The Structure of Lexical Variation

Cognitive Linguistics Research

5

Editors

René Dirven

Ronald W. Langacker

Mouton de Gruyter
Berlin · New York

The Structure of Lexical Variation

Meaning, Naming, and Context

Dirk Geeraerts
Stefan Grondelaers
Peter Bakema

1994
Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter & Co., Berlin

⊗ Printed on acid-free paper
which falls within
the guidelines of the ANSI
to ensure permanence and durability

Library of Congress Cataloging-in-Publication Data

Geeraerts, D.

The structure of lexical variation : meaning, naming, and context /
by Dirk Geeraerts, Stefan Grondelaers, Peter Bakema.

p. cm. — (Cognitive linguistics research ; 5)

Includes bibliographical references and index.

ISBN 3-11-014387-9 (alk. paper)

1. Lexicology. 2. Language and languages — Variation. I. Grondelaers, Stefan, 1966 — . II. Bakema, Peter, 1967 — .
III. Title. IV. Series.

P326.G44 1994

413'.028 — dc20

94-12628

CIP

Die Deutsche Bibliothek — Cataloging-in-Publication Data

Geeraerts, Dirk:

The structure of lexical variation : meaning, naming, and context / by
Dirk Geeraerts; Stefan Grondelaers ; Peter Bakema. — Berlin ; New
York : Mouton de Gruyter, 1994

(Cognitive linguistics research ; 5)

ISBN 3-11-014387-9

NE: Grondelaers, Stefan;; Bakema, Peter;; GT

© Copyright 1994 by Walter de Gruyter & Co., D-10785 Berlin.

All rights reserved, including those of translation into foreign languages. No part of this book
may be reproduced or transmitted in any form or by any means, electronic or mechanical,
including photocopy, recording, or any information storage and retrieval system, without
permission in writing from the publisher.

Printing: Arthur Collignon GmbH, Berlin

Binding: Dieter Mikolai, Berlin

Printed in Germany

Acknowledgements

The present study contains the results of a research project that ran from 1990 to 1993 at the Research Centre for the Semantics of Syntax and the Lexicon of the Department of Linguistics of the University of Leuven. The project was supported by grant OT90/7 of the research council of the University of Leuven, and by grant 2.0078.90 of the Belgian National Science Foundation (NFWO).

The division of linguistic labor among the authors was as follows. Dirk Geeraerts was responsible for the inception, definition, planning, and supervision of the research. Peter Bakema and Stef Grondelaers compiled the corpus and prepared the materials used in the various analyses brought together in this book; database management was Stef Grondelaers's special task. The text of the book was written by Dirk Geeraerts.

For various kinds of practical help along the way, the authors owe a special gratitude to Lieve Herten, Jan Willems, Vincent de Keyzer, Willy Smedts, Paul Bijnens, Fred Truyen, Dirk Speelman, and Eliane Mahy. For critical comments with regard to earlier versions of the text, thanks are due to Patricia Defour, Karoline Claes, Petra Campe, Maarten Lemmens, Arthur Mettinger, Ron Langacker, René Dirven, and John Taylor. The present text is probably not the best book they can imagine, but their constructive criticism has certainly made it a better one.

Contents

1.	Varieties of variation	1
2.	Methods and materials	17
2.1.	Selected sources and database structure	18
2.2.	Characteristics of the corpus	32
2.3.	Points of methodology	37
3.	Semasiological variation	45
3.1.	Types of prototypicality	45
3.2.	Non-discreteness of word meanings [1]: definability	56
3.3.	Non-discreteness of word meanings [2]: uncertainty of membership status	76
3.4.	Non-equality of word meanings: salience effects	89
3.5.	The influence of contextual variation	105
4.	Onomasiological variation	117
4.1.	Non-discreteness in lexical fields: demarcation problems	118
4.2.	Non-equality in lexical fields: entrenchment	134
4.3.	The influence of contextual variation	146
5.	Formal variation	155
5.1.	The influence of prototypicality	156
5.2.	The influence of entrenchment	169
5.3.	The influence of contextual variation	177
6.	Ten theses about lexicology	189
	References	197
	Index of subjects	205

Chapter 1

Varieties of variation

Deciding what to wear is one thing – but deciding how to name what you are wearing is no less a matter of choice. Suppose you are putting on a pair of trousers made of strong blue cloth, such as are worn especially for work or as an informal kind of dress. Various lexical alternatives then suggest themselves: *jeans*, *blue jeans*, *trousers*, *pants*. But the options do not have the same value. *Jeans* and *blue jeans*, to begin with, have another meaning than *trousers* and *pants*: jeans are a type of trousers, whereas *trousers* names *all* two-legged outer garments covering the lower part of the body from the waist down, regardless of the specific kind involved. (In the technical terms of lexical semantics, *jeans* is a hyponym, or subordinate term, of the more general, superordinate term *trousers*.) *Pants*, on the other hand, represents a more complicated case than *trousers*, because it may be used both for the general class of trousers, and for a man's underpants. (In this case, *pants* is technically speaking a synonym of *underpants*.) The latter kind of usage, however, appears to be typical for British English. At the same time, *pants* in its more general reading is an informal term in comparison with *trousers* (but then again, this is a stylistic difference that occurs specifically in British English).

All the data in this example, summarized in Figure 1(1), have been taken from the first edition of the *Longman Lexicon of Contemporary English* (1981). Precisely because they involve lexical and semantic variation, it may well be the case that the data in the figure do not adequately capture the intuitions of all native speakers of English: the variation may be even more extensive than suggested here. The point about Figure 1(1), however, is not to achieve descriptive completeness with regard to *pants* and its cognates, but to illustrate the various types of variation that have to be taken into account in descriptive lexicological research.

The various *kinds of lexical variation* involved in the example, then, may be systematically distinguished in the following way. First, there is the fact that words may mean several things, as with the more restricted and the

more general reading of *pants*. Second, the same kind of referent may be named by various semantically distinct lexical categories, as illustrated by the choice between *jeans* / *blue jeans* and *trousers* / *pants*: even though *jeans* and *pants* are not synonyms, there are situations in which both are appropriate names for a particular garment. In fact, any time *jeans* is

	British English		American English
	formal	informal	
two-legged outer garment (in general)	<i>trousers</i>	<i>pants</i>	<i>pants</i> / <i>trousers</i>
men's under- wear	<i>underpants</i>	<i>pants</i>	<i>underpants</i>
trousers made of strong blue cloth	<i>jeans</i> / <i>blue jeans</i>		<i>jeans</i> / <i>blue jeans</i>

Figure 1(1)

Sample lexical data on *pants* and cognate terms

appropriate, the hyperonymous term *pants* will be suited as well; the reverse, of course, is not the case. Third, the same kind of referent may be named by various words, which may or may not differ from a semantic point of view; this type of variation, then, encompasses the previous one. The choice between *trousers* and *pants* (in its general reading), for instance, may be influenced by considerations of formality and stylistic appropriateness, but does not involve denotational semantic differences of the type distinguishing *jeans* and *trousers*. Even though they do not have precisely the same stylistic value (at least in British English, *pants* is more informal than *trousers*), *trousers* and *pants* (in its general reading) are equivalent as far as their meanings are concerned. Therefore, in a situation in which a particular

garment may receive the name *jeans* or *pants* or *trousers*, the pairs of alternatives have a different status. In choosing between *jeans* and *trousers*, for instance, the choice is not just between words, but between different semantic categories. In choosing between *trousers* and *pants*, on the other hand, the choice is between words that are semantically equivalent, but that are invested with different stylistic values. Finally, the stylistic distinction that exists between *trousers* and *pants* is an example of a more general contextual type of variation, involving the fact that a specific lexical phenomenon (such as a preference for expressing a particular meaning by means of one item rather than another) may be subject to the influence of contextual factors, like a speech situation asking for a particular style, or geographical distinctions among groups of speakers.

The purpose of the present study is to explore the structural characteristics of these varieties of lexical variation taken by themselves, and of the way in which they interact with each other. Notice, in this respect, that the four types interlock and overlap in intricate ways. Contextual variation, for instance, is not restricted to the formal side of the language, but touches upon the semantic phenomena as well. In the example contrasting *trousers* and *pants* (in its general reading), the contextual, stylistic variation involves words that are otherwise semantically equivalent. However, the meaning variation exhibited by *pants*, also correlates with contextual factors of a geographical nature: contextual variation (the fourth type mentioned above) may crosscategorize with the semantic variation mentioned as the first type above. What we will try to do, then, is not just to study each variational perspective in its own right, but to disentangle the interaction between the various types of variation. Studying one of them separately, indeed, cannot be done properly if the question is not asked to what extent the phenomenon in question might be influenced by any of the others.

In order to make the research more manageable, let us introduce a number of *terminological distinctions*. We will use the following terms to refer to the four different kinds of variation that we have informally identified above.

Semasiological variation:

the situation that a particular lexical item may refer to distinct types of referents.

Onomasiological variation:

the situation that a referent or type of referent may be named by means of various conceptually distinct lexical categories.

Formal variation:

the situation that a particular referent or type of referent may be named by means of various lexical items, regardless of whether these represent conceptually different categories or not.

Contextual variation:

the situation that variational phenomena of the kind just specified may themselves correlate with contextual factors such as the formality of the speech situation, or the geographical and sociological characteristics of the participants in the communicative interaction.

These concepts are illustrated in Figure 1(2) on the basis of the *pants/trousers*-example as described in Figure 1(1). The figure may be read as follows. Semasiological variation involves the situation that one word may possess diverse semantic values, as when *pants* may either be synonymous with *trousers* 'two-legged outer garment covering the lower half of the body', or with *underpants* 'a man's short undergarment worn below the waist'. Onomasiological variation involves the situation that the same thing may be identified as a member of different categories. In a given situation, for instance, a particular pair of trousers might be referred to either as a member of the category *trousers/pants*, or as a member of the subordinate category *jeans/blue jeans*. Semasiological and onomasiological variation are both forms of conceptual (or "semantic") variation: they involve differences of categorization. Semasiological and onomasiological variation study lexical categorization from different perspectives: the semasiological approach takes its startingpoint in the words naming a conceptual category, while the onomasiological approach takes its startingpoint in the things categorized. Contextual variation involves speaker-related and situation-related differences, such as the stylistic differences distinguishing *pants* (in its general reading) and *trousers* in British English. The geographical differences between British English and American English also fall within the class of contextual variation. As explained before, contextual variation is not necessarily restricted to cases such as the *pants/trousers*-example in the figure, which does not involve semantic differences: contextual variation and conceptual variation of the semasiological or onomasiological kind may clearly crosscategorize. Formal variation basically involves the situation that a particular entity may be referred to by means of different words. These different words may express a conceptual distinction, in which case we get onomasiological variation, or they may not, in which case we get, for instance, "pure" geographical variation.

conceptual variation	semasio-logical variation	<i>pants</i> (1) trousers (two-legged garment etc.) (2) men's underwear	
	onomasio-logical variation	<i>jeans/blue jeans</i> or <i>trousers/pants(1)</i>	formal variation
contextual variation		<i>pants(1)</i> (informal British English) versus <i>trousers</i> (less informal British English)	

Figure 1(2)

An illustration of the major terminologically distinct forms of lexical variation

The terminological distinctions illustrated in Figure 1(2) invite a number of comments that will help to describe the *theoretical background* of the study presented here. The terminological pair *onomasiology* / *semasiology* is a traditional one in European lexicology and lexicography. (See, for instance, Kronasser 1952 and Quadri 1952 for detailed overviews of the achievements of early semasiology and onomasiology, respectively.)

The pair *onomasiology* / *semasiology* is generally regarded as identifying two different perspectives for studying the relationship between words and their semantic values. The semasiological perspective takes its startingpoint in the word as a form, and describes what semantic values (as dependent variable) the word (as independent variable) may receive. The onomasiological perspective takes its startingpoint on the level of semantic values and describes how a particular semantic value (as independent variable) may be variously expressed by means of different words (as dependent variables). In actual practice, onomasiological research is rather concerned with sets of related concepts than with a single semantic category; as such, it traditionally coincides with lexical field research. In this respect, it is worthwhile to

note that the introduction of the onomasiological perspective into lexicological research was a typical aspect of the structuralist phase in the development of lexical semantics, which followed upon the prestructuralist phase dominated by historical semantics. (See Geeraerts 1988a, 1986b for the background of this division in periods.) While the prestructuralist phase in the history of lexical semantics had a predominantly semasiological focus (concentrating as it did on the changes of meaning of individual words), the structuralist stage stressed the necessity of complementing the semasiological perspective with an onomasiological one. (On the complementarity of both perspectives, see for instance Baldinger 1980.)

Now, given the traditional terminological distinction between semasiology and onomasiology, our own use of these terms will be slightly different with regard to the tradition. First, the semasiological perspective traditionally only deals with semantic variability in the form of polysemy: the fact that words may have different meanings (like "trousers worn by women" and "trousers in general" in the case of *pants*). We will use the term in a broader way, including types of variability that would traditionally be considered cases of referential rather than semantic variation. Think, for instance, of the fact that jeans may have various widths ranging from tight-fitting to decidedly wide. There is no reason to suppose that the word *jeans* has, say, two distinct meanings: "tight-fitting trousers made of strong blue cloth, such as are worn especially for work or as an informal kind of dress", and "wide trousers made of strong blue cloth, such as are worn especially for work or as an informal kind of dress". Such a distinction is traditionally considered a case of mere referential variability (or vagueness), and as such, structurally irrelevant: the classical idea is that only semantic variation of the polysemic kind is structurally important and hence worthy of linguistic scrutiny. In line with the general trend in prototype theory and cognitive semantics we will, however, include this type of referential variability in the realm of semasiological research. (We will, in fact, concentrate on it, rather than on the traditional topic of polysemy.)

There are two reasons for this extension of the domain of semasiology. First, the distinction between vagueness and polysemy appears to be less strict than has been traditionally assumed. There is now evidence (which we will not repeat here; see Geeraerts 1993) that the various criteria that have been invoked as operational tests for the distinction between vagueness and polysemy yield contradictory results, and furthermore, that they may be subject to contextual variation. Second, referential variation can be shown to be structurally relevant. This is, in fact, one aspect of the data that we will

be presenting ourselves. In section 5.1., for instance, we will show that differences of referential salience influence choices among lexical alternatives: a word is used more often as a name for a referent of a particular type when that referent occupies a statistically prominent position within the referential range of application of the item. And in section 3.2., referential distinctions will turn out to be definitionally important when items cannot receive classical definitions in terms of a necessary-and-sufficient set of definitional features. Throughout the study, these and similar observations will justify broadening the scope of semasiology to include referential, non-polysemic variability.

With regard to *onomasiology*, on the other hand, our use of the term is more restricted than is usual in structural semantics. Traditionally, the distinction between a conceptual interpretation of the onomasiological perspective (involving a choice among distinct conceptual categories) and a purely formal interpretation (involving a choice among various word forms, regardless of their conceptual status) is not prominently present in lexicological theorizing. Lexical field analysis tends to describe systems of related alternative words in their mutual relations, but largely ignores the question when or why one of the alternatives within the system rather than another is chosen as a name for a particular type of referent. We would like to suggest, on the other hand, that the latter question can only be answered properly if a distinction is maintained between the semantic and the formal aspects of onomasiology in the broad sense. There are, in fact, two aspects to the onomasiological selection of a name for a referent: there is the choice of a conceptual category for identifying or describing the referent, and there is the choice of a lexical item naming that category. For instance, when you decide to identify or describe a particular garment as a member of the category "jeans", there is a formal choice to be made between *jeans* and *blue jeans*. Against this background of a conceptual and a formal interpretation of onomasiology in the broad, traditional sense, we propose a terminological distinction between *onomasiology* (in a restricted sense), involving categorial, conceptual variation in naming referents, and *formal variation*, involving the selection of different word forms regardless of the question of categorial variability as manifested in lexical fields.

Given, then, the terminological distinctions introduced above, the *scope of the following study* can be described with more precision with regard to each of the four types of variation. In a succinct and highly abstract way (to be illustrated and made more concrete presently), the crucial points can be formulated as follows.

With regard to *semasiological* and *onomasiological* variation, we will show that both the (onomasiological) structure of lexical fields and the (semasiological) structure of the range of application of a single word are pervasively characterized by two phenomena: non-discreteness and non-equality (or, to put it more positively, by flexibility and salience effects).

On the *semasiological* level, non-discreteness will show up in the absence of definitions in terms of necessary and sufficient attributes, and in differences of membership status. Non-equality will show up in definitional clustering and differences of definitional weight, and in differences of membership salience.

On the *onomasiological* level, non-discreteness will show up as the absence of a mosaic-like lexical field structure. Non-equality will show up in the fact that various categories may have various degrees of entrenchment, entrenchment being defined as onomasiological salience.

With regard to *formal* variation, we will show that the choice of a particular word for naming a referent is determined by three factors. Apart from the contextual influence of "sociolinguistic" variables like geographical distinctions or register, lexical choices are determined by the semasiological and onomasiological characteristics of the referents involved: a referent (or set of referents) is expressed more readily by a category of which it is a central member, and it is expressed more readily by an item with a higher entrenchment value.

With regard to *contextual* variation, we will show that the semasiological, onomasiological, and formal types of variation may all be subject to contextual or situational effects.

These phenomena will be illustrated by means of a case study involving the field of clothing terms in contemporary Dutch. From general magazines and fashion magazines that appeared in the year 1991, we have collected roughly 9000 instances of words identifying garments. Only lexical items have been included of which the actual referent could be identified on a picture accompanying the text. For each occurrence of the item *broek* 'trousers' that is included in the database, for instance, we know on the basis

of visual information whether it names a pair of jeans, or a pair of knickerbockers, or a pair of bermuda trousers, or whatever other type of trousers. This referential information is included in the database by means of a componential description identifying referential characteristics such as the length, width, material etc. of trousers. It should be clear how this organization of the database allows for a systematic investigation into the various forms of variation that interest us here. Studying semasiological variation methodologically involves questions of the type "Given a word like *trousers*, what are the referents that it occurs with?". Research into formal and onomasiological variation, on the other hand, will take its startingpoint in questions of the form "Given a particular type of garment (like a two-legged outer garment covering the body from the waist down, made of strong blue cloth), what are the words that alternatively name it?". And an inquiry into contextual variation will open with questions such as "Is there a systematic difference between the semasiological range of application of *trousers* in Dutch magazines published in The Netherlands and its range of application in Dutch magazines published in Belgium?".

The major findings of the investigation can be illustrated by means of a brief overview of the main body of the book. In the chapter on semasiological variation (chapter 3), we will disentangle the various prototypicality effects that characterize semasiological structures. Prototypicality, in fact, comes in many forms. It may involve the fact that it is impossible to give a definition of a word that conforms to the classical ideal of being formulated in terms of necessary and sufficient conditions. Or it may involve the fact that genuine and persistent doubts exist about whether a particular entity is a member of a given category. Or it may involve the fact that some of the members of a category carry more weight within the structure of the category, in the sense of being central cases of the set. The latter type of prototypicality will be particularly important for the architecture of our investigation: when we come to the study of formal variation, we will show that the selection of a word for naming a particular garment directly correlates with the semasiological structure (in terms of salience and centrality) of the category represented by that word. Chapter 3, incidentally, will also establish that the semasiological structure of lexical items is subject to contextual variation. We will present examples where the use of a word in Belgian Dutch differs significantly from its use in Netherlandic Dutch.

The chapter on onomasiological variation (chapter 4) basically discusses the internal structure of lexical fields (where lexical fields can be loosely regarded as sets of alternatives for naming particular entities). Most impor-

tantly, it will be shown that the items in a field are characterized by different degrees of onomasiological salience, just like the members of a category are characterized by different degrees of prototypical, semasiological salience. The concept of onomasiological salience that we will introduce (and operationally define) in chapter 4 is a generalization of the notion "basic level category" as formulated by Berlin (1973, 1974, 1978). According to Berlin's approach, the basic level in a hierarchical taxonomy is the one that will normally be used: in a taxonomy including the superordinate term *garment*, the generic terms *skirt*, *trousers*, and *suit*, and the subordinate terms *wrap-around skirt*, *pleated skirt*, *legging*, and *jeans*, the intermediate level including *skirt*, *trousers*, and *suit* would probably be the basic level, because *skirt*, *trousers*, and *suit* are more obvious and more usual names to identify garments than *garment* or *wrap-around skirt* and *pleated skirt*. The basic level, in other words, is defined in terms of naming preferences: given a particular referent, the most likely name for that referent from among the alternatives provided by the taxonomy will be the name situated at the basic level. In this sense, the basic level is onomasiologically salient: within the lexical field defined by the taxonomy, the generic level specifies a set of preponderant items. In a more psychological vein, basic level categories can be said to have a high cognitive entrenchment: they are firmly engrained in the mental lexicon. Chapter 4, then, will show that onomasiological entrenchment is not (or at least not purely) a matter of taxonomical levels, but rather involves individual categories regardless of the taxonomical level they belong to. We will be able to show, for instance, that the entrenchment of *jeans* is roughly of the same magnitude as that of *broek* 'trousers', in spite of the fact that *jeans* is a subordinate term of *broek*. Furthermore, we will clarify that entrenchment values are subject to contextual variation. In particular, the onomasiological salience of certain items will appear to be different in specialized fashion magazines in comparison with the general magazines.

The chapter on formal variation (chapter 5) deals with the crucial question of which factors determine the choice of a lexical item over potential alternatives. For one thing, it will turn out that contextual factors are important: certain items are more typically Belgian Dutch than Netherlandic Dutch, for instance. This is, to be sure, not a shocking conclusion: already in the *pants/trousers*-example with which we started this chapter, differences between stylistic registers and differences between British English and American English could be observed. More important, however, will be the recognition that the semasiological and onomasiological structures described in chapters 3 and 4 have an impact on the phenomenon of lexical choice.

Starting from semasiological salience effects, it will be shown that there exists a tendency for entities to be preferentially named by means of a category to which they typically belong. Starting from onomasiological salience, there will appear to be an independent tendency for entities to be named by the items with the highest entrenchment value of a set of potential alternatives. For instance, when something is both a *wikkelrok* 'wrap-around skirt' and a *plooirok* 'pleated skirt', an existing preference for calling such a garment either *wikkelrok* or *plooirok* reflects the relationship between the (independently established) entrenchment values of both categories. Apart from contextual effects, then, lexical choices are determined by the semasiological and onomasiological characteristics of the items involved: a referent (or set of referents) is expressed more readily by a category of which it is a central member, and it is expressed more readily by an item with a higher entrenchment value.

In order to specify what we believe to be the *innovative character* of the picture we are painting, it will be useful to compare our approach with two main approaches to lexicology and lexical semantics: the structuralist one (as represented by Lyons 1963, Pottier 1964, Lehrer 1974, or Coseriu & Geckeler 1981) and the cognitive one (as represented by the work of Vandeloise 1986, Taylor 1989, Brugman 1989, Cuyckens 1991, and others). For three reasons, our general perspective on the semantics of the lexicon is a cognitive one.

First, as mentioned already, we do not believe that it is either possible or useful to disregard referential variation when doing semantic analysis. In line with the general tenets of structuralism, structural semantics has tended to stress the mutual distinctiveness of lexical categories at the expense of their internal structure and their referential connections. In the structuralist approach, the way in which lexical items differ with regard to each other is the most important aspect of semantic analyses, because it is precisely these external boundaries of an item with regard to other items that determine its position within the structure of the lexicon; a structure, in a sense, is nothing else than a set of mutual delimitations. This perspective has led to a relative disregard of the structured variation that may exist within lexical categories – a situation that has been radically reversed by the prototype-theoretical interest in structured polysemy and the structural relevance of referential multiplicity.

The general framework of the present study is cognitive for yet another reason: the features in terms of which we will describe the internal structure of lexical categories are the ones that have been brought to the fore precisely

by prototype-theoretical approaches to lexical semantics. Flexibility (non-rigidity) and salience (non-equality) are no phenomena that were central to the structuralist approach. Although it is a bit of an exaggeration to claim (as is sometimes done by proponents of the cognitive approach) that the prototype-oriented conception of the semantic structure of linguistic categories is an absolute novelty in the history of lexical semantics, it is at least true to say that prototype theory has revived the interest in phenomena that were in focus in the prestructuralist stage of the development of lexical semantics, but that were relegated to the background of the attention in the structuralist era. (Again, see Geeraerts 1988a on these historical connections.)

Last but not least, our approach is a cognitive one because it focuses on questions of categorization: what is the internal structure of lexical categories, and how can the same entity be alternatively categorized? The relationship between these two fundamental questions is aptly defined by Kleiber (1991: 35) in the following way.

La question *Pourquoi appelle-t-on x Z?* est en fait une question ambiguë. Elle correspond, soit à une interrogation qui porte sur le choix du terme Z par rapport aux termes qui ne conviennent pas à x, soit à une interrogation qui concerne le choix du terme Z par rapport à des catégories ou dénominations qui conviennent également à x.

[The question *Why is x called Z?* is in fact an ambiguous one. It involves either an investigation into the choice of the term Z in comparison with alternative terms that do not fit x, or an investigation that envisages the choice of Z in comparison with categories or names that equally apply to x.]

Kleiber's first question is the semasiological one: what are the conditions for x to fall within the range of application of Z? Why can x be categorized as a Z at all? What are the restrictions on the use of Z that allow x as a member of Z but exclude y? Kleiber's second question is the onomasiological one: what are the conditions for Z to be used as a name for x rather than W (given that x is a member of both Z and W)? What are the restrictions on the use of Z and W that favor the selection of one at the expense of the other? Crucially, these are both questions about categorization: questions about the definition and the internal structure of categories, and questions about the choice among alternative categories. And it is precisely because of its emphasis on categorization that the investigation is a cognitive one: if anything

at all, Cognitive Linguistics is a theory about categorization in and through language (see a.o. Lakoff 1987: xi-xvii; Taylor 1989: vii-viii).

Against the background of the overall cognitive orientation of the present study, three specific characteristics have to be mentioned as additions to the mainstream of prototype-oriented research within the cognitive tradition. One characteristic follows in a straightforward manner from structural semantics. The systematic addition of an onomasiological perspective to the predominantly semasiological perspective of prototype theory corresponds with the legacy of structuralist field theory. In this sense, our research links up with and elaborates on recent attempts (like those of Lehrer 1990 and Schmid 1993a) to introduce the major findings of prototype theory into the tradition of lexical field theory. At the same time, of course, the onomasiological, field-theoretical questions will be tackled from an eminently cognitive point of view: we will show that the structural features of non-rigidity and non-discreteness that characterize the semasiological structure of lexical categories, also fundamentally shape the structure of lexical fields. Specifically, as we mentioned above, the importance that we will attach to onomasiological salience effects in lexical fields rests on a generalization of Berlin's concept of basic level.

A second innovation with regard to cognitive lexicology in its mainstream form is the addition of a contextual perspective to the study of lexical variation. Cognitive linguistics has a lot of attention for cultural differences in the relationship between language and the world, but the variation that may exist within a single linguistic community has not often been investigated from a cognitive point of view. Broadening the scope of cognitive lexicology in this way implies that a connection with sociolinguistics is made: the "contextual" variation to be considered refers precisely to the kind of speaker-related and situation-related variables that are customary within sociolinguistics. Such a connection, if successful, is not just important for Cognitive Linguistics alone, given that the study of semantic variation is a traditionally neglected area of sociolinguistics (see Heath 1988). This extension of the scope of the lexicological investigation links up with existing work in anthropological linguistics such as Dougherty (1978), Kempton (1981), and MacLaury (1991b), where differences of conceptualization within a single community are explicitly envisaged.

Third, the kind of non-elicited referential data that we will use is new within the Cognitive Linguistics paradigm. Prototype-theoretical research within the psycholinguistic and anthropological linguistic traditions has mainly been based on elicitation techniques like experiments and question-

naires. The research that has been done within the field of theoretical linguistics, on the other hand, has been based mainly on introspection or text corpora. Our own intention is to avoid both the problems that come with the use of elicited data (like the observer's paradox) and the disadvantages of using mere text corpora (viz. that no independent access to the referents of the linguistic utterances is guaranteed). How exactly we will try to achieve this will be explained in more detail in the next chapter, but it may be pointed out here that the use of this kind of referentially enriched non-elicited data is, to our knowledge, a novelty within Cognitive Linguistics. The use of referential data as such (that is, data about the range of actual denotata that a word may refer to) is not new: in this respect, our investigation may be placed in the tradition of the lexicological studies carried out by Labov (1973, 1978), and in that of the anthropological color terminology research exemplified by the work of Berlin & Kay (1969), and the early work of Rosch (Heider 1972, Heider & Olivier 1972). We will broaden the scope of this denotational line of research by applying it to strictly non-elicited language materials.

All in all, the approach that is probably closest to our own within the landscape of present-day lexicology is the one presented by Robert MacLaury in his work on Mesoamerican color terminology (1987, 1991b, 1992, 1991a). The features that we share with his approach are: an attempt to combine research into prototypicality effects within individual words with field research into the relationship between various related items; an interest in salience relations among alternative denominations for the same referents; and an explicit consideration of the social variation that exists among speakers of the same community. Still, there are some important distinctions between MacLaury's approach and the one presented here. Ours, in fact, is both broader and narrower in scope than MacLaury's. It is narrower in scope because it does not extend towards the diachronic domain. While MacLaury tries to formulate a universal model of color category evolution, we will have no claims to make about patterns of historical evolution. This has primarily something to do with the fact that our synchronic data do not allow an unambiguous historical interpretation, and secondarily with the fact that MacLaury's explanatory model of color term evolution takes its startingpoint in physiological salience effects in the field of color perception. The field of clothing terms obviously does not rely on any specific, physiologically universal perceptual capacity of human beings, and it is therefore impossible to posit a universal perceptual startingpoint for an evolution leading to the synchronic situation that is described by the investigation. Our ap-

proach is broader than MacLaury's because it attempts to define a more explicit and more sophisticated model of lexical variation. While MacLaury's primary focus lies with the specifics of color categorization and color term evolution, our main centre of interest lies with the structure of lexical variation. This means, among other things, that we will (as already described) devote explicit attention to the way in which the various forms of variation interact with each other, and that we will maintain a systematic distinction between the formal and the conceptual aspects of onomasiological variation – features that are absent in MacLaury's model. The model MacLaury develops contains generalizations about color categorization, whereas the present study develops a model for the general architecture of lexical variation.

The innovative features mentioned above do not mean, to be sure, that our approach is not subject to important *restrictions*. Three of them should be mentioned explicitly from the outset. First of all, there is a restriction on the scope of the investigation. Given a division of the field of semasiological variation into referential and polysemous variation, we will be concerned almost exclusively with referential variation. The semasiological variation that we will consider is not of a kind that would spontaneously or traditionally be classified as being polysemic. There is a lot of variation in the use of the clothing terms that we will consider, but even in the face of this variation, the term would not normally be considered polysemic. At the same time, a complete theory of lexical variation will obviously have to include a theory of polysemy, and the recent resurgence of the interest in semantic relations like metaphor and metonymy (among many others, see Lakoff & Johnson 1980, Paprotte & Dirven 1985, Sweetser 1990) proves that Cognitive Linguistics is in the process of developing such a theory. This implies, in other words, that the step we will try to take here towards a comprehensive model of lexical variation will eventually have to be followed by an additional step, in which the scope of the investigation is broadened to include variation of a polysemic kind. A second major restriction is of a methodological kind. Although we will often use quantitative data to bring home a particular point, we will not try to develop a statistically sophisticated approach for studying the various forms of variation that we will identify. The kind of analyses that we have to offer remain informal in fundamental respects, but we hope that they will suffice to establish why it may be interesting to try to tackle the field of lexical variation with more refined techniques for quantitative data analysis. The third restriction is again a methodological one, but it has to do with the scope of the analyses rather than with their

kind. As we are primarily interested in sketching a framework for the study of lexical variation, the analyses we will present will consist of illustrative cases studies: exemplariness rather than exhaustiveness will be our guiding-line. This means, for instance, that not all of the lexical items in the field that we will investigate will be presented with an equally detailed descriptive analysis. It will become clear soon enough in the course of the text that such exhaustiveness is beyond the limits of a medium-sized book like the present one. More importantly, a complete coverage of all aspects of the field with the same kind of descriptive depth is not necessary for our purpose. What we primarily try to do is to present a descriptive framework for the study of lexical variation: a systematically interrelated set of questions to ask, and a number of analytic tools for answering them. What is essential, we feel, is the coherence of the framework, the importance of the questions, and the relevance of the tools, rather than the specifics of the field under investigation.

In various respects, then, this study can only be a modest and moderate first step towards the development of a cognitive semantic theory of variation within lexical fields. In spite of its limitations, however, we hope that it may be appreciated as a fair contribution to the attainment of that goal. Over and above the specific innovative features mentioned above, we believe that it is the comprehensiveness of the analytic framework that we describe, that might determine its potential importance. As far as we can judge, the model presented here is the first to bring together the various forms of lexical variation that may be distinguished into a single systematic study. If there is anything of lasting value in the present approach at all, we think it will be our attempt to sketch a way of studying various related forms of lexical variation in their mutual interdependencies.

Chapter 2

Methods and materials

The three major kinds of linguistic method mentioned by Labov (1972) also apply in the realm of lexicology: lexical analysis may be based on *introspection*, on the *elicitation* of data by means of surveys and experiments, and on the *observation* of non-elicited language use. The first method is illustrated by the work of Anna Wierzbicka, who has vigorously defended it in a number of recent publications (1991, 1985). The second method is the usual one in psycholinguistics, as in Eleanor Rosch's work on prototypicality, which has given such a strong impetus to the development of cognitive semantics (see the following chapter); it is also the method used by the lexicological work of William Labov himself and his students (Labov 1973, 1978; Kempton 1981). The third method may be best exemplified by the traditional approach of lexicography; large-scale reference works like Murray's *Oxford English Dictionary* (or, more recently, the *Collins Cobuild* dictionary) are often based on a huge corpus of textual quotations.

For one general methodological reason, the present study will opt for the latter, corpus-based approach. Such an approach, in fact, minimizes the danger of methodological distortions. Specific elicitation techniques, for instance, may guide the informant towards a particular kind of answer – if only through his conscious awareness of the researcher's presence. Admittedly, a careful experimental design will often succeed in reducing the danger to negligible proportions, but even so, circumventing the problem by focusing on non-elicited language is at least as attractive. More importantly, what people think they do with words is not necessarily the same as what they actually do, in the sense that our conscious awareness of the flexibility with which we use the lexical resources of the language may well be rather restricted. Empirical evidence to this effect may be found in Geeraerts (1988b), where a corpus-based analysis of the nineteenth-century readings of the Dutch words *vernien* and *vernietigen* is compared with the treatment that both words receive in the nineteenth-century dictionaries of Dutch (which were invariably based on introspection). It appears that the intro-

spective method succeeds very well in pinpointing the prototypical core of the items in question, but is hardly able to capture the peripheral uses to which the core meanings appear to give rise in actual usage. Precisely because the relationship between core and periphery is of central importance for a cognitive approach to the lexicon, we will try to avoid the limitations of the introspective approach.

Given this general preference for a corpus-based approach, the present chapter will do three things: first, specify how the corpus has been compiled, second, give a global description of the resulting database, and third, discuss some of the theoretical aspects of the methodology adopted here in more detail.

2.1. Selected sources and database structure

A major drawback of traditional corpus-based lexicology is the absence of a *direct, independent acquaintance with the referents* of the expressions being studied. When, for instance, lexicographical reference works are compiled on the basis of a corpus of quotations, there is basically only the textual context to support the interpretation of the word forms. This may suffice to establish that *coat* refers to a particular type of outer garment with sleeves, covering at least the upper part of the body, but it will probably be insufficient to determine what different types of coats exist, how they vary in length, whether *coat* can be applied to such a garment when it has a hood, and so on. But precisely by choosing the lexical field of clothing terminology as the domain of research, this drawback of traditional textual corpora can be avoided. In fashion magazines and other periodicals, in fact, names for garments often refer to items of clothing that are represented by means of photographs or pictures. By compiling a corpus of such names, the referential range of application of a particular term may be established on independent grounds, without having to resort to mere textual interpretation. In this respect, then, the database for the present study is restricted to names for clothing types with independently describable referents.

In actual practice, material has been collected from twelve *magazines*. As to their type, three classes of magazines may be distinguished within the selection: fashion magazines, women's magazines, and glossy magazines. The magazines of the first class contain only or predominantly articles about

clothing; specifically, they may contain contributions with patterns and instructions for making particular garments on one's own. Women's magazines contain articles about fashion next to various other contributions; they do not contain patterns. These characteristics also hold for the glossy magazines, but the latter typically have a more sophisticated outlook than the women's magazines. Characteristically, the glossies have a lower frequency of appearance (monthly instead of weekly), and are markedly more expensive. While the women's magazines feature articles on all practical matters, the contributions in the glossies are largely restricted to "lifestyle" issues like art and culture, fashion, travel, and gastronomy. Moreover, some glossies are specifically intended for men, while the other class by definition primarily addresses an audience of women.

It is important to note that the three classes are situated on a "specialization dimension"; or at least, the fashion magazines are more specialized with regard to the field of clothing than the women's and glossy magazines, in the sense that they address a semi-professional audience of lay(women) who may engage in making clothes themselves. Fully specialized professional periodicals, addressing an expert audience of tailors, manufacturers, and shopkeepers, have not been included in the database, because the available publications did not yield a sufficient number of clothing terms conforming to the requirement of referential identifiability. (Perhaps surprisingly, illustrations are relatively scarce in this type of publication. Although it could be surmised that illustrations might be merely redundant for the experts, this probably does not apply to the new trends in fashion, where visual information about new developments is likely to be just as important for experts as for laymen.)

Apart from the classification according to the specialization dimension, the twelve magazines may be classified on the basis of their geographical status. One criterion for making such a geographical classification distinguishes between magazines with an editorial office in The Netherlands, and those with an editorial office in Belgium. This criterion does not guarantee, to be sure, that the editors and journalists in question are exclusively Netherlandic or Belgian; it happens to be the case that Belgian journalists work in The Netherlands and vice versa. The point of the geographical classification is not, however, to investigate whether the language of the editors of the magazines is representative for that of Belgium or The Netherlands at large, but merely to check whether any significant distinction at all is associated with it. The geographical distinction can be perspectivized in another way as well: rather than the editorial office's location (the production side of the

communicative process), the intended audience (the reception side of the communication) can be envisaged. A distinction can then be maintained between magazines that are exclusively distributed in Belgium, those that are exclusively distributed in The Netherlands, and those that are distributed in both countries.

The geographical variation in the data calls for a specific terminological convention. The word *Dutch* may be used as the name of the language under investigation here (regardless of whether it is spoken in Belgium or in The Netherlands), but also as the adjective corresponding with the geographical name *The Netherlands*. In order to keep both readings well separated, we will use *Dutch* only as the name of the language. When the notion "belonging to or coming from The Netherlands" is meant, we will use the term *Netherlandic*. According to this convention, the geographical variation in our sources involves the distinction between Netherlandic Dutch and Belgian Dutch. For additional accuracy, it may be noted that Dutch is the standard language only in the northern part of Belgium (Flanders), while the southern part (the Walloon area) uses French. In this respect, it might have been possible to utilize the term *Flemish Dutch* rather than *Belgian Dutch*. However, in order to maintain the parallelism with *Netherlandic Dutch*, we prefer to refer to the country (Belgium) rather than to the federal state (Flanders).

The relationship among the sources used in compiling the database is represented in Figure 2.1(1). Along the horizontal dimension, a distinction is maintained between the fashion magazines (*Burda*, *Knip*), the women's magazines (*Libelle_b*, *Feeling*, *Flair*, *Margriet*, *Libelle_n*), and the glossies (*Avenue*, *Esquire*, *Avantgarde*, *Cosmopolitan*, *Man*). The vertical dimension represents the geographical status of the sources on the basis of their distribution. Whereas *Libelle_n* is only distributed in The Netherlands, and *Libelle_b* only in Belgium, all other magazines are distributed in both countries. (The two *Libelles* have independent and separate editors. The name *Libelle_n* will be used to refer to the *Libelle* that is published and distributed in The Netherlands, while *Libelle_b* refers to the Belgian version.) The dotted square indicates which magazines have a Belgian editorial office. Note that a relatively high number of glossy magazines is included in the database to compensate for the fact that the frequency with which clothing terms appear in these sources is much lower than in the other periodicals. Only issues from the 1991 volume of the magazines have been used, but in each case, the entire year has been covered (so that, for instance, no bias in favor of summer or winter clothing was created).

Figure 2.1(1)

Overview of the sources used in compiling the database

In order to keep the corpus manageable, not all words that fall within the field of clothing terminology have been included in the database. The following systematic *restrictions on the referential range of the corpus* have been applied. First, the corpus is restricted to items that name types of garments, so that names for types of cloth, patterns and decorations, sewing techniques, parts of garments, colors, accessories etc. have not been included. Second, clothing for special purposes and special occasions (sportswear, working clothes, uniforms, evening dresses etc.), and clothing for children has been excluded. The restriction is a contextual one: when, for instance, jump-suit-like garments are presented in the magazines as something that can be worn in the same circumstances as ordinary suits and dresses, they obviously have to be included. Third, underwear and lingerie, overcoats and raincoats, and garments consisting of more than one piece (like suits) have not been incorporated. All in all, the database covers the following types of standard, functionally unmarked garments for adults: jackets, pullovers, cardigans, shirts, trousers, skirts, and dresses.

It should be noted that the referential restriction entails that certain types of semantic variation (more particularly, of the homonymic sort) remain out of sight. *Rok*, for instance, refers not only to skirts, but also to tailcoats, which are not included in the database because they belong to the domain of

functionally specialized, formal clothing. Similarly, *hemd* may mean 'shirt' as well as 'undershirt'; the latter reading is not included in the database because it belongs to the domain of lingerie and underwear. (When, however, items of clothing that would normally be considered pieces of underwear appear in the magazines as garments that can be worn as upperwear, they are obviously included in the database. An undershirt-like piece of clothing worn with the same function and in the same circumstances as an ordinary t-shirt or a blouse is then *de facto* not considered a piece of underwear.)

The classification into subfields also determines the referential part of the structure of the database. Generally speaking, the referential description takes the form of a *componential analysis*. All referents are assigned a componential description, the first feature of which is based on the aforementioned classification into subfields. More specifically, the following features are used:

[A]

– *Jack*-like garments: informal jackets, blousons; garments covering the upper part of the body, typically with a loose-fitting blousing cut, without lapels, and with a front fastening that can be fastened up to the neck

[B]

– *Colbert*-like garments: formal jackets; garments covering the upper part of the body, usually worn as the top half of a suit, typically with lapels and a front fastening that cannot be closed up to the neck

[C]

– *Vest*-like garments: cardigans; garments covering the upper part of the body, worn on top of a shirt, made of wool or a similar warm and supple material, with separate front panels

[D]

– *Trui*-like garments: pullovers, sweaters, jumpers, jerseys, slipovers; garments covering the upper part of the body, made of wool or a similar warm and supple material, with at most a partial fastening at the front

[E]

– *Hemd*-like garments: shirts, t-shirts, blouses; garments covering the upper part of the body, made of light material, constituting the first layer of clothing above the underwear

[F]

– One-piece garments: all garments (with the exclusion of dresses)

that cover both the lower and the upper part of the body, such as jump-suits and dungarees

[G]

– *Jurk*-like garments: dresses; garments covering the lower and the upper part of the body, worn by women

[H]

– *Broek*-like garments: trousers; garments covering the lower part of the body, divided into two legs

[I]

– *Rok*-like garments: skirts; garments covering the lower part of the body, not divided into two legs, worn by women.

Except for [F], each of these classes is constituted round a highly salient standard type of garment that is named by a high frequency lexical item. In the case of *vest* ‘cardigan’, *trui* ‘pullover’, *hemd* ‘shirt’, *jurk* ‘dress’, *broek* ‘trousers’, and *rok* ‘skirt’, the lexical correspondence with the English translations is straightforward. In the case of *jack* ‘blouson’ and *colbert* ‘jacket (as of a suit)’, however, Dutch draws a salient distinction between informal and more formal jackets that seems to be less outspoken in English, where the fact that both types fall within the concept *jacket* seems to be more preponderant than the separate status of *blouson*. As a terminological note, it should be mentioned that Dutch *broek* is a general name for all kinds of two-legged garments covering the lower part of the body; as such, it is a hyperonym of items like *bermuda*, *legging*, *short*, and *jeans*. At the same time, *broek* is the regular name for one particular type of two-legged garment covering the lower part of the body, viz. the default case with long legs. Although it seems that *broek* can be used more easily in the hyperonymous reading than its translational equivalent *trousers* (which seems to be more strongly restricted to the hyponymous reading), a hyperonymous application is not excluded in English. It is present, for instance, in dictionary definitions of *shorts* as “trousers reaching only to knees or higher”. In most cases where we translate *broek* as ‘trousers’, the reference is to the hyperonymous reading rather than to the hyponymous default case.

The use of the major features [A]–[I] is motivated by the fact that the semantic dimensions that are needed for the more detailed description of the referents differ from one basic type to the other. Specifically, some features are only relevant with regard to restricted subfields. For instance, a “wrap-around” type of fastening occurs with skirts but is irrelevant for trousers, and turtleneck collars can only be found on jumpers. In the case of zippered

informal jackets, it is important to distinguish between jackets that have asymmetrical front panels (like the typical *vliegeniersjack* 'aviator's jacket') and those that do not; in the case of zippered skirts, trousers, and dresses, the distinction is obviously irrelevant. In this sense, assigning major features is a matter of efficiency: a single componential structure, based directly on features like width, length, type of fastening, type of collar, presence of sleeves, and so on, would yield a more complex and cluttered description than is the case when the referents are assigned a global feature of the type indicated above.

However, because the definition of the global features as given above is based primarily on the prototypical, unmarked cases of each subfield, borderline cases might cause classificatory problems. This is specifically the case when it is intuitively unclear whether the lexical item that is used as the reference point for the delimitation of the subfields, can be used as a name for the garment in question. For instance, as will be described in full detail in section 3.3., it is not immediately obvious whether people classify culottes (*broekrok*) as a kind of trousers or as a kind of skirt; so should they receive the feature [H] or [I]? Note, however, that the features needed for an adequate description of culottes are the same as for the description of trousers: culottes have separate legs, the length and width of which may be relevant for the definition of the category. As such, culottes receive the feature [H], because the features used for the description of trousers are precisely the ones that are relevant for the description of culottes. More generally, a uniform solution for borderline cases has been sought on the basis of descriptive efficiency.

Next to the global features [A]-[I], the componential analysis consists of specific features that hold within each of the major subfields as identified in the previous pages. As it would take too long to present the full system for the componential description of the garments at this point, we shall restrict the presentation to a single subfield, viz. that of the trousers. Additional information about the componential system will be given whenever necessary further on in the text. In general, the set of relevant features has been determined on the basis of a preliminary inspection of the selected sources, and on the basis of existing descriptions of clothing types. Specifically, Van Domselaar & Horsten (1990) has been particularly useful, because these course materials for the professional training of entrepreneurs and retailers in the textile industry contain line drawings of various sorts of garments. Five dimensions were selected for the descriptions of the [H]-subfield, with the following values.

LENGTH

- [1] The garment does not reach further down than the groins
- [2] The garment reaches down to the thighs
- [3] The garment reaches down to the knees
- [4] The garment reaches down to the calves
- [5] The garment reaches down to the ankles

WIDTH AND CUT

- [1] The garment is tight-fitting
- [2] The garment has a straight cut, neither tight-fitting nor wide
- [3] The garment has a loose, wide cut
- [4] The garment is tight-fitting round the hips but has gradually widening legs
- [5] The garment is loose-fitting round the hips but has straight or tight-fitting legs
- [6] The garment has a regular straight cut as far down as the knees, but has widening legs below the knees
- [7] The garment has a loose cut from the hip to the knee, but has tight-fitting legs below the knees

END OF LEGS

- [1] The ends of the legs exhibit no special features
- [2] The ends of the legs fit tightly round the person's legs by means of an elastic band
- [3] The ends of the legs fit tightly round the person's legs by means of tied laces
- [4] The ends of the legs fit tightly round the person's legs by means of buttons
- [5] The ends of the legs have an elastic band that fits under the feet

MATERIAL

- [1] The trousers are made of cotton or linen or a similar, relatively smooth cloth
- [2] The trousers are made of wool or a similar, relatively coarse or fluffy material
- [3] The trousers are made of denim
- [4] The trousers are made of corduroy
- [5] The trousers are made of silk or a similar smooth and shining material
- [6] The trousers are made of smooth, shiny leather or an imitation of it
- [7] The trousers are made of rough, mat leather or an imitation of it
- [8] The trousers are made of stretch

DETAILS

- [1] Seams and/or pockets are strengthened by metal buttons
- [2] The waist part of the garment has folds, pleats or creases
- [3] The garment has a very low crotch
- [4] The garment is made up of several visible strips of cloth
- [5] The legs have sharply pressed creases
- [6] The garment has an elastic band in the waist

Given this system for the description of trousers and related garments, the items included in Figure 2.1.(2), which are based on the illustrations found in the Detex courseware as compiled by Van Domselaar & Horsten (1990), may be described by means of the following componential descriptions:

<i>hot pants</i>	111?.
<i>shorts</i>	221?.
<i>bermuda</i>	321?.
<i>knickerbocker</i>	332?.
<i>kuitbroek</i>	411?.
<i>spijkerbroek, jeans</i>	52131
<i>pantalon</i>	521?5
<i>bandplooibroek</i>	551?2
<i>hippiebroek</i>	541?.
<i>olifantbroek</i>	561?.
<i>jodhpur</i>	571?4
<i>harembroek</i>	533?.
<i>Turkse broek, drollenvanger</i>	531?3
<i>leggings</i>	5116.
<i>skibroek</i>	5156.
<i>joggingbroek</i>	532?6
<i>broekrok</i>	4331?.

The order of the featural dimensions in these descriptions is the same as that in the overview given above. The configuration [111?], for instance, contains the information that the clothing type in question is characterized by value 1 ("does not reach further down than the groins") on the first dimension (LENGTH). Points indicate that the dimension is irrelevant, i.e. that it does not receive a value for the item under description. This is mostly the case for the dimension DETAIL (the last one in the row). The question marks

bandplooibroek

hippiebroek

pantalon

olifantbroek

jodhpur

harembroek

Figure 2.1(2)
Sample garment types in the [H]-subfield

hot pants

short

bermuda

knickerbocker

broekrok

turkse broek

Figure 2.1(2) - continued
Sample garment types in the [H]-subfield

kuitbroek

legging

skibroek

joggingbroek

spijkerbroek

Figure 2.1(2) - continued
Sample garment types in the [H]-subfield

indicate that the exact dimensional value is difficult to determine for the referent under consideration. The dimension MATERIAL, for instance, mostly does not receive a specific value, because the line drawings used as the basis for the description do not contain a clue as to the materials used. In fact, the values on the MATERIAL dimension given to *jeans*, *leggings*, and *skibroek* are not strictly warranted, as they do not follow straightforwardly from the drawings. Merely in order to illustrate the descriptive system, they have been included as default options. In the actual database, question marks will obviously be less numerous on the MATERIAL dimension than in this sample, because the photographs in the magazines are usually precise enough to allow an identification of the materials used. It should also be clear that the descriptions given here are not general descriptions (let alone definitions) of items like *bermuda* or *skibroek*; they describe the specific referents included in the figure, not the lexical item as such. The whole point of the componential description is, in fact, to get an idea of the referential range of application of words like *bermuda* or *skibroek* by collecting a large number of tokens of those words.

In addition to the global features and the specific features, the componential analysis contains two fields for general features that may be relevant for any type of garment: one field specifies whether the garment is worn by a man or a woman, and the other is an open text field for any type of comment or remark; in this way, potentially relevant information that is not yet included in the componential system may be incorporated into the description. Specifically, the commentary field may contain information about the color and pattern of the garments. Because the range of potential values on these dimensions is so large ("plain", "striped", "floral", "geometrical", etc.) that an initial determination of a fixed number of values is impossible, the inclusion of an open text field is an obvious step to take. It should be clear from the addition of such a "wastebasket" field, that the formalized componential analysis is not to be considered the *nec plus ultra* of the description. As we will see later on (most specifically in section 3.2.), the initial componential description as included in the database will have to be critically interpreted. In specific cases, moreover, the initial componential description will have to be revised in the course of the analysis, when features that had not been included, unexpectedly appear to be relevant after all.

To illustrate, let us have a brief look at an example that will be treated in more detail in section 3.4.. In general, allocation of a referent to either the C- or B-type depends on two factors. B-type garments are jackets of the formal type; they are invariably longer than the waist-line, reaching down to the

hips, they always have revers and long sleeves, and they are never knitted. The C-types, on the other hand, refer to jacket- and cardigan-like garments that lack these characteristics. Introspectively, one would not expect C-type garments to occur in the referential realm of items like *colbert* and *blazer*, which typically refer to the standard formal jackets referred to as B-types. In actual practice, C-type pieces of clothing peripherally show up in the semasiological range of application of *colbert* and *blazer*. This intrusion of C-type garments into the semasiological realm of *colbert* and *blazer* is rather unexpected; intuitively, one would not expect knitted garments or jackets without sleeves to be called either *blazer* or *colbert*. But because the dimensions used for the componential description of the B-type garments do not coincide entirely with those used for the description of the C-type garments, all referents of *blazer* and *colbert* classified as C-types have to be reinvestigated in the original magazine photographs to check the dimensions that are absent in the componential system for the C-types. If not, it would be impossible to arrive at a uniform graphical, schematic representation of the semasiological structure of the lexical items *colbert* and *blazer*.

The database compiled on the basis of the magazines obviously does not consist of the referential, componential description only. Apart from a reference specifying the location of a term in the magazine sources, each database record should obviously also contain a *lexical description* (as distinct from the referential description in componential form) of the recorded cases. Items of clothing are often not just named by means of a single lexical item like *broek* 'trousers' or *jurk* 'dress', but the latter may constitute the head of a full noun phrase like *getailleerde jurk met wijde hals* 'waisted dress with a wide neckline'. The way in which the garments are named, then, is captured in the database by means of four fields. NAME1 specifies adjectival premodifiers like *getailleerd* 'waisted', whereas NAME4 registers prepositional postmodifying phrases like *met wijde hals* 'with a wide neckline'. NAME3 and NAME2 specify the head of the noun phrase, depending on whether it is, respectively, a compound noun or not. (Loanwords that are morphologically complex in the original language, like English *sweatshirt* or French *deux-pièces* 'two-piece', are treated as simplex words.) Although most of our attention will go towards the major categories with which a referent is named (represented by NAME2 and NAME3), the modifying elements recorded in NAME1 and NAME4 will have a role to play in chapter 5.

2.2. Characteristics of the corpus

The entire database totals 9205 records. The distribution of the material over the various sources is given in Figure 2.2(1), which is built up according to the same principles as Figure 2.1(1). For each major group of magazines that may be distinguished, Figure 2.2(1) specifies the number of records belonging to that set of sources. The total number of records for the magazines that are distributed both in The Netherlands and in Belgium, for instance, is 3589. Similarly, there are 3949 records for the magazines that have a Belgian editorial office.

Figure 2.2(1)

Distribution of the material over the various sources

The distribution is not an even one, in the sense that the harvest of clothing terms is not the same for each magazine or group of magazines. To be sure, this is not a cause for concern, because the possible effects of the distributional asymmetries may be easily controlled for by taking into account relative frequencies. More importantly, is there any way in which the representative quality of the data can be measured? One way of answering the question is to have a look at the *saturation* of the corpus. When the database is a fairly representative reflection of the actual situation, the relative increase in the number of lexical types that are added to the

material with each new portion of records should be low, or rather, it should diminish with the growth of the database. In Figure 2.2(2), the increase of the number of different lexical types in the database is charted for every successive 1000 records. (The final portion actually represents the increase from 9000 to 9205 records. Spelling variants like *blouse* and *bloese* have been treated as one type.) Figure 2.2(3) presents the same data in graphical form.

	Name2	Name3	Total
1-1000	47	101	148
1-2000	56	205	261
1-3000	63	283	346
1-4000	65	375	440
1-5000	71	444	515
1-6000	73	529	602
1-7000	74	589	663
1-8000	77	633	710
1-9000	79	690	769
1-9205	80	705	785

Figure 2.2(2)

Increase of lexical types per successive set of 1000
records (actual numbers)

It is immediately obvious that there is a dramatic distinction between the saturation of the set of simplex terms and that of the compound items. The set of NAME2s exhibits only a marginal increase after the first few thousand records have been added to the data. The NAME3s, on the other hand, are characterized by a steady increase at a high rate. For instance, the "new type"/"new token" ratio for the first thousand records is 0.1 for NAME3s: on the average, a new NAME3 is introduced with every tenth new record. It is still 0.057 for the last full portion of thousand records (between 8000 and 9000), which means that every twentieth new record yields a new NAME3 type. By contrast, the ratio is 0.47 and 0.002 respectively for the set of NAME2s, which means that between the 8000th and 9000th record, a new

NAME2 only appears with every 500th new record. The distinction between free words and compounds may be further illustrated by considering the average number of records per lexical type for each category. For the free words, this is 60.85, against a mere 4.63 for the compound items. Further, a comparison between the frequency distributions of both classes of words shows that they are characterized by markedly different patterns. Disregarding types that are mere spelling variants of items that occur elsewhere in the corpus, the set of NAME2s that occur only once in the database represents 23.7% of the total set of NAME2 types; by contrast, the NAME3s with frequency of occurrence 1 constitute no less than 67.7% of the total set of types on NAME3 level.

Figure 2.2(3)

Increase of lexical types per successive set of 1000 records (graphical representation)

In short, there appear to be major distributional differences between the class of NAME2s and the class of NAME3s, differences that can be summarized in the statement that *the database approaches the saturation point for name2s, but is hardly saturated for name3*. The difference is related straightforwardly to the formal distinction between NAME2 and NAME3. Because the latter contains compound lexical items, it is no surprise that many of the NAME3s are in fact ad hoc formations, incidental coinages rather than well-established, conventional names. Typical categories of such low frequency incidental names involve brand names (*501 jeans*, *Chanel-jasje* 'Chanel jacket'), types of motif or material (*bloemenjack* 'flower blouson', *stipblouse* 'polka-dot blouse', *donsjack* 'fluff jacket', *chiffon-*

blouse 'chiffon blouse'), and unusual indications of function (*weekendjasje* 'weekend jacket', *nazomerjasje* 'Indian summer jacket', *citybermuda* 'bermuda shorts for the city'). Although the presence of these incidental coinages is morphologically of great interest (because they provide an opportunity for a corpus-based investigation into the mechanisms of morphological productivity), morphological analyses will not be the main focus of the following chapters. It should be noted that the lowest frequency items in the corpus do not always fall in the category of "incidentally coined compounds". Other categories of low frequency items include rare spelling variants (like *pull-over* instead of the regular spelling *pullover*), incidental abbreviations of full forms (like *pull* for *pullover*), uncommon loanwords (like *débardeur* instead of *slipover*), and borderline cases like *bustier*, which primarily names a type of underwear, but which is included in our database as a name for a lingerie-like blouse.

The polylexical expressions in the database (i.e. those expressions that contain a premodifying NAME1 or a postmodifying NAME4) generally have low frequencies. The highest frequencies occur with *rechte rok* 'straight skirt' (38 occurrences) and *witte blouse* 'white blouse' (33 occurrences). In what follows, we will concentrate on single words (NAME2s and NAME3s), but polylexical expressions of the *rechte rok*-type will not be ignored; they will play an important role in sections 4.2., 5.1., and 5.2..

The saturated nature of the database (at least with regard to the NAME2s) inspires a reasonable degree of confidence with regard to its representativity. This confidence is strengthened when a comparison is made between our database and the text corpus of present-day Dutch compiled by the Instituut voor Nederlandse Lexicologie in Leiden. (For details of the comparison, see Geeraerts & Bakema 1993.) At the time of making the comparison, the Leiden database, which covers all domains of general language use, had a size of forty-two million tokens. According to international standards, this is a decidedly large corpus. (The Collins Cobuild corpus, for instance, totals 7.3 million tokens.) Still, the frequencies of the clothing terms in the Leiden corpus are consistently and considerably lower than those in our database. For instance, the total frequency of all the items that appear in our material with a record frequency of more than ten is 7884. In the Leiden corpus, the same items total only 3827 instances. If, then, it would probably be considered methodologically safe to base a linguistic investigation on a corpus of forty-two million tokens, our own database should inspire an even higher degree of methodological confidence.

It should be added, however, that there are two respects in which the da-

tabase is manifestly subject to *limitations* as to its representativity. To begin with, the number of records referring to garments worn by men is only 977 (or 10.6 %). Apparently, the sources we have used are primarily concerned with clothes for women, which means that the descriptive results of our study are subject to an important limitation: they probably do not paint an entirely trustworthy picture of the lexical situation involving men's wear. For most of the analyses that we will present, this restriction will not be crucial, but section 3.4. will contain a number of examples for which a careful consideration of the gender dimension will turn out to be quite revealing.

More generally, it should be borne in mind that the results presented here do not claim to be representative for the clothing vocabulary of present-day Dutch at large. The study deals only with the domain of written discourse in certain kinds of magazines, but says nothing at all about how Dutch speakers may use words like *colbert*, *rok*, or *blouse* in everyday spoken discourse (or, for that matter, in any other style, register, or discursive context). The purpose of the investigation is not to draw an exhaustive picture of Dutch clothing vocabulary in all circumstances and contexts, but to show by example how the various types of variation involved in painting such a complete picture may be approached in a methodologically rigorous way. The distinction, for instance, between the written discourse of the magazines and everyday spoken discourse belongs to the same type of sociolinguistic variation that is represented in the study by the distinction between Netherlandic and Belgian sources, and by the distinction between various source groups. This means, obviously, that various kinds of follow-up studies may be envisaged on the basis of the present results. We are, for instance, currently engaged in a follow-up study involving the use of clothing terminology on the labels used in shop windows to identify garments. (Among other things, the study reveals that the distinction between Belgian Dutch and Netherlandic Dutch is much bigger in the shop window material than in the magazines material analyzed here.)

Turning from the low frequency range of the corpus to the high frequency items, the *twenty-five most frequent names* in the corpus (both NAME2s and NAME3s) are distributed as follows over the major subfields of the material:

- [A] *jack* 'blouson, informal jacket' (162)
- [B] *jasje* 'jacket' (605) – *blazer* 'jacket' (242) – *colbert* 'jacket' (166)
- [C] *vest* (319) 'cardigan' – *vestje* '(small) cardigan' (123) – *cardigan* 'cardigan' (94)

- [D] *trui* 'pullover' (661) – *truitje* '(small) pullover' (97)
 [E] *blouse* 'blouse' (649) – *t-shirt* 't-shirt' (283) – *overhemd* 'shirt' (224) – *shirt* 'shirt' (148) – *topje* 'top' (120) – *hemd* 'shirt' (97)
 [G] *jurk* 'dress' (271) – *jurkje* '(small) dress' (96)
 [H] *broek* 'trousers' (638) – *bermuda* 'bermuda shorts' (170) – *jeans* 'blue jeans' (146) – *legging* 'tight-fitting women's trousers' (142) – *shorts* 'shorts' (83) – *broekrok* 'culottes' (80)
 [I] *rok* 'skirt' (542) – *rokje* '(small) skirt' (184).

The figures between brackets specify the frequencies of the items. The translations are only rough ones; a detailed analysis of, for instance, the distinction between *colbert* and *blazer* will be given further on. In what follows, these twenty-five items will be the basis for the analysis; this does not mean that other expressions will not be envisaged, but merely that the high frequency items constitute the best startingpoint for closer scrutiny.

2.3. Points of methodology

The approach introduced in the previous two sections raises a number of points of a general methodological nature. Successively, we will now deal with the following questions. First, is using such a structuralist tool as componential analysis consistent with the overall cognitive orientation of our study? Second, what are the advantages of looking closely at the way in which lexical items are actually used, rather than just relying on introspection? Third, to what extent is the approach sketched here an objective one? And fourth, are there any major alternatives for the componential system used here?

To begin with, then, a methodological note about the *compatibility of a featural description and a cognitive, prototype-based approach* to lexical semantics is in order. Against the background of the development of linguistic semantics, the cognitive approach has often been primarily defined in contrast to the componential model of semantic analysis that was current in transformational grammar and that is stereotypically associated with Katz & Fodor's analysis of *bachelor* (Katz & Fodor 1963); in an early defense of a prototypical approach, Fillmore (1975) referred to this approach as the "checklist theory" of meaning. The cognitivists' reaction against this featural

approach had, however, the negative side-effect of creating the impression that prototypical theories rejected any kind of componential analysis. This is a misconception for the simple reason that there can be no semantic description without some sort of decompositional analysis: it is hard to conceive of any form of comparative analysis that does not involve breaking down the comparanda into components and characteristics. As a heuristic tool for the description and comparison of lexical meanings, then, a componential analysis retains its value (a value that, incidentally, it did not acquire with the appearance of componential analysis as an explicit semantic theory, but which had been obvious to lexicographers from time immemorial). Rather, the difficulties with the neostructuralist kind of feature analysis that grew out of structuralist field theory lie elsewhere; it is not the use of decomposition as a descriptive instrument that causes concern, but the status attributed to the featural analysis.

Specifically, featural definitions are classically thought of as criterial, i.e. as listing attributes that are each indispensable for the definition of the category in question, and that taken together suffice to delimit that category from all others. In contrast, the cognitive approach claims that there need not be a single set of defining attributes that conforms to the necessity-cum-sufficiency requirement. In this sense, the componential description illustrated here is only the first step in the process of semantic analysis. It is not primarily used to define lexical items, but merely to chart their referential range of application. Whether a classical, criterial definition for the items may be extracted from the description of their ranges is a different question, which will constitute the second step in the analysis, to be taken in section 3.2.. Obviously, then, the componential descriptions used here are not invested with specific psychological status. When a componential configuration like [H52131] is used in the description of *spijkerbroek* or *jeans*, it is not implied that the meaning of those words is mentally stored in precisely such a componential format, nor is it suggested that understanding *spijkerbroek* or *jeans* consists of building up the meaning of those words by mentally assembling specific featural values on dimensions like "tightness of fit" and "length of legs".

It can hardly be sufficiently stressed that the type of componential analysis used here deliberately and consciously involves the description of things rather than meanings. The variety that exists between the entities that occur in the referential range of application of a lexical item is charted in featural form. Only after this first step is taken can the question be broached whether there is sufficient communality among those referents to formulate a defini-

tion of (the meaning of) the item. The question whether componential analysis does indeed (in accordance with its structuralist self-conception) constitute an analysis of the linguistic meaning of lexical items, or whether it is rather a description of entities in the world, is not new in the history of lexical semantics; see, for instance, Guiraud's critical insistence on the necessity for the adherents of componential analysis to prove that what they have to offer is more than just a referential description (1975: 101). But what is a critical question in the context of a structuralist conception of componential analysis is here avoided by straightforwardly embracing the referential nature of the componential description. In the present framework, the referential (or, if one wishes, "encyclopaedic" rather than "linguistic") status of the componential analysis is not something to be shunned, but is rather accepted as a preliminary but methodologically indispensable step of the semantic analysis.

But even within the broadly defined framework of Cognitive Linguistics, the referential, usage-based investigation advocated here is not necessarily accepted unconditionally. In actual practice, lexical studies with a cognitive semantic orientation exist both in the form of introspective analyses, and in the form of corpus-based research. Schmid (1993b: 272) even considers the corpus-based approach in work such as that of Rudzka-Ostyn (1988, 1989), Schulze (1988, 1991), Dirven (1985, 1990), and Geeraerts (1983, 1990) to be typical of the European branch of the Cognitive Linguistics movement, in contrast with the more introspectively conducted studies of American researchers of a Cognitive Linguistic persuasion. From a more theoretical point of view, explicit attention for the way words are actually used would seem to follow straightforwardly from Langacker's characterization of Cognitive Linguistics as a usage-based model that rejects the Chomskyan neglect of linguistic performance (1987: 46). However, as we already mentioned in the introduction to this chapter, Wierzbicka (1985) has coupled a prototype-oriented form of lexical research with an explicit defence of the introspective method. What *advantages*, then, *are there to a referential, usage-based approach* that avoids relying exclusively on introspection?

An answer was already formulated in the introduction to the present chapter: the introspective method may succeed rather well in pinning down the prototypical core of the items under investigation, but is hardly able to capture the peripheral uses to which the core meanings appear to give rise in actual usage. As an illustration, consider Wierzbicka's definition of *dress* (1985: 382):

A KIND OF THING MADE BY PEOPLE FOR WOMEN AND GIRLS TO WEAR. IMAGINING THINGS OF THIS KIND PEOPLE COULD SAY THESE THINGS ABOUT THEM:

they are made to be worn on the body, below the head, to cover most of the body

so that all the parts of a woman's body which people think should not normally be seen are covered with that one thing

and to protect most of the body with undesirable contact with the environment

and to cause the woman wearing it to look good

they are made in such a way that when they are on the body the lower half surrounds the lower half of the woman's body from all sides

so that the legs are not separated from one another

and so that the genital area of the woman's body seems to be hidden

and so that women wearing things of this kind look different from men

things of this kind are thought of as something suitable for women to wear in most kinds of places and in most kinds of circumstances.

When we have a look at the actual garments that occur in our database as instances of *jurk* (the Dutch equivalent of *dress*), we find cases in our material that do not conform to the description. If, for instance, "covering most of the body" is interpreted as "covering more than 50% of the body", then a number of very short summer dresses with open backs and low necklines do not display the feature in question. And if "the parts of a woman's body which people think should not normally be seen" include the upper part of the thighs, then dresses with long side slits contradict the image. Furthermore, some dresses have such wide armholes and such a plunging décolletage that they could not normally be worn without exposure of the breasts (unless they are worn with an additional t-shirt or blouse underneath).

The comparison shows, in other words, that the description proposed by Wierzbicka may well be adequate for the majority of cases in the range of *dress*, but does not really cover *all* possible instances. Admittedly, such a comparison is risky for at least two reasons. First, we start from the assumption that English *dress* and Dutch *jurk* are equivalent as far as their referential range of application is concerned. As long as we do not have a similar corpus-based analysis of *dress* as the one we have presented for *jurk*,

the comparison will have to remain a conditional one. Second (and more importantly), it is not even certain that Wierzbicka actually intends the definition to apply to all the cases in the extension of *dress*. By introducing the phrase "imagining things of this kind people could say these things about them", the perspective is shifted from the objective features of the things that are being called *dresses* to the subjective image that people say they have about dresses when they are asked for it. In a sense, Wierzbicka defines *dress* by referring to what people think dresses are. And if what people think dresses are only involves the central cases of the category "dress", then, of course, it makes no sense to complain that the description of this mental image does not apply to non-prototypical dresses: it never intended to do so anyway. On this reading of Wierzbicka's view, its reference to subjective images could be construed as implying a conscious restriction of the description to the prototypical core of the category. And because introspection probably does work efficiently for retrieving such prototypical images, the introspective method may be salvaged.

It is not quite clear, however, whether this interpretation of Wierzbicka's position is a valid one. On the one hand, she argues that "a valid definition must be empirically adequate, that is, it must be phrased in such a way that it covers the entire range of use of a given word, expression, or construction" (1989: 738). On the other hand, if a definition such as that of *dress* is to be applicable to all things that may be called by that name, people should be able to assert all the characteristics mentioned in the definition any time they see a dress. But surely, when imagining a less prototypical kind of dress than the kind whose features are included in the definition, people will not imagine it *as* a prototypical case. What people could say about dresses changes when peripheral members of the category are at stake: default dresses, for instance, may well cover most of the body, but that feature may be suppressed when a fancy type of summer dress is involved.

Even if, however, we accept that definitions of the kind illustrated above are explicitly restricted to the prototypical core of the categories, a counter-argument may be advanced in favor of a method based on observing actual usage. Note that it remains a matter to be settled empirically whether the lexical knowledge that people have in their minds is indeed restricted to a mental image of the core of the category in question. It is not *a priori* given that the idea of a category that people may introspectively retrieve from memory is an adequate reflection of the extent of that person's actual knowledge of the category. On the contrary, if it is part of his knowledge to produce or accept an application of *dress* to non-prototypical cases, then he

"knows" more about the category than would be included in his introspectively retrieved idea of the category. That knowledge, to be sure, is not necessarily conscious knowledge; it is less "knowledge *that* (lexical item *x* may refer to entities with such and such characteristics)" but rather "knowledge *how* (lexical item *x* may be successfully used)". In order, then, to get a better grasp on the lexical "knowledge how", usage-based investigations of the type illustrated in this monograph are vital, precisely if it is suspected that conscious knowledge may only partially cover the full extent of a person's "knowledge how". In our own research, we will not be making any psychological claims about how the lexical facts that may be observed in the course of an investigation into actual vocabulary use are stored in the brain. We rather see the present study as providing the groundwork for such an investigation: it specifies what people do with words, but an investigation into the mental representations and procedures that they use for doing those things may well require psychological modes of research that go beyond the purely linguistic methods used here. In the interdisciplinary framework of a cognitive investigation into natural language, the type of usage-based linguistic inquiry illustrated here has a legitimate role to play next to psychological and neurophysiological types of research.

To complete our argumentation for an approach that is not exclusively based on introspection, it should be emphasized that this methodological preference does not imply that our own endeavours are completely free of intuitive aspects, in the sense that the researcher's own understanding of the instances of language use under investigation is entirely ignored or suppressed. More precisely, *the referential approach does not entail that the investigation proceeds in a purely objective fashion*, without any recourse to interpretative activities on the part of the investigator. The point may be illustrated by considering the initial selection of the descriptive features included in the componential system. The choice of those features is not dictated automatically by the referents of the words themselves. In principle, an infinite number of characteristics could possibly be included in the descriptive framework. In the case of trousers, for instance, it would be possible to refer to the presence of lining in the legs, to whether the hip pockets have a flap or not, or to the number of nooses in the waist intended to hold a belt. The fact that, in actual practice, we have decided not to include these features in the componential system is determined by assumptions about their relevance for the description. As we remarked earlier, we have tried to avoid excessive bias in the selection of the features on the basis of a preliminary inspection of part of the selected sources, and on the basis of existing de-

scriptions of clothing types (in particularly, the Detex courseware). But we do not want to deny that our own pre-existent knowledge of the field of Dutch clothing terms, and our own intuitions about what would be pertinent features for describing that field, have played their role in the choice of a particular componential system. In this, as in so many other aspects, lexical semantics is basically a *hermeneutic* enterprise (see Geeraerts 1992). Lexical description does not simply consist of recording referents, but of trying to determine what features of the referents motivate or license the use of a particular item. The importance of this interpretative side of the lexical method would probably become even more apparent when a different lexical field is investigated. Even more than for words like *dress* and *trousers*, the question for words like *democracy* and *art* is not just what entities they refer to, but what features they refer to *in* those entities.

Given, then, that the componential system taken as a startingpoint here broadly speaking embodies a hypothesis about what might constitute appropriate referential features for describing the field of clothing terms, there exists an obvious possibility of taking issue with this hypothesis and suggesting *alternative descriptive models*. It might be argued, specifically, that next to the material features chosen here, more attention should be devoted to functional features, such as the occasion on which the garment is meant to be worn (work, formal social events, sports activities, etc.), or the age and social status of the intended wearer. (Again, see Wierzbicka 1985 on the importance of such functional notions.) As a matter of principle, we absolutely do not want to deny that such features are part and parcel of our conceptualization of clothing items. However, we have consciously restricted the investigation to those features that could be reliably identified from the photographs in the magazines, because that was the most economical way of building a database of spontaneous language use that was large enough to allow for quantitative analysis. Given the methodological framework of the study, we feel that such a restriction is legitimate for two reasons. First, we have tried to diminish the impact of variation on the functional level by restricting the database to garments worn in functionally unmarked circumstances. As mentioned earlier, clothing for special purposes and special occasions (sportswear, uniforms, evening dresses etc.), has been excluded, and only clothing for adults has been included. (In fact, the models appearing in the magazine photographs mostly belong to the group of young adults, aged between 20 and 40.) Second, any type of functional vestimentary difference will be reflected by visual, material features of the type used here. The formality of a particular type of jacket, for instance, is not just a question of

when it can be worn, but also of what it looks like: the formality is not an independent characteristic next to and apart from the visual features of the jacket. Because of this connection between functional and material features, we feel confident that functional differences that are not filtered out by the restriction of the database to functionally unmarked cases, will at least be represented indirectly in the componential descriptions.

On the most fundamental level, however, we feel that a componential system with material features as used here should not be defended purely on *a priori* grounds. Basically, we have opted for this approach because it enables us to do a number of things that would otherwise be more difficult to achieve: compiling a database of spontaneous language use that is large enough to allow for quantitative treatment would have been much more difficult on the basis of less easily identifiable functional features. But still, the ultimate proof of the pudding will have to be in the eating. The initial methodological assumption that the descriptive model taken as a starting-point is a helpful one, will only be fully confirmed when linguistically significant results can be extracted from it. Even though we have reason to believe already that a material set of features will be descriptively useful, our attitude at this point is essentially just to see how far we can get with it. And we can get a long way, we hope to show.

Chapter 3

Semasiological variation

The purpose of this chapter is to demonstrate the importance of prototype theory for the semasiological analysis of lexical items, but such a purpose cannot be fulfilled without a critical enquiry into the way in which the notion of “prototypicality” may be used. In fact, the current appeal of prototype theory should not obscure the fact that the exact definition of prototypicality is not without problems. The first section of this chapter will try to analyze the sources of the confusion surrounding the definition of prototypicality by making clear that prototypicality is itself, in the words of Posner (1986), a prototypical concept. In the subsequent sections, the importance of the various conceptions of prototypicality identified in the course of this analytical exercise will be applied to the clothing terms database. Two major types of prototypicality effects will be considered: non-discreteness (involving absence of classical definability and degrees of category membership), and non-equality (involving salience effects). The final section considers the question whether these prototypicality effects are subject to contextual, source-related variation.

3.1. Types of prototypicality

As a first step towards a classification of types of prototypicality, let us have a look at four characteristics that are frequently mentioned (in various combinations) as typical of prototypicality. In each case, a quotation from early prototype studies is added to illustrate the point.

- (i) *Prototypical categories cannot be defined by means of a single set of criterial (necessary and sufficient) attributes*: “We have argued that many words ... have as their meanings not a list of necessary and

sufficient conditions that a thing or event must satisfy to count as a member of the category denoted by the word, but rather a psychological object or process which we have called a prototype" (Coleman & Kay 1981: 43).

(ii) *Prototypical categories exhibit a family resemblance structure, or more generally, their semantic structure takes the form of a radial set of clustered and overlapping meanings*: "The purpose of the present research was to explore one of the major structural principles which, we believe, may govern the formation of the prototype structure of semantic categories. This principle was first suggested in philosophy; Wittgenstein (1953) argued that the referents of a word need not have common elements to be understood and used in the normal functioning of language. He suggested that, rather, a family resemblance might be what linked the various referents of a word. A family resemblance relationship takes the form AB, BC, CD, DE. That is, each item has at least one, and probably several, elements in common with one or more items, but no, or few, elements are common to all items" (Rosch & Mervis 1975: 574-575).

(iii) *Prototypical categories exhibit degrees of category membership; not every member is equally representative for a category*: "By prototypes of categories we have generally meant the clearest cases of category membership defined operationally by people's judgements of goodness of membership in the category ... we can judge how clear a case something is and deal with categories on the basis of clear cases in the total absence of information about boundaries" (Rosch 1978: 36).

(iv) *Prototypical categories are blurred at the edges*: "New trends in categorization research have brought into investigation and debate some of the major issues in conception and learning whose solution had been unquestioned in earlier approaches. Empirical findings have established that ... category boundaries are not necessarily definite" (Mervis & Rosch 1981: 109).

As a first remark concerning these characteristics, it should be noted that they are not the only ones that may be used in attempts to define the prototype conception of categorization. Two classes of such additional features

should be mentioned.

On the one hand, there are characteristics that do not pertain (as the four mentioned above) to the structure of categories, but that rather pertain to the epistemological features of so-called non-Aristotelian categories. For instance, the view that prototypical categories are not "objectivist" but "experiential" in nature (Lakoff 1987) envisages the epistemological relationship between concepts and the world rather than the structural characteristics of those concepts. In particular, it contrasts the allegedly classical view that "categories of mind ... are simply reflections of categories that supposedly exist objectively in the world, independent of all beings", with the view that both categories of mind and human reason depend upon experiential aspects of human psychology. Such an epistemological rather than structural characterization of natural concepts also has a methodological aspect to it; it entails that prototypical categories should not be studied in isolation from their experiential context. While such an epistemological or methodological conception of prototypical categorization is extremely valuable, we shall take a structural point of view in the following pages; we shall try to determine whether it is possible to give a coherent, structurally-intrinsic characterization of prototypical categories.

On the other hand, there are structural characteristics of prototypical concepts that can be reduced to the four basic structural features mentioned above. For instance, in Geeraerts (1985a, 1986a) the flexibility of prototypical concepts is stressed, together with the fact that a distinction between semantic and encyclopaedic components of lexical concepts cannot be maintained in the case of prototypical concepts (1985b). But the flexibility of prototypical categories is linked in a straightforward manner with the fourth characteristic: uncertainties with regard to the denotational boundaries of a category imply that it need not be used in a rigidly fixed manner. Similarly, the absence of a clear dividing line between encyclopaedic and purely semantic information follows from this very flexibility together with the first and second characteristic. The possibility of incorporating members into the category that do not correspond in every definitional respect with the existing members entails that features that are encyclopaedic (non-definitional) with regard to a given set of category members may turn into definitional features with regard to a flexibly incorporated peripheral category member. The resemblance between central and peripheral cases may be based on allegedly encyclopaedic just as well as on allegedly "semantic" features. In short, features of prototypicality that are not included among the ones mentioned in (i)-(iv) may often be reduced to those four, and this in

turn justifies a preliminary restriction of the discussion to the latter.

	Nonequality (differences in structural weight)	Nonrigidity (flexibility and vagueness)
Extensionally	(iii) degrees of representativity	(iv) absence of clear boundaries
Intensionally	(ii) clustering of overlapping senses	(i) absence of clas- sical definitions

Figure 3.1(1)
Characteristics of prototypicality

A second remark with regard to the four characteristics involves the fact that they are systematically related along two dimensions. On the one hand, the third and the fourth characteristic take into account the referential, *extensional* structure of a category. In particular, they have a look at the members of a category; they observe, respectively, that not all referents of a category are equal in representativeness for that category, and that the denotational boundaries of a category are not always determinate. On the other hand, these two aspects (centrality and non-rigidity) recur on the *intensional* level, where the definitional rather than the referential structure of a category is envisaged. For one thing, non-rigidity shows up in the fact that there is no single necessary and sufficient definition for a prototypical concept. For another, family resemblances imply overlapping of the subsets of a category. To take up the formulation used in the quotation under (ii) above, if there is no definition adequately describing A, B, C, D, and E, each of the subsets AB, BC, CD, and DE can be defined separately, but obviously, the "meanings" that are so distinguished overlap. Consequently, meanings exhibiting a greater degree of overlapping (in the example: the senses corresponding with BC and CD) will have more structural weight than meanings that cover peripheral members of the category only. In short, the clustering of meanings that is typical of family resemblances implies that not every

meaning is structurally equally important (and a similar observation can be made with regard to the components into which those meanings may be analyzed). The systematic links between the characteristics mentioned at the beginning are schematically summarized in Figure 3.1(1).

As a third remark, it should be noted that the four characteristics are often thought to be co-extensive, in spite of incidental but clear warnings such as Rosch & Mervis's remark that a family resemblance structure need not be the only source of prototypicality (1975: 599). Admittedly, it is easy to consider them to be equivalent; already in the quotations given above, partial reasons for their mutual interdependence can be found. More systematically, the following links between the four characteristics might be responsible for the idea that prototypicality necessarily entails the joint presence of all four.

First, linking the first to the second characteristic is the argument mentioned above: if there is no single definition adequately describing the extension of an item as a whole, different subsets may be defined, but since the members of a category can usually be grouped together along different dimensions, these subsets are likely to overlap, i.e., to form clusters of related meanings.

Second, linking the second to the third characteristic is the idea that members of a category that are found in an area of overlapping between two senses carry more structural weight than instances that are covered by only one meaning. Representative members of a category (i.e., instances with a high degree of representativity) are to be found in maximally overlapping areas of the extension of a category. (In the example, A and E are less typical members than B, C, and D, which each belong to two different subsets.)

Third, linking the third to the fourth characteristic is the idea that differences in degree of membership may diminish to a point where it becomes unclear whether something still belongs to the category or not. Categories have referentially blurred edges because of the dubious categorial status of items with extremely low membership degrees.

And fourth, linking the fourth to the first characteristic is the idea that the flexibility that is inherent in the absence of clear boundaries prevents the formulation of an essence that is common to all the members of the category. Because peripheral members may not be identical with central cases but may only share some characteristics with them, it is difficult to define a set of attributes that is common to all members of a category and that is sufficient to distinguish that category from all others.

These circular links between the four characteristics are, however, mis-

leading. A closer look at some (familiar and less familiar) examples of prototypicality reveals that they need not co-occur.

BIRD

The category *bird* (one of Rosch's original examples of prototypicality) shows that natural categories may have clear-cut boundaries. At least with regard to our own, real world, the denotation of *bird* is determinate; educated speakers of English know very well where birds end and non-birds begin. They know, for instance, that a bat is not a bird but that a penguin is. Of course, the principled indeterminacy described by Waismann (1952) as "open texture" remains: when confronted with an SF creature (a post-World War III mutant) that looks like a bird but talks like a man, we would not be sure whether it should be called a bird or not. A boundary problem that is typical for a prototypical organization of the lexicon would then arise. As it functions now, however, in present-day English, *bird* is denotationally clearly bounded, the archaeopterix notwithstanding. As has been remarked elsewhere (Lakoff 1987), the existence of prototypicality effects in clearly bounded concepts such as *bird* implies that a strict distinction has to be made between degree of membership and degree of representativity. Membership in the category bird is discrete; something is or is not a bird. But some birds may be birdier than others: the swallow does remain a more typical bird than the ostrich.

RED

Color terms such as *red* constituted the startingpoint for prototypicality research; drawing on the views developed in Berlin & Kay (1969), Rosch's earliest work is an experimental demonstration of the fact that the borderline between different colors is fuzzy (there is no single line in the spectrum where red stops and orange begins), and of the fact that each color term is psychologically represented by focal colors (some hues are experienced as better reds than others) (Heider 1972; Heider & Olivier 1972). These prototypical characteristics on the extensional level are not matched on the definitional level. If *red* can be analytically defined at all (i.e., if it does not simply receive an ostensive definition consisting of an enumeration of hues with their degree of focality), its definition might be "having a color that is more like that of blood than like that of an unclouded sky, that of grass, that of the sun, that of ... (etc., listing a typical exemplar for each of the other main colors)". Such a definition (cf. Wierzbicka 1985: 342)

does not correspond with either the first or the second characteristic mentioned above.

ODD NUMBER

Armstrong, Gleitman & Gleitman (1983) have shown experimentally that even a mathematical concept such as *odd number* exhibits psychological representativity effects. This might seem remarkable, since *odd number* is a classical concept in all other respects: it receives a clear definition, does not exhibit a family resemblance structure or a radial set of clustered meanings, does not have blurred edges. However, Lakoff (1987) has made clear that degrees of representativity among odd numbers are not surprising if the experiential nature of concepts is taken into account. For instance, because the even or uneven character of a large number can be determined easily by looking at the final number, it is no wonder that uneven numbers below 10 carry more psychological weight: they are procedurally of primary importance.

VERS

As shown in Geeraerts (1987), the first characteristic mentioned above is not sufficient to distinguish prototypical from classical categories, since, within the classical approach, the absence of a single definition characterized by necessity-cum-sufficiency might simply be an indication of polysemy. This means that it has to be shown on independent grounds that the allegedly prototypical concepts are not polysemous, or rather, it means that prototypical lexical concepts will be polysemous according to a definitional analysis in terms of necessary and sufficient conditions (the classical definition of polysemy), but univocal according to certain other criteria. These criteria may be found, for instance, in native speakers' intuitions about the lexical items involved, intuitions that may be revealed by tests such as Quine's (1960) or Zwicky & Sadock's (1975). In this sense, the first characteristic has to be restated: prototypical categories will exhibit intuitive univocality coupled with analytical (definitional) polysemy, and not just the absence of a necessary-and-sufficient definition.

Once this revision of the first characteristic is accepted, it can be demonstrated that the first and the second criterion need not co-occur. Lexical items that show clustered overlapping of senses may either conform or not conform to the revised first characteristic. An example of the first situation

is the literal meaning of *bird*, an example of the second situation the Dutch adjective *vers*, which corresponds roughly with English fresh (except for the fact that the Dutch word does not carry the meaning "cool"). Details of the comparison between both categories may be found in the paper mentioned above; by way of summary, Figures 3.1(2) and 3.1(3) represent the definitional analysis of both items.

a. "Being able to fly"

e. "Not domesticated"

b. "Having feathers"

f. "Being born from eggs"

c. "Being S-shaped"

g. "Having a beak or bill"

d. "Having wings"

Figure 3.1(2)

A definitional analysis of *bird*

These figures, incidentally, exemplify a type of representation that will be used extensively further on in this study. Each of the labeled boxes represents a descriptive feature (or combination of features) that seems important in the semantic structure of the item. Within each box, instances of use, or members, of the category under consideration are mentioned. The overlapping of boxes illustrates that not all instances of a category share all the features that are relevant for the category. In the case of *bird*, for instance, it appears that there exist birds that cannot fly in the proper sense of the word, that do not have regular feathers (but are merely covered with a kind of down), or that do not even have wings. The shaded area in the figure indicates the prototypical core of the category, where the overlapping of the

subsets as represented by the labeled boxes is maximal.

The distinction in intuitive status between *vers* and *bird* can be demonstrated by means of the Quinean test (roughly, a lexical item is ambiguous if it can be simultaneously predicated and negated of something in a particular context). Thus, taking an example based on the corresponding ambiguity in the English counterpart of *vers*, it would be quite normal to

Figure 3.1(3)

A definitional analysis of *vers*

state that the news meant in the sentence *there was no fresh news from the fighting* is fresh in one sense ("recent, new") but not in another ("in optimal condition"): it makes sense to say that the news is at the same time fresh and not fresh. By contrast, it would be intuitively paradoxical to state that a penguin is at the same time a bird and not a bird (disregarding figurative extensions of the semantic range of *bird*). Nevertheless, the definitional analyses in Figures 3.1(2) and 3.1(3) make clear that both concepts exhibit prototypical clustering. In both cases, too, the structural position of the instances just discussed (news, penguin) is not in the central area with maximal overlapping. Finally, neither *bird* nor *vers* can receive a classical definition in terms of necessary-and-sufficient attributes. In the case of *vers*, the necessity requirement is not met with: there is no single feature that is common to all the members of the category. In the case of *bird*, the sufficiency requirement is not met with: the features that can be cited as common to all birds do not suffice to distinguish birds from other species (like the duck-billed platypus). In short, then, the revised version of the first characteristic need not coincide with the second characteristic: both *bird* and

vers are characterized by the clustered overlapping of definitional subsets, but in the latter case the absence of a classical definition does not correspond with intuitive univocality, whereas in the former case it does.

	<i>bird</i>	<i>vers</i>	<i>red</i>	<i>odd number</i>
Analytic polysemy coupled with intuitive univocality	+	–	–	–
Clustering of overlapping subsets	+	+	–	–
Degrees of representativity	+	+	+	+
Fuzzy boundaries	–	+	+	–

Figure 3.1(4)
The prototypicality of “prototypicality”

The insight derived from a closer look at the four examples just described may be summarized as in Figure 3.1(4). It is now easy to see to what extent “prototypicality” is itself a prototypical notion. For one thing, the examples brought together in Figure 3.1(4) exhibit a family resemblance structure based on partial similarities. For instance, the set of prototypical concepts characterized by clustering of senses overlaps with the subset characterized by fuzzy boundaries (because of *vers*), and so on. For another, some concepts are more typically prototypical than others, in the sense that they exhibit more of the “prototypical” characteristics. (*Bird* and *vers* are more prototypical than *red*. Notice, in particular, that the category *fruit* makes a good candidate for prototypical prototypicality, in the sense that it seems to combine all four characteristics. It shares the prototypical characteristics of *bird*, but in addition, the dubious membership status of things such as coco-

nuts and, perhaps, tomatoes, seem to point out that the denotational boundary of *fruit* is less clear-cut than that of *bird*.)

However, although the examples considered above do not have a set of attributes in common, they do share a single feature, viz. degrees of membership representativity. It is highly doubtful, though, whether this feature alone suffices to distinguish prototypical concepts from classical concepts. If the possibility of a single necessary-and-sufficient definition is one of the features par excellence with which the classical conception has been identified, it might be claimed that degrees of representativity are entirely compatible with the classical conception of categorization. It is, in fact, in that sense that Armstrong, Gleitman & Gleitman (1983) deal with a category such as *odd number*. The experiments used by Rosch to measure degrees of representativity are not, they claim, indicative of prototypicality since they occur with classical, rigidly definable concepts such as *odd number*. However, such an argumentation partly begs the question, to the extent that it presupposes that prototypicality should be defined in terms of non-classical definability alone. If that assumption is itself questioned, the Armstrong, Gleitman & Gleitman results basically show that a number of characteristics that were thought to coincide in the concept of prototypicality need not in fact always co-occur. At the same time, the debate over the status of *odd number* shows that the concept "prototypical concept" has no clear boundaries: given the dissociation of the features that were originally thought to coincide, it is not immediately clear whether a concept such as *odd number* should be included in the set of prototypical concepts or not.

Of course, contrary to the situation in everyday speech, such a boundary conflict should not be maintained in scientific speech. A discipline such as linguistics should try to define its concepts as clearly as possible, and the purpose of this section is precisely to show that what has intuitively been classified together as instances of prototypical categories consists of distinct phenomena that have to be kept theoretically apart. In line with prototype theory itself, however, such an attempt at clear definition should not imply an attempt to define the "true nature" or the "very essence" of prototypicality. Determining an "only true kind" of prototypicality is infinitely less important than seeing what the phenomena are and how they are related to each other by contrast or similarity. In this respect, the foregoing analysis corroborates Wierzbicka's remark that there are "many senses" to the notion *prototype*, and that "the notion *prototype* has been used in recent literature as a catch-all notion" (1985: 343). However, a more systematic analysis than Wierzbicka's reveals that this very multiplicity of usage also supports

cognitive semantics, in the sense that it shows that the same categorization principles may guide common sense and scientific thinking: the concept of prototypicality has been used in the same loose and clustered way that prototype theory pinpoints as a major structural characteristic of everyday categories.

The practical consequences of this insight for an investigation into the structure of lexical variation will be clear. The study of semasiological variation in a prototype-oriented framework will have to distinguish systematically between the various types of prototypicality that may be recognized. In the following sections, then, the various prototypicality effects that were brought together in Figure 3.1(1) will be investigated separately. In section 3.2., intensional non-discreteness will be discussed; it will be shown that the clothing terms database does indeed contain cases for which classical definitions in terms of necessary and sufficient conditions cannot be given. Section 3.3. focuses on extensional non-discreteness. It will be shown that the corpus contains categories whose membership boundaries are not clear; moreover, this unclarity will be shown to affect the intensional discreteness (i.e. the definability) of the categories in question. In section 3.4., intensional and extensional non-equality will come to the fore. It will be shown that membership salience effects may occur both with classically definable and not classically definable categories.

3.2. Non-discreteness of word meanings [1]: definability

This section will do two things: present a methodology for determining whether lexical categories can be classically defined or not, and show that the clothing terms database does indeed contain both classically and non-classically definable elements. The polysemy problem that was discussed in the previous section with regard to the issue of classical definability will not play a role for our database; we will only investigate monosemous items. This does not mean, to be sure, that definability questions can be straightforwardly settled; we will rather try to show that answering such questions not only requires a systematic approach, but also involves a number of unexpected difficulties.

Classical definability implies that a definition can be found that characterizes all the members of the category to be defined, and only those. The

definition has to be *general*, in the sense that it applies to all the members of the category, and it has to be *distinctive*, in the sense that it adequately distinguishes the category from all others. For instance, let us assume that we are trying to define the category "bird" (as a biological species). We will then have to list the attributes that all birds have in common, if there are any; further, we will have to make out whether this list of attributes (or any subset of it) suffices to distinguish birds from mammals, reptiles, and fishes, to say the least. As illustrated in the previous section, the attributes that one would be inclined to mention as general characteristics of birds, often do not have the required commonality. On the other hand, the attributes that do seem to be general among birds do not suffice to distinguish birds from other species; even when the features in question are taken together, the duck-billed platypus is a counterexample to the alleged definition. It may be useful to point out that there are various other ways of terminologically indicating the classical nature of definitions. One is to say that classical definitions define *all and only* the members of the category, while another is to say that they *uniquely* define the category. More importantly, however, it has to be noted that applying the definition meets with particular problems in the case of our material. Before turning to an actual example, we will consider each of the two requirements in more methodological detail.

The first part of the joint requirement of generality and distinctiveness would seem to be easy to check: our componential description of the referents of each item allows us to check whether there are any attributes that these members have in common. There is an important reason, however, for rejecting such a straightforward and mechanical procedure. The descriptive features that define the various configurations in the referential range of a lexical item cannot be taken at face value, but have to be *interpreted* (as an automatic consequence of which, the issue of classical definability cannot be settled mechanically). There are basically two forms of interpretation to be taken into account. For ease of reference, they will be called the quantitative and the qualitative one. The *quantitative* interpretation involves numerical dimensions, i.e. dimensions whose values constitute a graded continuum. The crucial point here is to see that it is not the individual value of a specific referent with regard to that dimension that is definitionally important, but rather the range of values with which the dimension occurs. If, for instance, a dimension like WIDTH receives the values [2], [3], and [4] in the semasiological range of application of an item, we should not say that the item has no common feature on the dimension WIDTH, but we should rather say that the width of the referents of the item in question ranges from value [2] to [4].

Although the presence of the values [2], [3], and [4] would superficially suggest that the referents of the item do not have common characteristics as far as their width is concerned, they do upon closer inspection: all of them fall within the range defined by the interval [2]-[4]. On the other hand, a *qualitative* reinterpretation of the superficially given values involves hidden variables. In particular, whereas all the dimensions in the database are visual ones, there may be covert dimensions of a functional nature. For instance, if the MATERIAL dimension of an item features the values [silk] and [cotton], there is again, superficially speaking, no common characteristic. If, however, both silk and cotton are used as light materials serving the purpose of keeping the person cool in warm weather, the common functional feature [light and cool] reduces the original variation on the MATERIAL dimension to epiphenomenal status.

The distinctiveness criterion for classical definability should be handled with equal care. To begin with, notice that the distinctiveness requirement crucially involves *negative* evidence. If a definition is to hold for all and only the members of a particular category, the definition should not apply to any specific thing that does not belong to the category. The distinctiveness of the definition is contradicted, in other words, if we can find a referent that falls within the scope of the definition but that falls outside the scope of the category. This does not mean, to be sure, that the items falling within the scope of the definition could never occur as members of other categories than the one to be defined. For instance, let us define the attributes *plusquint* and *deciminus* of natural numbers. A natural number is *plusquint* if it is larger than five; it is *deciminus* if it is smaller than ten. Both definitions are classical: they are as mathematically precise as you can get. At the same time, both categories naturally overlap: the natural numbers 6, 7, 8, and 9 fall within both categories. This means that the number 7 may sometimes be called a *plusquint* and sometimes a *deciminus*. Suppose further that we have actually encountered both ways of speaking in our corpus of mathematical language, and that we are trying to define the word *deciminus*. We have noted that all *deciminus*es share the property of being smaller than ten, and therefore propose to define *deciminus* accordingly; we also notice, however, that 7 is sometimes called a *plusquint*. When confronted with such a *plusquint* instance of 7, could we then say (repeating the sentence introduced above) that 7 falls within the scope of the definition of *deciminus*, but that it falls outside the scope of the category (because it is not then called a *plusquint*)? Of course not: falling outside the scope of the category means *never* occurring within it. The number 7 is not a counterexample to the pro-

posed definition of *decimus* because it is occasionally called a *plusquint*, but it would be a counterexample if it were never called a *decimus*. In the same way, the duck-billed platypus is only a counterexample to the classical definability of *bird* because it is never categorized as a bird. At the same time, the fact that not *all* decimuses can be called *plusquint* implies that the definition of the latter category should not be so broad as to include the entire range of application of *decimus*; it must include those decimuses that are sometimes called *plusquint*, but it must exclude those that are never so called.

As a practical consequence of this observation, we will have to check any alleged classical definition of a lexical item against the words with which the item referentially overlaps. In particular, the definition should not be overly general, in the sense that the entire overlapping category (rather than just the intersecting part) is drawn into the category to be defined. The relationship between two items that share referents may, however, take other forms than the kind of *overlapping* (partial co-referentiality, semi-synonymy) that is illustrated by the *plusquint/decimus* case. Systematically, there are three other relations to be envisaged. When the items are *synonymous*, no problem arises when the definition of the definiendum covers the entire range of application of the second item. Similarly, when the definiendum is a *hyperonym* of the second item, the definition may (in fact, must) cover all the referents of the second item. But when the definiendum is a *hyponym* of the other word, a definition that exceeds the referential boundaries of the hyponymous item will have to be rejected. To summarize, the distinctiveness criterion does not apply to the synonyms and the hyponyms of the definiendum; in the case of overlapping and hyperonymous categories, it should only be applied to the overlapping and hyperonymous categories as a whole, not to those subsets of the latter that they share with the definiendum.

Now that we have a better idea of how the classical definability of lexical items can be established, actual examples can be considered. In the following pages, the lexical item *legging* will be considered in detail as an illustration of the methodology outlined above. The other lexical items that will be analyzed further on in this section will be treated more succinctly; rather than concentrating on the analytical procedure itself, we will then concentrate on the results of the analysis.

The semasiological information that can be extracted from the database takes the form of a list of referential descriptions of the form illustrated in Figure 3.2(1). The figure lists all the referential configurations with which the item *legging* occurs, together with their respective frequencies. Incom-

plete records (like records that contain question marks, see the introduction of the componential system in section 2.1.) have been left out of consideration.

<i>Configuration</i>	<i>Frequency</i>
H3118.v	3
H4118.v	36
H41186v	3
H4211.v	1
H5118.v	58
H51186v	2
H5128.v	1
H5154.v	1
H5211.v	3
H52115v	1
H52116v	1

Figure 3.2(1)
The semasiological range of *legging*

Finding out whether *legging* can be classically defined would now seem to follow a straightforward procedure: first, it would have to be established whether there are any characteristics that are common to all referents of *legging*, and second, it would have to be investigated whether the resulting definition is sufficient to distinguish *legging* from all other categories that are neither hyponyms nor synonyms of *legging*. On the basis of this procedure, *legging* would definitively turn out not to be classically definable. A glance at the figure suffices to appreciate that the only truly general characteristics of all the listed instances of *legging* are the fact that they all involve trouser-like garments, as represented by the feature [H], and the fact that they are worn by women, as represented by the feature [v]. At the same time, the database contains various trousers worn by women that are never called *legging*; specifically, wider and shorter types of trousers for women fall outside the category.

But obviously, we have not yet subjected *legging* to the quantitative and

qualitative reinterpretation process that we described above. Let us now try to establish whether we can salvage the classical definability of *legging* by using a more refined approach. A quantitative reinterpretation is important for the dimensions WIDTH and LENGTH. The width of the referents of *legging* varies between the values [1] and [2], which is to say that leggings are either tight or narrow. The dimension LENGTH has a range between values [3] and [5]; the referents of *legging* reach down at least to the knee, but they may also cover the entire leg down to the ankle. The impact of a qualitative reinterpretation can be appreciated when we have a look at the dimension MATERIAL. The fact that the predominant value on this dimension, viz. [8], refers to stretchy fabrics suggests that there is a causal connection with the dimension WIDTH: leggings are mostly tight or narrow precisely because they are made of elastic material. Could it be the case, then, that the feature [elastic] allows for a reduction of the variation on the dimension MATERIAL? A renewed consultation of the original pictures on which the database records were based, reveals that this is indeed the case. On the one hand, the records that contain the value [1] on the dimension MATERIAL appear to be made of a finely woven tricot that is at least moderately elastic. On the other hand, the single record that features the value [4] for MATERIAL involves a stretchy, very tight-fitting corduroy. In other words, although not all leggings have the same degree of elasticity, they do share a certain amount of stretchiness. The resulting picture of the common characteristics of all instances of *legging* may be summarized as follows:

a *legging* is a two-legged outer garment for women covering the lower part of the body from the waist down, ranging in length from the knee to the ankle, and made from elastic materials such that the width of the legs ranges from tight to narrow.

The next step involves checking whether this set of common features is sufficiently distinctive to act as a definition of *legging*. In order to get an idea of the lexical items that have to be included in an analysis of the distinctiveness of the definition of *legging* that was given above, Figure 3.2(2) lists the onomasiological alternatives with which *legging*-configurations occur in the database. For each of the various configurations that are situated within the semasiological range of *legging*, Figure 3.2(2) specifies the other lexical items referring to that configuration, together with their frequency.

<i>Configuration</i>	<i>Item / Frequency</i>	
H3118.v	broekje	1
H3118.v	kniebroekje	1
H3118.v	legging	3
H3118.v	piratenbroek	1
H3118.v	wielrennersbroek	1
H4118.v	broek	12
H4118.v	broekje	1
H4118.v	caleçon	1
H4118.v	denimbroek	1
H4118.v	jeans	1
H4118.v	kuitbroek	1
H4118.v	legging	36
H4118.v	leggings	9
H4118.v	piratenbroek	3
H4118.v	stretchbroek	1
H4118.v	tricotbroek	1
H41186v	broek	2
H41186v	caleçon	1
H41186v	legging	3
H41186v	leggings	2
H4211.v	broek	17
H4211.v	caleçon	1
H4211.v	kuitbroek	1
H4211.v	legging	1
H4211.v	pantalon	1
H4211.v	streepbroek	1
H5118.v	broek	12
H5118.v	broekje	1
H5118.v	legging	58
H5118.v	leggings	23
H5118.v	skibroek	3
H5118.v	stretchbroek	1
H5118.v	stretchleggings	1
H5118.v	tricotbroek	5
H51186v	broek	6

H51186v	caleçon	2
H51186v	legging	2
H51186v	leggings	7
H51186v	pantalon	1
H5128.v	broek	1
H5128.v	joggingbroek	1
H5128.v	legging	1
H5128.v	skibroek	2
H5154.v	legging	1
H5211.v	bloemenbroek	1
H5211.v	broek	74
H5211.v	caleçon	1
H5211.v	jeans	11
H5211.v	legging	3
H5211.v	pantalon	2
H52115v	broek	24
H52115v	legging	1
H52115v	pantalon	1
H52115v	tricotbroek	1
H52116v	legging	1
H52116v	broek	3

Figure 3.2(2)

The onomasiological alternatives for *legging*

It should be noted that the figure is already a revised version of the overview that may be automatically retrieved from the database. Three kinds of elements have been discarded. First, incomplete configurations have been removed from the list. For instance, the configuration [H?2?1.v], which occurs once in the semasiological range of *legging* is listed as having *hot-pants* as an onomasiological alternative. Given, however, that no *legging* is ever as short as *hot-pants* are (which are typically hardly longer than the groin), the alternation is an artefact of the question marks in the configuration. Even if the question mark for the dimension LENGTH of the *hot-pants*-exemplar were to hide an unusually long specimen (as long as, say, ordinary shorts), the length of the allegedly corresponding *legging*-exemplar can be shown to be greater.

Second, a number of alleged cases of onomasiological alternation have to

be removed because the descriptive configurations on which they are based turn out to be insufficiently specific. For instance, [H52115v] occurs once with the alternative term "*sigarettenpijp*"-*broek* (literally 'cigarette-legged trousers'). Apparently, the motivation for the latter name is not the fact that the legs are straight and narrow like a cigarette. (They are narrow, but only moderately so; moreover, they are not specifically straight.) Rather, the motivation for "*sigarettenpijp*"-*broek* resides in the fact that the legs' ends have a rather long turnover that make the legs look like filtertip cigarettes. Because this defining characteristic of "*sigarettenpijp*"-*broek* does not surface as a definite value on any of the descriptive dimensions, the onomasiological correspondence could very well be a spurious one: there would only be an actual overlap between *legging* and "*sigarettenpijp*"-*broek* if any of the leggings that have the [H52115v]-configuration had a filtertip turnover. Upon inspection, this appears not to be the case.

Third, alternative names may have to be removed because they do not appear often enough to determine their actual semantic range. *Legging-broek* (literally 'legging-trousers'), for instance, occurs once as an alternative for the configuration [H5118.v], but it is unclear whether it should be regarded a synonym or a hyponym of *legging*. In the former case, the term would have the status of a specificational compound; given that *broek* is a general name for all kinds of two-legged garments covering the lower part of the body, *legging-broek* could then be interpreted as "a *broek* that is more specifically a *legging*". Taking into account, however, that *broek* may also be conceptualized in terms of its prototype (a possibility to which we will come back in more detail when we discuss *broekrok*), *legging* and *broek* may also be semi-synonyms: not all leggings are prototypical instances of *broek* (because leggings may be shorter than the prototypical pair of trousers, which reaches down to the ankle), and not all prototypical trousers are leggings (because the former are not as tight as the typical *legging*). In this case, *legging-broek* would refer to a particular type of *legging*, viz. one that is long enough and wide enough to fall within the set of prototypical trousers, and would therefore have to be considered a hyponym of *legging*. The single case of *legging-broek* that is available does not, however, allow us to settle the question.

The status of the remaining onomasiological alternatives in terms of their lexical relationship with regard to *legging* is specified in Figure 3.2(3). In accordance with the methodology set out above, we need not worry about a possible lack of distinctiveness of the definition of *legging* with regard to a hyponym like *stretchlegging* ('legging in stretchy fabric'), or with regard to

synonymous expressions like *leggings* and *caleçon*. (In passing, it may be noted that the latter term is geographically restricted to the Belgian sources, a point to which we will come back in section 5.3., where the influence of geographical factors on choices among synonyms is discussed. Also, note that the treatment of hyponyms is actually less obvious than suggested here, as will be discussed at length in the following section.) The hyperonym *broek* does not present a real danger either, because the definition of *legging* specifies an actual subset of the range of application of *broek* (which may, in particular, refer to two-legged garments that are much shorter and much wider than the referents of *legging*).

Synonyms	<i>leggings, caleçon</i>
Hyperonym	<i>broek</i>
Hyponym	<i>stretchleggings</i>
Semi-synonyms	<i>kuitbroek, kniebroekje, broekje</i> <i>jeans, denimbreek, stretchbroek, tricotbroek</i> <i>streepbroek, bloemenbroek</i> <i>piratenbroek, wielrennersbroek, skibroek, joggingbroek, pantalon</i>

Figure 3.2(3)

The lexical relations between *legging* and its onomasiological alternatives

As suggested in Figure 3.2(3), various subsets may be distinguished within the set of semi-synonyms. A first set comprises *kuitbroek* 'calf-long pair of trousers', *kniebroekje* 'knee-long pair of trousers', and *broekje* 'short pair of trousers, shorts', which each delimit a specific area within the range of application of *broek* on the basis of the length of the garment. The second set consists of *jeans*, *denimbreek*, *stretchbroek*, and *tricotbroek*, which each

refer to trousers made of a particular fabric. In the third set we find items that refer to a particular type of motif or decoration: *streepbroek* indicates the presence of stripes, and *bloemenbroek* signals the presence of flowers. The fourth subset contains *piratenbroek* 'pirate's trousers', *wielrennersbroek* 'cyclist's trousers', *skibroek* 'ski pants', *joggingbroek* 'jogging pants', and *pantalon* 'pair of trousers'. There is an interesting distinction between the fourth subset and the previous three, in the sense that the items in the latter may be adequately defined as a hyponym of *broek* by referring to a single dimension (length, material, and motif respectively). These dimensions determine the morphological structure of the words, in the sense that the first member of the compound refers to the specific value on the relevant dimension that is crucial for the item. For the items in the fourth subset, however, various dimensions have to be mentioned at the same time in order to specify their proper position.

In the cases involved in the first three sets, establishing the referential overlap with the semasiological range of *legging* is a relatively straightforward matter. The items are actual semi-synonyms to the extent that they may refer to pieces of clothing that are not *legging*-like on any dimension that is irrelevant for the item in question. For instance, referents of *kuitbroek* that are not tight-fitting enough to fall within the definition of *legging* are not so called either. *Kuitbroek* is characterized on the basis of the dimension LENGTH, and its specific value on this dimension happens to fall within the range of lengths that is definitional for *legging*. Because of its "unidimensional" nature, however, it is not definitionally specified with regard to other dimensions that are subject to restrictions in the case of *legging* (such as, in the example, WIDTH), and it may therefore refer to pieces of clothing that are definitely too wide for leggings. We have to make methodological allowances, though, for the fact that the number of records we have for the various items does not always suffice to establish their overlapping status beyond all doubt. That is to say, we may not have enough examples of an item like *bloemenbroek* to establish whether it does indeed occur with referents that are not leggings. In these cases, we have nevertheless listed the item as a semi-synonym (rather than discarding it as we did earlier with *legging-broek*) on the basis of the assumption that the morphological structure of the item is a good indication of its semantics – on the basis of the assumption, for instance, that trousers with a flower motif may be called *bloemenbroek* regardless of their other characteristics.

For the items of the fourth subset identified above, such an abductive underpinning of the classification on the joint basis of intuition and morpho-

logical structure is less obvious. Although most of the referents of *skibroek* are relatively tight-fitting, the widest ski pants are definitely wider than the widest referent of *legging*; *skibroek*, in other words, has a wider range on the dimension WIDTH than *legging*. The same criterion also distinguishes *joggingbroek* from *legging*. For *wielrennersbroek*, on the other hand, the most important distinctive dimension is LENGTH: whereas *legging* ranges from dimensional value [3] to [5], the referents of *wielrennersbroek* are never longer than the knee. Similarly, *piratenbroek* refers to garments that may be as long as the knees or the calves, but never as long as the ankles.

In short, the discussion of *legging* shows that a careful analysis of semasiological ranges and lexical relations is necessary to determine the classical definability of lexical items. The *legging*-example produces positive results, in the sense that the item in question appears to be classically definable. The next step will be to discuss a number of cases where the definitional analysis yields negative results, in the sense that the items in question cannot be defined on a classical basis. The discussion will be based on a subset of the field of clothing terminology. In particular, we will consider pieces of clothing that cover the upper part of the body, that can be entirely opened at the front, and that are never worn as the first layer of clothing above the underwear. The items with the highest frequencies in this subset are *jack*, *colbert*, *blazer*, *jasje*, and *vest_{nl}*. The distinction between *vest_{nl}* and *vest_b* is necessary because there is a marked difference between the ways in which the item *vest* is used in the Belgian and the Netherlandic sources; we will come back to this point below.

Figure 3.2(4) gives an overview of the ranges of application of the items. The figure is based on individual analyses of the items along the principles demonstrated in connection with *legging*. That is to say, the dimensions used in the figure do not necessarily reproduce the information structure of the database in a straightforward manner, but may be the result of a reinterpretation of the stored data or even a reconsideration of the original pictures. Dimensions that are not distinctive within the subset have been left out. For instance, all the types of clothing included in the subset are worn by men and women alike; accordingly, the dimension SEX has not been retained in the figure. A plus sign means that a particular dimensional value occurs within the range of application of the item; a minus sign indicates that it never occurs. Thus, plus signs on all values of a particular dimension mean that both values may occur. For instance, the referents of *jasje* may either occur with a type of fastening that can be fastened up to the neck, or with a type of fastening that stops on the chest somewhat lower than the neck; by contrast,

jack is never used as a name for garments that cannot be fastened entirely. The only dimensions in the overview for which the relevant values might have to be restated in terms of ranges are LENGTH and FASTENING, since both involve measures of length; for instance, the referents of *jasje* have a fastening whose length ranges from up to the chest to up to the neck. The other dimensions consist of discontinuous values.

	jack	colbert	blazer	vest _{nl}	jasje
<i>length</i>					
shorter than the waist	—	—	+	+	+
as long as the waist	+	—	+	+	+
lower than the waist	+	+	+	+	+
<i>cut</i>					
blousing	+	—	—	—	+
wide and straight	—	+	+	+	
narrow and straight	+	+	+	+	
waisted	—	+	+	+	+
<i>material</i>					
woven fabrics	+	+	+	+	+
knitted	—	—	+	+	+
leather	+	—	—	—	+
<i>fastening</i>					
up to the neck	+	—	—	+	+
lower than neck	—	+	+	+	+

Figure 3.2(4)

The semasiological ranges of *jack*, *blazer*, *colbert*, *vest_{nl}*, *jasje*.

A first thing to note is that the referential ranges included in Figure 3.2(4) suggest the existence of certain hyponymy relations. It appears, for instance, that all dimensional values that occur in the range of *jack* also occur in the range of *jasje*; at the same time, the latter item exhibits a number of dimensional values that are absent in the case of *jack*. In this particular case, the suggestion that *jack* is a hyponym of *jasje* (because the referential range of the latter word includes that of the former) is supported by the intuition that *jasje* is a cover-term for the entire set of items included in 3.2(4). However, the overview in the figure is not really a good way of settling the hyponymy relations among the items, because the referential ranges are being considered in terms of separate dimensions rather than dimensions in combination. Consider a fictitious case in which an item A is represented by the referential types [ac] and [bd], and an item B by the types [ad] and [bc]. In both cases, the first dimension ranges over the values [a] and [b], and the second dimension over the values [c] and [d]. Judging on the basis of an overview of dimensional ranges, then, A and B would be synonymous, since they have the same dimensional ranges. Judging on the basis of the dimensional values as they occur in combination, however, it becomes clear that there is neither a relationship of synonymy nor hyponymy between both items. It is therefore necessary to establish hyponymy relations on another basis than Figure 3.2(4) as such.

A corpus-based approach for the recognition of hyponymous relations may be established as follows. If A is a hyponym of B, B may occur as an alternative name for all referents of A. Of course, B need not be as frequent as A for the referential set in question, because A may be more entrenched than B (in the sense of “entrenchment” that will be discussed in more detail in section 4.2.). Also, it may be expected for statistical reasons that the less common referential types of A may not occur in the corpus with B as an alternative denomination; in actual practice, it may be sufficient to establish that B occurs as an alternative for the most common referents of A. In Figures 3.2(5) and 3.2(6), such overviews of onomasiological alternatives are given for *blazer* and *colbert*. Given the statistical margin that was just mentioned, it can be deduced from the figure that *colbert* is a hyponym of *blazer* and *jasje*, and that *blazer* is a hyponym of *jasje*. At the same time, of course, it should be established that there are cases of B that are not named by means of A (lest a situation of synonymy rather than hyponymy obtains). This type of information, however, can be safely derived from overviews like the one in Figure 3.2(4): for instance, the plus sign on the “knitted” value of the dimension MATERIAL for *blazer* as opposed to the minus sign for *colbert*

indicates that knitted referents of *blazer* never occur with the name *colbert*, for the simple reason that the referential range of *colbert* does not include knitted garments of any kind.

<i>Colbert</i> : configurations	Frequency for <i>colbert</i>	<i>Blazer</i> as alternative	<i>Vest</i> as alternative	<i>Jasje</i> as alternative
B1121m	1	—	—	—
B1122v	7	+	—	+
B2111m	2	+	—	+
B2121m	52	+	—	+
B2121v	4	+	—	+
B2122v	17	+	—	+
B2131m	5	—	—	+
B2211m	1	—	—	—
B2212m	1	—	—	—
B2221m	13	+	—	+
B2222m	1	—	—	+
B2222v	4	+	—	+
C3212m	1	—	—	—
C3311v	3	—	—	+
C3312v	1	+	+	+

Figure 3.2(5)

Blazer, *vest*_{nl} and *jasje* as onomasiological alternatives for *colbert*.

Repeating the procedure illustrated in 3.2(5) and 3.2(6) for the items *jack* and *vest* leads to the hyponymy relations that are charted in Figure 3.2(7). (The label “+lh” indicates that there is a relationship of hyponymy between the items in question, given that a lexical test of hyponymy as illustrated in 3.2(5) and 3.2(6) is used. The label “-lh” signals the absence of hyponymy according to the lexical criterion.) The definitional question regarding the five items can now be made more precise: can the items be classically defined on the basis of the overview in 3.2(4) without obscuring the lexical

<i>Blazer</i> : configurations	Frequency for <i>blazer</i>	<i>Vest</i> as alternative	<i>Jasje</i> as alternative
B1122v	10	—	+
B1222v	10	—	+
B2111m	2	—	+
B2111v	1	—	+
B2112v	17	—	+
B2121m	5	—	+
B2121v	14	—	+
B2122v	48	—	+
B2131v	3	—	+
B2132v	8	—	+
B2212v	12	—	+
B2221m	3	—	+
B2222v	22	+	+
B2232v	5	—	+
C2212v	1	+	+
C1312v	1	+	+
C2311v	1	+	+
C2312v	7	+	+
C2322v	1	—	+
C2332v	1	—	+
C2412v	1	+	+
C3212v	1	+	+
C3312v	14	—	+
C3332v	1	+	+
C3412v	4	—	+
C3432v	1	+	+

Figure 3.2(6)

*Vest*_n and *jasje* as onomasiological alternatives for *blazer*.

relations summarized in 3.2(7)? This involves reviewing all possible classical definitions of the items and checking whether they respect the relations in 3.2(7). Note, however, that only two of the four dimensions included in Figure 3.2(4) can be used to establish classical definability. Whereas the values on the LENGTH and the FASTENING dimension can be expressed in terms of ranges, any reference to the other two dimensions automatically

hyponym:	jack	colbert	blazer	vest _{nl}	
hyperonym:					
colbert	-lh
blazer	-lh				+lh
vest _{nl}	-lh				-lh
jasje	+lh	+lh	+lh	+lh	

Figure 3.2(7)

The hyponymy relations between *jack*, *blazer*, *colbert*, *vest_{nl}*, *jasje*.

involves disjunctive values; therefore, including these dimensions in the definition inevitably turns the definition into a non-classical one. Thus, it will be sufficient to investigate only the potential definitions that refer to the dimensions LENGTH and FASTENING. A further restriction follows from the specific position of *jasje*. Both from the point of view of the featural ranges given in 3.2(4) and from the point of view of the lexical relations specified in 3.2(7), *jasje* appears to act as a cover-term with regard to the other items: it is a hyperonym of the other four, and the row of plus-signs accompanying *jasje* in 3.2(4) indicates that it indeed includes all definitional possibilities that are relevant for the other items. In this sense, the search for classical, distinctive definitions may be confined to the other four items. In all of the

potential definitions mentioned below, then, *jasje* will be used as a cover-term; roughly, it may be defined as referring to garments that cover the upper part of the body, that can be entirely opened at the front, and that are never worn as the first layer of clothing above the underwear.

The definitions that are to be taken into account may be systematically grouped together in three sets: definitions that refer only to the dimension LENGTH, definitions that refer only to the dimension FASTENING, and definitions that involve both dimensions at the same time. For ease of reference, we will call *jack1* the definition that involves LENGTH, *jack2* the definition that involves FASTENING, etc.. All in all, the following twelve definitions have to be considered.

Jack1

Jasje whose length ranges from the region of the hip to the region of the upper part of the legs

Jack2

Jasje that can always be fastened up to the neck

Jack3

Jasje whose length ranges from the region of the hip to the region of the upper part of the legs, and that can always be fastened up to the neck

Colbert1

Jasje that reaches down to the region of the upper part of the legs

Colbert2

Jasje that can only be fastened as far as the chest (but not up to the neck)

Colbert3

Jasje that reaches down to the region of the upper part of the legs and that can only be fastened as far as the chest (but not up to the neck)

Blazer1

Jasje whose length ranges from the waist to the region of the upper part of the legs

Blazer2

Jasje that can only be fastened as far as the chest (but not up to the neck)

Blazer3

Jasje whose length ranges from the waist to the region of the upper

part of the legs and that can only be fastened as far as the chest (but not up to the neck)

*Vest*₁

Jasje whose length ranges from the waist to the region of the upper part of the legs

*Vest*₂

Jasje with a fastening whose length ranges from the up to the chest to up to the neck

*Vest*₃

Jasje whose length ranges from the waist to the region of the upper part of the legs and that has a fastening whose length ranges from up to the chest to up to the neck

These twelve definitions may occur in 81 (=3⁴) combinations. That is to say, the general definability question boils down to 81 questions of the type: if *jack* is defined as *jack*₁, *colbert* as *colbert*₃, *blazer* as *blazer*₂, and *vest* as *vest*₁, do the lexical relations that follow from these definitions then conform to the actual relations that are summarized in Figure 3.2(7)? Or, in other words, is there any combination of the twelve classical definitions that makes the right predictions about the attested lexical relations? Rather than considering all 81 possibilities separately, it can be shown in the following way that the question has to be answered in the negative.

First, consider all combinations of two elements from among the set of definitions that refer only to LENGTH. The co-occurrence of *jack*₁ and *colbert*₁ is to be excluded, because this would counterfactually imply that *colbert* is a hyponym of *jack* (as the range of LENGTH for *jack* as defined includes the range as defined for *colbert*). The co-occurrence of *jack*₁ and *blazer*₁ is to be excluded because it would imply that *jack* is a hyponym of *blazer*. The co-occurrence of *jack*₁ and *vest*₁ is to be excluded because it would imply that *jack* is a hyponym of *vest*. The co-occurrence of *colbert*₁ and *vest*₁ is to be excluded because it would imply that *colbert* is a hyponym of *vest*. And the co-occurrence of *blazer*₁ and *vest*₁ is to be excluded because it would imply that *blazer* and *vest* are synonymous.

Second, consider all combinations of two elements from among the set of definitions that refer only to FASTENING. The co-occurrence of *jack*₂ and *vest*₂ has to be excluded because it implies that *jack* is a hyponym of *vest*. The co-occurrence of *colbert*₂ and *blazer*₂ has to be excluded because it implies that *colbert* is a synonym of *blazer*. The co-occurrence of *colbert*₂

and *vest*₂ has to be excluded because it implies that *colbert* is a hyponym of *vest*. And the co-occurrence of *blazer*₂ and *vest*₂ has to be excluded because it implies that *blazer* is a hyponym of *vest*.

Third, consider all combinations of two elements from among the set of definitions that refer to both LENGTH and FASTENING. The co-occurrence of *jack*₃ and *vest*₃ has to be excluded because it implies that *jack* is a hyponym of *vest*. The co-occurrence of *colbert*₃ and *vest*₃ has to be excluded because it implies that *colbert* is a hyponym of *vest*. And the co-occurrence of *blazer*₃ and *vest*₃ has to be excluded because it implies that *blazer* is a hyponym of *vest*.

jack	colbert	blazer	vest
—	1	1	—
2	2	—	—
2	—	2	—
3	3	—	—
—	3	3	—
3	—	3	—

Figure 3.2(8)

Allowed combinations of classical definitions of *jack*, *colbert*, *blazer*, *vest*.

The alternatives that remain at this point are summarized in Figure 3.2(8). It is now immediately obvious that there can be no combination of four classical definitions that respects the existing restrictions, if only because all possible definitions of *vest*₁₁ have already been ruled out. We may conclude, then, that there is no set of classical definitions for *jack*, *colbert*, *blazer*, and *vest* that sufficiently distinguishes the items among each other and that respects the lexical relations that appear to exist among them. Whether this is the dominant situation in the lexicon is difficult to say on the basis of our material; after all, we have only been able to examine a few lexical categories. One general conclusion, at least, is that indeed *not all lexical categories can be classically defined*. In addition, it is worthwhile to point out that the definability issue seems to be strongly influenced by the

specific subfield of the field of clothing terminology that is being considered. As we will have occasion to discuss in more detail later on, the “skirts”-subfields contains classically definable categories, whereas the subfield consisting of shirts, blouses, t-shirts, and their likes is as unclassical as the subfield analyzed in the previous pages.

But although the extent of the phenomenon may well be less massive than early prototype research has tended to suggest, establishing the absence of classical definability clearly requires more sophisticated argumentation than either proponents or adversaries of the classical model usually exhibit. In this respect (and this is entirely in line with the overall focus of our study), we find the *methodological conclusions* to be derived from the foregoing more important than determining the scale of the non-classical definability of the lexicon. What we have tried to illustrate (apart from the fundamental fact that non-classical definability is a real phenomenon) is the importance of a rigorous procedure in answering definability questions. Apart from the importance of distinguishing between polysemy and univocality (see the previous section), there are two requirements that discussions of definability should strictly adhere to. First, the generality of classical definitions implies that no disjunctive features are included in the analytical definitions, and second, the distinctiveness of classical definitions implies that they should make the right predictions about the lexical relations among the items involved. As both the *legging*-example and the *jasje*-subfield shows, complying with these demands requires a careful, step by step procedure that stands in sharp contrast with the usual loose-handedness followed in these matters.

3.3. Non-discreteness of word meanings [2]: uncertainty of membership status

Membership unclarities may arise both with regard to individual members of a category and with regard to sets of referents. In the former case, the question might be, for instance, whether referential type [B1122v] belongs to the category *colbert* or not. In the latter case, the question involves hyponymy when the set of referents itself constitutes a lexical category. The former type of referential unclarity need not bother us too much: as soon as the corpus shows that a particular type of referent is named by means of a spe-

cific item, we may conclude that the referent in question *eo ipso* belongs to the category represented by that item (except perhaps in the case of apparent mistakes and confusions). The referent may be a very peripheral member of the category, but it is a member nonetheless. In the hyponymous case, however, matters are more complicated. In some cases, deciding on the hyponymous status of one item with regard to another may be hampered by the absence of sufficient examples in the corpus; we encountered an example of this situation in the previous section when the relationship between *legging* and *legging-broek* was discussed. As opposed to this type of statistical uncertainty, we will try to show that there exist unclarities about hyponymous relations that are an actual feature of the real language situation. That is to say, it can be maintained that some potentially hyponymous relations are indeterminate in the language users' mental lexicon itself.

To begin with, let us establish that *a semantic analysis alone cannot establish hyponymy*. Consider an example in which *blurg* is a potential hyperonym, characterized by the combination of features ABC, and in which *plurk* is a potential hyponym, characterized by the combination ABCD. If *plurk* is not a hyponym of *blurg*, defining *blurg* as either A, AB or ABC does not yield a classical definition, because the definition then wrongly includes all plurks. Conversely, defining *blurg* as ABC automatically turns *plurk* into a hyponym (given that we take ABCD to be the definition of *plurk*). There is an obvious circularity here: on the one hand, how *blurg* should be defined (in particular, whether it can be defined in a classical way) depends on the hyponymy of *plurk*, but on the other hand, the hyponymy of *plurk* depends on the definition of *blurg* that is chosen as a startingpoint. Preconceived ideas about hyponymy (in fact, the very definition of hyponymy in terms of semantic inclusion) would seem to suggest that the relation between *plurk* and *blurg* is by definition hyponymous. However, it could very well be the case that *blurg* combines all ABCs that are not Ds; within the set of ABCs, *blurg* and *plurk* are then complementary rather than hierarchically ordered as a hyperonymous/hyponymous pair.

What we need, therefore, in order to settle the issue of classical definability, is indubitable evidence to prove that a particular type of referent *never* belongs to the category to be defined: *plurk* and *blurg* are complementary categories if plurks are never called blurgs, and vice versa. Such evidence might be adduced by asking people whether any particular plurk is an example of the category *blurg*; if the answer is invariably negative, a maximum degree of certainty is achieved. In a corpus of non-elicited material, on the other hand, hyponymy obtains if, for instance, all referents that

are at one time designated with the item *plurk* also occur with the name *blurg* (but not vice versa). Although we would thus seem to have an operational test for hyponymy, a major difficulty now has to be taken into account. The categorial judgements that we would like to rely on need not yield clear-cut results, either because (in the case where we rely on informants' judgements) the informants hesitate or disagree among each other, or because (in the case where we rely on a corpus) the referential range of the potential hyponym is only partially covered by the potential hyperonym. In other words, the categorial judgements in question may be non-dichotomous: the hyponymous status of a particular category with regard to another one may be a matter of degree. In such a case, we have a definitional problem because the hyponymy question cannot be decided univocally. If there simply is no clear answer with regard to the question whether a particular category constitutes a subset of another category, the issue of classical definability may turn out to be similarly undecidable. Or rather, a category may turn out to be difficult to define, simply because the referential range of the category is unclear.

This situation can be illustrated on the basis of the relationship between *broek* 'trousers', *rok* 'skirt', and *broekrok* 'culottes, pantskirt, divided skirt'. The main point of the discussion is double: first, to demonstrate how difficult it may be to prove a hyponymous relationship between categories, and second, to spell out the definitional consequences of the undecidability. Initially, we may define *broek* as a "two-legged outer garment covering the lower part of the body from the waist down", and *rok* as an "outer garment for women covering the lower part of the body from the waist down, with no separate coverage of the legs". Depending on the taxonomical relationship between *broekrok*, *rok*, and *broek*, however, the definitions may have to be refined. Systematically, there are four situations to be considered: *broekrok* is a hyponym of *broek* but not of *rok*, *broekrok* is a hyponym of *rok* but not of *broek*, *broekrok* is a hyponym of both *rok* and *broek*, and *broekrok* has a separate status, being a hyponym of neither *rok* nor *broek*. In the following overview, the definitions of the three terms are given in such a way as to maintain classical definability.

(1) *Broekrok* is a hyponym of *broek* but not of *rok*

– *Broek*: two-legged outer garment covering the lower part of the body from the waist down

– *Rok*: outer garment for women covering the lower part of the body from the waist down, with no separate coverage of the legs

– *Broekrok*: two-legged outer garment for women covering the lower part of the body from the waist down, with legs that are so wide that the impression is created that there is no division between the legs; *broek* that looks like a *rok*

(2) *Broekrok* is a hyponym of *rok* but not of *broek*

– *Broek*: two-legged outer garment covering the lower part of the body from the waist down, with legs that are not so wide as to create the impression that there is no division between the legs

– *Rok*: outer garment for women covering the lower part of the body from the waist down, such that no division between the legs is visible

– *Broekrok*: two-legged outer garment for women covering the lower part of the body from the waist down, with legs that are so wide that the impression is created that there is no division between the legs

(3) *Broekrok* is a hyponym of both *rok* and *broek*

– *Broek*: two-legged outer garment covering the lower part of the body from the waist down

– *Rok*: outer garment for women covering the lower part of the body from the waist down, such that no division between the legs is visible

– *Broekrok*: two-legged outer garment for women covering the lower part of the body from the waist down, with legs that are so wide that the impression is created that there is no division between the legs

(4) *Broekrok* is a hyponym of neither *rok* nor *broek*

– *Broek*: two-legged outer garment covering the lower part of the body from the waist down, with legs that are not so wide as to create the impression that there is no division between the legs

– *Rok*: outer garment for women covering the lower part of the body from the waist down, with no separate coverage of the legs

– *Broekrok*: two-legged outer garment for women covering the lower part of the body from the waist down, with legs that are so wide that the impression is created that there is no division between the legs

In the first configuration, culottes are taken to be a kind of trousers. In order to correspond with the traditional semantic definition of hyponymy, *broekrok* will have to be defined in such a way that it is a specification of the more general definition that holds for *broek*. This can be achieved by adding to the definition of *broek* that culottes are worn only by women, and

that they look like skirts: their legs are so wide as to obliterate the impression that they are separate legs. In the second configuration, *broekrok* sides with *rok* rather than with *broek*, and the definition of *rok* will have to be so wide as to include culottes. Obviously, this cannot be achieved by referring to the objective presence or absence of a division separating the legs: if skirts are defined as lacking separate legs, culottes would not be skirts. The alternative is to define *rok* on the basis of a visual image: regardless of the actual presence or absence of separate legs, the things that may be called *rok* (including culottes) generally create the impression that there is no such division. The third configuration is a straightforward combination of the previous two. The fourth combination, finally, presupposes that *broekrok* is a category with a separate status on the same level as *broek* and *rok*. *Broekrok* and *rok* can then be distinguished by the fact that the former has separate legs, whereas the latter does not. *Broek* and *broekrok*, on the other hand, are distinguished by the fact that the latter looks like a skirt, whereas the former has legs that are not so wide as to create the impression that they are not there.

These definitions, however, only pertain to the prototypical cases of the various categories. It is, for instance, still an empirical question to be settled independently whether the legs of the referents of *broek* are indeed never so wide as to make the garment look like a skirt. This is undoubtedly the case for typical trousers, but does it hold for all of them? In a similar way, do all culottes (that is to say, all referents of *broekrok*) actually create the impression of being skirts? An inspection of the pictorial material on which the database is based soon reveals that this is not the case: some culottes are not so wide that their legged nature always remains hidden. This observation, then, calls for a revision of the definitions. Let us suppose that the distinction between, for instance, the category *broek* and the category *broekrok* were to be described in terms of the different ranges that they allow on the dimensions WIDTH and LENGTH, in the following way:

- *Broek*: two-legged outer garment covering the lower part of the body from the waist down, with legs ranging in width from tight-fitting to very wide, and ranging in length from the groin to the ankle
- *Broekrok*: two-legged outer garment for women covering the lower part of the body from the waist down, with legs ranging in width between loose-fitting and very wide, and ranging in length between the thighs and the ankles

Regardless of the definition of *rok* that would complete the set, these definitions are only compatible with those taxonomical configurations in which *broekrok* is a hyponym of *broek*. Or, to be more precise, they are only incompatible with the other two situations if it is not accepted that *broek* and *broekrok* cannot be classically defined. If these definitions are descriptively adequate, but if *broekrok* is not a hyponym of *broek*, then *broekrok* and *broek* cannot be classically defined in such a way that their taxonomical distinctness is captured by the definitions. This shows, in other words, that the intensional issue of classical definability may depend on the extensional issue of membership status.

But how then can we answer the question what exactly the taxonomical relationship between *broekrok*, *broek*, and *rok* might be? There are various kinds of support for the view that the relationship is an extremely unclear one. In general, three types of evidence might be considered: the formal structure of the item, the distribution of *broekrok*, *rok*, and *broek* in the corpus, and native speakers' intuitions. We will now demonstrate that none of these is sufficient to settle the matter.

To begin with, let us note that the morphological structure of *broekrok* as a specificational compound with *rok* as its formal head, does not sufficiently justify the conclusion that things called *broekrok* are instances of the category *rok*: jellyfish is not a kind of fish, and similar examples of exocentric compounds are not difficult to find. The morphological structure does not establish the hyponymy, but rather, the presence or absence of a hyponymous relationship determines whether we are dealing with an ordinary specificational structure or not. In addition, it may be noted that an analysis of *broekrok* as a copulative compound (in which case *broekrok* would be a hyponym of both *broek* and *rok*) is precluded for formal reasons. Whereas the members of Dutch copulative compounds (like *priester-dichter* 'priest-poet' and *hotel-restaurant*) typically retain an independent stress pattern, *broekrok* has a unified stress pattern with main stress on the first syllable.

In the second place, let us try to have the corpus decide the question, following the procedure that was introduced in the previous section. A hyponymous relationship between, for instance, *broekrok* and *rok* would imply that the members of the *broekrok* category could in principle also be called *rok*, i.e., that they would also occur with the name *rok* in our corpus. If the same kind of referent that is called *broekrok* can also be called *rok*, the latter name is likely to show up in our corpus. To be sure, there is no reason to suppose that it will occur just as often as *broekrok*. This has something to do with the entrenchment effects that we introduced in chapter 1 and that

will be discussed in more detail in section 4.2.: if culottes constitute a well-entrenched category, they will rather be named as a category of their own (i.e. with a name like *broekrok*) than as a member of a superordinate category like *rok* that contains many other types of skirts besides culottes. Still, the superordinate category is not likely to be completely absent from the corpus. To mention only one example: while *jeans* is both an intuitively indubitable hyponym of *broek* and a well-established category in its own right, we find that of the 109 prototypical members of the *jeans*-category that occur in the corpus, 13 bear the name *broek*.

Configuration	<i>broek</i>	<i>broekrok</i>	<i>broek</i> in %
2,3	2	11	15
2,4	0	7	0
3,3	1	18	5
4,3	7	24	22
4,4	0	10	0
5,3	28	6	82

Figure 3.3(1)

Distribution of *broek*-denominations over the semasiological range of *broekrok*

In this respect, we first have to note the absence of *rok*-denominations for culottes. This implies that *rok* is not a likely hyperonym for *broekrok*. Apparently, pantskirts are not spontaneously categorized as skirts – in spite of the fact that the formal structure of the compound *broekrok* seems to suggest the opposite. The item *broek* does, on the other hand, occur as an alternative name for the garments that fall into the *broekrok*-category: we have encountered 38 instances of *broek* for things that are indubitably culottes. The distribution of the *broek*-denominations over the examples of culottes is given in Figure 3.3(1). The figure has been composed in the following way. First, a survey of the various referents of the *broekrok*-category that occur in the corpus is used to determine the major types of variation that *broekrok*-referents exhibit. Specifically, a classification is imposed on the material in terms of the length and the width of the garments, because these two charac-

teristics are precisely the ones that may play a role in the definition of *broekrok*. For each of the six categories, the frequency with which they occur in the referential range of *broekrok* is indicated in the third column of the figure. In the second column, the frequencies with which members of the six categories are called *broek* is given. The fourth column specifies the frequency of *broek* as a percentage of the sum of *broek*- and *broekrok*-examples.

The hypothesis that the distribution of *broek* over the referents of *broekrok* reflects the hyponymous relationship of the latter with regard to the former could be corroborated by the presence of two different patterns. On the one hand, there could be an even distribution of *broek* over the main referential subtypes of the *broekrok*-category. This is a situation that could be considered the ideal reflection of hyponymy in a corpus: all major types of the potential hyponym can be named by the potential hyperonym, and each type receives the hyperonym with roughly the same relative frequency. On the other hand, there might be an asymmetrical distribution, to the extent that the less frequent examples of the potential hyponym would receive the potential hyperonym more often as an alternative term. This situation would not be surprising from the point of view of a prototype-oriented theory of categorization: if the more peripheral instances of a category are by definition the ones that deviate from the central tendencies of the category, they are also more likely to be named by another term (if an adequate one is available – but obviously, a hyperonym is such an adequate term). As Figure 3.3(1) shows, the actual distribution of *broek* over the referents of *broekrok* is characterized by neither of these patterns: the *broek*-percentages are neither roughly uniform, nor are they straightforwardly inversely correlated with the frequency of the various *broekrok*-types. In this respect, the hypothesis that *broek* acts as a hyperonym of *broekrok* seems to be disconfirmed by the fact that the distribution of *broek* over the *broekrok*-referents does not correspond to the expected pattern (or, to be more precise, to neither of the two patterns that are compatible with the hyperonymous status of *broek*).

There is, however, still another hypothesis to be considered. The highest *broek*-rate in 3.3(1) involves what might be called the “standard”, prototypical type of trousers: long legs down to the ankles, not too wide but neither tight-fitting. Suppose, then, that the distribution of *broek* over the *broekrok*-referents is determined by the extent to which the *broekrok*-referent in question conforms to the prototypical type of trousers. In that case, rather than centre-periphery effects in *broekrok* itself, it is the central tendencies of

broek that would determine the choice of the alternative name. Figure 3.3(2) demonstrates that the hypothesis is correct. According to the hypothesis (and given the fact that the [length 5, width 3] configuration embodies the central tendency of *broek*), the *broek*-percentages are expected to diminish from the top to the bottom of the figure, and from the left to the right. Because the general distributional pattern clearly corresponds to the expectations, we may conclude that rather than dissimilarity with regard to the *broekrok*-prototype (as in the hypothesis considered above), it is similarity with regard to the *broek*-prototype that is the main factor in explaining the *broek*-distribution.

	length 5	length 4	length 3	length 2
width 3	82%	22%	5%	15%
width 4	—	0%	—	0%

Figure 3.3(2)

Frequency of *broek*-denominations in the semasiological matrix of *broekrok*

We might conclude from this observation that *broek* is being used less as a hyperonym of *broekrok* than as a concept that is situated on the same level, and that partially overlaps with it. Although this would seem to settle the issue of the relationship between *broekrok* and *broek* in favor of the situation in which *broekrok* has a separate status, it is important to note that this is a misleading way of rendering the situation. If a distinction is maintained between *broek* in its prototypical reading (the reading whose importance we can establish in connection with *broekrok*) and *broek* in the broader, hyperonymous reading that we were primarily interested in, we may note that the distribution of *broek* over the *broekrok*-range is primarily determined by the prototypical reading of *broek*, but that does not give us sufficient information about the larger, hyperonymous reading of *broek* that is our primary concern. The evidence, then, is inconclusive: the distributional data that we encounter in the corpus so to speak deal with a different reading of *broek*

than the one we are interested in.

The same inconclusiveness comes to the fore when a third kind of evidence is considered, viz. native speakers' intuitions. Simply asking people whether a *broekrok* is a *broek* or a *rok* leads to widely divergent answers. Most conspicuously, the answer is seldom given immediately, but only comes after a period of hesitation, a request for further clarification, or a confession of uncertainty. Apparently, the question itself is an unexpected one; people do not have a conception of the taxonomic status of *broekrok* that is firmly entrenched in their mental lexicon (in the way in which, for instance, the hyperonymy of *vehicle* with regard to *car* is well established). Rather, it seems as if people begin to think about an answer only when the question is asked. The position of *broekrok* in the taxonomy of garments is not, it seems, given in advance, but has to be computed on the spot; moreover, the results of the computation are far from uniform.

On an anecdotal level, other observations point in a similar direction. In the conservative Protestant community of Doornspijk in The Netherlands, a vigorous discussion took place in the course of 1991 about the question whether women were allowed to wear pantskirts in church. The crucial text is Deuteronomy 22:5 ("A woman shall not wear an article proper to a man, nor shall a man put on a woman's dress"), but the whole discussion obviously involved the question whether culottes are trousers (men's wear) or skirts. Here is how the church council motivated its point of view (reproduced from an article in the journal *De Telegraaf*):

Het boosaardig hart heeft een vrouwelijke broek bedacht in de vorm van een broekrok. De ene keer lijkt het meer op een rok en de andere keer meer op een broek, maar wij behoren ons te onthouden van alle schijn des kwaads. De man zal zich als man kleden, en de vrouw zal vrouwenkleding dragen, niet alleen op de Dag des Heeren, maar ook op school en in het gezin.

[The evil heart has invented female trousers in the form of culottes. Sometimes, they look rather like a skirt, and at other times, rather like a pair of trousers, but we should avoid all semblance of evil. A man shall be dressed as a man, and a woman shall wear women's clothes, not just on the Day of the Lord, but also at school and at home.]

In the end, the conservative views of the minister and the church council won out, but the very fact that the discussion took place reveals that the hyponymous status of *broekrok* with regard to *broek* is not as obvious as that of,

say, *stallion* with regard to *horse* or *house* with regard to *building*.

The same conclusion was reached on the basis of a small survey. We asked 256 first year law students to choose between statements to the effect that culottes are a kind of trousers, a kind of skirt, or a category in its own right. The students were informed beforehand that our interest lay in taxonomical relations of the kind that exist between concepts like *stallion* and *horse*. Figure 3.3(3) contains an overview of the results. (Note that the "n"-column specifies the absolute number of replies.) Although there is a majority for an independent status of *broekrok* (which in itself may be a sign of uncertainty), the fundamental fact is the heterogeneous nature of the results. There is no clearly dominant answer of the kind that would appear in cases like *stallion/horse*; there is no agreement about the taxonomical status of *broekrok* in the way there would be about *stallion* or *mare* with regard to *horse*.

	Male		Female		Total	
	n	%	n	%	n	%
trousers	27	23,9	33	23,1	60	23,4
skirt	24	20,	22	16,1	46	18
a category of its own	55	48,7	86	60,1	141	55,1
no opinion	8	7,1	1	0,7	9	3,5
total	114	100	142	100	256	100

Figure 3.3(3)

A survey of taxonomical judgements about *broekrok*

All in all, we believe that there is sufficient evidence to suggest that the taxonomical status of *broekrok* (in particular, its hyponymous relationship

with regard to *broek* or *rok*) is far from clear. Language users apparently do not possess a stable, clear-cut idea of the relationship between *broekrok* and *broek*, in contrast with the fixed taxonomical relationship that exists between *lion* and *animal* or *car* and *vehicle*. The crucial point, we ought to emphasize, is the absence of a *stable* taxonomical relationship. We do not claim that people cannot assign a taxonomical status to *broekrok* when asked for it, but rather that such an assignment is not a permanent aspect of their lexical knowledge (in the way that it would be for *lion* with regard to *animal* or *car* with regard to *vehicle*). Native speakers of Dutch do not have a readily available conception of the taxonomic status of *broekrok* that is firmly entrenched in their mental lexicon, but rather begin to think about an answer only when the question is asked. Moreover, the results of such an *ad hoc* search for an answer suggest that the actual taxonomical statements that people come up with are quite flexible: there is an outspoken tendency to think of *broekrok* as a category of its own, but an allocation to the domains of *broek* or *rok* is far from excluded. Ultimately, this point illustrates the *flexibility* of categorization in general: depending on the perspective taken, the referents of *broekrok* may be categorized in different ways, but there is, taxonomically speaking, no single dominant perspective of the kind one would expect when starting from the taxonomical model provided by such clear cases as *stallion/horse* or *lion/animal*. (In this respect, it would be interesting to investigate in further research what contextual factors might prime categorizing culottes as instances of *broek* or *rok*.)

There are three major implications to be retained from this observation. First, it establishes that *extensional non-discreteness* (in the sense of indeterminacy about category membership) does indeed play a role in the semasiological structure of the lexicon. Second, it shows that such extensional non-discreteness may influence the issue of *intensional non-discreteness*. As we discussed earlier in this section, whether *broekrok* is classically definable or not depends in part on the taxonomical relationship between *broekrok*, *broek*, and *rok*. But if that relationship itself is unstable, so is the definability issue. And third, the existence of taxonomical instability suggests that a purely *relational conception of lexical semantics* is based on an overly optimistic view of the nature of lexical relations. Such a relational conception of semantics is not an uncommon one, not in the least because it is a cornerstone of John Lyons's conception of lexical semantics:

I consider that the theory of meaning will be more solidly based if the meaning of a given linguistic unit is defined to be the set of (para-

digmatic) relations that the unit in question contracts with other units of the language ..., without any attempt being made to set up "contents" for these units (Lyons 1963: 59)

The question *What is the sense of x?* ... is methodologically reducible to a set of questions each of which is relational: *Does sense-relation R_i hold between x and y?* (Lyons 1968: 444).

Through well-known books like Lyons (1968) and (1977), this relational conception of structural semantics became very influential; Cruse's textbook on lexical semantics (1986), for instance, is largely devoted to a detailed investigation of the various "sense-relations" (like hyponymy, antonymy, and synonymy) that constitute the basic semantic links between lexical items.

In itself, the indeterminacy surrounding the relationship between *broekrok* and *broek* does not establish that the relational approach is misguided; rather, it merely indicates that the approach should not assume that all sense-relations are necessarily well-defined. This conclusion itself is quite important against the background of the history of lexical semantics. Lyons's relational approach crucially refers to a distinction between "sense" and "reference". The latter involves the relationship between words and the extra-linguistic entities (things, events, actions, qualities etc.) that they stand for, while the former indicates "its place in a system of relationships which it contracts with other words in the vocabulary" (1968:427). In line with the central tenets of linguistic structuralism, the proper focus of semantics as a linguistic enterprise is on sense rather than reference, because it is precisely such a system of relationships that constitutes the *structure* of the language.

From such a point of view, the entire cognitive, prototype-oriented conception of word meaning might easily be interpreted as based on an unacceptable confusion of sense and reference. In particular, the non-discreteness effects involving category membership would seem to involve the indeterminacy of reference rather than the concept of sense. Lyons, in fact, explicitly accepts the existence of "indeterminacy of reference" (1968: 412), but since the linguistically crucial phenomena involve sense-structures rather than referential usages, referential vagaries need not undermine the ideal of a classically well-defined description of the sense of lexical items. Or, to put it informally, reference may be fuzzy, but sense is neat. The *broekrok*-example, however, shows that sense-relations may be subject to indeterminacy just like referential relations: the taxonomical sense-relations between *broekrok* and *broek* appear to be insufficiently determinate to

establish whether there is a relationship of hyponymy between both items. As such, there is no reason to suppose that a classically discrete form of semantic analysis can be maintained by restricting linguistic semantics to the notion of "sense".

3.4. Non-equality of word meanings: salience effects

In the previous sections, we have discussed the extensional boundaries of lexical categories (involving questions of fuzzy membership) and their intensional boundaries (involving questions of definability). We will now have to look more closely into the categories themselves, discussing the question whether they have an internal structure in terms of differences in salience between various members or groups of members of the category. Against the background of the classification of prototypicality effects presented in section 3.1., we will now deal with prototypicality in terms of salience rather than fuzziness. We will argue that this type of prototypicality is indeed widespread among the lexical items that we are dealing with.

As a first step, let us note that some features (i.e., some dimensional values) are more salient within a category than others. The values on the componential dimensions are not uniformly distributed over the referents of the category. This can be easily illustrated when we have a look at the various dimensions that are relevant for the structure of *legging*. Leaving out of consideration all records with incomplete componential configurations (see section 3.2.), we retain precisely 110 examples of *legging*. On the dimension LENGTH, value [3] occurs three times, value [4] forty times, and value [5] sixty-seven times. On the dimension WIDTH, value [1] occurs one hundred and four times, while value [2] is present only two times. And on the dimension END OF LEGS, value [1] can be found one hundred and eight times, whereas [2] and [5] each occur in only one record. In all of these cases, then, there is a marked skewing of the frequency distribution of the dimensional values.

A second step involves switching to an extensional mode of enquiry, taking into account the frequency of occurrence of the various types of referents that occur in the referential range of the category. Following the representational format of Figure 3.1(2) (see also Geeraerts 1989 for more examples of this kind of representation), Figure 3.4(1) charts the structure of

legging. The boxes indicate the various features that seem relevant in the structure of the item. Each box represents a specific feature; it contains the referential configurations that exhibit the feature represented by the box, together with the absolute frequency with which that configuration occurs in the range of *legging*. The configurations are simplified in the sense that features that are irrelevant for the internal structure of the category have

a Reaching down to the ankles or the calves

b Tight-fitting

c Made of elastic material

d Without fastening on the end of the legs

Figure 3.4(1)

The semasiological structure of *legging*

been left out. For instance, the [v]-feature is not mentioned, because all leggings are worn by women. The figure establishes that there is a *correlation between intensional and extensional salience*: the salient intensional

elements (the descriptive attributes with the highest frequency) co-occur in the most salient extensional elements (the category members with the highest frequency). The category as a whole appears to be structured in terms of a maximally overlapping high frequency core region surrounded by a peripheral area with low extensional frequency and decreasing intensional overlapping.

The relationship between the salience of intensional features and the frequency of occurrence of members of a category is not new in the literature on prototypicality. It plays a major role, in fact, in Eleanor Rosch's original development of the prototypical model of categorization. In Rosch & Mervis (1975), for instance, the idea takes the form of the hypothesis

that the more an item has attributes in common with other members of the category, the more it will be considered a good and representative member of the category (1975: 582).

There are, of course, differences between the type of analysis presented here, and Rosch's original studies. Most importantly, there is a methodological distinction in the sense that Rosch used an experimental method whereas the present study is based on non-elicited material. The correlation between co-occurrence of attributes and membership frequency established by Rosch involves, on the one hand, a set of typicality ratings for the members of a category, and on the other, a calculation of the number of attributes shared by particular groups of members of the category (based on an experimental task in which subjects are asked to list attributes for the members of the category). It then appears, for instance, that the most typical members share a high number of attributes, whereas less typical members share less attributes. Translating this approach to the present study, the typicality measure corresponds with the frequency with which particular referents occur within a category. The co-occurrence of attributes, on the other hand, corresponds with the overlapping of intensionally delimited subsets in diagrams like Figure 3.4(1). The fact, then, that the definitionally central area of an item's field of application (i.e. the maximally overlapping area in figures like 3.4(1)) contains the highest percentage of category instances, may be likened to the Roschian correlation between co-occurrence of attributes and typicality.

At this point, the question arises whether the existence of such a correlation is the usual situation in the lexicon. How widespread is the phenomenon? Figures 3.4(2) to 3.4(9) present further examples of the same structural

Figure 3.4(2)
The semasiological structure of *colbert*

Figure 3.4(3)
Relative frequencies within the semasiological structure of
colbert

Figure 3.4(4)
The semasiological structure of *blazer*

Figure 3.4(5)

Relative frequencies within the semasiological structure of
blazer

phenomenon. (Still further examples can be found in section 3.5.) Figure 3.4(2) represents the semasiological structure of *colbert*; the figures between brackets indicate the frequency of the configurations in question. Figure 3.4(3) repeats the same structure, but specifies the relative frequencies per area in the figure. Figures 3.4(4) and 3.4(5) do the same thing for *blazer* as what 3.4(2) and 3.4(3) do for *colbert*.

The featural dimensions and values used in the description of C- and B-type of referents are as indicated in the following list. In general, allocation of a referent to either the C- or B-type depended on two factors. B-type garments are jackets of the formal type; they are invariably longer than the waist-line, reaching down to the hips, they always have revers and long sleeves, and they are never knitted. The C-types, on the other hand, refer to jacket- and cardigan-like garments that lack these characteristics. The intrusion of C-type garments into the semasiological realm of *colbert* and *blazer* is rather unexpected; introspectively, one would not expect knitted garments or jackets without sleeves to be called either *blazer* or *colbert*. (There are, actually, more of such surprises in the data. In general, the amount of variation found in the corpus surpasses by far what is intuitively expected.) Because the dimensions used for the componential description of the B-type garments do not coincide with those used for the description of the C-type garments, a specific difficulty came to the fore when drawing diagrams like those in Figures 3.4(2) and 3.4(4). For instance, because the dimension POSITION OF THE BUTTONS was selected for inclusion in the diagram but was only systematically recorded in the data base for those garments classified as B-types, all referents classified as C-types had to be reinvestigated to check the position of the buttons. (Traditionally, the position of the buttons is gender-related: women's garments have the buttons on the left panel, and men's garments on the right panel. It appears from the data, however, that the dimension is not redundant with regard to the sex of the person wearing the item under consideration, as represented by the [m] "male" and [v] "female" features in the componential configurations.)

B-type referents:

WIDTH

- [1] The garment is waisted.
- [2] The garment is loose fitting.

FASTENING

- [1] The jacket has a single-breasted button fastening.
- [2] The jacket has a double-breasted button fastening.

FABRIC

- [1] The jacket is made of a thin, creased or wrinkled fabric.
- [2] The jacket is made of a relatively thick and smooth fabric.
- [3] The jacket is made of a thick and coarse fabric.

POSITION OF THE BUTTONS

- [1] The buttons are located on the right front panel from the wearer's point of view.
- [2] The buttons are located on the left front panel from the wearer's point of view.

C-type referents:

LENGTH

- [1] The garment is not longer than the waist.
- [2] The garment is roughly as long as the waist.
- [3] The garment is longer than the waist.

FASTENING

- [1] The garment does not have a fastening; the panels cannot be attached to each other.
- [2] The garment has a zipper fastening.
- [3] The garment has a full, single-breasted button fastening.
- [4] The garment has a full, double-breasted button fastening.

MATERIAL

- [1] The garment is made of a relatively thick and smooth fabric.
- [2] The garment is made of coarsely knitted material.
- [3] The garment is made of finely knitted material.
- [4] The garment is made of a towelling-like material.

SLEEVES

- [1] The garment does not have sleeves.
- [2] The garment has long sleeves.

The distinction between *colbert* and *blazer* appears to reside in two points. First, their prototypical centres are very similar, except for the fact that *blazers* appear to be worn predominantly by women, while *colberts* are worn predominantly by men. The predominance of "male" garments in the structure of *colbert* is even more outspoken if one takes into account the relatively unrepresentative character of our database with regard to men's wear (to which we drew the attention in section 2.2.). Simplifying somewhat, one could say that *colbert* and *blazer* prototypically refer to exactly the same type of garment, except that the first item is used when the jacket is worn by a man, and the second when it is worn by a woman. This is a simplification, to be sure, because there is an "objective" feature on the basis

a As long as the calves or longer

b Straight

c Without fastening at the end of the legs

d Made of relatively light material

Figure 3.4(6)

The semasiological structure of *broek*

of which the prototypes differ, viz. the typical position of the buttons for men's wear and women's wear. A second feature distinguishing *colbert* and *blazer* involves the periphery rather than the core of the items: there is more variation in the periphery of *blazer* than in that of *colbert* (which may have something to do with the fact that clothing for women is generally more varied in appearance than that for men).

In the next set of examples, only the relative frequencies of the various areas in the diagrams have been indicated. While the structure of *t-shirt* is

- a Made of non-spongy, non-fluffy,
non-woolly material,
and having a full row of buttons
- b With long sleeves
- c Waist-long, or worn tucked in at the waist
- d With a stiff collar

Figure 3.4(7)
The semasiological structure of *overhemd*

very outspoken, the relationship between *hemd* and *overhemd* is more or less like that between *blazer* and *colbert*. The core area of both items is the same, consisting of standard shirts with long sleeves, a full set of buttons, a more or less stiff collar, and made of a fine material. There is some evidence (which we have not included in the figures) that – just as with *blazer* and *colbert* – sex plays a role in differentiating between both prototypes. Of all

- a Made of non-spongy, non-fluffy, non-woolly material
- b With long sleeves
- c Waist-long, or worn tucked in at the waist
- d With a full row of buttons
- e With a stiff collar

Figure 3.4(8)
The semasiological structure of *hemd*

the “male” referents in the range of *overhemd*, almost 94% is to be found in the central area, versus about 75% of the “female” referents; in this sense, it would seem that *overhemd* is, in its central application, a term that is more specifically used to refer to shirts worn by men. *Hemd*, on the other hand, is not outspokenly “female”, but rather neutral: 66.6% of the male and 68.7% of the female referents are found in the core area. Due to the underrepresentation of male clothing in the database as a whole, this distinction should be treated with some care. But even if it should not stand up to further scrutiny,

- a Without collar
- b With only a partial or no fastening
- c Made of non-spongy, non-fluffy, non-woolly material
- d Without buttons or other type of fastening
- e At least waist-long or worn tucked in
- f With short sleeves

Figure 3.4(9)

The semasiological structure of *t-shirt*

another difference between *hemd* and *overhemd* is quite obvious in Figures 3.4(7) and 3.4(8): like *blazer* with regard to *colbert*, *hemd* has a larger range of application, with more internal variation, than *overhemd*. The *broek*-example, finally, involves what is to some extent a case of auto-hyponymy: while the item serves as a superordinate term for any two-legged

garment covering the lower part of the body (in the sense in which leggings or bermudas or jeans are all trousers), the central area of *broek* consists of what is, loosely speaking, a hyponym of that larger reading (viz. the regular long pair of trousers). The gradual transitions between the core area and the periphery preclude treating both kinds of usage (the "hyperonymous" and the "hyponymous" one) as clearly separate meanings of the item *broek*.

Taken together, the examples of semasiological structures presented above inspire the following two major *conclusions*. First, the correlation between extensional and intensional salience effects appears to be quite common in the lexical field of clothing terms. Even though the sample of ten items presented here and in the following section is not very large, the fact that the items under investigation belong to different classes strongly suggests that the coupling of intensional and extensional salience is a pervasive structural characteristic of lexical items. Notice, for one thing, that the phenomenon occurs both with classically definable terms like *legging*, and with items like *colbert* and *blazer*, that cannot receive a classical definition according to the argumentation of section 3.2.. This observation establishes that the phenomenon is not just a side-effect of the definitional structure of an item (and more specifically, of the absence of a classical definition), but may in fact appear as a phenomenon in its own right. For another, notice that intensional and extensional salience phenomena affect both a subordinate term like *legging* and its superordinate term *broek*.

As a second major conclusion, let us notice that the correlated salience structure of the items allows for many variations. There is, for instance, a lot of diversity in the depth of the salience structure, in the sense that the relative frequency of the prototypical centre may be greater in some items than in others. The structure of *blazer*, for instance, is "flatter" than that of *colbert*: the core area of the former word represents 24.7% of all cases, whereas that of the latter item represents 46.1%. Also, the prototypical centre may be clearly confined to one central area of the diagram (as with *hemd* or *legging*), or it may extend over more than one square in the figure (as with *t-shirt*). At the same time, the transition towards the periphery of the category may be very gradual in terms of frequency, as in the case of *blazer*, or it may be very abrupt, as in the case of *overhemd*.

Although the examples presented in this section demonstrate the pervasiveness of intensional and extensional salience effects in the semasiological structure of lexical items, a conservative approach to lexical semantics might claim that these phenomena are not of crucial importance for the linguistic study of the lexicon. From a traditional structuralist point of

view, the crucial point in lexical studies is the structure of the lexicon, conceived of as the sum of all mutual delimitations between lexical items; what is structurally important is the way in which lexical items mutually delimit each other. Within such a perspective, it might be argued that the internal salience structure of lexical categories is irrelevant for a linguistic approach to the lexicon. The fact, for instance, that various referents of a category occur with different frequencies (and as such, exhibit varying degrees of salience) might be considered a question of language use rather than linguistic structure (or, some might say, a question of pragmatics rather than semantics). In this respect, it may be useful to round off this section by making clear that there are three major *reasons for considering the internal make-up of categories a relevant aspect of linguistic lexicological structure* rather than just a usage-based, pragmatic side-effect.

In addition to the three points mentioned below, it might be suggested that psycholinguistic experiments like Rosch's establish the cognitive, psychological reality of the prototypicality effects. This is undoubtedly the case, but using such a recognition as an argument in favor of an incorporation of prototypicality into linguistic studies smacks of circularity. The argument presupposes that linguists agree that such psychological data should be described and explained within the linguistic study of the lexicon, whereas the invocation of the distinction between language structure and language use precisely tends towards the position that there is a proper linguistic way of studying the lexicon that is different from the psycholinguistic approach. In traditional structuralist terms, focusing on the systematic structure of the lexicon establishes the methodological autonomy of linguistic lexical semantics: the linguistic study of the lexicon has its own subject matter, defined as the structural relations that delimit the vocabulary items with regard to each other, and its own methodology, defined as the description of those relations. It is therefore important to find reasons for incorporating semasiological salience effects into lexical semantics that may appeal to proponents of the structuralist methodology within their own framework. The following three arguments, then, seem incontrovertible even for traditionally minded structuralist semanticists.

In the first place, the absence of classical definability turns the internal structure of lexical items into a structurally relevant feature. For items that cannot be described by means of a single definition in terms of necessary-and-sufficient attributes, family resemblance structures of the kind illustrated by the figures in this section replace classical definitions. To the extent that definitions are crucial for the structuralist conception of the lexicon

as a system of elements delimiting each other's range of application, non-classical definability will simply have to be accepted as a structural feature of the lexicon. And the non-classically definable items precisely take the form of overlapping subsets clustered round prototypical core regions.

In the second place, the internal structure of the words is necessary to describe their mutual differences in a proper way. In the *colbert/blazer*-example, for instance, merely staying at the outer boundaries of the items leads to the conclusion that *blazer* is a hyperonym of *colbert*, since all the possible features of referents named *colbert* are also possible in the range of application of *blazer*, but not the other way round. (See section 3.2. for a discussion along these lines.) Such an approach does not, however, get down to one crucial difference between both items, viz. the fact that *colbert*-referents are typically jackets for men, and *blazer*-referents typically jackets for women. This conclusion would probably be reached when the structural approach were to be restricted to the mutual delimitation of the prototypical centres of the items rather than their ranges of application as a whole, but then, of course, the second crucial difference (the greater variability of *blazer*) tends to be lost out of sight. In short, descriptive adequacy of the type aspired at by structural semantics requires that the entire semasiological structure of an item (the prototypical centre together with the outer boundaries) be taken into account.

In the third place, the internal prototypical structure of lexical items has important formal consequences, as will be demonstrated in section 5.1.. On the one hand, it can be shown that the choice of the modifiers accompanying the head of the noun phrase is influenced by the semasiological structure of the item: there is an inverse correlation between the salience of a particular feature for the semantics of a particular lexical item, and the frequency with which that feature is expressed as a modifying element in a noun phrase with the items as its head. On the other hand, the choice of the head word itself is influenced by the salience effects: it can be shown that referents are preferentially named by means of a category to which they typically belong. The point about these observations is that formal considerations tend to play a major role in autonomistic approaches to language. The Saussurean doctrine of the unity of the linguistic sign (which is the ultimate cornerstone of structural semantics) implies that the elements of the linguistic system have to be described as pairings of meaning and form. Specifically, methodological license in the field of semantics has to be avoided by restricting the description to those aspects of meaning that correlate with formal phenomena. But if, consequently and conversely, formal phenomena can be invoked to

establish the legitimacy of incorporating a particular type of semantic phenomenon into the description, the data to be presented in section 5.1. demonstrate the linguistic importance of the prototypical salience effects described in the present section.

3.5. The influence of contextual variation

Among the twenty-five most frequent items in the corpus that we have singled out for closer scrutiny, only *vest* and *shirt* exhibit significant contextual variation, the former along the geographical dimension, and the latter along the specialization dimension. With regard to *vest*, there is a distinction between the way in which *vest* is used in the Belgian sources and the way in which it is used in the Netherlandic sources: The kind of graphical representation that was used in the previous section for the totality of the materials that are available for one particular item, can also be applied to subsets of the material defined in terms of groups of magazines sharing a specific variable. In Figures 3.5(1) and 3.5(2), for instance, the available data for *vest* are distributed over the geographical dimension. Figure 3.5(1) presents the analysis of the Netherlandic material, i.e., the instances of *vest* as found in the magazines *Burda*, *Knip*, *Margriet*, *Libelle*_{nl}, and the glossies. Figure 3.5(2) presents the Belgian material, as found in *Libelle*_b, *Flair*, and *Feeling*. Figure 3.5(3) collapses both figures, specifying the relative frequencies for each set of sources of the various areas in the figure.

The major distinction between both subsets resides in the fact that *vest* in the Belgian sources contains references to A- and B-type garments which are entirely absent in the Netherlandic sources. This finding corresponds with the observation that in substandard and dialect varieties of Belgian Dutch, *vest* is the most widespread term for jackets; it is, for instance, a standard entry in the many normative phrase books and purity of language reference works that are being produced in Flanders.

It should be noted that the distinction between the frequencies recorded in Figure 3.5(1) and those recorded in Figure 3.5(2) can be proven to be statistically significant, whereas the distinction between, for instance, the corresponding frequencies for the specialized sources (the fashion magazines *Burda* and *Knip*) and the general magazines is not significant. (As the fashion magazines belong to the group of Netherlandic sources, the comparison

		knitted	
waist-long			B1222(1) B2222(1) A214(1) A484(1)
	C2442(1) C2112(2) C2242(1)	C3432(1) C3422(2) C1222(1)	B2232(3)
	C2232(1) C2132(1) C2122(2)	C3322(11) C3332(7) C1322(2) C1332(1)	C3331(1) B1132(3) B2131(1)
	C2331(2)	C2322(19) C2332(15)	C3311(2) B1121(1) B2122(3)
one row of buttons	C2311(1)	C3342(1) C3312(3)	C3211(1) C3111(1)
		long sleeves	

Figure 3.5(2)

Frequencies within the semasiological range of *vest* in the Belgian sources *Libelle*_b, *Flair*, and *Feeling*

should obviously be restricted to the Netherlandic general magazines, in order to avoid an interference from the geographical dimension.) This statistical approach can obviously be generalized, by comparing various subsets among the sources used in the study. Figure 3.5(4) lists the frequencies for the relevant dimensions in the structure of *vest* as they appear in various sets of sources.

source group	1	2	3	4	5	6
feature						
long-sleeved	75	120	29	81	53	51
waist-long	49	74	16	49	29	28
knitted	73	94	25	59	48	40
one-row fastening	77	117	28	82	49	52
non-C-type	15	1	0	1	10	1

Figure 3.5(4)

Frequencies within the semasiological range of *vest*

Systematically, Figure 3.5(4) maintains a distinction between the following source groups (see section 2.1. and Figure 2.1(1) for more information about the background of these distinctions):

- group 1: magazines that are produced in Belgium: *Flair*, *Feeling*, *Libelle*_b
- group 2: magazines that are produced in The Netherlands: *Margriet*, *Libelle*_{nl}, *Burda*, *Knip*, *Avenue*, *Esquire*, *Avantgarde*, *Cosmopolitan*, *Man*
- group 3: fashion magazines produced in the The Netherlands: *Burda*, *Knip*
- group 4: general women's magazines produced in The Netherlands: *Margriet*, *Libelle*_{nl}
- group 5: general women's magazines distributed exclusively in Bel-

gium: *Libelle*_b

– group 6: general women's magazines distributed exclusively in the Netherlands: *Libelle*_{nl}.

The impact of the geographical dimension can be measured in three different ways: by comparing all journals produced in Belgium with all journals produced in The Netherlands (regardless of whether they are fashion magazines, women's magazines, or glossies); by comparing all journals produced in Belgium (which are all general women's magazines) with all women's magazines produced in The Netherlands; and by comparing the journals that are exclusively distributed in Belgium with those that are exclusively distributed in The Netherlands. Along the specialization dimension, the Netherlandic fashion magazines can be compared with the Netherlandic women's magazines. (A comparison between the women's magazines and the glossies has not been made, as the absolute frequencies in the latter group appeared to be too low for statistical safety.)

The statistical measure we have used to determine the variation is a standard χ^2 -test. For readers who are less familiar with statistics, it may be noted that a χ^2 -test, roughly, provides an answer to the following question: given that there is a difference between two samples a and b, can we conclude that the samples are taken from different populations? For instance, we may note that the figures we find for *vest* in our database are different in the Netherlandic sources as compared with the Belgian sources. Those figures constitute a sample: we have not been able to collect all instances of *vest* in the relevant population, if only because the clothing terms that we have recorded were restricted to those accompanied by photographs. So how can we know that the differences in the sample reflect differences in the actual populations, rather than just being due to chance? The χ^2 -test provides an answer to the question by indicating a significance level for the attested differences. When, for instance, the test yields the conclusion that samples a and b differ significantly at the 0.1 level, this means that we can be 90% certain that the differences in the samples reflect actual differences in the populations from which the samples are taken. The next step, of course, concerns the question what level of significance is required. If it is found that you can be 90% certain about the differences, is that sufficient, or should a higher level of significance be attained? The degree of certainty that is commonly accepted as the minimum level in linguistic studies is set at 95%, i.e. a significance level of 0.05 on the χ^2 -test.

Here are the results of the χ^2 -tests, in the order just described. First, let

us compare all journals produced in Belgium with all journals produced in The Netherlands (source group 1 versus source group 2). The distinction between *Flair*, *Feeling*, *Libelle_b* and *Burda*, *Knip*, *Margriet*, *Libelle_n*, *Burda*, *Knip*, *Avenue*, *Esquire*, *Avantgarde*, *Cosmopolitan*, *Man* is significant at the 0.01 level. Second, let us compare all general women's magazines produced in Belgium with all general women's magazines produced in The Netherlands (source group 1 versus source group 4). The distinction between *Flair*, *Feeling*, *Libelle_b* and *Margriet*, *Libelle_n* is significant at the 0.01 level. Third, let us compare the journals that are exclusively distributed in Belgium with those that are exclusively distributed in The Netherlands (source group 5 versus source group 6). The distinction between *Libelle_b* and *Libelle_n*, however, is not significant at the 0.05 level that is usual in linguistic analyses, but only at the 0.11 level; this suggests that the geographical differences that are found in the first two tests are mainly due to the other sources, rather than to *Libelle_n* and *Libelle_b*. Finally, let us compare the Dutch fashion magazines with the Dutch women's magazines (source group 3 versus source group 4). The distinction between *Burda*, *Knip* and *Margriet*, *Libelle_n* is not significant at the 0.05 level, which shows that there is no major variation within the Dutch sources.

source group	1	2	3	4
feature				
length 4 or 5	8	89	2	75
width 1	8	86	2	71
elastic cloth	8	85	2	70
no band on end of legs	8	85	2	4

Figure 3.5(5)

Frequencies within the semasiological range of *legging-leggings-caleçon*

For further comparison, an example may be considered where no signifi-

cant geographical variation occurs. Figure 3.5(5) gives the frequency of occurrence in various groups of sources of the different sets of configurations that can be distinguished within the category represented by the synonyms *legging*, *leggings*, *caleçon*. The source groups are identified by the same numbers as in Figure 3.5(4). A χ^2 -test does not yield any significant results between *Flair*, *Feeling*, *Libelle*_b on the one hand and *Burda*, *Knip*, *Margriet*, *Libelle*_n on the other (source group 1 versus source group 2). Neither is there a statistically significant distinction between *Flair*, *Feeling*, *Libelle*_b and *Margriet*, *Libelle*_n (source group 1 versus source group 4), nor between *Burda*, *Knip* and *Margriet*, *Libelle*_n (source group 3 versus source group 4).

source group	A	B
feature		
long-sleeved	21	17
waist-long	24	53
with a stiff collar	14	0
with a full row of buttons	17	4
made from non-spongy etc. material	23	36

Figure 3.5(6)

Frequencies within the semasiological range of *shirt*

Next to vest, we should mention the item *shirt* as an example of the influence of the "specialization" dimension that is present in the structure of the sources used in compiling the database. Figure 3.5(7) presents the structure of *shirt* when contextual variation is not taken into account. Surprisingly, the item appears to have two different prototypical centres. This double prototype has to be interpreted against the background of the formal origin of the word. One highly salient area of the range of application of *shirt* cor-

- a Made from non-spongy, non-fluffy, non-woolly material
- b With long sleeves
- c Waist-long, or worn tucked in at the waist
- d With a full row of buttons
- e With a stiff collar

Figure 3.5(7)
The semasiological structure of *shirt*

responds with the central area of *hemd* and *overhemd*: a shirt with long sleeves, a full row of buttons, and a stiff collar. The other salient area corresponds with the central area of *t-shirt*: short sleeves, no collar, no buttons. The first observation can be explained by the fact that the loanword *shirt* is the regular translation of *hemd* and *overhemd* (and hence has the same central application as these items). The second observation can be

explained by assuming that *shirt* may also act as an abbreviation of *t-shirt*. Although it would not be theoretically excluded to have categories with multiple salient areas, there is an additional factor involved in the case of *shirt*: the different prototypes appear to be correlated with different source groups. This becomes clear when it is recognized that the configurations in the *overhemd*-like prototype of *shirt* exclusively belong to the glossies, whereas the configurations in the *t-shirt*-prototype come from the other source groups. (In fact, the latter are almost exclusively Netherlandic sources; in the Belgian sources, *shirt* hardly ever appears.) Figure 3.5(6) presents the frequency data in the same way as in Figures 3.5(4) and 3.5(5). Group A consists of the glossies, group B of the other sources. A χ^2 -test indicates a significant distinction at a 0.001 level.

To conclude, this section has demonstrated how the presence of contextual, source-related semasiological variation in the database can be investigated. The actual contextual variation appeared to be restricted to just a few cases, but the examples illustrated both variation along the geographical dimension and variation along the specialization dimension. The chapter as a whole has been concerned with the structural characteristics of the semasiological range of application of lexical items. Specifically, it was argued that these semasiological structures are characterized by various kinds of prototypicality effects. In the opening section, a classification of those phenomena was presented by distinguishing between non-equality effects (salience) and non-rigidity effects (flexibility), both from an intensional and from an extensional perspective. In the following sections these various prototypicality effects were investigated separately. In section 3.2., intensional non-discreteness was discussed; it was shown that the clothing terms database does indeed contain cases for which classical definitions in terms of necessary and sufficient conditions cannot be given, next to classically definable lexical items. Section 3.3. focused on extensional non-discreteness. It was indicated that the corpus contains categories whose membership boundaries are not clear; moreover, this unclarity was shown to affect the intensional discreteness (i.e. the definability) of the categories in question. In section 3.4., intensional and extensional non-equality came to the fore. It was demonstrated that membership salience effects are a pervasive semasiological feature, in the sense that they may occur both with classically definable and not classically definable categories.

In the following chapters, the semasiological findings of this chapter will be extrapolated to the domains of onomasiological and formal variation. In chapter 4, it will be argued that non-equality/salience and non-rigid-

ity/flexibility structurally characterize lexical fields no less than they do individual lexical items. In chapter 5, it will be proved that the semasiological structures uncovered in the present chapter are one of the constituent factors determining which lexical item is chosen to name a particular type of referent. The startingpoint will then be the semasiological structures introduced and schematically represented in section 3.4..

Chapter 4

Onomasiological variation

Semantic fields are supra-lexical categories. The field of all words referring to skirts, for instance, provides a lexical map of the category “skirt”: items like *pleated skirt*, *miniskirt*, and *wrap-around skirt* name things that belong together in the category “skirt”. In the tradition of French structural linguistics, the categorial feature that all items in a field share is called the *archiséme*; an item such as *skirt*, whose lexical meaning coincides with the archiséme, is called an *archilexème*. (For this terminology, see e.g. Pottier 1964 and Rastier 1987.) To be sure, not every archiséme need be expressed by an archilexème. The lexical field of all verbs relating to figures of speech and expressions (such as *allegorize*, *satirize*, *apostrophize*, *personify*) does not correspond with a single archilexic item; whereas *skirt* is a hyperonym for *pleated skirt* etc., no similar hyperonym is available for the set *allegorize*, *satirize*, *apostrophize* etc.. In both types of semantic field, however, the shared archiséme captures the categorial status of the field: there is a category of types of skirts, and there is a category of verbal activities involving figures of speech.

If semantic fields are conceptual-linguistic categories, do they exhibit the same kind of “non-standard” characteristics that we pointed at earlier in the case of individual words? If so, we should be able to find evidence for the non-discreteness of semantic fields, and for the non-equality of the elements in the field. That is to say, the two major structural characteristics whose presence was revealed by a prototype-theoretical approach to the semasiology of individual words, would then reappear at the onomasiological level. The observation that non-discreteness and non-equality are among the basic structural characteristics of categories at the level of individual lexical categories, leads to the hypothesis that the same features may characterize categories at the supra-lexical level, i.e. semantic fields. In this chapter, we will try to make clear that this hypothesis is indeed borne out by the data. In both cases, indications for the importance of the phenomena in question may be found in the existing literature. But in both cases also, we will go beyond the

existing ideas, and suggest a more systematic treatment of the impact of non-equality and non-discreteness in lexical fields.

4.1. Non-discreteness in lexical fields: demarcation problems

Non-discreteness in lexical fields may roughly appear in two distinct forms: among fields, and among the elements within one field. The first kind involves questions like the following: where does one field end and where do the neighboring ones begin? Can a lexical item belong to several fields at the same time? Is there a strict borderline between two fields, or is there a transitional zone between them? The second kind of non-discreteness revolves round the question how neatly the lexical categories within one field are separated one from the other. Do they cut up the semantic space of the field into clearly separated subfields, or is the structure of the field a much more fuzzy one? The relationship between both phenomena will be clear. On the one hand, the first question involves the *external boundaries of the field*, taken as a category of its own. If fields are conceptual-linguistic categories with the same characteristics as lexical categories, at least some fields will be difficult to delineate strictly from adjacent fields. On the other hand, the second question involves the *internal structure of the field*, thought of as a set of smaller categories on the lexical level. If each of these categories is itself a non-discrete entity with possibly fuzzy boundaries, the structure of the field to which they belong will be affected: rather than sharp divisions between the individual items within the field, a more blurred picture with variously overlapping subfields emerges.

To appreciate more fully how both characteristics differ from the classical conception of lexical fields, a brief look at the history of lexical field research may be useful. The first empirical study in which the structurally inspired semantic program was worked out, was conducted by Jost Trier in his monograph *Der Deutsche Wortschatz im Sinnbezirk des Verstandes. Die Geschichte eines sprachlichen Feldes* (1931), in which he examined the evolution of the terminology for a particular set of intellectual characteristics from Old High German up to the beginning of the 13th century; Trier (1932) deals with the further evolution in Middle High German. Trier started from the fundamental insight that in the case of abstract notions such as intellectual characteristics (where there are no tangible or easily identifiable extra-

linguistic referents), only a mutual demarcation of the notions concerned could give a decisive answer regarding their exact value. Words should not be considered in isolation, but they should be considered in their relationship with regard to semantically related words: the semantic demarcation of words is always a demarcation relative to other words (the notion "demarcation" itself seems to imply that at least one other entity is involved). The image on which this conception is based is that of a *mosaic*: the conceptual substance of language is divided into a number of adjoining small areas, in the way a mosaic divides two-dimensional space.

As an illustration of Trier's approach, we can have a look at Trier (1934), in which a subarea of the vocabulary concerning intellectual properties is dealt with (viz. the words denoting knowledge). At the beginning of the 13th century, courtly language possesses three core notions referring to knowledge: *wisheit*, *kunst* and *list*. The distinction between the latter two reflects the class structure of the feudal society: *kunst* conveys the knowledge and skills of the courtly knight (viz. courtly love, the chivalric code of honor and the "artes liberales"), whereas *list* is used to indicate the knowledge and the skills of those who do not belong to the nobility (such as the technical skills of the craftsmen). *Wisheit* is a general term which may be used in connection with noblemen as well as in relation to ordinary citizens; it is predominantly employed in the religious-ethical sense, in a way that is similar to Latin *sapientia*. In a sense, *wisheit* indicates that one has the appropriate skills and the required knowledge to occupy one's appointed position in society (whatever that might be). The general concept *wisheit* indicates that the distinct spheres of the aristocratic *kunst* and the civil *list* are embedded in a common religious world order.

A century later, the structure of the field has undergone considerable changes. *List*, which gradually acquires a pejorative meaning (conveying "cunning") is replaced by *wizzen*, which does not, however, have precisely the same meaning as the earlier term *list*. *Kunst* and *wisheit* as well have acquired a different scope. *Wisheit*, for instance, has ceased to be a general term; it conveys a particular type of knowledge, in the sense that it refers to religious knowledge in a maximally restricted sense (the knowledge of God). *Kunst* and *wizzen*, on the other hand, indicate higher and lower forms of profane knowledge, without there being any reference to social class distinctions. *Wizzen* refers to technical skills (e.g. the skills of a craftsman), whereas *kunst* denotes pure forms of science and art. The example as a whole, summarized in Figure 4.1(1), demonstrates how the structure of lexical fields develops from one synchronic period to another.

Figure 4.1(1)

Shifts in the German field of intellectual abilities according to Trier

How did lexical field analysis evolve after Trier? On the one hand, the structuralist foundations of his approach were enthusiastically adopted on a large scale; on the other, critical comments were formulated which led to alternatives for Trier's specific realization of the lexical field approach. Two important points of criticism will be mentioned below. They do not, however, cover the complete range of critical reactions with regard to Trier's views. It should be noted, for instance, that the descriptive, philological aspect of Trier's study has been criticized as well (among other things because the texts on which his study is based apparently cannot be considered representative for Old High German and Middle High German at large). For a systematic overview of the criticisms against Trier's field conception, see Geckeler (1971: 115-167). As a historical note, it should be mentioned that Trier's mosaic image as well as his use of the term *field* to refer to a collection of semantically related words that mutually determine each other's value, were borrowed from Ipsen (1924) (a paper in which, in fact, the field concept plays only a minor role). Furthermore, it may be noted that already in the nineteenth century, studies can be found that use the intuition of a structured lexical field as the basis for semantic research (see Coseriu &

Geckeler 1981: 20 and Öhman 1953). It is only in the structuralist era, however, that this approach achieved methodological supremacy.

Trier's use of the mosaic image was not a happy one. To begin with, the image suggests that the mosaic covers the whole surface of the field, i.e. that there are no *gaps in the lexical field*, that no pieces are lacking in the mosaic. This *Lückenlosigkeit* (absence of gaps) is obviously contradicted by the linguistic facts, as will be obvious from Figure 4.1(2), which reproduces part of an analysis of English cooking terms presented in Lehrer (1974): some of the systematically present semantic possibilities are simply left unfilled (in the figure, there is no word for the preparation of food in a pan without water and oil, nor for cooking with oil on a flame). Such examples are innumerable, and the conception of a closed system has been generally abandoned.

	Conducted heat (oven)	Radiated heat (fire)	Contact heat (pan)
with water vapor- ized	steam		
without oil not vapor- ized	boil		
with oil, without water	(oven-fry)		fry
without water, without oil	roast		
	bake	broil	

Figure 4.1(2)
English cooking terms (after Lehrer 1974)

A further assumption that can be deduced from the image of the mosaic, is that fields are, internally as well as externally, clearly delineated, i.e. that the words in a field, like mosaic pieces, are separated by means of sharp lines, and that different fields link up in the same clear-cut way. The whole lexicon would then be an enormous superfield falling apart in huge but clearly delineated sets, which in turn break up in smaller field structures, and so on until we reach the ultimate level of the mosaic stone, the word. This *compartmentalization of the lexicon* was criticized from different angles. In particular, it was pointed out by various researchers that it might often be difficult to indicate exactly where a field ends. Discreteness will usually only be found in the core of a field, whereas there is a peripheral transition zone around the core where field membership is less clearly defined.

It is worth mentioning in this respect that Trier (1968), looking back on the development of lexical field theory, regrets that he failed to correct Ipsen's mosaic image. This could have avoided, he admits, unnecessary confusion with regard to the character of lexical fields. The image of the closely fitting word and field boundaries should be substituted by a star-like conception of lexical fields in which the centre of the field sends out beams that are able to reach other cores with their extreme ends. As a prime example of such a star-like conception of lexical fields, Ducháček's analysis of the semantic field of French words referring to beauty (1959) should be mentioned. Figure 4.1(3) reproduces part of Ducháček's representation of the structure of the field. The centre of the field lists a number of items (like *beau* and *beauté* themselves) that refer to the concept of beauty in its purest and simplest form. The outer circle specifies fields that are adjacent to the field of beauty. The items situated on the rays connecting the core of the field with the adjacent concepts refer to types of beauty, beautiful characteristics etc. that are associated with the neighboring concepts. Thus, for instance, *sex-appeal* and *séduisant* 'seductive' are placed on the line that radiates towards the field of *séduction*; sex-appeal is a kind of beauty tinged with the concept of seduction. Similarly, *sublime* names a type of beauty associated with *grandeur* 'magnificence', and terms like *merveilleux*, *fascinant*, and *charmant* name forms of beauty conceived of as a magic force (*magie*).

Although the external non-discreteness of fields was fairly generally recognized, representations like Ducháček's far from abound in lexical field theory. Rather than explicitly analyzing the transitions from one field to another (along the model furnished by Ducháček), lexical field researchers more often impose a practical restriction on the set of items to be discussed.

Figure 4.1(3)

The French field of the concept "beauty", according to Ducháček (1959)

In other words, representations like the one in Figure 4.1(2) have always been more popular among field theorists than representations like those in Figure 4.1(3). The internal non-discreteness of fields, on the other hand, received even less attention from classical lexical field theory. From one point of view, this is rather surprising, because Trier himself, in a critical reaction to the very rigid application of the mosaic picture by Jolles (1934), had stipulated as early as 1934 that

die Binnengrenzen [des Feldes], weit davon entfernt, als klare mathematische Grenzkonturen sich zu erweisen, in Wahrheit vielmehr Ueberschneidungszonen und schwankende Uebergangssäume darstellen. [The internal boundaries of fields, far from being mathematically clear demarcation lines, in reality rather constitute overlapping areas and fluctuating transitional zones (Trier 1934: 446).]

From another point of view, however, it is quite understandable that the

basic structuralist idea that the items in a field mutually determine each other's value would favor the idea of clear demarcation lines between the lexical categories in a field. In actual practice, at any rate, not much descriptive attention was devoted to the un-mosaic-like character of the inner boundaries within lexical fields. Probably the most outspoken descriptive study in this respect is Gipper's investigation into the German word *Sessel* 'armchair' and *Stuhl* 'chair' (1959), in which he empirically demonstrates (among other things) that the same piece of furniture may variously be called *Sessel* or *Stuhl* by the same informant.

But whereas observations like those of Gipper remained largely uninfluential in the framework of structuralist field theory, the internal non-discreteness of lexical fields comes to the fore naturally in the framework of prototype-theoretical approaches. If the boundaries of lexical items taken by themselves may be vague, then it is to be expected that no sharp division may exist between the items in a field. Rather than a strict separation of the *right/left*-type (where one term obviously excludes the other), multiple overlapping between the items in a field can be expected to be the rule, in the sense that items that are peripheral with regard to one particular category are likely to appear in other categories as well.

As a first step towards a *representational model* for such a conception of lexical fields, let us have a look at Lehrer's reconsideration (1990) of her own analysis of English cooking terms. The upper part of Figure 4.1(4) specifies part of Figure 4.1(2). The crucial point is the multiple applicability of *boil*: it may either refer to any instance of cooking with water (in which case it is a superordinate name with regard to *simmer* and *steam*), or it may refer to the typical action of boiling, i.e. cooking food that is largely submerged in water at a full boil (with rolling bubbles). In this case, *simmer* and *steam* are contrastive rather than subordinate terms: simmering is gentle rather than full boiling, and in the case of steaming, the food is not submerged. In the upper part of Figure 4.1(4), both applications of *boil* are treated as clearly separate meanings. From a prototype-theoretical point of view, however, this is debatable, because the various types of boiling could be seen as instances of the same category. Next to the (proto)typical case of boiling, we find peripheral instances; these peripheral instances can be alternatively lexicalized with *steam* and *simmer*. Lehrer's representation of this alternative analysis is reproduced in the lower part of Figure 4.1(4).

Such a representation is not, however, as informative as might be. For one thing, it contains no indication of the actual semantic differences among the items involved; it indicates the way in which the words divide a particu-

Figure 4.1(4)
English cooking terms (after Lehrer 1990)

lar semantic space among themselves, but does not specify the substance of that space. For another, it suggests that there are instances of *simmer* and *steam* that fall outside the semasiological range of application of *boil*, but it does not indicate precisely which ones. It may be useful, therefore, to explore other formalisms to replace the mosaic-like representations. We shall do so for the lexical field of Dutch terms referring to skirts, or at any rate, the most frequent ones in our corpus. The items in question are: *rok* 'skirt', *rokje* 'short skirt', *plooirok* 'pleated skirt', *wikkelrok* 'wrap-around skirt', *minirok* 'short skirt, miniskirt', *doorknooprok* 'button-through skirt', *klokrok* 'flared skirt'. The componential dimensions and features used for the description of the *rok*-subfield are the following.

LENGTH

- [1] The skirt is as long as the upper part of the thighs.
- [2] The skirt is as long as the lower part of the thighs.
- [3] The skirt is as long as the knees.
- [4] The skirt is as long as the calves.
- [5] The skirt is as long as the ankles.

WIDTH AND CUT

- [1] The skirt gradually narrows towards the hem and has a somewhat spherical shape.

[2] The skirt has a roughly rectangular shape.

[3] The skirt is wide and pyramid-shaped.

FASTENING

[1] The skirt has a partial fastening (buttons or zip fastener).

[2] The skirt has a full fastening (buttons or zip fastener).

[3] The skirt's front panels are wrapped.

[4] The skirt has an elastic band.

FABRIC

[1] The skirt is made of cotton or a cottonlike fabric.

[2] The skirt is made of flannel or a flannellike fabric.

[3] The skirt is made of denim.

[4] The skirt is made of corduroy.

[5] The skirt is made of silk or a silky fabric.

[6] The skirt is made of leather or a leatherlike fabric.

[7] The skirt is made of suede or a suedelike fabric.

[8] The skirt is made of a stretchy fabric.

FINISHING

[1] The fabric has been pleated into wide and sharp folds.

[2] The fabric has been pleated into narrow and sharp folds.

[3] There are vertical ruffles along the skirt which give it a wavy appearance.

[4] The skirt is composed of broad horizontal bands that give it a tiered appearance

[5] The skirt is smooth (i.e. unpleated and unruffled).

A first possibility for representing the structure of the field is illustrated in Figure 4.1(5). For each configuration in the field, it is indicated whether it belongs to the range of application of the items in question. The figure clearly shows that the field is not structured in mosaic-like fashion; rather, the items clearly encroach on each other's domain. This is hardly surprising: wrap-around skirts and short skirts are obviously not mutually exclusive; some skirts can be both at the same time, and so we find an area of overlap between *wikkelrok* and *rokje*. By fully specifying the configurations involved, we get a good idea of the semantic substance of the field.

A problematic aspect of Figure 4.1(5) involves the representativity of the listed configurations. The diagram suggests that, for instance, the referential range of *wikkelrok* is not entirely included in that of *rok*; there would be, then, instances of wrap-around skirts that are not skirts. This is implausible; there is no reason to doubt that *wikkelrok* is a hyponym of *rok*. Rather, we get a misleading picture because the configurations we use are in a sense too

Figure 4.1(5)

Dutch words referring to skirts, represented by referential configurations

specific. If the question is asked whether configuration [I32355] can be called *rok*, the absence of a positive answer may simply be due to quantitative restrictions on the corpus: we expect to find a *rok*-denomination sooner or later when more material is added to the database. When an intensional rather than an extensional perspective is taken, the problem dissipates. *Wikkelrok* can be simply (and classically) defined as a skirt with a wrap-around type of fastening, i.e. [I..3..]. In general, the definition of the various types of skirts considered here is relatively straightforward in comparison with the *jasje*-subfield analyzed in section 3.2.. Specifically, the items may be defined unidimensionally, i.e. by referring to just a single dimension in the componential structure: *rokje* and *minirok* refer to the dimension LENGTH, *doorknooprok* involves the dimension FASTENING, *plooirok* involves TREATMENT OF MATERIAL, and *klokrok* refers to WIDTH AND CUT.

Figure 4.1(6)

Dutch words referring to skirts, represented by definitions

As a consequence of the intensional reinterpretation of the data, (part of) the field structure can now be diagrammed as in Figure 4.1(6). (The relevant definition is added to each item. The figures in the overlapping areas indicate the number of referents in each area. For instance, while *wikkelrok* is present with 74 records, there are 18 cases in the area where *wikkelrok* and *klokrok* overlap. Note that these 18 cases are a subset of the 74 instances of *wikkelrok*. The overlapping figures only refer to the overlapping of two items. For instance, the fact that no figure is included in the area where

wikkelrok, *klokrok*, and *rokje/minirok* overlap does not necessarily mean that no records could be found there.)

Clearly, representations like those in 4.1(6) can only be given if the definitional structure of the field is as simple as in the case of skirts. When no classical definitions are possible, representations based on full configurations, as in 4.1(5), regain some of their attractivity. An example of such a field is that of shirt-like garments, the structure of which is given in Figure 4.1(7). The figure is built up on the same principles as Figure 4.1(5), but for reasons of graphical economy, the configurations have been numbered: a full specification would have cluttered the figure beyond recognition. The configurational value of the abbreviatory numbers is given in the legend accompanying the figure. The componential system used for the description of the shirt-field is as follows.

SHAPE

- [1] The garment covers the trunk below the shoulders.
- [2] The garment covers trunk and shoulders but leaves the arms uncovered.
- [3] The garment covers trunk, shoulders and upper arms, but leaves the lower arms uncovered.
- [4] The garment covers the trunk and the arms.

LENGTH

- [1] The garment does not cover the trunk below the midriff and cannot be tucked into skirt or trousers.
- [2] The garment is tucked into skirt or trousers.
- [3] The garment covers the hips.
- [4] The garment covers the thighs.

FASTENING

- [1] The garment does not have a fastening.
- [2] The garment has a partial fastening from the throat down.
- [3] The garment has a full fastening.
- [4] The garment's panels are wrapped.

FABRIC

- [1] The garment is made of a smooth, cottonlike fabric.
- [2] The garment is made of a flannellike fabric.
- [3] The garment is made of denim.
- [4] The garment is made of silk.

COLLAR

- [1] The garment is collarless.
- [2] The garment has a stand-up collar.

Figure 4.1(7)

Dutch words referring to shirt-like garments, represented
by referential configurations

- [3] The garment has a soft collar.
- [4] The garment has a stiff collar.
- [5] The garment has a lapel collar.

NECKLINE

- [1] The garment does not have a neckline because it does not cover the shoulders.
- [2] The garment has a V-shaped neckline.
- [3] The garment has a round neckline.
- [4] The garment's neckline has a rectangular shape.
- [5] The garment has a small, U-shaped neckline.
- [6] The garment has a wide, oval neckline.

POSITION OF THE BUTTONS

- [1] The buttons are located on the right front panel from the wearer's point of view.
- [2] The buttons are located on the left front panel from the wearer's point of view.
- [3] The position of the buttons is irrelevant because the garment does not have a fastening.

DETAILS

- [1] The garment has epaulettes.

The configurations that are replaced by numbers in Figure 4.1(7) are the following.

1.	E3231432v	20.	E1331122v	39.	E4211133v
2.	E3431412v	21.	E1414113v	40.	E4211163v
3.	E4231332v	22.	E2111123v	41.	E4311133v
4.	E4231411v	23.	E2131122v	42.	E1211143v
5.	E4231422v	24.	E2211123v	43.	E2111113v
6.	E4231432v	25.	E2211133v	44.	E2331133v
7.	E4234412v	26.	E2214163v	45.	E3111133m
8.	E4331411v	27.	E2221133v	46.	E3211123v
9.	E1111113v	28.	E2234132v	47.	E3211133m
10.	E1121113v	29.	E2312143v	48.	E3218233v
11.	E1124113v	30.	E2331132v	49.	E3221121v
12.	E1131112v	31.	E3111123v	50.	E3221233v
13.	E1134133v	32.	E3111163v	51.	E3222133v
14.	E1141123v	33.	E3131121v	52.	E3231132v
15.	E1211113v	34.	E3131142v	53.	E3231133v
16.	E1211413v	35.	E3311133v	54.	E3311123v
17.	E1231112v	36.	E3211163v	55.	E3211133v
18.	E1311113v	37.	E3131122v	56.	E3311163v
19.	E1331112v	38.	E3332322v	57.	E3314133v

58.	E3321323v	106.	E3331432v	154.	E4333422m
59.	E3331411v	107.	E3431432v	155.	E4334231v
60.	E3411133v	108.	E4231411m	156.	E4334431v
61.	E3411163v	109.	E4231412v	157.	E4334432v
62.	E3411233v	110.	E4231431m	158.	E4334532v
63.	E3414133v	111.	E4232431m	159.	E4431322v
64.	E3431133v	112.	E4232431v	160.	E4432431m
65.	E3441133v	113.	E4232432v	161.	E1111213v
66.	E4211133m	114.	E4233411v	162.	E1214133v
67.	E4211133v	115.	E4331412v	163.	E1331522v
68.	E4211163v	116.	E4331422v	164.	E2131522v
69.	E4212163v	117.	E4334411m	165.	E2134422v
70.	E4214133v	118.	E4334411v	166.	E2214123v
71.	E4221123v	119.	E4334412v	167.	E2214133v
72.	E4311143v	120.	E4431311m	168.	E2231122v
73.	E4311163v	121.	E4431431v	169.	E2231412v
74.	E4311363v	122.	E4431432v	170.	E2231522v
75.	E4314133v	123.	E4434332v	171.	E2314133v
76.	E4314163v	124.	E3231131v	172.	E2331322v
77.	E4321322v	125.	E3233421v	173.	E2331522v
78.	E4411133v	126.	E3331431v	174.	E3231122v
79.	E2131132v	127.	E4131411v	175.	E3131322v
80.	E2231132v	128.	E4131432v	176.	E3131522v
81.	E2311133v	129.	E4221423v	177.	E3141123v
82.	E2411143v	130.	E4231222v	178.	E3231312v
83.	E3221133v	131.	E4231322v	179.	E3231322v
84.	E3221322v	132.	E4231431v	180.	E3231332v
85.	E3311133m	133.	E4231432m	181.	E3231412v
86.	E4221133m	134.	E4233431v	182.	E3231422v
87.	E4231132v	135.	E4233432v	183.	E3231522v
88.	E4231311m	136.	E4234411m	184.	E3234322v
89.	E4232411m	137.	E4234421m	185.	E3234522v
90.	E4311133m	138.	E4234422v	186.	E3321123v
91.	E4321413v	139.	E4234431m	187.	E3331122v
92.	E4331311m	140.	E4234431v	188.	E3331132v
93.	E4331411m	141.	E4234432v	189.	E3331162v
94.	E4332411m	142.	E4321411v	190.	E3331312v
95.	E4333111v	143.	E4324422v	191.	E3331322v
96.	E4411323v	144.	E4331122v	192.	E3331512v
97.	E4411523v	145.	E4331322v	193.	E3331521v
98.	E4414133v	146.	E4331421v	194.	E3331522v
99.	E4414323v	147.	E4331431m	195.	E3334122v
100.	E2331411m	148.	E4331431v	196.	E3334412v
101.	E3231411m	149.	E4331432v	197.	E3334522v
102.	E3231431m	150.	E4331522v	198.	E3421142v
103.	E3234122v	151.	E4332431v	199.	E3431122v
104.	E3331412v	152.	E4332432v	200.	E3431132v
105.	E3331422v	153.	E4333411m	201.	E3431162v

202.	E3431332v	217.	E4232132v	232.	E4312123v
203.	E3431522v	218.	E4232232v	233.	E4324413v
204.	E3434122v	219.	E4232522v	234.	E4331312v
205.	E3434132v	220.	E4233412v	235.	E4334122v
206.	E3434432v	221.	E4234122v	236.	E4334132v
207.	E4131122v	222.	E4234212v	237.	E4341323v
208.	E4131412v	223.	E4234322v	238.	E4344113v
209.	E4134122v	224.	E4234411v	239.	E4344323v
210.	E4224413v	225.	E4234522v	240.	E4431122v
211.	E4231122v	226.	E4241322v	241.	E4431312v
212.	E4231311v	227.	E4241522v	242.	E4431412v
213.	E4231312v	228.	E4241523v	243.	E4434412v
214.	E4231421v	229.	E4242122v	244.	E4434512v
215.	E4231433v	230.	E4244523v	245.	E4434522v
216.	E4231522v	231.	E4311423v	246.	E4332411v

Note that Figure 4.1(7) is strikingly different in comparison with 4.1(5). In the latter case, the central position of *rok* in the figure is an indication of its superordinate status (a status that is intuitively beyond doubt, and that is confirmed by a definitional analysis). In the former case, on the other hand, there is no single item that is structurally similar to *rok* (and neither is there an intuitively plausible candidate for archilexic status). Rather, the field is highly fragmented, with multiple overlappings among the categories. By contrast, the *rok*-field is structured in such a way that *rok* itself is the only item that overlaps with all others. We may conclude, then, that even given the statistical margin of error that has to be observed with extensional representations like those in 4.1(5) and 4.1(7), such diagrams are able to reveal major differences in the structures of distinct fields.

To round off the discussion, let us note that even in those cases where no classical definitions are readily available, the diagrams can be intensionally enriched by drawing various circles for a single item, rather than representing each category by a single circle. If, for instance, the field relationship between *blazer* and *colbert* is to be charted, diagrams of the kind introduced in section 3.4. could be combined. Although the resulting picture would, of course, turn out to be quite cluttered, this type of picture would have the advantage of representing both the structure of the lexical field and the prototypicality structure of the items in the field. (For representational attempts in the same direction, see the work of Schmid 1993b and MacLaury 1992.)

Ultimately, which type of field representation should be chosen will depend both on the kind of items involved (whether they can be classically defined or not, for instance), and on the purpose of the representation. We feel, in other words, that it is neither necessary nor desirable to select a

single representational format as the only useful one. Rather than to argue for a specific type of graphical representation, our purpose has been to show that there are various methods for charting the structure of lexical fields that all point in the direction of their non-classical, un-mosaic-like character.

4.2. Non-equality in lexical fields: entrenchment

In the literature, differences of onomasiological salience have so far been described primarily in terms of the *basic level hypothesis*. The background of the hypothesis is the ethnolinguistic recognition that folk classifications of biological domains usually conform to a general organizational principle, in the sense that they consist of five or six taxonomical levels (Berlin, Breedlove & Raven 1973; Berlin 1974, 1978). Figure 4.2(1) illustrates the idea with two sets of examples. The highest rank in the taxonomy is that of the “unique beginner”, which names a major domain like *plant* and *animal*. The domain of the unique beginner is subdivided by just a few general “life forms”, which are in turn specified by “folk genera” like *pine*, *oak*, *beech*, *ash*, *elm*, *chestnut*. (The “intermediate” level is an optional one.) A folk genus may be further specified by “folk specifics” and “varietal taxa”. To the extent that the generic level is the core of any folk biological category, it is the basic level:

“Generic taxa are highly salient and are the first terms encountered in ethnobiological enquiry, presumably because they refer to the most commonly used, everyday categories of folk biological knowledge” (Berlin 1978: 17).

The generic level, in other words, is onomasiologically salient: within the lexical field defined by the taxonomy, the generic level specifies a set of preponderant items. In this sense, the basic level embodies a set of naming preferences: given a particular referent, the most likely name for that referent from among the alternatives provided by the taxonomy will be the name situated at the basic level.

Apart from embodying a concept of onomasiological salience, basic level categories are claimed to exhibit a number of other characteristics. From a psychological point of view, they are conceptualized as perceptual and func-

tional gestalts. From a developmental point of view, they are early in acquisition, i.e., they are the first terms of the taxonomy learned by the child. From a linguistic point of view, they are named by short, morphologically simple items. And from a conceptual point of view, Rosch et al. (1976) claim that the basic level constitutes the level where prototype effects are most outspoken, in the sense that they maximize the number of attributes shared by members of the category, and minimize the number of attributes shared with members of other categories. An objectivist explanation for this situation is suggested: the basic level is purported to be the level where reality itself exhibits a maximal correlation of attributes; basic level categories are the categories where reality itself maximizes the internal coherence and the external distinctiveness of categories.

<i>Rank</i>	<i>Ethnobiological examples</i>	
KINGDOM	plant	animal
LIFE FORM	tree	fish
INTERMEDIATE	evergreen	fresh water fish
GENERIC	pine	bass
SPECIFIC	whitepine	black bass
VARIETAL	Western white pine	large mouthed bass

Figure 4.2(1)
Examples of taxonomical organizations

Applying the basic level model to the field of clothing terminology, it seems intuitively plausible that items like *broek* 'trousers', *rok* 'skirt', *trui* 'sweater', *jurk* 'dress' are basic level categories: their overall frequency in the corpus is high, and they typically have the monomorphemic form of basic level categories. A further extrapolation yields the right-hand side of Figure 4.2(1), in which *kledingstuk* 'garment' is considered a unique beginner in contrast with, say, *gereedschap* 'utensils' or *speelgoed* 'toys'.

Although Figure 4.2(1) suggests that the basic level model is a plausible framework for the field of clothing terminology, there are at least two *reasons for being careful* about the basic level model of onomasiological salience. In the first place, the model presupposes the existence of a clear taxonomical organization of the lexicon that seems to rely on a mosaic-like picture of conceptual structures. Even apart from the general point, made in the previous section, that such a picture may not be the most adequate one, there are specific reasons for questioning the neatness of the division in levels that is presupposed by the basic level model. For one thing, uncertainties about inclusion relations undermine the stability of the taxonomical hierarchy. If it is indeterminate (as we have tried to show in section 3.3.) whether *broekrok* is a hyponym of *broek* or of *rok*, it is also unclear whether *broekrok* is to be situated one level below these terms, or rather on the same level. For another, the lexicon is not a single taxonomical tree with ever more detailed branchings of nodes, but is rather characterized by multiple overlapping hierarchies. Consider, for instance, how an item like *dameskledingstuk* 'woman's garment, item of clothing typically or exclusively worn by women' would have to be included in a taxonomical model of the lexicon. As Figure 4.2(2) shows, such a classification on the basis of gender-specificity cross-classifies with a classification based on functional gestalts like *broek* and *rok*. So can we say that *dameskledingstuk* belongs to the same level as *broek* and *rok*? Whereas the latter items would probably be basic level terms, this could hardly be the case for *dameskledingstuk*. But how can the level of *dameskledingstuk* be determined at all, if it does not fit into the same taxonomical hierarchy as *broek* and *rok*? In a basic level model of onomasiological salience, the lower degree of salience of *dameskledingstuk* would have to result from its taxonomical position with regard to the level where *broek* and *rok* are situated; but this taxonomical position is unclear because *dameskledingstuk* cross-classifies with the *broek/rok*-classification.

A second, even more important reason for being suspicious about the basic level hypothesis as a model of the distribution of degrees of onomasiological salience involves the observation that differences of onomasiological

preference also occur among categories on the same level in a taxonomical hierarchy. The basic level model contains a hypothesis about alternative categorizations of referents: if a particular referent (a particular piece of clothing) can be alternatively categorized as a garment, a skirt, or a wrap-around skirt, the choice will be preferentially made for the basic level category "skirt". But analogously, if a particular referent can be alternatively categorized as a wrap-around skirt or a miniskirt, there could just as well be a preferential choice: when you encounter something that is both a wrap-around skirt and a miniskirt, what is the most natural way of naming that referent? At this point, lexical field representations like those in Figures 4.1(5) and 4.1(6) prove their value: the areas of overlap among items such as *rokje* and *plooirok* reveal where alternative categorial choices within the same hierarchical level (that of the hyponyms of *rok*) occur.

Figure 4.2(2)

An example of taxonomical cross-classification

If, then, we have to reckon with intra-level differences of salience next to inter-level differences, the concept of onomasiological salience has to be generalized in such a way that it relates to individual categories at any level of the hierarchy (or what is left of it when all forms of hierarchical fuzziness

are taken into account). Terminologically, this concept of *generalized onomasiological salience* can be equated with Langacker's notion of *entrenchment* (1987: 59-60). Langacker introduces the concept in connection with the process of unit formation: a particular linguistic construct (such as a new compound, or the use of a word in a new reading) may gradually transcend its initial incidental status by being used more often, until it is so firmly entrenched in the grammar or the lexicon that it has become a regular well-established unit of the linguistic system. Metaphorically, entrenchment is a form of conceptual "wiring in": a well-entrenched concept is more firmly anchored in the language user's knowledge of the language. An extended study of entrenchment phenomena in syntax is provided by Deane (1992: 187-236). The difference between the approach followed by Deane and ours resides in the fact that we will try to develop a quantifiable, operational criterion for measuring entrenchment values in a corpus of observed language use. Such a quantitative, usage-based measure of entrenchment is absent in Deane's approach, although some of the indices of entrenchment that he uses (such as the ease with which a particular phrase is topicalized) are in principle amenable to quantitative treatment.

The generalized concept of entrenchment, defined as onomasiological salience, may be operationally defined as the ratio between the frequency with which the members of a lexical category are named with an item that is a unique name for that category, and the total frequency with which the category occurs in the corpus. For instance, the lexical category "bird" will be highly entrenched if, of a total of 100 references to birds, 60% or so occurs with the name *bird* rather than with hyperonyms like *animal* or *vertebrate*, and hyponyms like *blackbird*, *sparrow*, or *gull*. Because we are dealing with the entrenchment of lexical categories rather than mere word forms, any occurrences of synonyms of *bird* would obviously have to be incorporated into the numerator of the ratio to be computed. On the basis of this basic operational definition, four alternative measures of entrenchment may be defined, depending on whether an intensional or an extensional approach is taken, and depending on whether a monolexical or a polylexical measure is calculated.

When an *extensional* startingpoint is taken, the calculation is based on the list of componential configurations that occur in the semasiological range of application of the category under consideration; this is the same kind of list as in Figure 3.2(2). In an *intensional* perspective, a definition of the category under investigation is used to determine which referents in the corpus fall within the definition. It should be clear from the previous section

that these two approaches need not coincide. There are, for instance, configurations within the corpus as a whole that do fall within the definitional range of *rok* 'skirt', but that do not fall within the extensional range of *rok* as it happens to occur in our material. (See the discussion of Figures 4.1(5) and 4.1(6) in the previous section.) From a theoretical point of view, a definitional approach would seem to be preferable, but at the same time, the problems we have identified in chapter 3 with regard to the notion of lexical definition lead to a certain amount of caution. To the extent that some categories are more difficult to define (more specifically, to the extent that they cannot receive a classical definition), the configurational, extensional method wins in attractiveness.

In order to explore the relationship between both methods for measuring entrenchment, we may have a further look at the *rok*-field examined in the previous section. Figure 4.2(3) specifies the entrenchment ratios of *rok* and a number of its hyponyms. It may be noted that *rokje* and *minirok* occur two times in the figure. On the one hand, they are treated separately as distinct items; on the other hand, they occur together as a single category. This reflects the fact that the decision to treat two items as synonymous may differ according to whether an intensional or an extensional perspective is taken. From a definitional point of view, *rokje* and *minirok* cover the same definitional domain (in terms of relevant values on the dimensions WIDTH and LENGTH); from a configurational point of view, however, their ranges are not sufficiently similar to claim full synonymy (see Figure 4.1(5)). Regardless of the question which measure of synonymy would ultimately have to be preferred, it seems rational to maintain the distinction between both approaches in the context of a comparison between the configurational and the definitional measure of entrenchment. (Note that the configurational entrenchment measure for *rokje/minirok* is calculated by lumping together the configurational ranges of both items. The definitional entrenchment measures for *rokje* and *minirok* are calculated on the basis of the definition of *rokje/minirok*.) Another methodological point to be mentioned concerns the treatment of incomplete componential configurations. Because the calculation of the configurational entrenchment measure is based on semasiological ranges of the type illustrated in Figure 3.2(2), the same kind of caution introduced there has to be applied here. It will be remembered from the discussion in section 3.2. that configurations containing question marks may pose problems of interpretation, and were systematically removed from the lists. The same procedure has been followed here, though only, of course, for the configurational measure.

Item	Configurational entrenchment		Definitional entrenchment	
rok	62.53	rank 1	54.85	rank 1
rokje-minirok	47.89	rank 2	33.84	rank 3
doorknooprok	47.50	rank 3	35.59	rank 2
rokje	43.71	rank 4	31.45	rank 5
wikkelrok-omslagrok- overslagrok	39.02	rank 5	32.43	rank 4
plooirok	26.98	rank 6	16.90	rank 6
klokrok	26.31	rank 7	3.41	rank 7
minirok	18.03	rank 8	2.39	rank 8

Figure 4.2(3)

Configurational and definitional entrenchment values in
the field of skirts

Two important observations can be made with regard to Figure 4.2(3). First, the definitional entrenchment value is always lower than the configurational one. This is as might be expected, given the fact that a configurational approach tends to miss a number of configurations that lie within the definitional range of the category (see the previous section); as these missing referents can only be added to the denominator of the entrenchment measure, the overall entrenchment value will obviously be lower. Second, the rank order of the categories on the configurational and definitional measures is roughly the same. There are two switches between the ranks 2 and 3 and between 4 and 5, but as these involve pairs of categories whose entrenchment values on both counts are close to each other, the differences in rank order may be considered insignificant. When *rokje/minirok*, *rokje*, and

minirok are removed from the comparison, the rank orders of *rok*, *door-knooprok*, *wikkelrok* (etc.), *plooirok*, and *klokrok* correspond in a one-to-one fashion. Precisely because entrenchment values are primarily useful as a measure of the *relative* salience of categories in a field with regard to each other, we may conclude that the definitional and the configurational approach are both useful in calculations of onomasiological salience. In what follows, both measures will continue to be used, but because the extensional, configurational approach avoids the definability problems of the intensional, definitional approach, there will be a certain preference for the generality of the configurational measure.

The mathematical relationship between the definitional and the configurational calculation of entrenchment can be made more precise. In particular, it can be demonstrated that the differences between the configurational and the definitional measures of entrenchment depend on two factors: first, the extent to which the actual range of occurrence of an item is representative of its virtual semasiological range as delimited by a definition, and second, the number of incomplete configurations that has been removed from the semasiological list used as the basis of the configurational entrenchment calculation. The following proof of this relationship requires a rich terminological apparatus that will not be used any further in the text. In this sense, the demonstration may be easily skipped by those readers who are primarily interested in the main line of book, and less in the technical details. Let us start by introducing the following concepts:

Name frequency

= the frequency of occurrence in the corpus of the category under investigation (for instance, the number of times that the item *rok* occurs in the corpus)

Corrected name frequency

= the frequency of occurrence of the category under investigation after removing incomplete configurations

Type frequency

= the total number of instances of a category as delimited by a definition

Configurational frequency

= the cumulative frequency in the corpus as a whole of the configurations that occur in the semasiological range of an item

Naming frequency proportion

= the ratio between "corrected name frequency" and "name fre-

quency"

Referential frequency proportion

= the ratio between "configurational frequency" and "type frequency".

The "naming frequency proportion" gives an indication of the number of configurations that are lost out of sight by removing incomplete configurations from the semasiological ranges on the basis of which the configurational entrenchment value is calculated. The "referential frequency proportion", on the other hand, gives an indication of the number of potential referents of a category that are lost out of sight by taking an extensional rather than an intensional startingpoint. When the "referential frequency proportion" is high, the actual range of occurrence of an item is highly representative of the semasiological range as derived from a definition. Given the following further set of definitions

Configurational entrenchment

= corrected name frequency divided by configurational frequency

Definitional entrenchment

= name frequency divided by type frequency

Entrenchment proportion

= configurational entrenchment divided by definitional entrenchment,

it can be easily verified that the "entrenchment proportion" equals the division of "naming frequency proportion" and "referential frequency proportion". It follows that the difference between the configurational and the definitional measure of entrenchment will become smaller, first, when the actual range of occurrence of an item is highly representative of the semasiological range as delimited by a definition, and second, when only a minimal number of incomplete configurations has been removed from the semasiological list used as the basis of the configurational entrenchment calculation.

Next to the choice for either an intensional or an extensional approach to entrenchment, the choice between a *monolexical* and a *polylexical* variant has to be mentioned. So far, only separate lexical items have been included in the numerator of the entrenchment measure. It may be wondered, however, whether phrasal alternatives like *korte rok* for *rokje* 'short skirt', *geplooid rok* for *plooirok* 'pleated skirt', or *rok met omslag* for *wikkelrok* 'wrap-around skirt' should not be included in the calculation. After all, these are unique names just like *rokje*, *plooirok*, *wikkelrok* and the rest.

Figure 4.2(4) compares the monolexical and polylexical entrenchment values for the field of skirts. Only definitional entrenchment values are considered (and hence, no separate treatment is given to *rokje* and *minirok*). As a methodological specification, it should be mentioned that noun phrases like *geplooid rok* have only been included in the calculation if they occurred as such; according to the same principles used for the monolexical count, in fact, an expression like *korte geplooid rok* is to be considered a name for a subclass of pleated skirts, rather than for the entire class. The polylexical entrenchment values are obviously always higher than the monolexical ones, because they result from the addition of expressions like *geplooid rok* to the numerator of the monolexical entrenchment calculation. The monolexical and polylexical entrenchment values do not differ very much; the rank order of the categories is largely the same according to both counts. In what follows, focus will be on monolexical entrenchment, but the polylexical approach will play an important role in section 5.2., when the formal impact of entrenchment values will be discussed.

Item	Monolexical entrenchment		Polylexical entrenchment	
rok	54.85	rank 1	54.85	rank 1
doorknooprok	35.59	rank 2	35.59	rank 3
rokje-minirok	33.84	rank 3	37.43	rank 2
wikkelrok-omslagrok-overslagrok	32.43	rank 4	35.13	rank 4
plooirok	16.90	rank 5	27.53	rank 5
klokkrok	3.41	rank 6	9.40	rank 6

Figure 4.2(4)

Monolexical and polylexical definitional entrenchment values in the field of skirts

Having introduced the concept of entrenchment, there are four additional steps to be taken. One of these is relegated to section 5.2., where we shall explore the *importance of entrenchment for the phenomenon of lexical choice*. In particular, referents that lie in an area of overlap between two categories may be alternatively named by either of both. As such, the question arises whether the choice for one of the alternatives reflects the entrenchment of both. For instance, when something is both a *wikkelrok* 'wrap-around skirt' and a *plooirok* 'pleated skirt', is there a preference for calling such a thing either a *wikkelrok* or a *plooirok*, and does this preference reflect the relationship between the entrenchment measures of both categories? A second question to be dealt with involves the question whether there are any external, *contextual effects on entrenchment values*. For instance, do entrenchment values differ from one group of sources to another? This point will retain our attention in section 4.3.. Third, the question arises how the concept of entrenchment may be incorporated into the non-orthodox *graphical representations* of the structure of lexical fields that we introduced in section 4.1.. Representing the differences in entrenchment between the elements in the field is one further step away from mosaic-like representations. One way of achieving this is illustrated in Figure 4.2(5), in which Figure 4.1(6) is enriched on the basis of the definitional entrenchment values in Figure 4.2(3). (The thickness of the lines and the degree of foregrounding with which the categories are drawn reflects their relative entrenchment.)

Figure 4.2(5)

Graphical representation of the field of skirts, enriched with definitional entrenchment values

Item	Configurational entrenchment
broek	46.47
short-shorts	45.61
bermuda	50.88
legging-leggings-caleçon	45.50
jeans-jeansbroek-spijkerbroek	81.66
t-shirt	70.61
blouse-bloetze-bloes	61.52
overhemd	31.45
topje	29.62
shirt	29.06
hemd	22.31
overhemdblouse	12.74

Figure 4.2(6)

Configurational entrenchment values in the field of trousers and in the field of shirts

Finally, does the generalized notion of entrenchment as introduced in this section support the idea that the basic level model of onomasiological salience is of restricted value? Two sets of data suggest a *relativization of the basic level hypothesis*. First, the data in 4.2(3) and 4.2(4) show that entrenchment values on the same level of the taxonomical structure are not uniform, but may even differ with a factor 10 (compare the definitional entrenchment values of *klokrok* and *rokje*). As Figure 4.2(6) shows, the same lack of uniformity appears in other fields (viz. that of trousers and that of shirts). Second, the *broek*-field reveals that the entrenchment value of hyponyms may not only be of roughly the same magnitude as that of their common hyperonym, but may even exceed the latter: *jeans* is much more firmly entrenched than its superordinate *broek*. Both types of observation run contrary the predictions of the basic level hypothesis. On a vertical scanning of the taxonomy, the hypothesis suggests that basic level terms like *broek* 'trousers' and *rok* 'skirt' should always have a higher onomasiological

salience than their subordinates, but this is disconfirmed both by the fact that *short*, *bermuda* and *legging* have roughly the same entrenchment as *broek*, and by the fact that the entrenchment of *jeans* is much higher than that of *broek*. On a horizontal scanning of the taxonomy, the basic level hypothesis suggests that co-hyponyms should have entrenchment values of the same magnitude, but this is disconfirmed in all three fields whose entrenchment structure has been considered here. We conclude, in short, that the basic level model as a hypothesis about the distribution of differences of onomasiological salience in lexical fields is not universally valid.

4.3. The influence of contextual variation

In the previous section, entrenchment measures were calculated on the basis of the entire corpus. Now, could it be the case that entrenchments differ from one group of sources to the other? Already in the ethnolinguistic literature on basic levels, suggestions are made to the effect that the basic level in ethnobiological taxonomies need not invariably be the generic one. Rosch et al. (1976), for instance, point out that individual differences in familiarity and expertise may influence the salience of a particular taxonomical level. In particular, they suggest that there are two potential candidates for basicness: the life form level of *bird* may be basic for some, while the generic level of *gull*, *chicken*, and *turkey* may be basic for others. Dougherty (1978) has generalized this suggestion. After contrasting evidence to the effect that the basic level in Tzeltal folk taxonomies of plants is the generic one, with evidence suggesting that the basic level is suprageneric for urban Americans, she concludes:

The limits on the degree of specificity or generality at which the basic category cuts can occur must be considerably wider than has been postulated. Among a people for whom distinct kinds of trees are generally of minor importance, the most salient contrast in their classification of the botanical domain might fall at the life-form level where *tree*, *vine*, and *bush* contrast, or at an intermediate level where *leaf-bearing*, *needle-bearing*, and *frond-bearing* contrast. For the denizen of a future colony in outer space the contrast of *plant* and *animal* might well be more salient than any subordinate categories. Among

people for whom the distinctions among tree genera are consistently significant the most salient category cuts will tend to be folk generic. And more specific distinctions may be the most salient categories for highly significant members of the domain (Dougherty 1978: 77).

In our case, it seems unlikely that major cultural differences would influence entrenchment values along the geographical axis. The dimension of specialization, however, provides a good startingpoint for investigating the effect of differences in expertise. In fact, it may be noted that *two opposite hypotheses* about the effects of expertise have been formulated. On the one hand, it is a natural suggestion that a higher level of expertise and domain familiarity will lead to a downward shift of the basic level in the taxonomy: expert knowledge of the field would seem to result in a greater salience of the lower levels in the taxonomy. On the other hand, it has been suggested by Cruse that

a speaker may use underspecification to suggest that he is an expert in a particular field, or has at least an everyday familiarity with some class of things. He may, for instance, refer to a diamond as *a stone*, or a horse as *an animal*, or a violin as *an instrument* (1977: 163).

Cruse does not present his thesis in terms of the ethnolinguistic basic level hypothesis, but his notion of "neutral level of specificity" is an exact parallel of the concept of basic level. Perhaps (but Cruse does not develop his thesis) the idea behind the hypothesis may be expressed as follows: for the expert, diamonds are such a salient member of the extension of *stone* that *stone* will naturally refer to diamonds. In other words, the semasiological salience of diamonds within the extension of *stone* affects the onomasiological salience of that word.

Figure 4.3(1) presents the entrenchment values of a number of hyperonymous categories in various subparts of our corpus. The figures between brackets indicate the frequency of the lexical items. The geographical distinction between the sources is based on the location of the editorial office of the magazines; see section 2.1.. The entrenchment values are definitional ones, but the mean results of a configurational count have been included for comparison. (The mean configurational result for the glossies is mentioned between brackets, because removing incomplete configurations from the configurational count leads to very low absolute frequencies. In this sense, the configurational result for the glossy magazines may not be statistically

reliable.) Two important observations stand out: the entrenchment of the hyperonymous categories is generally higher in the specialized fashion magazines, whereas the geographical dimension only yields significant differences for the item *jasje* 'jacket'. Let us now have a closer look at each of these observations in turn.

	Fashion magazines	General magazines (NL)	General magazines (B)	Glossies
rok	75.00 (237)	56.52 (91)	41.04 (197)	48.38 (15)
broek	40.13 (183)	28.78 (137)	35.89 (243)	28.79 (74)
trui	66.66 (102)	48.25 (180)	44.29 (322)	42.99 (46)
jurk	77.96 (92)	42.57 (43)	38.81 (111)	24.65 (18)
jasje	48.49 (210)	9.37 (55)	32.08 (290)	12.86 (40)
mean	70.04	45.72	55.38	[5.27]

Figure 4.3(1)

Definitional entrenchments and mean configurational entrenchment for generic concepts in subsets of the corpus

The observation that the entrenchment values of the hyperonyms is higher in the specialized magazines seems to support the hypothesis formulated by Cruse: a rising level of expertise is reflected by an increased tendency to use unmarked higher level items. The observation is further illustrated by Figure 4.3(2), in which the entrenchment values of a number of hyponyms of one of the items in Figure 4.3(1) are compared. Although the results are somewhat skewed by the low frequencies of *wikkelrok*, *doorknooprok*, and *klokrok*, the general picture is as expected: the entrenchments of the hyponyms are lower in the specialized magazines. There is, however, a problem with an unmitigated endorsement of Cruse's hypothesis. The semasiological explanation of the onomasiological differences in salience, in fact, does not apply to our material. Note that the reasoning goes like this: for a diamond merchant, "diamond" is a semasiologically central sense of *stone*. So, if he wishes to express that something is a diamond, he can use the word *stone*, because to him, "diamond" is the primary referent of *stone*. Such an explanation is less

obvious when various applications of the same category are involved: it is counterintuitive to say that an expert may use *rok* 'skirt' to express the notion "miniskirt" because that is, to him, the central case of the category *rok*, and at the same time, to say that he may use *rok* to express the notion "pleated skirt" because that is *also* a central case in the structure of *rok*. In such a situation, *rok* no longer differentiates between the hyponymous concepts. Using *rok* to express both the concept "miniskirt" and the concept "pleated skirt" rather than, respectively, *minirok* and *plooirok* themselves is an inefficient creation of polysemy. Therefore, rather than assuming that *rok* when used by experts is polysemous between "miniskirt" and "pleated skirt" (and, in fact, the other hyponymous concepts of "skirt"), we may assume that the word is merely vague, and that it is actually being used in its superordinate reading.

	Fashion magazines	General magazines (NL)	General magazines (B)
rokje-minirok	11.50	39.72	41.46
plooirok	16.12	34.61	24.13
wikkelrok-	27.77	25.00	33.33
omslagrok-			
overslagrok			
doorknooprok	21.87	16.66	52.38
klokkrok	2.40	3.22	4.45

Figure 4.3(2)

Definitional entrenchments for *rok*-hyponyms in subsets of the corpus

But this leaves the question unanswered why expert sources exhibit a tendency to use the vague, superordinate concept more than the lower level concepts. It turns out that the difference can be explained by taking into account *pragmatic contextual variation involving the speech situation* (rather than the more or less permanent effects of familiarity and expertise of the speakers that were the startingpoint of the discussion). The explanation involves two observations. First, when one referent is referred to more than

once in the same context, more hyperonyms are being used in the second and following instances of naming the referent than when it is first mentioned. Second, there are more cases of multiple reference in the fashion magazines than in the general sources. Taken together, these observations explain the differences between the entrenchment values in the fashion magazines and the general sources. An example of such multiple naming is the following text (*italics indicate the relevant items*):

De warme dagen staan weer voor de deur ! Een kuitlange *doorknooprok* is dan onmisbaar. Als het wat koeler is, zijn *leggings* de ideale oplossing; de *rok* wordt dan niet dichtgeknoopt. De wijde blouse met de grote splitten opzij – voor de vlotte knoop – kan prima met de *rok* gecombineerd worden (*Burda* 9:83).

[Summer days are here again. A *button-through skirt* as long as the calves becomes indispensable. On cooler days, *leggings* provide warmth and the *skirt* is left unbuttoned. The loose-fitting blouse with the long slide slits – for making an elegant knot – nicely combines with the *skirt*.]

	Fashion magazines		General magazines (NL)	
	First occurrence	Second occurrence	First occurrence	Second occurrence
rok	67.95	83.77	51.62	70.83
trui	56.64	68.39	43.87	49.10
jasje	46.78	47.45	73.58	15.19
broek	30.40	41.93	28.54	41.85
jurk	51.19	88.25	36.31	60.43
mean	50.59	65.96	33.54	47.48

Figure 4.3(3)

Definitional entrenchments of generic concepts in first and following instances of use

	Fashion magazines		General magazines (NL)	
	First occurrence	Second occurrence	First occurrence	Second occurrence
rok	55.16 (187)	44.84(152)	64.08 (116)	35.92 (65)
trui	57.59 (110)	42.41 (81)	53.89 (235)	46.11 (201)
jasje	45.16 (266)	54.84 (323)	71.66 (473)	28.34 (187)
broek	58.80 (318)	41.20 (222)	76.89 (426)	23.11 (128)
jurk	59.71 (83)	40.29 (56)	52.75 (67)	47.25 (60)
total	53.61 (964)	46.39 (834)	67.26 (1317)	32.74 (641)

Figure 4.3(4)

Frequency of first and following instances of reference to generic categories

The first step in the explanation is corroborated by the data in Figure 4.3(3). The shift from subordinate to superordinate terms can be explained in at least two ways. First, when a referent is first mentioned, the identificational requirements are highest, and a more specific description may be used to ensure maximal identifiability of the referent in question. Once an item has been identified, however, a less specific name may be used because the reader may need less identificational help. Second, a Gricean tendency towards maximizing information might be invoked in combination with a principle of economy. Regardless of whether the additional information is necessary to identify the referent, there might be a tendency to conform to the maxim "be as informative as necessary". But once the information has been specified, it need not (on economic grounds) be repeated every time the referent is mentioned. Our data do not allow us to make a definitive choice between both alternatives, but note that the identificational hypothesis is

mainly relevant when naming a referent at the generic level is insufficient to identify the garment. That is only the case when, for instance, more than one skirt appears in the picture to which the text refers, and that is an extremely rare situation in our corpus.

The second step in the explanation is corroborated by the data in Figure 4.3(4). (The numbers between brackets specify absolute frequencies. The main numbers give the percentages of first and following occurrences within one group of sources.) The higher number of second references in the fashion magazines is easily explained by the fact that these sources provide more information about the referents than the general sources. Their higher level of specialization is reflected by the presence of more specific descriptions of the garments; in particular, they often contain patterns with detailed instructions.

	General magazines (B)	General magazines (NL)
jasje	42.15	24.50
blazer	19.17	39.47
colbert	12.96	24.06
vest	27.12	49.10
jack-blouson	58.53	68.75

Figure 4.3(5)

Geographical differences in the configurational entrenchment of *jasje* and its hyponyms

The second major observation made with regard to Figure 4.3(1) involved a case of *geographical variation*. The observation that the entrenchment value of *jasje* is higher in the Belgian sources is corroborated by the data in 4.3(5), in which the entrenchment values of a number of hyponyms of the *jasje*-category are distributed over the geographical dimension. Note that the data in 4.3(1) are based on a definitional entrenchment measure, and those in 4.3(5) on a configurational measure (due to problems with the definability of the hyponyms – compare section 3.2.) Although the data concerning *vest* should be handled with care (as *vest* does not cover precisely the

same ground in Belgium and in The Netherlands – compare section 3.5.), a consistent pattern emerges: while the entrenchment of *jasje* is higher in the Belgian sources, that of the hyponyms is each time higher in the Netherlandic sources than in the Belgian ones. This is precisely as might be expected. If the higher entrenchment of *jasje* in the Belgian sources indicates a relative preference for a “hyperonymous” identification of jacket-like garments, the hyponymous terms are likely to suffer a correlated drop in attractiveness (and vice versa for the Netherlandic sources).

It can be deduced from Figures 4.3(1) and 4.3(2) that major geographical differences of entrenchment like those noted for *jasje* do not seem to be extremely frequent. In particular, there is no general tendency for Belgian Dutch to prefer the hyperonymous level. (We suggest that the results for *doorknooprok* in 4.3(2) are a statistical distortion due to the low absolute frequencies of the item in question.) To summarize, then, contextual differences of entrenchment do occur, but in the form of permanent, speaker-related effects of a sociolinguistic or geographical nature they do not appear to be widespread. Rather, the major differences that we have found appear to involve pragmatic, situational factors rather than sociolinguistic or geographical ones.

Chapter 5

Formal variation

In the previous chapters, we have been concerned with semantic phenomena: the semasiological variation that exists within the range of application of a lexical item, and the onomasiological variation that shows up in the fact that a particular referent or set of referents may be alternatively named by means of distinct but related semantic categories. In this chapter, we turn towards aspects of formal variation: the fact that several lexical items may be used to name a referent. Such formal variation will, of course, have to be studied from the “sociolinguistic” perspective of the influence of what we have called “contextual” factors, but the influence of semantic factors on the matter of lexical choice should be envisaged as well. Therefore, before treating the influence of contextual variation in the final section of this chapter, we will consider the question to what extent semasiological and onomasiological variation of the kind discussed in the previous chapters has an effect on the formal variation within the lexicon. In addition to the question whether these semantic factors influence the lexical choices made in naming a referent, we will consider the question whether the form of lexical expressions correlates with these semantic factors. (It is, for instance, a plausible question whether categories with high onomasiological entrenchment values are less often named by means of compounds and derivations than categories with low entrenchment measures, or whether peripheral, non-prototypical members of an item’s range of application are more readily identified by means of a polylexical expression consisting of a noun plus modifiers.)

A consideration of these semantic factors as determinants of formal phenomena is not just important for descriptive completeness, but has a distinct theoretical importance. It is customary in linguistic analyses to prove the importance of particular notions by pointing out that they determine, or at least correlate with certain formally observable characteristics of the language. Therefore, if there should be any doubt left about the importance of non-classical lexical characteristics like prototypicality and onomasiological

salience, the respectability of these semantic notions may be established by pointing out that they have specific reflections on the formal side of the language.

5.1. The influence of prototypicality

It is an intuitively plausible idea that a word will be used more often for naming a particular referent when that referent is a member of the prototypical core of that word's range of application. When, for instance, a particular referent belongs to the core of item *x* but to the periphery of *y*, it is to be expected that *x* will be a more likely name for that referent than *y*. This is not, to be sure, necessarily so. In principle, the likelihood of the appearance of *x* could be just as big as that of the appearance of *y*. For instance, the overall onomasiological entrenchment of *y* (as discussed in section 4.2.) might raise the onomasiological attractiveness of that item to such a degree that *x* becomes the less likely name. In this section, the relation between the semasiological status of a referent within the range of application of an item, and the onomasiological status of that item with regard to that referent will be investigated in detail. (The additional effect of entrenchment on onomasiological naming preferences will be considered in the next section.)

Terminologically speaking, what is at stake here is the correlation between the semasiological status of a referent (or a group of referents) with regard to a lexical item, and the onomasiological status of that item with regard to the referent (or group of referents) – the relationship, in other words, between the prototypicality structure of an item and its onomasiological cue validity structure. Cue validity as defined in psychological research is the ratio between the frequency with which a cue is associated with a category, and the total frequency of the cue in the material. Thus, the cue validity of a referent (the cue) with regard to a word is the ratio between the number of times the referent is named with that word, and the global frequency of the referent. Cue validity may then be interpreted as an indication of the probability that a particular word will be used as a name for a particular referent: given a particular referent, what is the chance that a particular item will be used as a name for that referent? Two theoretical points have to be clarified here.

First, the concept of cue validity may be used in various ways in the

study of prototypicality phenomena. In particular, note that the relationship between intensional and extensional prototypicality effects that was the subject matter of section 3.4., may be described in terms of cue validity effects. If the frequency with which a particular referent r receives the name x is taken as the relevant cue, then the ratio between that frequency and the total frequency of x is a cue validity measure: it specifies the probability that r will occur in the semasiological range of x . This probability is, of course, an indication of the extensional prototypicality structure of x : the referents with the highest cue validity are the ones with the highest relative frequency within the range of application of x , i.e., they are the extensionally central members of x . What we did in 3.4., then, was to show that this kind of extensional prototypicality, defined as a particular cue validity measure, correlates with an intensional kind of prototypicality, defined in terms of the attributes shared by the referents under investigation (or, if one wishes, in terms of the overlapping of the definitional subsets that may be defined within the semasiological range of the item). What we propose to do in this section, on the other hand, differs fundamentally from the way in which the notion of cue validity was used in section 3.4. Specifically, the cue validity measure is defined differently. If the frequency with which a particular referent r receives the name x is taken as the relevant cue, we will now consider the ratio between that frequency and the total frequency of r , rather than the ratio between the frequency of the pair x/r and the total frequency of x . In this way, we get an onomasiological rather than a semasiological cue validity measure. This onomasiological cue validity measure will then be correlated with the semasiological prototypicality structure of the items as discussed in section 3.4.: we will investigate whether the onomasiological cue validity of the central subsets of lexical items is higher than that of peripheral subsets.

Although an onomasiological cue validity measure occurs only rarely in the original psycholinguistic literature on prototypicality, it is not entirely absent. In particular, attention may be drawn to Experiment 2 from Rosch & Mervis (1975). Whereas Experiment 1 as described in the article characteristically involves the semasiological correlation between an extensional measure of membership frequency and an intensional measure of family resemblance among the attribute sets of the category members, Experiment 2 onomasiologically involves alternative denominations for the same item. For given concrete nouns like *collie*, for instance, subjects were asked to list three superordinate categories (like *dog*, *animal*, or *pet*). A weighted measure for category dominance was introduced; it would indicate, for instance,

that *dog* is a more dominant category with regard to *collie* than *pet*. In the terms used earlier, this could be expressed by saying that the onomasiological cue validity of *collie* with regard to *dog* is higher than with regard to *pet*: *dog* is a more likely category for talking about collies than *pet*. In the experiment reported on by Rosch & Mervis, a correlation was found between this onomasiological cue validity measure ("category dominance" in their terms), and prototypicality, i.e. membership typicality. The items that are most representative for a category are strongly dominated by that category, i.e. they more easily receive that category as a name than any alternative category for which they are less representative.

The hypothesis tested in this section basically involves the same correlation. The difference with the original Rosch & Mervis research resides in two points of methodology. First, we use non-elicited corpus material rather than experimentally obtained data. And second, the stimuli that triggered the denominations recorded in the corpus consist of actual referents (garments) rather than words like *collie*. While the first distinction merely characterizes our research as a typically linguistic rather than psycholinguistic one, the second point is of more importance, because it avoids a potential source of distortion in the original Rosch & Mervis design. Starting from verbal stimuli like *collie* implies that *collie* itself is not included in the range of possible onomasiological alternatives for naming collies. But precisely because this is likely to be one of the most "dominant" alternatives, it may be worthwhile to include it in the investigation.

A second preliminary remark involves the relationship between the notion of entrenchment as used in the previous chapter, and the onomasiological cue validities studied here. Although there is an obvious similarity between both (both measure categorial frequencies in relation to global frequencies), there are two differences. A minor difference is the fact that entrenchment relates to concepts (words plus their synonyms), whereas onomasiological cue validities will only be computed for single words. In most cases to be discussed, the distinction will not be relevant, because the items under investigation do not occur with clear synonyms. A major difference is the fact that entrenchment values relate to categories as a whole, whereas the onomasiological cue validities of this section relate to subsets of a category (individual referents or groups of referents). In this sense, the cue validities are a kind of distributed entrenchments – distributed, that is, over the various subsets of a lexical item.

(Added) set of configurations	Cumulative semasiological frequency per set	Cumulative onomasiological cue validity per set
H5118v, H4118v	99	46.47
H5211v, H4211v	105	25.17
H5128v, H5154v, H3118v	110	25.52

Figure 5.1(1)

Comparison between prototypicality and onomasiological cue validity for the item *legging*

In Figures 5.1(1) through 5.1(5), the hypothesis that there is a correlation between the prototypicality structure of an item and the onomasiological cue validity of its subsets is tested on the basis of five case studies, involving the items *legging*, *colbert*, *blazer*, *vestn*, and *vestb*. The items have been chosen from among those for which box representations were drawn in sections 3.4. and 3.5.. On the basis of these diagrams, sets of referential configurations can be distinguished according to their centrality in the semasiological structure of the item. In Figure 5.1(1), for instance, three subsets are singled out by starting from the maximally overlapping area in Figure 3.4(1), and gradually enlarging this core area towards less central cases. The more we move away from the core of the category, the less attributes the items in the subsets have in common; the definitions become more diffuse, disjunctive, family-resemblance-like. At the same time, the relative frequency within the item of each additional set of configurations diminishes. For each of the subsets that are distinguished in this manner, Figures 5.1(1)-5.1(5) give the cumulative frequency per set within the item under consideration, and the

(Added) set of configurations	Cumulative semasiological frequency per set	Cumulative onomasiological cue validity per set
B2121m	53	72.60
B2121v, B2122v, B2131m B2111m	82	33.60
B1121m B1122v, B2212m, B2211m, B2221m, B2222m, B2222v	113	29.89
C3312v, C3311v, C3212m	118	24.48

Figure 5.1(2)

Comparison between prototypicality and onomasiological cue validity for the item *colbert*

onomasiological cue validities of the cumulated sets. It is crucial for an understanding of the figures that the semasiological frequencies and the onomasiological cue validities involve cumulated sets of configurations. In Figure 5.1(1), for instance, the cue validity value 25.17 relates to the set constituted by the configurations [H5118v], [H4118v], [H5211v], and [H4211v] together. The cue validity value 25.52 is calculated after adding [H5128v], [H5154v], and [H3118v] to the previous set.

If there is a correlation between the prototypicality structure of an item (either extensionally in terms of the frequencies of member configurations, or intensionally in terms of the overlapping of definitional subsets), and the

(Added) set of configurations	Cumulative semasiological frequency per set	Cumulative onomasiological cue validity per set
B2122v	48	39.02
B2132v, B2112v, B2212v, B2232v, B2222v	114	36.36
C2332v, C1312v, C2311vv B1122v, B1222v	137	22.82
C2212v, C2312v, C3312v, B2121v, B2121m, B2221m, C3432v, C3332v, C3412v, C3212v, C2412v	189	19.16

Figure 5.1(3)

Comparison between prototypicality and onomasiological cue validity for the item *blazer*

tendency to use that item for naming particular referents or groups of referents, then the onomasiological cue validities of the subsets distinguished in Figures 5.1(1)-5.1(5) should gradually diminish. The rightmost column of the figures shows precisely what is expected: the farther one moves away from the centre of the category, the lower becomes the onomasiological cue validity of the set in question. By incorporating referents that are less core-like, the likelihood of a referent being named by the item diminishes. The

(Added) set of configurations	Cumulative semasiological frequency per set	Cumulative onomasiological cue validity per set
C2332, C2322	35	53.03
C3332, C3322	52	58.42
C2312, C3312, C3342, C3422, C3432	78	49.68
C3242 C1331, E3331122	81	49.09

Figure 5.1(4)

Comparison between prototypicality and onomasiological cue validity for the item *vestn*

effect occurs both in the case of items that do have a classical definition, like *legging*, and items that do not, like *colbert* (see section 3.2.).

Although the overall picture arising from Figures 5.1(1)-5.1(5) supports the initial hypothesis, there is an important nuance to be added. Notice, to begin with, that the changes in the cue validity values are not a straightforward consequence of the changes in the semasiological prototypicality structure. For instance, the spectacular decrease of the cue validity measure from the first to the second set of configurations distinguished in the case of *colbert* is not repeated in the case of *vestn*, although the semasiological position of the subsets is roughly similar. The central set of *colbert* represents 43,2% of the entire range of that item, whereas the central set of *vestn* represents 44,9% of the total. The second sets represent 69,4% and 64,1% respectively. Both in terms of maximal overlapping and in terms of relative

(Added) set of configurations	Cumulative semasiological frequency per set	Cumulative onomasiological cue validity per set
C2322, C2332	34	26.98
C3322, C3332, C1322, C1332	65	25.17
C2232, C2132, C2122, C2312, C3432, C3422, C1222, C3342, C3312	81	31.39
C2331, C2311, C3311, C2442, C2112, C2242, B2232, C3331, B1132, B2131, B2122	101	27.82
C3211, C3111, A214, A484, B1222, B2222	107	27.36

Figure 5.1(5)

Comparison between prototypicality and onomasiological cue validity for the item *vest*_b

frequencies, then, the structural position of the subsets under consideration is similar. Still, the change in the onomasiological cue validity from the first subset of *vest_{nl}* to the second far from mirrors the corresponding change in the case of *colbert*. On the contrary, there even is a slight increase of the cue validity (a phenomenon that can also be observed in the case of *vest_b*). One conclusion to be drawn from this observation is that the onomasiologically relevant prototypical core of *vest_{nl}* and *vest_b* had perhaps better be defined in terms of the first two subsets in the figures rather than in terms of the first subset alone. Taking a slightly wider perspective, however, leads to a more far-reaching conclusion: the cue validity structure of some items is much flatter than that of others. On the one hand, *vest_{nl}* and *vest_b* have a very flat structure, in the sense that the onomasiological attractiveness of the central portions of the item (whether defined in terms of the first subset distinguished in the figure, or in terms of the first and second subset) does not differ dramatically from the cue validity of the item as a whole. On the other hand, *colbert* exhibits a steep structure, with major differences between the maximum and minimum cue validity values; *legging* and *blazer* occupy a middle position. It appears, then, that the effect of semasiological prototypicality on onomasiological naming preferences is not always equally strong. We have found no cases for which the cue validity of the core of the item is decidedly lower than that of the word as a whole, but there do appear to be cases, like *vest_{nl}* and *vest_b*, in which the difference between the cue validity of the core area and that of the entire range of application is minimal.

The positive correlation between prototypicality and onomasiological cue validity establishes that *there exists a tendency for referents to be preferentially named by means of a category to which they typically belong*. The choice of a particular expression for naming a particular type of referent involves more, however, than merely the choice of a major category. Taking into account that referents are often named by means of full noun phrases rather than just a single noun, the question arises whether the choice of the modifiers accompanying the head of the noun phrase is also influenced by the semasiological structure of the item. Intuitively, this is a plausible idea: given that, for instance, the characteristic of having two legs is an intrinsic feature of *broek* 'trousers', it is unlikely that *broek* will be modified by an adjective expressing the concept "two-legged"; conversely, since trousers do not have a typical color, adjectives indicating the color of a particular pair of trousers may be expected to occur relatively frequently. Formulating this intuitive expectation in a generalized way, *an inverse correlation may be expected between the prototypicality of a particular feature for a particu-*

lar item and the frequency with which that feature is expressed in a noun phrase having the item as its head. This general hypothesis may be specified in two ways, according to whether it refers to featural dimensions as a whole, or to the specific dimensional values that occur on that dimension.

In the first case, a distinction is made between those dimensions that play an intrinsic role in the semasiological structure of an item (such as two-leggedness or length in the case of *broek*), and those that do not (such as color in the case of *broek*). Adjectives (or other modifying phrases) expressing concepts such as "red" or "blue" are then lumped together in contrast with adjectives expressing concepts such as "two-legged" or, on the dimension LENGTH, "short" or "long". The hypothesis obviously specifies that modifiers in polylexical expressions identifying a particular type of referent will more often express accidental dimensions such as color than intrinsic dimensions such as two-leggedness. Clearly, the intrinsicness of a dimension may be a matter of degree. To keep matters simple, however, the following calculation will only contrast dimensions that are clearly accidental for a particular item, with dimensions that clearly do play a role in the definition or the prototypical structure of an item.

In the second case suggested above, the attention is focused on the intrinsicness of the various values of those dimensions that do play a structural role in the semasiological make-up of an item. For instance, given that LENGTH is structurally important for *broek*, it may be observed that the members of the category *broek* are typically long; for adults, short trousers are not prototypical. It may be hypothesized, then, that the prototypical, more intrinsic value "long" will be expressed less often in polylexical expressions than non-prototypical values such as "short".

The first hypothesis is tested in Figure 5.1(6). For each of the items *broek*, *legging*, *colbert*, and *blazer*, the degree of polylexical expression of three structurally important dimensions is compared with that of one structurally irrelevant characteristic, viz. color. (The structural importance of the dimensions in question is reflected by the fact that they exhibit a marked asymmetry in the frequency with which their various values occur. For instance, the prototypical *legging* is tight and as long as the ankle. By contrast, there is no typical color for the referents of *legging*. With regard to the MOTIF-dimension, a plain specimen of the clothing type is always the dominant, unmarked case.) The figure next to the name of the item involved specifies the number of referents in the semasiological range of the item. The figure in the second column from the left specifies the percentage of polylexically expressed dimensions in relation to the total number of referents of

Dimension	% of polylexical expressions	Examples
<i>blazer</i> (n=242)		
width	8.3	ruimvallende blazer
fastening	2.1	blazer met een dubbele knoopsluiting
motif/pattern	7.9	blazer met krijtstrepen, blazer met bloemdessin
color	23.6	bruine blazer
<i>colbert</i> (n=166)		
width	4.2	getailleerd colbert
fastening	4.8	double breasted colbert
motif/pattern	7.8	geruit colbert, pied-de-poule colbert
color	22.3	rood colbert
<i>legging</i> (n=142)		
length	0.7	kuitlange legging
width	2.8	strakke legging
motif/pattern	21.1	gestreepte legging, legging met bloemen
color	26.1	zwarte legging
<i>broek</i> (n=638)		
length	6.3	korte broek
width	17.2	nauwsluitende broek
motif/pattern	5.6	broek van ruitjesstof
color	23.2	broek in pasteltinten

Figure 5.1(6)

Polylexical expression of dimensions with and without prototypical values

Dimensional attribute	Referents with the attribute (%)	Polylexical expressions (%)	Examples
<i>blazer: width</i>			
waisted	12.9	34.6	getailleerde blazer
wide	87.1	1.7	ruimvallende blazer
<i>blazer: fastening</i>			
double breasted	35.0	4.8	blazer met een dubbele knoopsluiting
single breasted	65.0	0.0	—
<i>blazer: motif/pattern</i>			
with pattern or motif	20.2	34.7	blazer in pepita-ruit
plain	79.8	1.0	effen blazer
<i>colbert: width</i>			
waisted	6.9	45.5	getailleerd colbert
wide	93.1	0.7	ruimvallend colbert
<i>colbert: fastening</i>			
double breasted	20.5	17.4	double breasted colbert
single breasted	79.5	3.4	single breasted colbert
<i>colbert: motif/pattern</i>			
with pattern or motif	26.5	27.3	geruit colbert

plain	73.5	0.8	effen colbert
<i>legging: length</i>			
knee/calf	36.4	2.5	kuitlange legging
ankle	63.6	0.0	—
<i>legging: width</i>			
straight	4.9	14.3	rechte legging
tight	95.1	1.5	strakke legging
<i>legging: motif/pattern</i>			
with pattern or motif	45.1	45.3	gedessineerde legging
plain	54.9	0.0	—
<i>broek: length</i>			
groin/thigh/knee	3.8	33.3	knielange broek
calf	15.5	12.9	drievierde broek
ankle	80.7	3.2	lange broek
<i>broek: width</i>			
tight	13.0	19.1	strakke broek
wide	36.7	15.6	wijde broek
straight	50.2	3.2	rechte broek
<i>broek: motif/pattern</i>			
with pattern or motif	21.2	24.4	broek met krijtstreep
plain	78.8	0.8	effen broek

Figure 5.1(7)

Polylexical expression of dominant and subordinate dimensional values

the item, i.e., it specifies the number of times the dimension is expressed as a modifier in a polylexical expression. Examples of the expressions involved are given in the rightmost column of the figure.

The second hypothesis may be confirmed on the basis of the data given in

Figure 5.1(7). For each of the structurally asymmetrical dimensional values included in Figure 5.1(6), the number of referents with that value is specified as a percentage of the total number of referents of the item, together with the number of times the value is explicitly expressed as a modifier in a polylexical expression (given as a percentage of the number of referents with that value). In some cases, dimensional values that are marked separately in the componential descriptions of the referents are lumped together to bring out the dimensional asymmetry more clearly; this is the case with some of the dimensions that form a continuum (such as LENGTH in the case of *legging*, where two values are taken together).

Already at a first glance, the hypothesis appears to be confirmed. On the dimension WIDTH of *blazer*, for instance, there is a marked dominance of wide jackets in contrast with waisted ones; as expected, polylexical expressions specifying the non-prototypical value (examples are given in the right hand column of the figure), are relatively more numerous than expressions specifying the prototypical value. This pattern occurs throughout the data in the figure.

Figure 5.1(7) also allows us to explain the apparent exceptions in 5.1(6). The two dimensions in question are precisely the ones that exhibit a less asymmetrical distribution of values than the other ones in the figure; the unmarked values of the dimension MOTIF of *legging* and the dimension WIDTH of *broek* represent 54.9 and 50.2 percent of all relevant cases, which is much less than the average of the unmarked values on the dimensions in 5.1(7) (viz. 75.1%). If this observation is interpreted as an indication of the fact that the dimensions MOTIF of *legging* and WIDTH of *broek* have less impact on the determination of the unmarked, prototypical members of the items in question, it is only to be expected that they will be polylexically expressed more often than the other dimensions, whose default option carries more structural weight.

5.2. The influence of entrenchment

The onomasiological cue validity measure that was introduced in the previous section gives an indication of existing naming preferences: when an item *x* has a higher cue validity with regard to a referent or set of referents *r* than *y* or any of the other names that apply to *r*, *r* will be named more often by

means of x than by any of the alternative terms. The results of the previous section show, then, that lexical choices are determined by the semasiological status of the referents to be named with regard to the various names that apply to it: in many cases at least, r has a preferential tendency to be named by x if r is a central member or subset of x . In plain language: when you have to name something, you preferentially choose those items of which the thing to be named is a typical representative. Intuitively, however, it is implausible that this is the only factor involved in making lexical choices. For instance, if a particular piece of clothing is an impeccable representative of the category of pleated skirts, the name *rok* 'skirt' may still be given more often than the degree of prototypicality with regard to *plooirok* 'pleated skirt' would seem to warrant, merely because *rok* as such is a much more frequent category than *plooirok*. The prototypicality-based attractivity of *plooirok* is then so to speak overruled by the attractivity of *rok*. In more technical terms, the higher onomasiological entrenchment (as discussed in 4.2. and 4.3.) of *rok* with regard to *plooirok* interferes with the prototypicality effects of the previous section.

Can it be established, then, that *differences of entrenchment influence lexical choices*? The data in Figure 5.2(1) may help to show that this is indeed the case. The basic idea is to investigate naming practices with regard to the set-theoretical sections of two items, i.e. the set of referents where both items overlap. The hypothesis is, obviously, that the relationship between the entrenchments of both items correlates positively with the relationship between the frequencies with which the overlapping area is named by one or the other item. For statistical security, the items have been chosen in such a way that the absolute frequency in the overlapping area is relatively high. (Because the entrenchment values are based on a configurational calculation, *rokje* and *minirok* are not treated as synonyms. The entrenchment values are given in percentages. Because we are dealing with entrenchment values, x and y are categories rather than simple lexical items, i.e., the entrenchment values pertain to sets of synonymous lexical items. Note, in particular, that the row comparing *legging* and *broek* actually involves data for *legging*, *leggings*, and *caleçon*.) The following abbreviations are used:

f_r = the total number of referents in the overlapping area r

f_{rx} = the number of times the referents in the overlapping area r are named by the item identified by x

f_{ry} = the number of times the referents in the overlapping area r are

named by the item identified by y

E_x = the entrenchment of item x

E_y = the entrenchment of item y .

x	y	f_{rx}	f_{ry}	E_x	E_y	E_x/E_y	f_{rx}/f_{ry}
hemd	overhemd	42	32	31.45	22.31	1.409	1.312
overhemd	blouse	43	128	31.45	61.52	0.511	0.335
hemd	blouse	22	85	22.31	61.52	0.362	0.258
topje	blouse	11	26	29.62	61.52	0.481	0.423
minirok	rokje	8	21	18.03	43.71	0.412	0.380
blazer	colbert	135	119	18.84	24.53	0.767	1.133
short(s)	bermuda	72	62	45.61	50.44	0.904	1.161
broekrok	bermuda	30	84	32.75	50.44	0.649	0.357
plooirok	rok	17	40	26.98	62.53	0.432	0.425
legging	broek	173	152	45.50	46.47	0.978	1.138

Figure 5.2(1)

Frequencies and entrenchments in overlapping areas

The hypothesis mentioned above can now be translated into the statement that a high, positive correlation may be expected between f_{rx}/f_{ry} and E_x/E_y . Note that the ratio f_{rx}/f_{ry} measures the relationship between the onomasiological cue validities of x and y . In fact, the onomasiological cue validity of x with regard to r equals f_{rx} divided by the total number of times that the members of r occur in the corpus, and the onomasiological cue validity of y with regard to r equals f_{ry} divided by the total number of times that the members of r occur in the corpus; as the denominator in both cases is identical, f_{rx}/f_{ry} specifies the ratio between both cue validities. Given that the onomasiological cue validity value measures naming preferences, the correlation between f_{rx}/f_{ry} and E_x/E_y specifies the extent to which entrenchment values influence lexical choices. The hypothesis that there is a high, positive correlation between f_{rx}/f_{ry} and E_x/E_y is confirmed: a correlation of 0.87641 with a significance level of 0.0043 is found.

The results of the previous section and this one can be summarized in the following two statements. On the one hand, if a lexical item w has a particu-

lar referent r as one of its core members, w will be a preferred name for r . On the other hand, if a lexical item w is more strongly entrenched than any of the alternative names for a referent r of w , w will be a preferred name for r . The summary is somewhat inaccurate to the extent that the distinction between items and categories (i.e. items plus their synonyms) is disregarded, but this formulation has been chosen to bring out the similarity between both results. In both cases, in fact, factors have been identified that determine the lexical choices made in naming a particular referent or set of referents. Roughly, *an item w is more readily chosen as a name for a referent r to the extent that (onomasiologically speaking) w is more strongly entrenched than its alternatives, and to the extent that (semasiologically speaking) the prototypical structure of w includes r as one of its more central members.*

The next step in the investigation will obviously be to determine the relationship between both factors. We will not endeavor to determine a quantitative measure for the relative strength of each factor, but a single case study may illustrate the phenomena involved. Using the same terminology as before, the following prediction may be formulated, given the idea that prototypicality and entrenchment both influence the choice of a particular expression: if a set of referents r that constitutes the area of overlap of the items x and y contains members that are prototypical for x next to members that are not prototypical for x , the ratio fr_x/fr_y will be higher for the former subset and lower for the latter subset, in comparison with the ratio fr_x/fr_y as calculated for r as a whole. The easiest way of testing the prediction is to have a look at one of the cases in Figure 5.2(1) in which r coincides with the referential range of x as a whole, i.e., a case in which x and y form a hyponymous/hyperonymous pair. As a case in point, Figure 5.2(2) list the results for *legging-leggings-caleçon* in comparison with the hyperonym *broek*. In accordance with the hypothesis, core referents of *legging-leggings-caleçon* (identified by means of the configurations [H4118v] and [H118v]) are more often named by means of any of these three items, and less often by means of the hyperonym; peripheral referents, on the other hand, are more often identified by means of *broek*. This is, needless to say, a straightforward extrapolation of the results obtained in the previous section. There, it was shown that peripheral members of an item's referential range of application are more likely to be named by any of the alternative terms that are available for them; in Figure 5.2(2), this result is specified with regard to one of those alternative terms, viz. a hyperonym.

Differences of onomasiological entrenchment do not only influence the choice of a name for a particular referent, they also have an effect on the

formal characteristics of those names. As was mentioned in section 4.2., the concept of a basic level organization of ethnobiological classifications included the idea that the basic level terms can be linguistically characterized as primary lexemes, i.e. relatively short, monomorphemic lexemes. Although we rejected the basic level hypothesis as such as an adequate model of entrenchment (as operationally defined in terms of onomasiological salience), the idea in itself that there is a correlation between the entrenchment of a category and the formal characteristics of its linguistic expression still stands. Even though there is not a single taxonomical level where all the highest entrenchment values are situated, more strongly entrenched categories could still correspond with "primary" lexemes. We will now present two kinds of data to support this hypothesis.

	frx	fry	frx/fry
<i>legging-leggings-caleçon</i> as a whole	173	152	1.138
core	153	34	4.500
periphery	20	120	0.166

Figure 5.2(2)

Naming frequencies in subsets of an overlapping area

In the first place, the relationship may be studied between the proportion of simplex forms available for a particular category (defined over the total number of unique expressions available for that category, such as it constitutes the numerator of the entrenchment ratio), and the entrenchment value itself. The expectation is that the proportion of simplex forms (in contrast with polymorphemic items such as compounds and derivations) falls as the entrenchment values fall. In Figure 5.2(3), fifteen categories have been brought together: five for which the configurational entrenchment value exceeds 50%, five for which it lies between 40 and 50%, and five for which it is less than 40%. The rightmost columns of the figure specify the average percentage of simplex forms for the five items of each class, and the overall

Category	Entrenchment value	Average % of simplex forms	Overall % of simplex forms
jeans-jeansbroek- spijkerbroek t-shirt blouse-bloetze-bloes rok bermuda	81.56 70.61 61.52 54.85 50.88	95.48	61.08
rokje doorknooprok broek short(s) legging	47.89 47.50 46.47 45.61 45.50	59.88	40.34
wikkelrok-omslagrok- overslagrok overhemd topje shirt plooirok	39.02 31.45 29.60 29.06 26.89	20.00	6.67

Figure 5.2(3)

Frequency of simplex forms in relation to configurational entrenchment values

percentage of simplex forms per class; the latter is obtained by treating each set of five cases as if it were a single category. In the first row, for instance, *jeans* is a simplex form, whereas *jeansbroek* and *spijkerbroek* are counted as polymorphemic items. The average percentage of simplex forms is obtained by averaging the frequencies with which simplex forms appear in each of the five cases separately. The overall percentage (in the rightmost

column) is obtained by adding up the frequencies of the simplex forms *jeans*, *t-shirt*, *rok* etc., and dividing that cumulative frequency by the total number of names (both simplex and complex ones) in the group.

The figure shows clearly that *more highly entrenched categories are more likely to be named with simplex forms*. (It should be remarked that there might be a problem with the decision to treat a particular item as polymorphemic or not. In the present sample, for instance, *t-shirt* could be a case for doubt. However, although it is a complex form in English, and even though the loanword *shirt* does occur in Dutch, *t-shirt* itself can hardly be considered a compound from the point of view of Dutch, if only because the *t*-element is invariably pronounced as /ti/, as in English, rather than as the corresponding Dutch /te/.)

In the second place, for those categories for which polylexical entrenchment values were calculated in section 4.2. (see Figure 4.2(4)), a three-way classification may be taken into account between simplex forms, polymorphemic items (either compounds or derivations), and polylexical expressions. For each of these three classes, the proportion they occupy within the total set of expressions that uniquely refer to the category under consideration may be listed. *Rok*, for instance, has no synonyms, so that the "simplex" class totals 100% of the expressions for the category "skirt" that together define the numerator of the polylexical entrenchment measure. The entrenchment value for "short skirt", on the other hand, takes into account the frequencies for a compound (*minirok*), a derivation (*rokje*), and a polylexical expression (*korte rok*). Assuming that polylexical identification is one more step away from "primariness", the expectation is that the number of compounds and derivations, and the number of polylexical expressions will rise as the entrenchment values diminish. As shown in Figure 5.2(4), the expectation is by and large confirmed by the data: *more highly entrenched categories are less likely to be named by means of polylexical expressions*. Most tellingly, the spectacular drop of the entrenchment value for *klokrok* is mirrored by a steep rise in the number of polylexical expressions. Because the data in Figure 5.2(4) are based on a smaller sample, they should be treated as complementary with regard to those of 5.2(3); but although they may be less outspoken than those of 5.2(3), they do strongly suggest that less entrenched categories are more readily expressed by means of polylexical expressions. (Note that the polylexical entrenchment values as given in Figure 4.2(4) and repeated in 5.2(4) are based on a definitional rather than a configurational measure of entrenchment.)

Category (cf. 4.2(4))	Poly- lexical en- trench- ment	% com- pounds and deri- vations	% poly- lexical expres- sions	Relevant polylexical expressions
rok	54.85	0	0	—
rokje/minirok	37.43	90.4	9.6	korte rok
doorknooprok	35.59	100	0	—
wikkelrok- omslagrok- overslagrok	35.13	92.3	7.7	rok met omslag, rok met overslag
plooirok	27.53	89.5	10.5	geplooide rok, rok met plooien
klokrok	9.40	36.4	63.6	klokkende rok, ruimvallende rok, ruime rok, wijde rok

Figure 5.2(4)

Percentages of polylexical expressions in relation to
polylexical entrenchment values

To round off this section, let us briefly spell out the systematic connections among the various lines of enquiry that we have so far pursued. First, we have systematically distinguished between the internal, semasiological structure of lexical items, and the supralexical, onomasiological semantic structures that exist within the lexicon as a whole. Second, we have shown that the infralexical and the supralexical semantic structures are characterized by the same design features: both exhibit flexibility and salience effects. Third, the present chapter has revealed that these design features (and specifically, the semasiological and onomasiological salience phenomena) have an identifiable impact on language use: both the selection of an expression from a set of alternatives, and the form that the selected expressions take, appear to be influenced by the infralexical and supralexical salience structure of the lexicon.

5.3. The influence of contextual variation

In the previous sections of this chapter, we have been able to show that the choice of a particular lexical item as a name for a particular type of referent is influenced by two kinds of semantic factors. Semasiologically, there appears to be a tendency for referents to be named preponderantly by means of a lexical item of which they are a core exemplar. Onomasiologically, the overall entrenchment values of alternative lexical categories influence the choice for one or the other. The question now arises whether there is anything else to lexical choice except the selection of semantic alternatives. In particular, can it be shown that contextual differences determine which of a number of lexical alternatives is actually chosen? Or, to put the question in a slightly different way, can it be shown that lexical alternatives are invested with contextual values rather than just semantic values? Up to a point, the question has already been answered, to the extent that there appeared to exist contextual influences both with regard to the semasiological (section 3.5.) and with regard to the onomasiological (section 4.3) characteristics of the items. But are there any contextual differences over and above the differences of a semasiological or onomasiological nature? Are there any "pure" contextual differences?

In this section, we will present two case studies showing that there are indeed such pure contextual differences. The first case study corresponds with the vertical dimension of contextual variation as represented in Figure 2.1(1): we will show that the difference between Belgian Dutch and Netherlandic Dutch actually shows up in a number of naming patterns. The second case study involves the horizontal dimension of Figure 2.1(1): we will show that stylistic differences among magazines influence the choice of diminutive forms.

In general, *geographical variation* of a formal kind shows up when the *distribution of synonymous forms* exhibits significant differences between both geographical areas. For instance, in the Netherlandic sources, turtleneck sweaters occur 25 times with a category-specific name, i.e. a name that uniquely identifies the category "turtleneck sweater"; in all 25 cases, the name is *coltrui*. In the Belgian sources, 51 turtleneck sweaters with a unique name can be found, but there is variation in the names: in 20 cases (39.21%), the name is *coltrui*, whereas the other 60.79% have *rolkraagtrui*. A comparison of the distributional frequencies reveals that the differences are significant according to a χ^2 -test, with $p < 0.001$. *Rolkraag-*

trui, in other words, is a typically Belgian Dutch word; it is invested with a specific sociolinguistic connotation.

In order to establish the existence of "pure" formal geographical variation in a methodologically sound way, comparisons such as these are subject to a number of restrictions. In the first place, contextual variation of another kind than the geographical one has to be excluded. In particular, the possible effect of the specialization dimension has to be neutralized. This can be easily achieved by restricting the analysis to the central area of Figure 2.1(1), viz. the general magazines. In the turtle-neck example as in the following ones, the Belgian sources comprise *Flair*, *Feeling*, and *Libelle*_B, whereas the Netherlandic sources comprise *Margriet* and *Libelle*_{NL}.

In the second place, one may wonder how synonymous the compared items have to be. In many cases, absolute synonymy will not be guaranteed. Consider the case of *broek* and *pantalon*, which can both be used to designate long trousers, but which are nevertheless not fully synonymous: as we know, *broek* is a general name for all kinds of trousers; *pantalon*, on the other hand, is restricted to long trousers. Given this hyperonymous rather than synonymous relationship, would it still be acceptable to compare both items as we did for *rolkraagtrui* and *coltrui*? In general, the answer has to be negative, because differences in the distributional patterns of both geographical areas might reveal differences of categorization rather than just differences of lexical preference. For instance, if in one area *broek* occurs relatively more often as a name for long trousers than *pantalon*, whereas the other area shows a different pattern, this could simply mean that long trousers are more readily identified as a category of their own in the latter than in the former. Put more generally, we have to beware of the fact that the results of the lexical comparison may be influenced by hidden forms of geographical variation involving, for instance, differences of categorization. Specifically, as we have already seen in section 4.3. that the entrenchment values of certain categories may be different in both countries, we will have to make sure that such differences do not interfere with the analysis undertaken here. For instance, it would be misleading to simply compare *blazer*, *colbert*, and *jasje* as alternative names for formal jackets, as we know that the relationship between the onomasiological entrenchment of the hyperonymous category *jasje* and that of the hyponymous categories *blazer* and *colbert* differs significantly in the Belgian and Netherlandic sources. If we were to find, then, that in Belgium *jasje* is used relatively more often for formal jackets than in The Netherlands, this would primarily reflect the higher entrenchment value of the category named by *jasje*, rather than a

purely formal preference for *jasje* rather than *colbert* or *blazer*. At the same time, this line of reasoning opens up a possibility in which comparisons of not strictly synonymous items turn out to be acceptable. If it can be established that no differences of relative entrenchment (or similar semantic factors) influence the results, the synonymy criterion may be relaxed. We will not pursue this line of investigation, though, and restrict ourselves to cases where the referential synonymy is maximal.

Taking into account these precautionary measures, Figure 5.3(1) charts some more examples establishing the existence of purely formal variation along the geographical dimension. In all the examples, the differences are significant at the 0.001 level according to a χ^2 -test.

Items	Belgian sources	Netherlandic sources
<i>caleçon</i> <i>legging</i> <i>leggings</i>	40 (38%) 26 (24.7%) 39 (37.3%)	— 91 (100%) —
<i>blouzon</i> <i>jack</i>	13 (46.4%) 15 (53.6%)	— 85 (100%)
<i>jeans</i> <i>spijkerbroek</i>	64 (97%) 2 (3%)	38 (70.4%) 16 (29.6%)

Figure 5.3(1)
Examples of significant formal variation
along the geographical dimension

It will be recalled from the discussion in section 2.1., that the geographical variation in the corpus could be defined in two different ways: either by taking into account the place of publication of the magazines (which is what we have been doing so far), or by taking into account the distributional scope of the magazines. In the latter option, *Feeling*, *Flair*, and *Margriet* (which are distributed in both countries) occupy a transitional

position in comparison with *Libelle*_b on the one hand and *Libelle*_{nl} on the other, which can be found only in Belgium and The Netherlands, respectively. The question arises, then, whether the differences between Belgium and The Netherlands will be more outspoken if we consider only the magazines that are at the extremes of the distributional continuum, viz. *Libelle*_b and *Libelle*_{nl}. Figure 5.3(2), which gives an overview of the absolute frequencies of the items per magazine, reveals that this is indeed the case. Except for the pair *jeans/spijkerbroek*, the lexical pattern in *Libelle*_b when considered separately is more markedly "Belgian" than if *Libelle*_b, *Feeling*, and *Flair* are taken together, as in Figure 5.3(1). In fact, the frequency of the typically Belgian items *caleçon*, *blouson*, and *rolkraagtrui* is equal or almost equal to 100% in *Libelle*_b. The figure also shows that the magazines with binational distribution (*Feeling*, *Flair*, *Margriet*) occupy a middle position between the strictly Belgian and Netherlandic magazines *Libelle*_b and *Libelle*_{nl}. There are, however, clear individual differences among the magazines in this group, in the sense that the "middle" position is most clearly occupied by *Flair*.

	<i>Libelle</i> _b	<i>Feeling</i>	<i>Flair</i>	<i>Margriet</i>	<i>Libelle</i> _{nl}
caleçon	12	20	8	—	—
legging	1	—	25	31	60
leggings	—	—	39	—	—
blouson	5	5	3	—	—
jack	—	1	14	33	52
jeans	28	7	29	11	27
spijkerbroek	—	1	1	12	4
coltrui	—	2	18	12	13
rolkraagtrui	18	5	8	—	—

Figure 5.3(2)

Formal variation along the geographical dimension,
with each magazine taken separately

The examples presented so far establish the existence of geographical lexical variation in our corpus. Our earlier investigations into the presence of contextual variation revealed, however, that geographic variation is not the only type of contextual variation to be reckoned with. In section 4.3., for instance, we found that *stylistic variation* along the horizontal dimension of Figure 2.1(1) was no less a real phenomenon than geographic variation along the vertical dimension. The same point can be illustrated here when we have a look at the *distribution of diminutive forms* over the various sources and source groups used in compiling the database.

	Frequency of non- diminutives	Frequency of diminutive forms	Diminutivi- zation percentage
<i>Burda</i>	532	11	2%
<i>Knip</i>	220	29	11.6%
<i>Libellen</i> _{nl}	272	35	11.4%
<i>Margriet</i>	218	29	11.7%
<i>Flair</i>	291	206	41.4%
<i>Feeling</i>	100	26	20.6%
<i>Libelle</i> _b	426	79	15.6%
<i>Avantgarde</i>	8	4	33.3%
<i>Avenue</i>	18	7	28%
<i>Cosmopolitan</i>	20	20	50%
<i>Man</i>	44	2	4.3%
<i>Esquire</i>	6	4	40%

Figure 5.3(3)

Distribution of the diminutives *jurkje*, *rokje*, *bloesje*,
truitje

As a first approximation, Figure 5.3(3) lists the number of times that the items *jurk* 'dress', *rok* 'skirt', *bloes* 'blouse' (with spelling variants *bloeze* and *blouse*), and *trui* 'sweater' occur in the various sources, compared with the number of times in which the corresponding diminutive forms (*jurkje*, *rokje*, *bloesje*, *truitje*) occur. In *Burda*, for instance, the four items taken

together occur 532 times, and the corresponding diminutives occur 11 times. The rightmost column specifies a "diminutivization percentage", which is obtained by dividing the frequency of the diminutives by the sum of the frequency of occurrence of the diminutive and the non-diminutive forms.

As a second step in the analysis, we can now determine the average diminutivization percentages for the three major groups of magazines that can be distinguished along the horizontal dimension in Figure 2.1(1). For the fashion magazines *Burda* and *Knip*, the average is 5.05%. (Note that the average is not calculated by averaging the percentages from Figure 5.3(3), but by computing the diminutivization percentage for the sum of the real frequencies of the diminutive and non-diminutive items in each source group.) For the general magazines *Margriet*, *Flair*, *Feeling*, and both *Libelles*, the average is 22.3%. For the glossies *Avantgarde*, *Avenue*, *Cosmopolitan*, *Man*, and *Esquire*, the average is 27.8%. On the whole, then, there is an unmistakable relationship between the overall type of the magazine and the extent with which it uses diminutive forms. Diminutivization, in other words, seems to be a stylistic marker: in choosing between diminutives and their non-diminutive counterparts, particular groups of sources act differently than others.

It will also be noted, however, that there is considerable variation between the sources in one particular group. Within the group of glossy magazines, for instance, there is a considerable distance between the 50% of *Cosmopolitan* and the 4.3% of *Man*. This suggests that it may be necessary to have a closer look at the results, and specifically, that it may be useful to apply a more refined classification of the source groups than the one used so far. For one thing, it seems safer to exclude the results for *Esquire* from the analysis: the absolute frequencies that yield the 40% result are too low to be considered trustworthy. If we then have another look at the remaining set of glossy magazines, a relationship between the diminutivization percentage and the intended audience of the magazines suggests itself. As the name indicates, *Man* is a lifestyle magazine that specifically addresses a male audience. *Cosmopolitan*, on the other hand, addresses a female audience. *Avantgarde* and *Avenue*, finally, are neutral with regard to gender-specificity: they apparently aim at an audience of both men and women. Taking into account the intended audiences as indicated here leads to the discovery of a positive correlation between the extent to which a magazine aims at a female audience, and the extent to which it uses diminutive forms in comparison with the non-diminutive counterparts. Notice, in fact, that the diminutivization percentage for the glossy magazine with an intended

audience consisting exclusively of women (*Cosmopolitan*) is 50%, whereas that for *Man* is 4.3%; *Avantgarde* and *Avenue* have an average diminutivization percentage of 23.4%. There is, in other words, a diminutivization cline from women-oriented to male-oriented magazines, with the gender-neutral ones in the middle.

But of course, the importance of the intended audience has so far been established only within the group of glossy magazines. Can we find a similar distinction within the group of fashion magazines and the group of general magazines? The question is complicated by the absence of magazines for men in these two source groups. However, it is possible to detect another audience-related factor in the diminutivization percentages, involving the age of the intended readership. Within the group of general magazines, *Feeling* and *Flair* are typically (and explicitly) intended for younger women, whereas *Margriet* and both *Libelles* include a more mature audience. Similarly, within the group of fashion magazines, *Knip* has a decidedly younger profile within the set of fashion magazines than *Burda*. Characteristically, then, the average diminutivization percentage of *Feeling* and *Flair* is much higher (37.2%) than that of *Margriet* and the *Libelles* (13.5%), and analogously, the diminutivization percentage of *Knip* (11.6%) is higher than that of *Burda* (2%).

Summarizing, we find that the diminutivization percentage is affected by four different factors. First, there is a positive correlation between the glamorous character of the magazines and their diminutivization percentage: on the average, glossy magazines exhibit higher diminutivization percentages than either fashion magazines or general ones. Second, there is a negative correlation between the technicality of the magazines and their diminutivization percentage: the more technically specialized magazines (i.e. the fashion magazines) have less diminutives. Third, there is a positive correlation between the women-oriented nature of the magazines and their diminutivization percentage: the more a magazine aims at a female audience, the more diminutives it uses. And fourth, there is a positive correlation between the youthfulness of the magazines' profiles and their diminutivization percentage: magazines addressing an audience of younger women exhibit more diminutives than age-neutral ones. These tendencies clearly interact with each other. For instance, although the fashion magazines address a readership consisting exclusively of women, the positive effect of the female audience is overruled by the negative effect of the technicality and specialization dimension.

These observations lead to two further questions. To begin with, it will

have to be determined to what extent the distributional pattern that emerges from the foregoing observations might be the result of hidden semantic factors. Could it be that the presence of diminutive forms simply reflects the extent to which the various sources refer to small exemplars of the categories under investigation? Could it be, for instance, that magazines for younger women use the diminutive *rokje* more often simply because they more frequently talk about shorter skirts? We will show presently that there is indeed an influence of such referential factors, but that they only tell part of the story. In addition, the question arises what other factors might explain the distribution of the diminutives. Is the distributional pattern an arbitrary one, in the sense that the correlations that we find might just as well go in the other direction? If the use of diminutives is a stylistic marker, is there a particular motivation for the specific coupling of forms and stylistic values that we have encountered? Is it a coincidence, for instance, that technical specialization is signalled by less diminutivization rather than more diminutivization? We will show that there is a clear motivation for the distributional pattern, and that this motivation is, in fact, of a semantic nature. The kind of semantics at stake, though, is of an entirely different nature than the referential meaning aspects that are relevant for the first question. The relevant semantic phenomena for the second question involve the non-denotational meaning of the diminutive morpheme. This is not surprising, of course: stylistic differences in the distribution of an item are likely to be based on the stylistic meanings of that item.

The first question can be answered by checking whether there is a correlation between the number of small referents that the items under investigation occur with, and the extent to which the referents receive diminutivized names. In Figure 5.3(4), this correlation is computed by comparing the referential and lexical diminutivization percentages for four groups of sources, viz. the specialized fashion magazines (*Burda* and *Knip*), the glossies (*Avantgarde*, *Avenue* and *Cosmopolitan*), the general magazines for younger women (*Flair* and *Feeling*), and the general magazines that do not specifically address an audience of younger women (*Margriet* and both *Libelles*). The "referential diminutivization" percentage indicates the number of times the items in question refer to garments that could possibly be called small in a literal sense, and that could thus trigger the use of a diminutive form. The "lexical diminutivization" percentage specifies the proportion of diminutives in relation to the total set of instances under consideration. For instance, there are 115 instances of *jurk* or *jurkje* in *Flair* and *Feeling* taken together: 50 instances of *jurkje*, and 65 of *jurk*. Hence, the lexical diminu-

tivization percentage is $50/115=43.4\%$. Of the set of 115 referents, 60 involve garments that are not longer than the thigh; hence, the referential diminutivization percentage is $60/115=52.1\%$.

	Referential diminutivization	Lexical diminutivization
<i>rok/rokje</i> $r = 0.78$		
fashion magazines	35.9 %	6.5 %
glossies	66.6 %	33.3 %
younger women	66.2 %	49.3 %
women – general	34 %	26.7 %
<i>trui/truitje</i> $r = -0.67$		
fashion magazines	84 %	4 %
glossies	93 %	6.8 %
younger women	79.7 %	21.2 %
women – general	80.6 %	5.3 %
<i>jurk/jurkje</i> $r = -0.11$		
fashion magazines	66 %	0 %
glossies	62.1 %	59.4 %
younger women	52.1 %	43.4 %
women – general	48.1 %	17.6 %
<i>bloes/bloesje</i> $r = 0.11$		
fashion magazines	38.3 %	4.8 %
glossies	40 %	0 %
younger women	36 %	39.6 %
women – general	17.2 %	5.2 %

Figure 5.3(4)

Referential and lexical diminutivization percentages for
rokje, *truitje*, *jurkje*, *bloesje*

It should be noted that it is not always obvious to determine what constitutes a literally small case in these examples. For *rok* 'skirt' and *jurk* 'dress', the answer is fairly simple, since length may be assumed to play a dominant role. Consequently, we have calculated the referential diminutivization percentages on the basis of those garments that are not longer than the thigh. In the case of *trui(tje)* and *bloes(je)*, however, length is probably not the only relevant factor; specifically in the case of blouses, overall length may be considered relatively unimportant, since a majority of blouses in the set under investigation are worn tucked in at the waist. Referential diminutivization percentages for *trui(tje)* and *bloes(je)*, then, do not only take into account overall length, but also the length of the sleeves. (In the calculation for *trui(tje)*, cardigan-like referents have been left out of consideration.)

The correlation coefficients r in the figure are based on a linear regression. If the distribution of the diminutives could be explained on the basis of the referential characteristics of the garments, the correlation coefficient should be near to 1; a correlation coefficient of 1 means that all and only lexical diminutives refer to literally small garments. The actual results establish quite clearly that a referential explanation of the distribution of the diminutivization percentages can only be invoked in the case of *rok(je)*, where a relatively high positive correlation of 0.78 is found. In the other cases, however, the correlation coefficients are either low (in the neighborhood of 0) or relatively high but negative, as in the case of *trui(tje)*. The overall correlation coefficient that is obtained by considering the relationship between the two columns of sixteen percentages as a whole, is not higher than 0.005 %.

But if aspects of referential semantics do not play a dominant role in the explanation of the distributional pattern that emerged from Figure 5.3(3), what factors do? It can be shown that the non-referential, non-denotational semantic values of the diminutive morpheme explain most of its distributional characteristics. It would go beyond the scope of this investigation to present the semantic range of the diminutive in Dutch in full detail; see Bakema, Defour & Geeraerts (1993) for an extensive treatment. In the present context, it suffices to note that the diminutive morpheme has a number of closely connected connotational values that cluster round the notion of emotional evaluation. Three important aspects of the cluster may be mentioned separately. First, the emotional appreciation may be quite straightforward, both in a negative and in a positive direction. Depreciation shows up when a form like *een romannetje* 'a small novel' is used to signal low literary value rather than a restricted number of pages; a typical and

quite conventional formation in this respect is *stationsromannetje* 'pulp novel such as may be characteristically found in railway station bookstalls'. Appreciation, on the other hand, comes to the fore in forms like *mijn zusje* 'my little sister', which may be used to refer affectionately to one's sister even if she is older. Second, the diminutive expresses or suggests familiarity and informality: *een etentje* 'a small dinner' is not necessarily short, frugal, or gastronomically worthless, but it will rather take place in a sociable, sympathetic, intimate atmosphere. What the diminutive form expresses is not a lack of copiousness, but an emotional overtone of friendliness. Third, the diminutive serves the function of relativizing the importance of the referent – a function that obviously links up with the connotation of informality that was just mentioned. When someone says that he has had *een ongelukje* 'a small accident', the problem may be more than trivial, but the speaker tries to tone down its impact.

These three connotational values explain the distributional pattern observed above. In the first place, the negative correlation between the diminutivization percentage of the magazines and the degree in which they specifically address a male audience may be explained by the greater emotionality of female speech in Western cultures. A higher frequency of diminutives is, in fact, one of the often-mentioned characteristics of female speech. If the use of diminutives has emotional and familiarizing overtones, and if female speech is indeed characterized by greater emotionality and a greater insistence on interpersonal relations (like familiarity) than male speech, it comes as no surprise that the magazines in our sample that explicitly aim at a male audience, exhibit diminutivization percentages that are much lower than those of similar magazines with a female audience. Along the same line of thought, in the second place, the negative correlation between the technicality of the magazines and their diminutivization percentage may be explained by the more professional, business-like character of the fashion magazines. One of their functions, in fact, is to help their readers to make their own clothes (which is why they contain patterns); they do not just show clothes, but explain how they can be made. Fulfilling this function, then, requires objective information rather more than emotive appraisal, and this in turn may imply the use of less diminutives. By contrast, in the third place, the intention of giving objective, practical, applicable information is lowest in the glossies. Their primary purpose is to entertain and to suggest an entertaining, glamorous lifestyle. As they focus on the world of leisure and luxurious, carefree living, they not only evoke positive emotions of the kind that are typically expressed by the diminutive,

but also, they stand on the side of informality: if formality is commonly associated with constraints and restrictions, then the illusion of freedom that is offered by the glossy magazines naturally finds its expression in the informal linguistic register to which the use of the diminutive contributes. Finally, in the fourth place, the positive correlation between the youthfulness of the magazines' profiles and their diminutivization percentage may be related to the fact that informality is one of the hallmarks of contemporary popular youth culture.

To round off the discussion, it may be useful to glance back briefly at the various points in our investigation where we have dealt with contextual variation. In all three sections where the relevance of contextual factors was considered, both the geographical dimension and the specialization dimension as charted in Figure 2.1(1), appeared to be important. In section 3.5., it was shown that the semantic variation in the use of *shirt* correlates with the specialization dimension, whereas differences in the use of *vest* had to be situated along the geographic dimension. Section 4.3. revealed that differences of onomasiological entrenchment correlate with specialization: the entrenchment values of a number of hyperonyms is higher in the specialized fashion magazines. At the same time, the *jasje*-example showed that there may be geographical variation in the entrenchment value of an item. And in the present section, the distribution of referential synonyms appeared to correlate with the geographical dimension, whereas stylistic variation along the specialization dimension appeared to determine the frequency of diminutives. All in all, then, we may safely conclude that contextual variation permeates the structure of lexical variation.

Chapter 6

Ten theses about lexicology

The foregoing chapters have taken us on a tour through the intricate domain of lexical variation. Although the result of our journey had probably better be characterized as an explorer's sketch rather than a cartographer's map of the field, we hope that we have been able to indicate an interesting path for further investigation. The *main results* of what we have tried to show can be summarized in the following ten theses.

- [1] Studies of lexicological variation should distinguish between four major, interlocking types of lexical variation: semasiological, onomasiological, formal, and contextual variation.
- [2] The semasiological structure of single lexical items and the onomasiological structure of lexical fields are substantially characterized by two non-classical features: non-discreteness (demarcation problems) and non-equality (differences of salience).
- [3] On the semasiological level, non-discreteness may show up intensionally in the absence of definitions in terms of necessary and sufficient attributes, and extensionally in differences of membership status. Non-equality may show up intensionally in definitional clustering and differences of definitional weight, and extensionally in differences of membership salience.
- [4] The four non-classical semasiological characteristics need not co-occur, although they often are related. Specifically, we have illustrated cases where extensional non-discreteness determines intensional non-discreteness when membership relations are unclear, and cases in which extensional non-equality correlates with intensional non-equality (in the sense that maximal overlapping of definitional subsets of a category correlates with relative semasiological frequency).

[5] On the onomasiological level, non-discreteness shows up as the absence of a mosaic-like lexical field structure: lexical fields need not have clear boundaries, and the items in the field may exhibit multiple overlapping. Non-equality shows up in the fact that various categories may have various degrees of entrenchment, entrenchment being defined as onomasiological salience.

[6] The basic level hypothesis as a model of the distribution of entrenchment values over the lexicon is not universally valid. Rather, entrenchment values may be defined in terms of individual lexical items rather than taxonomical levels.

[7] Formal variation may be studied from two perspectives: the perspective of lexical choice, and the perspective of the internal structure of the chosen items. In the first case, the question is which factors determine the choice of one lexical expression rather than another as a name for a particular referent or set of referents. In the second case, the question is whether the form of the expression that is chosen can be related to the semasiological or onomasiological characteristics of the expression; this question is most relevant with regard to polymorphemic and polylexical expressions.

[8] Lexical choices are determined by the semasiological and onomasiological characteristics of the items involved: a referent (or set of referents) is expressed more readily by a category of which it is a central member, and it is expressed more readily by an item with a higher entrenchment value.

[9] The formal structure of lexical expressions is related to the semasiological and onomasiological characteristics of the categories involved. Semasiologically, the intrinsicness of a semantic dimension or dimensional value correlates inversely with the frequency with which it is expressed as a modifier in a polylexical expression. Onomasiologically, the entrenchment of a category correlates inversely with the frequency with which it is named by means of polymorphemic items.

[10] Contextual variation is basically of two kinds: it may involve permanent speaker characteristics, or it may involve situational

factors of a pragmatic nature. More importantly, it may involve each of the three kinds of structural variation (semasiological, onomasiological, formal). The most outspoken contextual effects in the corpus were found for pragmatic influences on onomasiological entrenchment, and for geographical influences on formal variation.

These findings are present in the following way in the structure of the book. Chapter 3 dealt with word meanings (semasiological variation), chapter 4 with lexical fields (onomasiological variation), and chapter 5 with naming and lexical choice (formal variation). In each chapter, the final section (3.5., 4.3., 5.3.) considered the influence of contextual variation. Within the chapter on formal variation, sections 5.1. and 5.2. dealt with the influence of semasiological and onomasiological factors respectively. Within the chapter on semasiological variation, sections 3.2. and 3.3. dealt with aspects of non-discreteness and flexibility, whereas non-equality and salience were treated in 3.4.. Within the chapter on onomasiological variation, 4.1. and 4.2. described non-discreteness/flexibility phenomena and non-equality/salience phenomena respectively. Bringing together semasiological and onomasiological variation under their common denominator as kinds of conceptual (or, if one wishes, semantic) variation, the structure of the book can be schematically represented as in Figure 6(1). The arrows specify where the influence of one type of variation on the other is treated. The lower part of the figure spells out the systematical relationship between the two chapters dealing with conceptual variation. The figure (which may be usefully compared with Figure 1(2) in the introductory chapter of the book) does not just give an overview of the way in which the various parts of the preceding text fit together, it also specifies the conceptual architecture, so to speak, of the investigation presented in the text: it indicates what the crucial types of variation are, what features pervasively characterize the structure of the two kinds of conceptual variation, and how the major forms of variation cross-categorize.

Taken together, the ten theses paint a picture of the structure of the lexicon that is larger in scope and stronger in coherence than has hitherto been usual in variational lexicology. Regardless of the descriptive qualities of the investigation and the potential importance of the specific empirical results we have obtained, we feel that the research presented in this book is methodologically important in the context of theoretical lexicology and lexical semantics at large. There are three main reasons why we feel this to be the case.

Figure 6(1)

The thematic organization of the book

First, our investigation combines, in what seems to be a natural and fruitful way, the legacy of (predominantly Continental) *structural semantics* with the new insights and methods that were developed in the context of *prototype semantics* in the last fifteen years. The structuralist tradition has stressed the importance of what might be called the “external” structure of

lexical categories: the fact that words do not exist in isolation, but are rather a part of associative and taxonomical groupings. Structuralist semantics insists that an adequate description of lexical items requires a description of their position within those lexical fields. On the other hand, the prototype-oriented tradition of research that developed within Cognitive Linguistics has stressed the importance of an investigation into the "internal" structure of lexical categories: the mutual relationship between the referents and meanings of each word taken separately. It insists that words cannot be described on the basis of distinctions with other words alone, but that the proper content of each word has to be studied on its own as well. What we have tried to show, then, is that both the internal and the external types of investigation are indispensable if we are to gain an adequate insight into the lexicon as a system of categories. Both field research and prototypicality research are an integral part of cognitive lexicology; words should be studied both in their lexical relationship to other words, and in their relationship to the world.

Second, our investigation adds a *contextual perspective* to the cognitive study of lexical variation. Although Cognitive Linguistics has a lot of attention for the cultural aspects of the relationship between language and the world, the variation that may exist within a single linguistic community has not often been investigated from a cognitive point of view. By systematically taking into account contextual variation involving speaker-related and situation-related variables, the scope of cognitive lexicology is broadened in the direction of sociolinguistics.

Third and foremost, we have systematically developed a *pragmatic, usage-based model* of lexicological research. The coupling of an onomasiological and a semasiological perspective does not merely imply the combination of an "external" and an "internal" conception of semantic structure, but it also embodies a shift from a preoccupation with structures to an interest in the way in which these structures are put to actual use. The questions we have asked are not just restricted to the traditional questions "What does lexical item x mean?", and "In what meaningful supra-lexical structures does x participate?". Rather, the insight into the semasiological and onomasiological structures of lexical knowledge that these questions lead to, naturally result in the question that was the main focus of chapter 5: "What are the factors that determine whether x is chosen as a name for a particular referent?". The change of perspective is perhaps best described as a shift from *meaning* to *naming*: the question is not just what semantic phenomena may be discerned within lexical items separately, or within the

lexicon as a whole, but also how these semantic phenomena (and other factors) determine how choices among lexical alternatives are made. The model we propose, in short, is comparable to the one recently suggested by Lehrer & Lehrer (1994), and in which they propose to describe the sense and reference of a word as an aggregation of various input vectors. Their model focusses on semasiological phenomena, whereas ours is more comprehensive by including onomasiological problems of naming next to problems of meaning. The basic idea is the same, though: various factors in combination determine the choice of an item as a name for a particular referent.

In spite of what we believe to be its innovative significance, however, we are well aware that our study is subject to a number of restrictions that should be overcome in the course of further investigations. There are three areas of research that call for an elaboration of our findings. Consecutively, they broaden the scope of the investigation towards areas and problems that lie further away from the present study.

In the first place, the methodological depth of the present study should be increased by bringing in more refined *statistical techniques*. Very often, our analyses have been informal, and where statistical data have been used, only low-level statistical methods have been invoked. Given the variational complexity of the data that we are dealing with, it is certainly worthwhile to try and apply more sophisticated quantitative techniques to the conceptual framework that we have developed. The first condition for such an elaboration of the quantitative approach will be a larger corpus. Even though our own set of materials is far from small in comparison with the tiny set of made up examples on which lexical analyses are all too often based, we estimate that an even larger corpus will be necessary to deal with the full complexity of the material in a mathematically refined way. Because this is an extremely time-consuming endeavor (compiling the corpus has taken about one third of the time necessary for the completion of this study), a restriction to one or two of the subsets that we have considered (like that of trousers, or that of skirts) would seem to be called for.

In the second place, the *empirical scope* of the study should be broadened. The elaboration can, of course, go in various directions. For one thing, the amount of contextual variation taken into account in the study could be increased. In particular, the relationship between technical sources and sources from the general language (like the ones used here) may be investigated in more detail: even though we have been able to range our sources on a technicality dimension, professional language in the strict sense has not been included. For another thing, the question arises whether the

patterns we have identified in the field of clothing terms also occur in other semantic fields. More specifically, while we have investigated a concrete field in which polysemy (in contrast with referential multiplicity) hardly plays any role, an extension of the model sketched here will have to include abstract domains and items that are clearly polysemous. Finally, it will be worthwhile to envisage an extension outside the domain of lexical semantics. If Cognitive Linguistics is right in claiming that there exist general principles of categorial organization that cross the line between syntax and the lexicon, then syntactic categories like "indirect object", "genitive", or "adverb" should be just as amenable to the approach sketched here as lexical categories like *broek*, *legging*, and *rok*.

In the third place, an extension from the theoretical domain to that of *applied linguistics* may be envisaged. The kind of study presented here is an example of fundamental research, i.e. theoretical research into the principles and patterns that structure a particular aspect of natural language. But if this kind of foundational research is to have more than a mere academic interest, its practical consequences have to be explored. Two fields of application naturally stand out: language technology and lexicography. With regard to the field of Artificial Intelligence and Natural Language Processing, the question will be how to translate the present view of the structure of lexicological variation into a formal model that can be incorporated into programmes for machine translation, expert systems, parsing programmes and the like. With regard to the field of dictionary making, it will have to be determined whether and how data of the kind unearthed here can be incorporated into lexicographical and terminographical reference works. What labels, for instance, would dictionaries have to use to describe the different kinds of variation, and how should entries be structured to adequately render the prototype-based semasiological structure of lexical items? And could onomasiological entrenchment measures be invoked to guide the selection of words to be incorporated into the dictionary?

Regardless of the tasks that still lie ahead of us, we have ultimately tried to achieve the following goals with the present study: first, to sketch a *descriptive framework* for the study of lexical variation by identifying the various phenomena that any truly comprehensive lexicological theory has to deal with; second, to develop and illustrate a number of *analytical techniques and representational mechanisms* that are useful for dealing in an insightful way with those phenomena; and third, to support or reject a number of *specific hypotheses* about the phenomena in question and their relations. We have tried to indicate, in other words, what to investigate in

variational lexicology, how to investigate it, and what kind of observations follow from the investigation. The last point is a theoretical one, in the sense that it involves the development of a specific theory of the structure of lexicological variation in terms of concepts such as prototypicality, entrenchment, and lexical choice. The second point is a methodological one, in the sense that it involves the development of specific methods for studying lexical variation. The first point is a demarcational one, in the sense that it tries to define the domain of the study of lexical variation as a specific sub-discipline of linguistics.

Now, whereas our theoretical conclusions may have to be rejected or supplemented on the basis of further research, and whereas our methodological suggestions will have to be refined and elaborated, the demarcation of the field of investigation of variational lexicology in terms of semasiological, onomasiological, formal, and contextual variation constitutes the hard core of our proposals. What we hope to have shown, then, is that this model of lexical variation is fruitful enough to inspire further development. There may be no final words on the topic of words, but

at best,
Reaching no absolute in which to rest,
One is always nearer by not keeping still

(Thom Gunn,
On the move).

References

- Armstrong, Sharon L., Lila R. Gleitman & Henry Gleitman
 1983 "What some concepts might not be". *Cognition* 13: 263-308.
- Bakema, Peter, Patricia Defour & Dirk Geeraerts
 1993 "De semantische structuur van het diminutief" [The semantic structure of the diminutive]. *Forum der Letteren* 34: 121-137.
- Baldinger, Kurt
 1980 *Semantic Theory. Towards a Modern Semantics*. Oxford: Basil Blackwell. Translation of *Teoría semántica* (1977).
- Berlin, Brent
 1978 "Ethnobiological classification". In Eleanor Rosch & Barbara B. Lloyd (ed.), *Cognition and Categorization* 9-26. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Berlin, Brent, Dennis E. Breedlove & Peter H. Raven
 1973 "General principles of classification and nomenclature in folk biology". *American Anthropologist* 75: 214-242.
 1974 *Principles of Tzeltal Plant Classification*. New York: Academic Press.
- Berlin, Brent & Paul Kay
 1969 *Basic Color Terms. Their Universality and Evolution*. Berkeley and Los Angeles: University of California Press.
- Brugman, Claudia M.
 1989 *The Story of Over. Polysemy, Semantics, and the Structure of the Lexicon*. New York: Garland.
- Coleman, Linda & Paul Kay
 1981 "Prototype semantics: the English word *lie*". *Language* 57: 26-44.
- Coseriu, Eugenio & Horst Geckeler
 1981 *Trends in Structural Semantics*. Tübingen: Gunter Narr.
- Cruse, D. Alan
 1977 "The pragmatics of lexical specificity". *Journal of Linguistics* 13: 153-164.
 1986 *Lexical Semantics*. Cambridge: Cambridge University Press.
- Cuyckens, Hubert
 1991 *The Semantics of Spatial Prepositions in Dutch. A Cognitive-Linguistic Exercise*. [PhD dissertation, Universitaire Instelling

Antwerpen.]

Deane, Paul D.

- 1992 *Grammar in Mind and Brain. Explorations in Cognitive Syntax*. Berlin: Mouton de Gruyter.

Dirven, René

- 1985 "Metaphor and polysemy". *Cahiers de l'Institut de Linguistique de Louvain* 11(3/4): 9-27.
 1990 "Metaphorical divergences in the evolution of Dutch and Afrikaans". In Dirk Geeraerts (ed.), *Diachronic Semantics (Belgian Journal of Linguistics 5)* 9-34. Brussels: Editions de l'Université de Bruxelles.

Dougherty, Janet W.D.

- 1978 "Salience and relativity in classification". *American Ethnologist* 5: 66-80.

Ducháček, Otto

- 1959 "Champ conceptuel de la beauté en français moderne". *Vox Romanica* 18: 297-323.

Fillmore, Charles

- 1975 "An alternative to checklist theories of meaning". *Proceeding of the Annual Meeting of the Berkeley Linguistics Society* 1: 123-131.

Geckeler, Horst

- 1971 *Strukturelle Semantik und Wortfeldtheorie*. München: Fink.

Geeraerts, Dirk

- 1983 "Prototype theory and diachronic semantics. A case study". *Indogermanische Forschungen* 88: 1-32.
 1985a "Diachronic extensions of prototype theory". In Geer Hoppenbrouwers, Pieter Seuren & Anton Weijters (ed.), *Meaning and the Lexicon* 354-362. Dordrecht: Foris Publications.
 1985b *Paradigm and Paradox. Explorations into a Paradigmatic Theory of Meaning and its Epistemological Background*. Leuven: Leuven University Press.
 1986a "Functional explanations in diachronic semantics". In Alain Bossuyt (ed.), *Functional Explanations in Linguistics (Belgian Journal of Linguistics 1)* 67-93. Brussels: Editions de l'Université de Bruxelles.
 1986b *Woordbetekenis. Een overzicht van de lexicale semantiek* [Word meaning. An overview of lexical semantics]. Leuven: Acco.

- 1987 "On necessary and sufficient conditions". *Journal of Semantics* 5: 275-291.
- 1988a "Cognitive grammar and the history of lexical semantics". In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics* 647-677. Amsterdam: John Benjamins Publishing Company.
- 1988b "Where does prototypicality come from?". In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics* 207-229. Amsterdam: John Benjamins Publishing Company.
- 1989 *Wat er in een woord zit. Facetten van de lexicale semantiek* [What there is in a word. Aspects of lexical semantics]. Leuven: Peeters.
- 1990 "The lexicographical treatment of prototypical polysemy". In Savas L. Tsohatzidis (ed.), *Meanings and Prototypes. Studies in Linguistic Categorization* 195-210. London and New York: Routledge.
- 1992 "The return of hermeneutics to lexical semantics". In Martin Pütz (ed.), *Thirty Years of Linguistic Evolution. Studies in Honour of René Dirven on the Occasion of his Sixtieth Birthday* 257-282. Amsterdam: John Benjamins Publishing Company.
- 1993 "Vagueness's puzzles, polysemy's vagaries". *Cognitive Linguistics* 4: 223-272.

Geeraerts, Dirk & Peter Bakema

- 1993 "Materiaalverzamelingsmethodes in lexicologie en lexicografie" [Data collection in lexicology and lexicography]. In Arend van der Veen (ed.), *Op je woorden passen. Voordrachten gehouden tijdens het symposium van het Instituut voor Nederlandse Lexicologie in Antwerpen op 18 januari 1993* 10-22. Leiden: Instituut voor Nederlandse Lexicologie.

Gipper, Helmut

- 1959 "Sessel oder Stuhl? Ein Beitrag zur Bestimmung von Wortinhalten im Bereich der Sachkultur". In Helmut Gipper (ed.), *Sprache, Schlüssel zur Welt. Festschrift für Leo Weisgerber* 271-292. Düsseldorf: Schwann Verlag.

Guiraud, Pierre

- 1975 *La sémantique*. Paris: Presses Universitaires de France. 8th ed. (1st ed. 1955).

Heath, Jeffrey

- 1988 "Lexicon". In Ulrich Ammon, Norbert Dittmar & Klaus Mattheier (ed.), *Sociolinguistics – Soziolinguistik* 2: 1153-1163. Berlin: Walter De Gruyter.

Heider, Eleanor R.

- 1972 "Universals in color naming and memory". *Journal of Experimental Psychology* 93: 10-20.

Heider, Eleanor R. & D. C. Olivier

- 1972 "The structure of the color space in naming and memory for two languages". *Cognitive Psychology* 3: 337-345.

Ipsen, Gunther

- 1924 "Der alte Orient und die Indogermanen". In Johannes Friedrich, Johannes B. Hofmann & Wilhelm Horn (ed.), *Stand und Aufgaben der Sprachwissenschaft. Festschrift für Wilhelm Streitberg* 200-237. Heidelberg: Carl Winter Universitätsverlag.

Jolles, André

- 1934 "Antike Bedeutungsfelder". *Beiträge zur Geschichte der deutschen Sprache und Literatur* 58: 97-109.

Katz, Jerrold J. & Jerry A. Fodor

- 1963 "The structure of a semantic theory". *Language* 39: 170-210.

Kempton, Willett

- 1981 *The Folk Classification of Ceramics. A Study of Cognitive Prototypes*. New York: Academic Press.

Kleiber, Georges

- 1991 "Hiérarchie lexicale: catégorisation verticale et termes de base". *Sémiotiques* 1: 35-57.

Kronasser, Heinz

- 1952 *Handbuch der Semasiologie*. Heidelberg: Carl Winter Universitätsverlag.

Labov, William

- 1972 "Some principles of linguistic methodology". *Language in Society* 1: 97-120.
- 1973 "The boundaries of words and their meanings". In Charles-James Bailey & Roger W. Shuy (ed.), *New Ways of Analyzing Variation in English* 340-373. Washington: Georgetown University Press.
- 1978 "Denotational structure". In Donka Farkas, Wesley M. Jacobsen & Karol W. Todrys (ed.), *Papers from the*

- Parasession on the Lexicon* 220-260. Chicago: Chicago Linguistics Society.
- Lakoff, George
1987 *Women, Fire, and Dangerous things. What Categories Reveal about the Mind*. Chicago and London: The University of Chicago Press.
- Lakoff, George & Mark Johnson
1980 *Metaphors We Live By*. Chicago and London: The University of Chicago Press.
- Langacker, Ronald W.
1987 *Foundations of Cognitive Grammar I. Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lehrer, Adrienne
1974 *Semantic Fields and Lexical Structure*. Amsterdam: North Holland Publishing Company.
1990 "Prototype theory and its implications for lexical analysis". In Savas L. Tsohatzidis (ed.), *Meanings and Prototypes* 368-381. London and New York: Routledge.
- Lehrer, Adrienne & Keith Lehrer
1994 "Fields, networks, and vectors". In Frank Palmer (ed.), *A Festschrift for John Lyons* [in press].
- Lyons, John
1963 *Structural Semantics*. Oxford: Basil Blackwell.
1968 *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
1977 *Semantics*. Cambridge: Cambridge University Press.
- MacLaury, Robert E.
1987 "Coextensive semantic ranges: different names for distinct vantages of one category". *Papers from the Regional Meeting of the Chicago Linguistic Society* 23: 268-282.
1991a "Prototypes revisited". *American Review of Anthropology* 20: 55-74.
1991b "Social and cognitive motivations of change: measuring variability in color semantics". *Language* 67: 34-62.
1992 "From brightness to hue. An explanatory model of color-category evolution". *Current Anthropology* 33: 137-186.
- Mervis, Carolyn B. & Eleanor Rosch
1981 "Categorization of natural objects". *Annual Review of Psychology* 32: 89-115.

Öhman, Suzanne

- 1953 "Theories of the 'Linguistic Field'". *Word* 9: 123-134.

Paprotté, Wolf & René Dirven (ed.)

- 1985 *The Ubiquity of Metaphor. Metaphor in Language and Thought*. Amsterdam: John Benjamins Publishing Company.

Posner, Michael

- 1986 "Empirical studies of prototypes". In Colette Craig (ed.), *Noun Classes and Categorization* 53-61. Amsterdam: John Benjamins Publishing Company.

Pottier, Bernard

- 1964 "Vers une sémantique moderne". *Travaux de Linguistique et de Littérature* 2: 107-137.

Quadri, Bruno

- 1952 *Aufgaben und Methoden der Onomasiologischen Forschung*. Bern: Francke.

Quine, Willard V.O.

- 1960 *Word and Object*. Cambridge, Mass.: The MIT Press.

Rastier, François

- 1987 *Sémantique interprétative*. Paris: Presses Universitaires de France.

Rosch, Eleanor

- 1978 "Principles of categorization". In Eleanor Rosch & Barbara B. Lloyd (ed.), *Cognition and Categorization* 27-48. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Rosch, Eleanor & Carolyn B. Mervis

- 1975 "Family resemblances: studies in the internal structure of categories". *Cognitive Psychology* 7: 573-605.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson & Penny Boyes-Braem

- 1976 "Basic objects in natural categories". *Cognitive Psychology* 8: 382-439.

Rudzka-Ostyn, Brygida

- 1988 "Semantic extensions into the domain of verbal communication". In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics* 507-553. Amsterdam: John Benjamins Publishing Company.

- 1989 "Prototypes, schemas, and cross-category correspondences: the case of *ask*". *Linguistics* 27: 613-662.

Schmid, Hans-Jörg

- 1993a "Cottage and Co.: can the theory of word-fields do the job?". In Peter Rolf Lutzeier (ed.), *Studien zur Wortfeldtheorie. Studies in Lexical Field Theory* 107-120. Tübingen: Max Niemeyer Verlag.
- 1993b *Cottage und Co., idea, start vs. begin. Die Kategorisierung als Grundprinzip einer differenzierten Bedeutungsbeschreibung*. Tübingen: Max Niemeyer Verlag.

Schulze, Rainer

- 1988 "A short story of down". In Werner Hüllen & Rainer Schulze (ed.), *Understanding the Lexicon. Meaning, Sense, and World Knowledge in Lexical Semantics* 395-414. Tübingen: Max Niemeyer Verlag.
- 1991 "Getting round to (a)round: towards the description and analysis of a 'spatial' predicate". In Gisa Rauh (ed.), *Approaches to Prepositions* 253-274. Tübingen: Gunter Narr.

Sweetser, Eve

- 1990 *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

Taylor, John R.

- 1989 *Linguistic Categorization. Prototypes in Linguistic Theory*. Oxford: Clarendon Press.

Trier, Jost

- 1931 *Der Deutsche Wortschatz im Sinnbezirk des Verstandes*. Heidelberg: Carl Winter Universitätsverlag.
- 1932 "Die Idee der Klugheit in ihrer sprachlichen Entfaltung". *Zeitschrift für Deutschkunde* 46: 625-635.
- 1934 "Das sprachliche Feld. Eine Auseinandersetzung". *Neue Jahrbücher für Wissenschaft und Jugendbildung* 10: 428-449.
- 1968 "Altes und neues vom sprachlichen Feld". *Duden-Beiträge zu Fragen der Rechtschreibung, der Grammatik und des Stils* 34: 9-20.

Vandeloise, Claude

- 1986 *L'Espace en Français*. Paris: Seuil.

Van Domselaar, Johannes & Henk J.J. Horsten

- 1990 *Kursus Textielwaren- en Textiel Artikelenkennis 2. Assortimentskennis Textiel*. Doorn: Stichting Detex.

Waismann, Friedrich

- 1952 "Verifiability". In Anthony G.N. Flew (ed.), *Logic and Language* 117-144. Oxford: Basil Blackwell.

Wierzbicka, Anna

- 1985 *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma.
1989 "Prototypes in semantics and pragmatics: explicating attitudinal meanings in terms of prototypes". *Linguistics* 27: 731-769.
1991 "Semantic rules know no exceptions". *Studies in Language* 15: 371-398.

Wittgenstein, Ludwig

- 1953 *Philosophische Untersuchungen – Philosophical Investigations*. Oxford: Basil Blackwell.

Zwicky, Arnold & Jerrold Sadock

- 1975 "Ambiguity tests and how to fail them". In John Kimball (ed.), *Syntax and Semantics* 4 1-36. New York: Academic Press.

Index of subjects

- Absence of lexical gaps 121
 - See also* Mosaic-like conception of lexical fields
- Age variation 183, 188
 - See also* Contextual variation; Speaker-related variation
- Alternative classification of source groups 182-188
- Alternative denomination (*see* Onomasiological alternatives in semasiological ranges)
- Ambiguity (*see* Polysemy)
- Analytical polysemy (*see* Definitional polysemy)
- Anthropological linguistics 13
- Applied linguistics 195
 - See also* Artificial intelligence; Lexicography; Natural language processing
- Archilexème 117, 133
 - See also* French structuralism
- Archiséme 117
 - See also* French structuralism
- Artificial intelligence 195
 - See also* Applied linguistics
- Auto-hyponymy 101-102
 - See also* Lexical relations
- Bandplooibroek* 26
- Basic level 10, 13, 134
 - characteristics 134-135
 - shift 147
 - expertise-related variation of the basic level 146-147
 - naming preferences 10, 134
 - See also* Taxonomy; Folk classification
- Basic level model of onomasiological salience 134-136, 190
 - See also* Problems with basic level model of onomasiological salience
- Bermuda* 23, 26, 30, 37, 145-146, 170-172, 173-175
- Bird* (*see* Prototypicality, examples)
- Blazer* 31, 36, 37, 67-76, 91-105, 133, 152-153, 159-164, 164-169, 170-172, 178, 179
 - bruine* 164-169
 - effen* 164-169
 - getailleerde* 164-169
 - in pepitaruit* 164-169
 - met een dubbele knoopsluiting* 164-169
 - met bloemdessin* 164-169
 - met krijtstrepen* 164-169
 - ruimvallende* 164-169
- Bloemenbroek* 61-67
- Bloemenjack* 34
- Bloes* 33, 145, 173-175, 181-183, 184-186
- Bloesje* 181-183, 184-186
- Bloeze* 145, 173-175, 181-183, 184-186
- Blouse* 33, 36, 37, 129-133, 145, 170-172, 173-175, 181-183, 184-186
- witte* 35
- Blouson* 152-153, 177-180
- Box diagram
 - bird* 52
 - blazer* 94-95
 - broek* 98
 - colbert* 92-93
 - hemd* 100
 - legging* 90
 - overhemd* 99
 - T-shirt* 101
 - vers* 53
 - vest* 106-108

- Broek* 10, 23, 31, 37, 61-67, 78-89, 98-105, 136, 137, 145-146, 147-153, 164-169, 170-172, 172-173, 173-175, 178, 195
drievierde 164-169
effen 164-169
in pasteltinten 164-169
knielange 164-169
korte 164-169
lange 164-169
met krijtstreep 164-169
nauwsluitende 164-169
rechte 164-169
strakke 164-169
Turkse 26
van ruitjesstof 164-169
wijde 164-169
- Broekje* 61-67
- Broekrok* 24, 26, 37, 64, 78-89, 136, 170-172
- Bustier* 35
- Caleçon* 61-67, 112, 145-146, 172-173, 177-180
- Cardigan* 36
- Categorization* 4, 7, 12-13, 178
 classical view 7, 8, 9, 47
 onomasiological variation as the result of alternative categorization 3, 4, 7, 12-13
See also Classical definability; Conceptual variation; Epistemological characteristics of natural categories; Experiential nature of natural categories; Onomasiological perspective; Onomasiological variation; Onomasiology; Semasiological variation; Prototypicality, varieties
- Centrality* (*see* Degrees of representativity)
- Chaneljasje* 34
- Checklist theory of meaning* 37
- Chiffonblouse* 34
- Chi square test* 105, 109, 110-112, 114, 179
- Citybermuda* 35
- Classical definability* 7, 8, 9, 38, 56-67, 128-129, 189
 examples 59-67, 125-129
 absence 7, 8, 45-56 *passim*, 67-76, 114
 combinations of classical definitions 74-75
 mechanical approach 57-58
 undecidability of classical definability (*see* Definitional consequences of membership uncertainty)
 salience effects in classical and non-classical categories 102
See also Corpus-based approach to classical definability; Definitional polysemy; Lexical Relations, importance for classical definability; Necessary and sufficient conditions; Prototypicality, varieties; Unidimensional definitions
- Classical view of natural categories* (*see* Categorization, classical view)
- Clustering* (*see* Family resemblance structure)
- Cognitive entrenchment* (*see* Entrenchment)
- Co-hyponymy* 146
See also Lexical relations
- Colbert* 22, 23, 31, 36, 37, 67-76, 91-105, 133, 152-153, 159-164, 164-169, 170-172, 178, 179
double breasted 164-169
effen 164-169
geruit 164-169
getailleerd 164-169
pied-de-poule 164-169
rood 164-169

- ruimvallend* 164-169
- single breasted* 164-169
- Collins Cobuild-corpus 35
- Color terms (*see* Prototypicality, examples)
- Coltrui 177-180
- Compartmentalization of the lexicon 122
 - See also* Mosaic-like conception of lexical fields; Trier's analysis of *Wisheit, Kunst* and *List*
- Componential analysis 9, 22-31, 37-44, 57
 - referential instead of semantic description 38-39
 - and prototype theory 37-39
 - psychological status 38
 - in Transformational Grammar 37-38
 - of the *broek*-subfield 24-30
 - of the *hemd*-subfield 129, 131
 - of the *jasje*-subfield 96-97
 - of the *rok*-subfield 125-126
 - of the *vest*-subfield 97
 - See also* Incompatible componential descriptions
- Compound expressions 31, 33-35
 - See also* Lexical description
- Compounds 155, 173-176
 - See also* Monomorphemic vs. polymorphemic names
- Conceptual variation
 - semasiological and onomasiological variation as conceptual variation 4, 5, 191-192
 - vs. formal variation (*see* Onomasiological vs. formal variation)
 - See also* Onomasiological variation; Semasiological variation; Variation, varieties
- Conceptualization 4, 7, 13
 - See also* Categorization
- Configurational entrenchment 138-142, 145, 152, 170-171, 173-174
 - See also* Entrenchment, measures
- Configurational frequency 141
 - See also* Entrenchment, measures
- Connotational meaning (*see* Non-denotational meaning)
- Contextual variation (*see* Situational variation; Speaker-related variation; Variation, varieties)
- Copulative compound 81
- Corpus of non-elicited material 8-9, 13-14, 17-44
 - statistics and figures 32-37
 - restrictions 21, 36, 97, 100
 - saturation 32-35
 - surprises 96
 - vs. elicitation techniques 14, 17, 91, 158
 - vs. introspection (*see* Usage-based approach vs. introspection)
 - vs. material from the Instituut voor Nederlandse Lexicologie 35
 - vs. text corpora 14, 17-18
 - See also* Alternative classification of source groups; Componential analysis; Corpus-based approach to hyponymy; Corpus-based approach to hyperonymy; Corpus-based approach to membership uncertainties; Direct knowledge of referents; Features; Frequency limitations; Lexical description; Garment types; Referential description; Source groups; Usage-based approach
- Corpus-based approach
 - to classical definability 56-76
 - to hyperonymy 69-76, 77-78, 81-

- 85
- to hyponymy 69-76, 77-78, 81-85
- to membership uncertainties 77-78, 81-85
- See also* Necessary and sufficient conditions
- Corrected name frequency 141
- See also* Entrenchment, measures
- Correlation intensional and extensional salience (*see* Salience effects)
- Cue validity 156-164, 171
 - correlation with salience effects 157, 159-164
 - vs. entrenchment 158
 - influence on lexical choice 156-164, 190, 191
 - in psychology 156
 - in psycholinguistics 157-158
 - intensional and extensional salience 156-157
 - onomasiological vs. semasiological cue validity 157
- Débardeur* 35
- Definitional analysis 51-54
 - See also* Box diagram; Classical definability; Definitional polysemy; Degrees of representativity; Family resemblance structure; Intuitive univocality; Semasiological structure; Tests to distinguish between referential variability and polysemy
- Definitional clustering (*see* Family resemblance structure)
- Definitional consequences of membership uncertainty 77-78, 81, 87
- Definitional entrenchment 138-142, 142-143, 144, 147-148, 148-149, 150-151
 - See also* Entrenchment, measures; Unidimensional definitions
- Definitional polysemy 51-54
 - See also* Classical definability; Intuitive univocality; Necessary and sufficient conditions; Polysemy; Referential variability vs. polysemy; Tests to distinguish between referential variability and polysemy
- Definitional vs. referential structure of a category 48
 - See also* Extensional non-discreteness; Extensional non-equality; Intensional non-discreteness; Intensional non-equality
- Definitional weight differences (*see* Family resemblance structure)
- Degrees of membership (*see* Membership uncertainties)
- Degrees of membership status (*see* Degrees of representativity)
- Degrees of representativity 45-56
 - passim, 90-91, 91-105, 156-157, 160, 189
 - correlation with family resemblance structure (*see* Salience effects)
 - See also* Prototypicality, varieties
- Demarcation problems (*see* Non-discreteness in lexical fields)
- Denimbroek* 61-67
- Denotation 50
- Derivation 155, 173-176
 - See also* Monomorphemic vs. polymorphemic names
- Detex course materials 24, 26, 43
- Deux-pièces* 31
- Diachronic domain 14
- Diminutives 181-188
- Diminutivization 182-188

- lexical diminutivization percentage 181-188
- referential diminutivization percentage 184-186
- See also* Non-denotational meaning; Stylistic variation; Pure stylistic variation
- Direct knowledge of referents 8, 14, 43-44
- Distinctiveness requirement 58-59, 61-67
- See also* Corpus-based approach to classical definability; Necessary and sufficient conditions
- Donsjack* 34
- Doorknoop* 125-129, 139-146, 147-153, 173-175, 175-176
- Drollenvanger* 26
- Ducháček's starlike conception of lexical fields 122-123
- See also* Non-discreteness of lexical fields
- Dutch
 - Belgian vs. Netherlandic 10, 20, 36, 65, 105-112, 114, 152-153, 177-180
 - See also* Geographical variation; Pure geographical variation
- Economy of expression 151
- See also* Pragmatic variation
- Elicitation 13-14, 17, 77
- See also* Questionnaires; Experiments
- Emotional evaluation 186-187
- See also* Diminutivization; Non-denotational meaning
- Encyclopedic vs. semantic features 47
- Entrenchment 8, 9, 10, 11, 13, 14, 69, 81-82, 136-137, 156, 190, 195, 196
- measures (*see* Configurational entrenchment; Definitional entrenchment; Polylexical entrenchment)
- influence on lexical choice (*see* Lexical choice)
- influence on formal structure of clothing names (*see* Monomorphemic vs. polymorphemic names; Monolexical vs. polylexical expression)
- hyponyms vs. hyperonyms 145-146
- in intensional field representation 144
- in syntax 138
- vs. cue validity 158
- vs. pure geographical variation 178-179
- contextual variation on entrenchment (*see* Geographical variation; Pragmatic variation; Stylistic variation)
- See also* Basic level model of onomasiological salience; Cue validity; Problems with basic level model of onomasiological salience; Unique name; Unit formation
- Epistemological characteristics of natural categories 47
- See also* Structural characteristics of natural categories
- Ethnolinguistics 146-147
- European vs. American Cognitive Linguistics 39
- See also* Usage-based approach vs. introspection
- Exocentric compound 81
- See also* Specificational compound
- Experiential nature of natural categories 47, 51
- See also* Epistemological characteristics of natural categories
- Experiments 13, 17, 157-158

- See also* Elicitation
- Expertise-related variation**
 of the basic level 146-147
 of lexical choice 183, 187, 194-195
See also Basic level, shift; Contextual variation; Stylistic variation; Technicality
- Extensional entrenchment** (*see* Configurational entrenchment)
- Extensional field representation** 125-128, 129-133
See also Intensional enrichment of extensional field representation; Non-discreteness in lexical fields, alternative representations
- Extensional non-discreteness** (*see* Membership uncertainties)
- Extensional non-equality** (*see* Extensional salience)
- Extensional salience** (*see* Degrees of representativity)
- External non-discreteness** (*see* Non-discreteness of lexical fields)
- Familiarity** 187
See also Diminutivization; Non-denotational meaning
- Family resemblance structure** 45-56
 passim, 89-91, 91-105, 159, 189
 alternative to classical definability 46, 48, 103-104
 examples (*see* Semasiological structure)
 correlation with degrees of representativity (*see* Salience effects)
See also Prototypicality, varieties; Relative feature salience
- Fashion magazines** 18-21, 32, 105-112, 112-114, 146-153, 181-188
- Features**
 specific 24-26, 30
 general 30-32
 graded 57-58, 169
 incompatibility of features (*see* Incompatible componential descriptions)
 intrinsicness of features (*see* Relative feature salience)
See also Componential analysis; Encyclopedic vs. semantic features; Global features; Quantitative interpretation of graded features; Qualitative interpretation of features
- Female speech** 187
See also Sex dimension; Sex variation
- First vs. following instances of naming** 150-152
See also Pragmatic variation
- Flexibility** (*see* Non-discreteness)
- Folk classification** 134, 146-147
See also Basic level; Taxonomy
- Folk genera** 134
See also Basic level; Taxonomy, levels
- Formal reflections of semantic phenomena**
 linguistic relevance 104, 155-156
See also Lexicological relevance of salience effects
- Formal structure of clothing names** (*see* Monolexical vs. polylexical expression; Monomorphemic vs. polymorphemic names)
- Formal variation** (*see* Formal structure of clothing names; Lexical choice)
- French structuralism** 117
See also Archiséme; Archilème
- Frequency as a prototypicality measure** 90-91, 102
See also Box diagram; Degrees

- of representativity; Prototypicality; Semasiological structure
- Frequency limitations 64, 77, 128
 - See also* Corpus of non-elicited material, restrictions
- Fruit 54-55
 - See also* Prototypicality as a prototypical notion
- Functional variation 43-44, 58, 61
 - See also* Hidden variation
- Fuzzy boundaries (*see* Membership uncertainties)
- Garment types 22-24
 - See also* Componential analysis of the *broek*-subfield, the *hemd*-subfield, the *jasje*-subfield, the *rok*-subfield, the *vest*-subfield
- General magazines 18-21, 32, 105-112, 112-114, 177-180, 181-188
- Generality requirement 56-58, 60-61
 - See also* Corpus-based approach to classical definability; Necessary and sufficient conditions
- Generic concepts (*see* Hyperonymy)
- Generic level 134, 146-147
 - See also* Basic level; Folk classification; Taxonomy, levels
- Geographical variation 1, 2, 3, 4, 9, 10, 14
 - distributional criterion 19-20, 179-180
 - editorial criterion 19, 147, 179-180
 - on entrenchment 152-153
 - on lexical choice (*see* Pure geographical variation)
 - on salience effects 105-112, 152-153
 - absence of geographical variation 111-112
 - See also* Contextual variation, Speaker-related variation
- Global features 22-24
 - See also* Referential classes of standard garments
- Glossy magazines 18-21, 32, 109-114, 146-153, 181-188
- Gricean maxims 151
 - See also* Economy of expression; First vs. following instances of naming; Pragmatic variation
- Harembroek* 26
- Hemd* 22, 23, 37, 98-105, 113, 129-133, 145, 170-172
- Hermeneutic aspects of lexical semantics 42-43
- Hidden variation 58, 61, 178, 184-186
 - See also* Functional variation; Qualitative interpretation of features
- Hippiebroek* 26
- Homonymic variation 21-22
- Hotpants* 26, 63
- Hyperonymy 1, 2, 10, 23, 59-67, 76-89, 102, 117, 119, 124, 133, 145, 146-153, 172-173
 - questionnaire-based approach 77, 81, 85-86
 - See also* Corpus-based approach to hyperonymy; Hyponymy; Lexical relations; Questionnaires; Taxonomy, relations
- Hyponymy 1, 4, 10, 23, 59-67, 76-89, 102, 124, 126, 136, 137, 139, 145, 146-153, 172-173
 - questionnaire-based approach 77, 81, 85-86
 - See also* Corpus-based approach to hyponymy; Hyperonymy; Lexical relations; Questionnaires; Taxonomy, relations
- Identificational requirement 151
 - See also* Economy of expression; First vs. following instances of naming; Pragmatic variation

- Incompatible componential descriptions 31, 91
See also Componential analysis
- Influence of extensional non-discreteness on intensional non-discreteness (*see* Definitional consequences of membership uncertainties)
- Instituut voor Nederlandse Lexicologie (*see* Corpus of non-elicited material)
- Intensional enrichment of extensional field representation 133
See also Extensional field representation
- Intensional entrenchment (*see* Definitional entrenchment)
- Intensional field representation 128-129
See also Entrenchment in intensional field representation; Non-discreteness in lexical fields, alternative representations
- Intensional non-discreteness (*see* Classical definability, absence)
- Intensional non-equality (*see* Intensional salience)
- Intensional overlapping (*see* Family resemblance structure)
- Intensional salience (*see* Family resemblance structure)
- Internal non-discreteness (*see* Non-discreteness in lexical fields)
- Introspection 14, 17-18, 39-43
See also European vs. American Cognitive Linguistics; Mental representations of word meaning; Usage-based approach vs. introspection; Wierzbicka's definition of *dress*
- Intuitive univocality 51-54, 76
See also Classical definability; Definitional polysemy; Quinean test of polysemy; Referential variability vs. polysemy; Tests to distinguish between referential variability and polysemy
- Jack* 22, 23, 36, 67-76, 152-153, 177-180
- Jasje* 36, 67-76, 128, 147-153, 178, 179, 188
- Jeans* 10, 23, 26, 30, 37, 38, 61-67, 82, 137, 145-146, 173-175, 177-180
501 jeans 34
- Jeansbroek* 145-146, 173-175
- Jodhpur* 26
- Joggingbroek* 26, 61-67
- Jurk* 23, 31, 37, 40, 136, 137, 147-153, 181-183, 184-186
- Jurkje* 37, 181-183, 184-186
- Kabelleggings* 61-67
- Klokrok* 125-129, 139-146, 147-153, 175-176
- Knickerbocker* 26
- Kniebroekje* 61-67
- Kostuum* 137
- Kuitbroek* 26, 61-67
- Language acquisition 135
See also Basic level characteristics
- Language use vs. language structure 103
See also Lexicological relevance of salience effects; Salience effects vs. relational conception of the lexicon; Structuralism
- Legging* 23, 30, 37, 59-67, 76, 77, 89-105, 112, 137, 145-146, 159-164, 164-169, 170-172, 172-173, 173-175, 177-180, 195
gedessineerde 164-169
gestreepte 164-169
kuitlange 164-169
met bloemen 164-169
rechte 164-169

- strakke* 164-169
- zwarte* 164-169
- Legging-broek* 64, 66, 77
- Leggings* 26, 61-67, 112, 145-146, 172-173, 177-180
- Lehrer's analysis of cooking terminology 121, 124-125
 - See also* Absence of lexical gaps; Lexical fields; Non-discreteness in lexical fields, traditional representations
- Lexical choice 3, 10, 176, 178, 190, 193-194, 196
 - influence of entrenchment 144, 155, 169-176, 177, 190
 - influence of prototypicality (*see* Cue validity)
 - interaction of entrenchment and prototypicality 172-173
 - contextual variation of lexical choice (*see* Age variation; Expertise-related variation of lexical choice; Pure geographical variation; Pure stylistic variation; Sex variation)
 - See also* Formal variation
- Lexical description 31
 - See also* Compound expressions; Postmodifiers; Premodifiers; Simplex expressions; Referential description
- Lexical fields 5, 7, 8, 9-10, 13, 38, 193
 - See also* Mosaic-like conception of lexical fields; Non-discreteness in lexical fields, traditional representations, alternative representations; Non-discreteness of lexical fields
- Lexical relations
 - importance for classical definability 59-67, 69-76, 77-89
 - stability 87
 - See also* Auto-hyponymy; Co-hyponymy; Hyperonymy; Hyponymy; Semi-synonymy; Synonymy; Taxonomy, instability, relations
- Lexicography 17, 18, 38, 195
 - See also* Applied linguistics
- Lexicological relevance of salience effects 102-105
 - See also* Formal reflections of semantic phenomena; Prototypicality; Salience effects vs. relational conception of the lexicon; Structuralism
- Lexicon 50, 136, 193, 195
- Linear regression 170-172, 184-186
- Loanword 31, 35, 113
- Lückenlosigkeit (*see* Absence of lexical gaps)
- Magazines with female audience 182-183
 - See also* Female speech; Sex-dimension; Sex variation
- Magazines with male audience 182-183
 - See also* Sex-dimension, Sex variation
- Membership uncertainties 45-56
 - passim*, 76-89, 114, 189
 - anecdotal evidence 85
 - influence on classical definability (*see* Definitional consequences of membership uncertainties)
 - formal considerations and membership 81
 - vs. sense and reference 87-89
 - and questionnaires 77-78, 81, 85-86
 - See also* Corpus-based approach to membership uncertainties; Morphological structure and membership; Onomasiological

- consequences of membership uncertainties; Prototypicality, varieties; Questionnaires
- Membership salience (*see* Degrees of representativity)
- Mental representations of word meaning 41-42
 - See also* Usage-based approach vs. introspection
- Metaphor/metonymy 15
- Minirok* 125-129, 139-146, 147-153, 170-172, 175-176
- Modifier selection (*see* Monolexical vs. polylexical expression)
- Monolexical vs. polylexical expression 35, 142-143
 - influence of relative feature salience 104, 164-169, 190
 - influence of entrenchment 175-176, 190
 - See also* Formal structure of clothing names; Formal variation; Modifier selection
- Monomorphemic vs. polymorphemic names 172-176, 190
 - See also* Formal variation; Morphological structure of lexical items; Simplex forms
- Monosemy vs. polysemy (*see* Referential variability vs. polysemy)
- Morphological structure of lexical items
 - on the basic level 135, 173
 - and classical definability 64, 66, 67
 - and membership 81
 - See also* Monomorphemic vs. polymorphemic names
- Morphology 35
- Mosaic-like conception of lexical fields 8, 119-122, 123, 125
 - See also* Compartmentalization of the lexicon; Trier's analysis of *Wisheit, Kunst* and *List*
- Mosaic-like picture of conceptual structures 136
 - See also* Problems with basic level model of onomasiological salience
- Mutual delimitations between lexical items (*see* Relational conception of the lexicon)
- Name frequency 141
 - See also* Entrenchment, measures
- Naming frequency proportion 141
 - See also* Entrenchment, measures
- Natural category 50
 - See also* Categorization; Classical definability; Epistemological characteristics of natural categories; Experiential nature of natural categories
- Natural language processing 195
 - See also* Applied linguistics
- Nazomerjasje* 35
- Necessary and sufficient conditions (*see* Generality requirement, Distinctiveness requirement)
- Neurophysiological research 42
- Non-classical characteristics (*see* Non-discreteness; Non-equality)
- Non-denotational meaning 186-188
 - See also* Diminutivization; Pure stylistic variation vs. hidden non-denotational variation
- Non-discreteness (*see* Onomasiological non-discreteness; Semasiological non-discreteness)
- Non-discreteness in lexical fields 117, 118, 123-134, 190
 - traditional representations 124-125
 - alternative representations (*see* Extensional field representation,

- Intensional field representation)
 - See also* Lehrer's analysis of cooking terminology; Non-discreteness of lexical fields
- Non-discreteness of lexical fields 118, 122-123
 - See also* Ducháček's starlike conception of lexical fields
- Non-equality (*see* Salience)
- Non-rigidity (*see* Non-discreteness)
- Non-standard characteristics (*see* Non-discreteness; Non-equality)
- Odd number (*see* Prototypicality, examples)
- Olifantbroek* 26
- Omslagrok* 139-146, 147-153, 173-175, 175-176
- Onomasiological alternatives in semasiological ranges 61-76, 82-85
- Onomasiological consequences of membership uncertainties 118
- Onomasiological non-discreteness (*see* Non-discreteness in lexical fields)
- Onomasiological non-equality (*see* Onomasiological salience)
- Onomasiological perspective 5, 7, 193
 - conceptual interpretation (*see* Onomasiology; Onomasiological variation vs. formal variation)
 - formal interpretation (*see* Formal variation)
 - as an addition to "standard" prototype theory 13
- Onomasiological salience 10, 176
 - generalized onomasiological salience (*see* Entrenchment)
 - intra-level differences 136-137, 145-146
 - See also* Basic level model of onomasiological salience; Problems with basic level model of onomasiological salience
- Onomasiological variation 1-16 *passim*, 117-153
 - vs. formal variation 2, 3, 4, 5, 7, 15
 - See also* Entrenchment; Non-discreteness in lexical fields; Non-discreteness of lexical fields
- Onomasiology 5, 7, 14
 - vs. semasiology 5-6
- Open texture 50
- Overhemd* 37, 98-105, 113, 114, 129-133, 145, 170-172, 173-175
- Overhemdblouse* 129-133, 145
- Overslagrok* 139-146, 147-153, 173-175, 175-176
- Pantalon* 26, 61-67, 178
- Partial co-referentiality (*see* Semi-synonymy)
- Pejorative meaning 119
 - See also* Trier's analysis of *Wisheit, Kunst* and *List*
- Philosophy 46
- Piratenbroek* 61-67
- Plooirok* 11, 125-129, 137, 139-146, 147-153, 170-172, 173-175, 175-176
- Polylexical entrenchment 142-143, 175-176
 - See also* Entrenchment, measures
- Polylexical vs. monolexical expression (*see* Monolexical vs. polylexical expression)
- Polymorphemic vs. monomorphemic names (*see* Monomorphemic vs. polymorphemic names)
- Polysemy 6, 11, 51-54, 56, 76, 149, 195
 - vs. referential variability (*see* Referential variability vs. polysemy)

- See also* Definitional polysemy; Quinean test of polysemy; Tests to distinguish between referential variability and polysemy; *Vers*
- Postmodifiers 31, 35
See also Lexical description
- Pragmatic variation
on entrenchment 148-152
See also Contextual variation; Economy of expression; First vs. following instances of naming; Gricean maxims; Identificational requirement; Situational variation
- Pragmatics vs. semantics 103
See also Salience effects vs. relational conception of the lexicon; Structuralism
- Premodifiers 31, 35
See also Lexical description
- Pre-structuralism 6, 12
- Problems with basic level model of onomasiological salience 136-137, 145-146
See also Entrenchment of hyponyms vs. hyperonyms; Mosaic-like picture of conceptual structures; Onomasiological salience, intra-level differences of; Taxonomy, instability; Taxonomies, cross-classification; Uncertainties about inclusion
- Prototype Theory 6, 11-16, 17, 18, 39-43, 45-46, 76, 83, 91, 124-125, 134-137, 138, 146-147, 157-158, 192-193
See also Categorization; European vs. American Cognitive Linguistics; Natural categories; Prototypicality; Semasiological variation
- Prototypicality 9, 14, 45-56, 193-196
on the basic level 135
as a catch-all notion 55
multiple prototypes 114
confusion around the notion of prototypicality 45, 49-56
as a prototypical notion 45, 54-56
examples (*see bird*; Color terms; *Odd number*; *Vers*)
varieties (*see* Classical definability, absence; Degrees of representativity; Family resemblance structure; Membership uncertainties)
See also Basic level characteristics; Box diagram; Frequency as a prototypicality measure; Prototype theory; Salience effects; Semasiological structure; Semasiological variation
- Psycholinguistics 17
- Psychological research 42, 103, 134
See also Basic level characteristics, Experiments; Psycholinguistics
- Pullover 35
- Pure geographical variation 177-180
editorial vs. distributional 179-180
vs. pure stylistic variation 178
vs. categorization differences 178
vs. entrenchment 178-179
See also Contextual variation; Geographical variation; Speaker-related variation
- Pure stylistic variation 178, 181-182, 183, 187-188
vs. pure geographical variation 178
vs. hidden referential variation 184-186
vs. hidden non-denotational

- variation 186-188
- See also* Contextual variation; Diminutivization as a stylistic marker; Non-denotational meaning; Speaker-related variation; Stylistic variation
- Qualitative interpretation of features 58, 61
- See also* Classical definability; Functional variation; Hidden variation
- Quantitative data analysis
 - limitations 15, 194
 - See also* Chi square test; Linear regression
- Quantitative interpretation of graded features 57-58, 60-61, 68
- See also* Classical definability; Features; Qualitative interpretation of features
- Questionnaires 13, 17
- See also* Elicitation; Hyperonymy, questionnaire-based approach; Hyponymy, questionnaire-based approach; Membership uncertainties and questionnaires
- Quinean test of polysemy 53
- See also* Tests to distinguish between referential variability and polysemy
- Radial set structure (*see* Family resemblance structure)
- Range of application (*see* Semasiological range)
- Referential approach (*see* Usage-based approach)
- Referential vs. definitional structure of a category (*see* Definitional vs. referential structure of a category)
- Referential description 22-30
- See also* Componential analysis; Corpus of non-elicited material; Features; Lexical description
- Referential frequency proportion 142
- See also* Entrenchment, measures
- Referential overlap (*see* Semi-synonymy)
- Referential range (*see* Semasiological range)
- Referential salience (*see* Semasiological salience)
- Referential variability
 - vs. polysemy 6, 11, 15, 53-54, 149, 195
 - See also* Definitional polysemy; Intuitive univocality; Quinean test of polysemy; Tests to distinguish between referential variability and polysemy
- Relational conception of the lexicon 87-89, 124, 139
- vs. salience effects 102-105
- See also* Structuralism
- Relative feature salience 104, 164-169, 190
- influence on monolexical vs. polylexical expression (*see* Monolexical vs. polylexical expression)
- See also* Features; Family resemblance structure
- Revision of semasiological range 63-64
- See also* Semasiological range
- Rok* 21, 23, 36, 37, 78-89, 125-129, 133, 136, 137, 139-146, 147-153, 170-172, 173-175, 175-176, 181-183, 184-186, 195
- geplooide* 142-143, 175-176
- klokkende* 175-176
- korte*, 142-143, 175-176
- korte geplooide* 142-143
- met omslag* 142-143, 175-176

- met overslag* 175-176
- met plooiën* 175-176
- rechte* 35
- ruime* 175-176
- ruimvallende* 175-176
- wijde* 175-176
- Rokje* 37, 125-129, 137, 139-146, 147-153, 170-172, 173-175, 175-176, 181-183, 184-186
- Rolkraagtrui* 177-180
- Salience** (*see* Onomasiological salience; Semasiological salience)
- Salience effects** 90-91
 - examples 91-105
 - in classical and non-classical categories 102
 - correlation with cue validity 157, 159-164
 - vs. relational conception of the lexicon 102-105
 - contextual variation on salience effects (*see* Geographical variation; Stylistic variation)
 - See also* Box diagram; Degrees of representativity; Family resemblance structure; Formal reflections of semantic phenomena; Frequency as a prototypicality measure; Lexicological relevance of salience effects; Prototypicality; Relational conception of the lexicon; Relative feature salience; Semasiological structure
- Saturation** (*see* Corpus of non-elicited material, saturation)
- Saussurean doctrine** 104
 - See also* Structuralism
- Semantic variation** 1, 155, 191
 - See also* Conceptual variation; Semasiological variation; Onomasiological variation; Variation, varieties
- Semasiological non-discreteness** (*see* Extensional non-discreteness, Intensional non-discreteness)
- Semasiological non-equality** (*see* Semasiological salience)
- Semasiological range**
 - blazer* 68
 - broekrok* 82
 - colbert* 68
 - jack* 68
 - jasje* 68
 - legging* 60
 - vest* 68, 109
 - See also* Onomasiological alternatives in semasiological ranges; Revision of semasiological range
- Semasiological salience** (*see* Extensional salience; Intensional salience)
- Semasiological structure**
 - blazer* 94-95
 - broek* 98
 - colbert* 92-93
 - hemd* 100
 - legging* 89-91
 - overhemd* 99
 - T-shirt* 101
 - vest* 106-108
 - See also* Box diagram
- Semasiological variation** 1-16 *passim*, 30-31, 45-116, 191-192
- Semasiology** 5, 6, 7, 8, 13
- Semi-synonymy** 59-67
 - See also* Lexical relations; Synonymy
- Sense vs. reference** 88-89
 - See also* Structuralism
- Sex dimension** 182-183
 - See also* Female speech; Magazines with female audience; Magazines with male audience
- Sex variation** 182-183, 187
 - See also* Contextual variation;

- Diminutivization; Female speech; Speaker-related variation
- Shirt* 37, 105, 112-114, 129-133, 145, 173-175, 188
- Short* 23, 145-146, 170-172, 173-175
- Shorts* 23, 26, 37, 145-146, 170-172, 173-175
- Sigarettenpijp-broek* 64
- Simplex forms 173-175
 See also Monomorphemic vs. polymorphemic names
- Simplex expressions 31, 33-35
 See also Lexical description
- Situational variation 8, 13, 149-153, 191, 193
 See also Contextual variation; Pragmatic variation
- Skibroek* 26, 30, 35, 61-67
- Sociolinguistic variation 8, 36
 See also Contextual variation
- Sociolinguistics 13
 See also Contextual variation
- Source groups 20-21
 See also Alternative classification of source groups; Fashion magazines; General magazines; Glossy magazines; Magazines with female audience; Magazines with male audience
- Speaker-related variation 8, 13, 153, 191, 193
 See also Age variation; Contextual variation; Geographical variation; Pure geographical variation; Pure stylistic variation; Sex variation; Stylistic variation
- Specialization dimension 19-20
 See also Expertise-related variation; Stylistic variation
- Specialization variation (*see* Expertise-related variation; Stylistic variation)
- Specialized sources (*see* Fashion magazines)
- Specificational compound 64, 81
- Spijkerbroek* 26, 38, 145-146, 173-175, 177-180
- Statistics (*see* Quantitative data analysis)
- Stipblouse* 34
- Streepbroek* 61-67
- Stress pattern 81
- Stretchbroek* 61-67
- Stretchleggings* 61-67
- Structural characteristics of prototypical categories 45-46, 47
 See also Prototypicality, varieties
- Structuralism 6, 11-13, 37-39, 87-89, 102-105, 117, 118-124, 192-193
 See also Checklist theory of meaning; Componential analysis, psychological status, referential instead of semantic description; Ducháček's star-like conception of lexical fields; Encyclopedic vs. semantic features; Formal reflections of semantic phenomena; French structuralism; Mosaic-like conception of lexical fields; Pre-structuralism; Quinean test of polysemy; Referential variability vs. polysemy; Relational conception of the lexicon; Saussurean doctrine; Sense vs. reference; Tests to distinguish between referential variability and polysemy; Trier's analysis of *Wisheit, Kunst* and *List*
- Stylistic variation 1, 2, 3, 4, 5, 10, 194-195
 on entrenchment 148
 on salience effects 105-112

- on lexical choice (*see* Pure stylistic variation)
- See also* Contextual variation; Speaker-related variation
- Supra-generic level as basic level 146
 - See also* Taxonomy, levels
- Supra-lexical category 117, 176, 193
 - See also* Lexical fields
- Surveys (*see* Questionnaires)
- Sweatshirt 31
- Synonymy 1, 2, 4, 59-67, 69, 74, 138, 177-180
 - extensional perspective 139
 - intensional perspective 139
 - See also* Lexical relations, Semi-synonymy
- Taxonomy 10, 134
 - instability 87, 136
 - levels 134, 146-147, 190
 - relations 78-89
 - cross-classification of taxonomies 136
 - See also* Basic level; Folk classification; Folk genera; Generic level; Hyponymy; Hyperonymy; Lexical relations; Membership uncertainties; Problems with basic level model of onomasiological salience; Supra-generic level as basic level
- Technicality 146-147, 183, 187
 - See also* Expertise-related variation
- Tests to distinguish between referential variability and polysemy 6, 51, 53
 - See also* Quinean test of polysemy; Definitional polysemy
- Text corpora 14
 - See also* Corpus of non-elicited material
- Topje* 37, 129-133, 145, 170-172, 173-175
- Tricotbroek* 61-67
- Trier's analysis of *Wisheit, Kunst* and *List* 118-122
 - See also* Lexical fields; Mosaic-like conception of lexical fields; Structuralism
- Trui* 22, 23, 37, 136, 137, 147-153, 181-183, 184-186
- Truitje* 37, 181-183, 184-186
- T-shirt* 37, 98-105, 113, 114, 129-133, 145, 173-175
- Type frequency 141
 - See also* Entrenchment, measures
- Typicality ratings 91
- Uncertainties about inclusion 136
 - See also* Hyponymy; Hyperonymy; Lexical relations, stability; Problems with basic level model of onomasiological salience; Taxonomy
- Uncertainty of membership (*see* Membership uncertainties)
- Unidimensional definitions 128-129
 - See also* Classical definability; Definitional entrenchment
- Unit formation 138
 - See also* Entrenchment
- Unique name 138, 142, 175, 177
 - See also* Entrenchment
- Usage-based approach 193-194
 - vs. introspection 17-18, 39-43
 - objectivity of the usage-based approach 42-43
 - See also* Corpus of non-elicited material; European vs. American Cognitive Linguistics; Introspection; Mental representations of word meaning; Wierzbicka's definition of *dress*
- Vagueness (*see* Referential variability)
- Variation

- varieties 1-16, 36, 189, 191-192, 195
- interaction 3, 15, 191-192
- See also* Contextual variation; Formal variation; Semasiological variation; Onomasiological variation
- Vers*
 - See also* Polysemy; Prototypicality, examples
- Vest* 22, 23, 36, 105-112, 152-153, 188
- Vest* 67-76, 159-164
- Vestje* 36
- Vliegeniersjack* 24
- Weekendjasje* 35
- Wielrennersbroek* 61-67
- Wierzbicka's definition of *dress* 39-40
 - See also* European vs. American Cognitive Linguistics; Usage-based approach vs. introspection
- Wikkelrok* 11, 125-129, 137, 139-146, 147-153, 173-175, 175-176