate for any word that has one, inheritance depends on the existence of that superordinate in the definition of the word. The existence of superordinate terms in both definitions and in **Gener** specifications calls into question whether that LF serves any real purpose that cannot as easily be satisfied by information in the Semantic Zone. So we have evidence that a paradigmatic relation is not arbitrary (but is

² Mel'čuk and Wanner (1996) suggest that this system might better be replaced with a system for default inheritance.

instead dependent on componential meaning), and therefore we must question whether paradigmatic relations must be explicitly represented by LFs.

Like the type-token relationship represented by **Gener**, many LFs representing meronymic (whole-part) relations seem redundant of information that is semantically retrievable. For example, if *admiral* is defined as ‘a chief in a sea-faring military organization’, and *navy* is defined as ‘a sea-faring military organization’, then the relation among them is self-evident. For this reason, hyponymy and meronymy are usually considered to be conceptual rather than lexical relations. Synonymy is another relation whose representation as an LF is entirely redundant of information in the semantic zone. If a word is the value of **Syn** for another word, then presumably we could derive that relation as easily by comparing the two words’ definitions. The same can be said for most of the remainder of paradigmatic relations.

If we consider MTT to be a text-generating system for machines, then specifying paradigmatic relations through LF specifications might be valuable in that it can generate vocabulary for text cohesion. It’s less clear that such paradigmatic information should be presented so crudely in a theory of language in the mind. When a computer generates text, it puts words together, but when people generate text, they put ideas together. The reason that people put words like *admiral*, *navy*, and *military* in the same text is because the *meanings* of the words go together, not just because the words are marked in the lexicon as “words that co-occur”. Thus, it does not seem that the case for representing paradigmatic information in the lexicon is very strong (with the possible exception, once again, of some opposite relations³), given that the information is entirely derivable from words’ semantic representations. However, if paradigmatic information were not present in the Semantic Zone, then its representation in LF specifications would not be redundant.

4 Might meaning be an emergent property of LF specifications?

While the last section suggested that paradigmatic LFs are not necessary because they redundantly represent semantic information, the other way of avoiding this redundancy would be to remove it from the Semantic Zone.

Certainly, there are cases in which representation of a word’s meaning is best done via LF specifications, for instance where the semantic representation of a word depends on that of others. In such cases, the LF information would not repeat information from the Semantic Zone, but instead would give access to such semantic information. For example, my vocabulary is studded with American and British dialectal variants. Since American English is my first language, it is likely that the semantic representation of the British variants is dependent on that of the American words. We might represent this situation by giving a definition in the Semantic Zone for *gasoline*, but leaving the Semantic Zone for *petrol* empty and instead representing the synonym (or translational equivalence) relation between *petrol* and *gasoline* via LFs, as shown in (6) and (7):

³ One reason why opposites are different from other types of paradigmatic relation is because opposition is a particularly binary relation. So, while a word may have several synonyms, it has *an* opposite on any particular semantic dimension. For example, while *sour* has several potential opposites (*bitter*, *salty*, *spicy-hot*, *sweet*), just one among that group can be *sour*’s opposite in any particular context. See Murphy, at press.

- (6) *gasoline*:
Semantic Zone: a type of petroleum-based fuel
Lexical Co-occurrence Zone: **Syn**_[UK] = *petrol*
- (7) *petrol*
Semantic Zone: { }
Lexical Co-occurrence Zone: **Syn**_[US] = *gasoline*

Some other relations lend themselves equally well to this treatment. For example, the meaning of *pride* ('a group of lions') can be represented through the specification of the **Sing** and **Mult** LFs in the entries for *pride* and *lion*, respectively, without recourse to the Semantic Zone at all. The pertinent question then becomes: does any information really need to be presented in any word's Semantic Zone?

LFs can be seen as creating a semantic network (Fontenelle 1997), just like the networks that represent the semantic component of the lexicon in thesaurus metaphor models. While Mel'čuk cautions repeatedly that LFs are not semantic units—they are not meanings unto themselves and "they do not and should not pursue semantic precision" (1996: 80), one is left to wonder whether the semantic precision Mel'čuk seeks might instead be an emergent property of that network. This was the original aim of another computational approach to language, WordNet.

When originally conceived, WordNet was intended to represent meaning associatively (Miller 1998a). No definitions were to be included in the database, just relational links among words and synonym sets. It was thought that the network would be rich enough that words' senses could be distinguished by their unique positions in the relational network. This assumption turned out to be wrong, and explanatory glosses were added to distinguish synonym sets. (This can be seen as roughly analogous to the mixed [definition/LF] representation of semantic information in MTT.) According to Miller (1998b), if more types of semantic relations had been included in WordNet, better differentiation of senses might have been achieved. Many (if not all) of the extra semantic relations Miller had in mind would be subdivisions of the semantic relations already included in WordNet—for example, different types of meronymy.

While MTT is a cross between the dictionary and thesaurus approaches to modeling a lexicon, it does the thesaurus part particularly well. The sheer number and varying types of lexical functions (both paradigmatic and syntagmatic) mean that each word's LF specifications are bound to be unique or at least only identical with those of its perfect synonyms. It may take a certain leap of faith into semantic holism (see Fodor et al. 1980; Fodor, Lepore 1992) to allow for the possibility that meaning is an emergent property of a semantic network. A smaller leap of faith would be necessary if direct semantic representation were available for some of the words in the network—a small set of unique beginners. If these words were defined, then semantic values for other words in the network would be rooted in a set of semantically basic subset of words. This notion is related to, but at odds with, the goal that semantic values be represented in MTT by a simpler subset of the language's lexicon (following in the footsteps of Wierzbicka's [e.g., 1996] work on lexical primes). While both approaches effectively define the bulk of the vocabulary in terms of a small subset of basic natural language vocabulary, they necessarily have different means for discovering semantic primitives. For Mel'čuk and Polguère (1987), these primitives are to be discovered through consistent application of the Decomposition Principle in the creation of Semantic Zone primitives. In

the network approach, the consistent description of LF values should have much the same effect. I have not found any serious consideration of this possibility in the MTT literature, the bulk of which is concerned with LFs to the exclusion of Semantic Zone definitions. Because of this, it is not clear if there is a principled reason for including both relational and non-relational information in the description of lexemes, and it is also not clear whether there are any serious repercussions for the model if the Semantic Zone were to be largely abandoned. If it were, then it might result in a need for further splitting of LF types (for example, more types of lexical opposition), but it might result in a more consistent and challenging theory.

5 Conclusion (or lack thereof)

In the foregoing, I have refrained from trying to convince you that MTT is an inappropriate way to model the lexicon and that instead one should follow my preferred views about the lexicon and paradigmatic relations within it. (That task is left to Murphy, at press.) Instead, I have explored MTT from the perspective of a foreigner—raising questions about its priorities and presuppositions (and along the way I have been impressed by the accomplishments of MTT researchers). It's my hope that these musings might inspire answers from those who may have come to take MTT for granted. As such, this paper has no conclusions, only questions. I hope that MTT researchers will find these questions interesting enough to propose answers to them in a similar forum in future.

Acknowledgements

I am particularly grateful to the anonymous referees for this paper, who made helpful, concrete suggestions for the improvement of an earlier draft.

References

- Alonso Ramos M., Tutin A. (1996), A classification and description of lexical functions for the analysis of their combinations, in Wanner (1996), 147-167.
- Bloomfield L. (1933), *Language*, New York, Holt.
- Chomsky N. (1965), *Aspects of the theory of syntax*, Cambridge MA, MIT Press.
- Cruse, D.A. (1986), *Lexical semantics*, Cambridge, Cambridge University Press.
- Fellbaum C. (1995), Co-occurrence and antonymy, *International Journal of Lexicography* 8, 281-303.
- Fellbaum C. (1998), *WordNet: an electronic lexical database*, Cambridge MA, MIT Press.
- Fodor J.A., Garrett M.F., Walker E.C.T., Parkes C.H. (1980), Against definitions, *Cognition* 8, 263-367.
- Fodor J., Lepore E. (1992), *Holism: a shopper's guide*, Oxford, Blackwell.
- Fontenelle, T. (1997), *Turning a bilingual dictionary into a lexical-semantic database*, Tübingen, Max Niemeyer.

- Gross D., Fischer U., Miller G.A. (1989), The organization of adjectival meanings, *Journal of Memory and Language* 28, 92-106.
- Heylen D. (1995), Lexical functions, generative lexicons, and the world, in P. Saint-Dizier, E. Viegas (eds.), *Computational lexical semantics*, Cambridge, Cambridge University Press.
- Jones S. (2002), *Antonymy: a corpus-based approach*, London, Routledge.
- Li C.N., Thompson S.A. (1981), *Mandarin Chinese: a functional reference grammar*, Berkeley, University of California Press.
- Mel'čuk I.A. (1996), Lexical functions: a tool for the description of lexical relations in a lexicon, in Wanner (1996), 37-102.
- Mel'čuk I.A., Polguère A. (1987), A formal lexicon in the Meaning-Text Theory (or how to do lexica with words), *Computational Linguistics* 13, 261-75.
- Mel'čuk I.A., Wanner L. (1994), Lexical co-occurrence and lexical inheritance: emotion lexemes in German, *Lexikos* 4, 86-161.
- Mel'čuk I.A., Wanner L. (1996) Lexical functions and lexical inheritance for emotion lexemes in German, in Wanner (1996), 209-78.
- Mel'čuk I.A., Žolkovskij A.K. (1970), Towards a functioning 'Meaning Text' model of language, *Linguistics* 57, 10-47.
- Miller G.A. (1998a), Foreword, in Fellbaum (1998), xv-xxii.
- Miller G.A. (1998b), Nouns in WordNet, in Fellbaum (1998), 23-46.
- Murphy M.L. (1995), *In opposition to an organized lexicon: pragmatic principles and lexical semantic relations*, PhD thesis, University of Illinois at Urbana-Champaign.
- Murphy M.L. (2000), Knowledge of words versus knowledge about words: the conceptual basis of lexical relations, in B. Peeters (ed.), *The lexicon/encyclopedia interface*, Amsterdam, Elsevier, 317-48.
- Murphy M.L. (2003), Is 'paradigmatic construction' an oxymoron? Antonyms as lexical constructions, presented at the seminar Lexical Semantics without Stable Word Meanings, Mitteleuropa Foundation, Bolzano, Italy, 8 April.
- Murphy M.L. (at press), *Semantic relations and the lexicon*, Cambridge, Cambridge University Press. (to appear November 2003)
- Percova N. (1996), RUSLO: an automatic system for derivation in Russian, in Wanner (1996), 307-18.
- Wanner L. (ed.) (1996), *Lexical functions in lexicography and natural language processing*, Amsterdam, John Benjamins.
- Žolkovskij A.K., Mel'čuk I.A. (1970), Sur la synthèse sémantique, *T.A. Informations* 2, 1-85.