

Lexical Creativity, Texts and Contexts

EDITED BY Judith Munat

John Benjamins Publishing Company

Lexical Creativity, Texts and Contexts

Studies in Functional and Structural Linguistics (SFSL)

Taking the broadest and most general definitions of the terms functional and structural, this series aims to present linguistic and interdisciplinary research that relates language structure — at any level of analysis from phonology to discourse — to broader functional considerations, whether cognitive, communicative, pragmatic or sociocultural. Preference will be given to studies that focus on data from actual discourse, whether speech, writing or other nonvocal medium.

The series was formerly known as *Linguistic & Literary Studies in Eastern Europe (LLSEE)*.

Founding Editor

John Odmark

Honorary Editors

Eva Hajičová
Charles University

Petr Sgall
Charles University

General Editors

Yishai Tobin
Ben-Gurion University of the Negev

Ellen Contini-Morava
University of Virginia

Editorial Board

Alexandra Y. Aikhenvald
La Trobe University

Joan Bybee
University of New Mexico

Nicholas Evans
University of Melbourne

Victor A. Friedman
University of Chicago

Anatoly Liberman
University of Minnesota

James A. Matisoff
University of California, Berkeley

Jim Miller
University of Auckland

Marianne Mithun
University of California, at Santa Barbara

Lawrence J. Raphael
CUNY and Adelphi University

Olga Mišeska Tomić
Leiden University

Olga T. Yokoyama
UCLA

Volume 58

Lexical Creativity, Texts and Contexts

Edited by Judith Munat

Lexical Creativity, Texts and Contexts

Edited by

Judith Munat

University of Pisa

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Lexical creativity, texts and contexts / edited by Judith Munat.

p. cm. -- (Studies in functional and structural linguistics, ISSN 0165-7712 ; v. 58)

Includes bibliographical references and index.

1. Lexicology. 2. Creativity (Linguistics) I. Munat, Judith.

P326.5.C74L49 2007

401'.4--dc22

2007020957

ISBN 978 90 272 1567 3 (Hb; alk. paper)

© 2007 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

List of tables	VII
List of figures	IX
List of contributors	XIII
Editor's preface	XV
Introduction	
Lexical creativity, textuality and problems of metalanguage <i>Leonhard Lipka</i>	3
Lexical creativity in discourse	13
How to do (even more) things with inonce words (other than naming) <i>Peter Hohenhaus</i>	15
The phonetics of 'un' <i>Jen Hay</i>	39
Lexical creativity in texts	
<i>The press</i>	
Tracing lexical productivity and creativity in the British Media: 'The Chavs and the Chav-Nots' <i>Antoinette Renouf</i>	61
<i>Cartoon art</i>	
Cathy Wilcox meets the phrasal lexicon: Creative deformation of phrasal lexical items for humorous effect <i>Koenraad Kuiper</i>	93
<i>Advertising and the media</i>	
Blendalicious <i>Adrienne Lehrer</i>	115

Electronic communication

- Keeping up with the times: Lexical creativity in electronic communication 137
Paula López Rúa

Fictional genres

- Lexical creativity as a marker of style in science fiction and children's literature 163
Judith Munat

Creative concept formation

- Dynamic creation of analogically-motivated terms and categories in lexical ontologies 189
Tony Veale
- Creative lexical categorisation in a narrative fiction 213
M^a Dolores Porto

Sociopolitical effects on creativity

- Occasional and systematic shifts in word-formation and idiom use in Latvian as a result of translation 239
Andrejs Veisbergs
- Critical creativity: A study of 'politically correct' terms in style guides for different types of discourse 263
Roswitha Fischer
- Name index 283
- Subject index 287

List of tables

Jen Hay

Comparison of degree of reduction of affixed and non-affixed forms	45
Wald statistics for model predicting the relative length of <i>un-</i> , over both affixed and unaffixed forms	46
Coefficients for model predicting relative length of <i>un-</i> over both affixed and unaffixed forms	46
Comparison of degree of vowel reduction in affixed forms by relative frequency of the derived form and the base	48
Wald statistics for model predicting the relative length of <i>un-</i> in affixed words	49
Coefficients for model predicting the relative length of <i>un-</i> in affixed words	49
Comparison of words containing <i>un-</i> and the lexical item <i>not</i> for early and late speakers	53
Comparison of <i>un-</i> prefixed adjectival forms and sequences of <i>not</i> + adj. which could be expressed with a prefixed form	54

Antoinette Renouf

Prefixes with growth or decline in corpus frequency ranking across 16 years	64
New prefixal formations in years of peak productivity for <i>cyber-</i> , <i>faux-</i> , <i>techno-</i> and <i>uber-</i>	65
Suffixes with growth/decline in frequency across 16 years	66
Some new suffixal formations in year of peak productivity for <i>-fest</i> and <i>-athon</i>	66
<i>Symbiosis</i> and variants with frequencies in <i>Independent/Guardian</i> , 1989–2005	67
Productivity for lexeme <i>synergy</i> , 1989–2005	68
Productivity for lexeme <i>NIMBY/nimby</i> , 1989–2005	69
Variants of <i>weapons of mass destruction</i> , Feb. 20, 2003	73
Creativity for <i>weapons of mass destruction</i> , Feb. 21, 2003	73

Decreased creativity of <i>weapons of mass destruction</i> , Feb. 28, 2002–May 29, 2004	74
Inflections of <i>chav</i> emerging during 2004–2005	76
Creative variants based on <i>chav</i> , 2004–2005	76
Evolution of <i>hoodie</i> references in <i>Independent/Guardian</i> data, 1989–2005	78
Collocational pattern for <i>hoodie</i> in Nov. 2004	78
Collocational profile for <i>hoodies</i> in Feb. 2005	79
Instances of <i>neet</i> since late 2004	80
Creativity for acronym <i>neet</i> since its inception in late 2004	80
<i>Tsar</i> and its inflections, 1989–2005	82
Significant left-hand collocates for <i>tsar</i> , 1996–2005	83
Significant left-hand collocates for <i>guru</i> , <i>supremo</i> and <i>pundit</i> , 1996–2005	83
Multi-word compounds based on <i>tsar</i> , 2000–2005	84
Puns on the vogue word <i>tsar</i> , 1996–2005	86
Koenraad Kuiper	
Number of PLIs per cartoon	103
Lexical substitutions	104
Relationships between lexical items	105
Lexical exchanges	106
The PLI deformation signature of Cathy Wilcox	109
Paula López Rúa	
Prefixation in on-line jargon	148
Suffixation in on-line jargon	149
Blends and clipped compounds: borderline cases	155
Tony Veale	
Cross-domain variants are sought for each member of five deity pantheons	207
M^a Dolores Porto	
Members of the lexical field of magician in <i>Forging the Darksword</i>	226
Frequency of the members in the semantic category of <i>MAGUS</i>	232
Roswitha Fischer	
Style guide recommendations for academic writing	270
Style guide recommendations for official documents	270
Style guide recommendations for news language	270

List of figures

Jen Hay

Model predictions showing the effect of the following syllable page	47
The effect of weight of the following syllable on the relative length of <i>un-</i>	50
The effect on relative length of <i>un-</i> when the following syllable carries lexical stress and when the word is utterance final	50
The effect of the phonotactics across the morpheme boundary on the predicted relative length of <i>un-</i>	51
Interaction between a speaker's period of birth and the frequency of word or base	52

Antoinette Renouf

Time series plot for <i>weapons of mass destruction</i> , 1989–2005	72
Time series frequency plot for <i>tsar</i> , 1989–2005	81

Koenraad Kuiper

Cathy Wilcox cartoon	108
Cathy Wilcox cartoon	110

Tony Veale

Structure of the Greek and Hebrew letters domain in WordNet	206
WordNet supplemented with new adhoc categories like Greek-letter, Hebrew-letter and 1st-letter	206

M^a Dolores Porto

Radial categories as seen by Cuenca and Hilferty (1999)	217
A partial representation of the senses of <i>ring</i> (Langacker 1991)	220
A partial representation of the meaning of <i>dead</i> as a category of senses	221
Acquisition of a new sense for the word <i>dead</i>	222
Recategorisation of the meaning of <i>dead</i> with the new sense as the central member of the category	224
Frequency of the members of the invented category FLOAT in <i>Forging the Darksword</i>	229
A schematic representation of the semantic category of HUMAN MOTION IN THE AIR in <i>Forging the Darksword</i>	231

List of contributors

Roswitha Fischer
Institut für Anglistik und Amerikanistik
Universität Regensburg
93040 Regensburg, Germany
roswitha.fischer@sprachlit.uni-regensburg.de

Jen Hay
Department of Linguistics
University of Canterbury
Private Bag 4800
Christchurch, New Zealand
jen.hay@canterbury.ac.nz

Peter Hohenhaus
Rauscherstr. 6/41
A-1200 Vienna, Austria
phohenhaus@yahoo.co.uk

Koenraad Kuiper
Department of Linguistics
University of Canterbury
Private Bag 4800
Christchurch 8020, New Zealand
kon.kuiper@canterbury.ac.nz

Adrienne Lehrer
Department of Linguistics
University of Arizona
Tucson, AZ 85721, USA
alehrer@u.arizona.edu

Leonhard Lipka
Rothschwaigestrasse 51

D-80997 München, Germany
Leonhard.Lipka@anglistik.uni-muenchen.de

Paula López Rúa
University of the Balearic Islands
Edifici Ramon Llull, Campus UIB
C. Valldemossa, km 7.5
07122 Palma de Mallorca (Balearic Islands)
Spain
paula.lrúa@hotmail.com

Judith Munat
Dipartimento di Anglistica
University of Pisa
Via Santa Maria 67
56126 Pisa, Italy
jemunat@gmail.com

Antoinette Renouf
School of English
University of Central England in Birmingham
Franchise Street
Perry Barr
Birmingham B42 2SU, UK
Antoinette.Renouf@uce.ac.uk

Dolores Porto Requejo
Dpto. Filología Moderna
Colegio Caracciolos
c/ Trinidad, 3
28801 Alcalá de Henares, Spain
mdolores.porto@uah.es

Tony Veale
School of Computer Science and Informatics
University College Dublin
Belfield, Dublin D4, Ireland
Tony.veale@UCD.ie

Andrejs Veisbergs
Visvalza 4a
LV 1050 Riga, Latvia
anveis@lanet.lv

Editor's preface

This collection of articles has its origin in a seminar I conducted at the 2004 ESSE Conference in Zaragoza (Spain), entitled “Lexical Creativity as a Feature of Textuality.” Two papers from that seminar, together with ten invited contributions make up the content of the present volume. This is a multi-faceted, predominantly synchronic investigation of lexical creativity viewed in relation to a variety of textual and contextual, including sociopolitical, constraints, in which nonce words and other new coinages are analysed in relation to co(n)text, speaker/writer intentions and the wider social environment.

The noted artist, Richard Serra, in speaking of his steel sculptures, has said that it is the *context* which makes him think and which sets the wheels in motion, serving as a stimulus to his monumental artistic constructions.¹ The role of context in lexical creativity is no less significant, and only by considering the specific (written or spoken) environments in which these novel formations appear can we fully comprehend the meanings intended by the speaker (a parallel with the creative artist) and the functions that they are intended to fulfil. Close observation of the context in which new lexical items occur will allow us to capture the clues that serve to interpret these unfamiliar and often semantically opaque words and expressions, and discover how these novel formations both reflect and implement the surrounding co(n)text. Thus the papers in this volume, while considering the way in which morphological theory affects the creation of novel words, also consider the way that these creative formations illuminate morphological theory.

The majority of coinages examined here will never find a permanent place in the lexicon because of their ephemeral nature, their utility being strictly bound to their original contexts of use. Nonetheless, these lexical items, whether creative manipulations of the rules or new formations based on canonical word-forming

1. These words are cited from the documentary film “Richard Serra. Thinking on your feet” made by Anna Tappeiner (Germany, 2005).

processes, all contribute to the construction of discourse meaning and text worlds in subtle and interesting ways.

Creativity, according to Pope (2005: xv) is always "creation from something," an on-going process in which the speaker/writer draws on a finite number of existing items from which an infinite number of fresh or imaginative solutions are 'created' to satisfy the communicative needs of the moment. This is creativity in the Chomskyan sense, where a limited set of rules and a finite set of terms allow humans to produce an infinite number of novel structures (be they words or longer utterances) that no one has previously produced. We might well reformulate the eternal question of creativity vs. productivity in the domain of word formation as originality informed by knowledge. Therefore, creativity is not to be seen simply as the product of 'genius' of a single user, but a collaborative effort between the speaker, who (re)combines the building blocks of the system, and the listener, who (re)constructs meaning by relying on linguistic as well as sociocultural knowledge (for a full discussion see Carter 2004). In the same manner, the reader approaches the interpretation of unfamiliar or novel words in written text.

Lamb (1998: 205) defines real creativity as the invention of new lexemes for new or old concepts, allowing us to build new concepts by drawing on and integrating ideas in our conceptual system that have not previously been connected; Benczes (2006: 7) similarly identifies creativity as our ability to make new associations between concepts based on "similarity, analogy or contiguity."

This overview of linguistic creativity is a synthesis of the approach taken in this volume. These data-based studies all aim to illuminate the way in which lexical creativity emerges from the interplay between a given communicative, textual or social environment and the speaker's manipulation of the system, as he coins new words, recombines existing expressions or forms new concepts.

Thus a children's story is seen to spawn novel lexical creations which are sharply distinct from those in the Science Fiction genre, in that they are determined by different textual worlds and reader-oriented agendas. Another study discusses the constraints operating on electronic communication to explain the unique types of word formations, abbreviations and acronyming used in text messaging or computer-mediated discourse. An in-depth study of British broadsheets over an extended period shows wordplay in the media as a self-generating creative process and documents the vast numbers of variations on existing words and expressions that are rapidly coined in the aftermath of certain media events. An investigation of the humorous effect achieved by the reformulation and manipulation of phraseological expressions sheds light on the semiotic interplay between visual image and text in cartoon art, while a study of ad hoc formations in interpersonal interaction shows that the metafunctions of novel formations in speech go well beyond the naming of new concepts or entities. And an examination of creative blends,

drawn principally from advertising and the media, illustrates novel wordplay as an attention-seeking device, at times bordering on a marketing strategy.

The area of concept creation is examined in two studies, one focussing on the generation of lexical analogies – similar to the creation of new noun compounds – in computer-based ontologies such as WordNet. The other concentrates on the way in which concept formation and semantic recategorisation serve to disambiguate neosemes and novel formations in the world of fantasy fiction.

The importance of sociopolitical attitudes in effecting language change, often conservative rather than innovative, is examined in relation to the imposition of 'politically correct' terms in government and other public documents; and political ideologies are identified as one of the principal motivating forces behind translation policies which have effected shifts in wordformation and idiom use in Latvian.

Also, an original diachronic study of recorded New Zealand speech offers insights into the phonotactics of compositionality, and suggests that morphological creativity may be reflected in phonetic detail, thereby providing a quantitative measure with which to distinguish stored words from creatively formed ones.

Lipka, in his introductory chapter, highlights and expands upon the various issues raised in these studies, adding numerous personal insights. He makes important connections between earlier contributions in the field (many of which his own) and the functional approach of the present volume. His introduction will enable the reader to appreciate the many links between the research presented here and earlier studies in the fields of morphology, word formation, lexicology, lexical semantics, discourse analysis and stylistics. While the rich bibliographies accompanying the single papers provide ample and up-to-date references to recent research in all of the foregoing fields, the older, classical works on word formation have not been ignored.

The original and thoughtful studies presented in this volume ultimately illustrate the way in which lexical creativity in both spoken and written discourse is determined by functional considerations relating to the particular contexts of occurrence, identifying the text (including co-text and wider social context) as the most significant stimulus to lexical innovation.

In the certainty that the issues raised in this volume will be of great interest to the morphological scholar, the student of lexicology, the stylistician and translator, as well as the researcher in computational linguistics and artificial intelligence, it is also my hope that the reader will find stimulating suggestions for further study and research.

Finally, I wish to express my special thanks to Peter Hohenhaus who provided numerous valuable suggestions and comments as the volume progressed and who was of paramount assistance in the course of the editing process.

References

- Benczes, R. 2006. *Creative Compounding in English*. Amsterdam: John Benjamins.
- Carter, R. 2004. *Language and Creativity: The Art of Common Talk*. London: Routledge.
- Lamb, S. 1998. *Pathways of the Brain: The neurocognitive basis of language*. [Current Issues in Linguistic Theory 170] Amsterdam: John Benjamins.
- Pope, R. 2005. *Creativity: Theory, History, Practice*. London and New York: Routledge.

Judith Munat
June 2007

Introduction

Lexical creativity, textuality and problems of metalanguage

Leonhard Lipka

1. Introduction and terminology

1.1 Creativity vs. productivity in word-formation (WF) patterns

The notion of productivity has long played a crucial role in morphological theory, e.g. in H. Marchand's theory of WF (Marchand 1969; Kastovsky 2005), in which he states that WF "studies the patterns on which language forms new lexical units, i.e. words" (1969:2). Examples of such patterns, categories or types are 'Noun + Adjective' (*waterproof, grass-green*), 'Noun + Noun', or derivational agent-nominalisations in *-er, -ist*. Bauer (1983:63) – following Lyons – defines **productivity** as rule-governed innovation, as opposed to **creativity**, which is said to be rule-changing. Here, we will follow this distinction but also use both terms synonymously, in a wider sense, where appropriate.

In general, various types (or categories) of processes for extending the lexicon may be distinguished, which I have referred to, globally, as **dynamic lexicology** (Lipka 2002a). This term renders J. Tournier's (1985:51) *dynamique lexicale* who distinguishes four categories, or macro-mechanisms of lexical innovation (see Lipka 2002a:109): 1. morpho-semantic neologism, i.e. basically WF, 2. semantic neologism, or semantic transfer (ST), viz. metaphor and metonymy (M & M), 3. morphological neologism, i.e. reduction processes like clipping, blending and acronymy and 4. external, i.e. loan processes.

Downing (1977) – in an article on compound nouns (discussed in Lipka 2002a:110f.) – argued that N+N is a practically unrestricted pattern of WF which can even serve purely deictic functions, as in her famous example *apple-juice seat*, while Clark and Clark (1979) postulated a class of innovative denominal verbs depending heavily on **context**, therefore called *contextuals* by them. These may be

interpreted as zero-derived verbs (conversions), parallel to suffixal derivatives, as in *to Houdini out of a closet*. All such instances of **lexical creativity** (in the widest sense) naturally depend strongly on specific *registers* or *text-types* (TTs) – see Lipka (1999a) – as investigated in detail in the contributions to the present volume.

1.2 The metalanguage of text linguistics (TL): Textuality, context vs. co-text, genre, text-type, register

As pointed out in Lipka (1999a) many terms used in the study of texts, such as cohesion and coherence, are used differently by different authors and therefore require a nominal definition. Since they must be explicitly defined they are notational terms (see Lipka 2002a). Thus textuality, according to Beaugrande and Dressler (1981: 10ff.), is based on seven standards or criteria which are seen as constitutive principles and these include intertextuality which, as the authors state, is responsible for text-types. Intertextuality refers to the factors which make one text dependent on others previously encountered (pre-texts). According to Brown and Yule (1983: 46) co-text, i.e. verbal context, must be distinguished from the wider context of situation. The recognition and inclusion of both is one of the essential features of Linguistic Pragmatics in the study of texts, in addition to the importance of speech acts (Searle 1969). As discussed in Lipka (1999a: 79f.), Esser (1991: 142) distinguishes genre, as a culture-specific, conventional, literary form of communication (German *Gattung*), from text-type (German *Textsorte*) on the grounds of purely linguistic characteristics of a text. In Lipka (2002a: 22f.) I define register, following Strang (1968) and Halliday (1978), as a combination of 1. field, 2. tenor and 3. mode of discourse.

In Lipka (2002b) I used serious and non-serious (= humorous) as the starting point for a binary distinction of texts. In this article I discussed several possible taxonomies of TTs. These may include relations of intersemioticity (coined in analogy to intertextuality) as in the lay-out, writing and illustrations of death notices. In certain death notices discussed in Lipka (2002b), humour may be in the eye of the beholder: shape is relevant in the *Book of Grave Humour* (Spiegel 1977), which itself is shaped like a gravestone (as is its sequel *Dead Funny* (Spiegel 1983)). Intersemioticity, i.e. visual aspects and devices, also plays a role in the textual examples analysed in Lipka (1999a) which use illuminated initial letters.

1.3 The categorisation and typology of texts

As discussed in Lipka (1999a: 80f.), Biber's (1988) study, based on the Lancaster-Oslo-Bergen or LOB Corpus of English and the London-Lund or LL Corpus, after

identifying the dimensions *written/spoken* and *oral/literate*, singles out four genres which show widely different situational characteristics but no one-to-one correspondence between speaking/writing and oral/literate: face-to-face conversation, academic expository prose, academic lectures and personal letters. He makes very fine distinctions – on purely linguistic criteria – between a variety of different text types and while they are clearly not all relevant for this volume, *intertextuality* certainly is since it is the basis for such lexical creativity as allusion, parody, word-play and humour, which are crucial in a number of chapters here. Obviously, there is no strict hierarchical sub-classification of texts but rather a cross-classification, and different parameters may be chosen as more or less important.

1.4 Devices, processes, domains and text-types (TTs)

All of the articles in this volume discuss creative phenomena in relation to specific domains of language use and a variety of TTs. Several of the titles make use of *attention-seeking devices* (ASDs) (see Lipka 2002a: 147, 189) of various sorts to hook the reader – verbal ASDs, such as allusions to and examples of the topic being discussed, as in *How to do things with words* or *Blendalicious*. In principle, such ASDs are widely employed in specific TTs such as advertising or media headlines. Humorous intentions of writers in non-serious TTs are the focus of several articles in this volume; in one case the domain or semiotic TT of cartoons is the object of analysis; and lexicalisation and institutionalisation (L & I – see Hohenhaus 2005, Lipka 2005) are focussed on in several other chapters. Non-lexicalised nonce-formations are discussed in relation to science fiction, the media, electronic communication, and children's literature. In addition, creative WF processes and patterns, loan words or formations, translation and language policy are dealt with in the single chapters. The contributions in this volume cover a wide range of creative devices and textual domains, and many privilege the functional aspects of WF.

2. Dynamic lexicology and neglected functions of WF

2.1 WF and dynamic lexicology in general

Lexical creativity is intrinsically connected with productive WF processes, but also involves other aspects of *dynamic lexicology* (DL) (see 1.1 and index in Lipka 2002a) such as metaphor, metonymy and loan words (see Lipka 1994). In my opinion, acronyms, clippings, blends – as in Marchand (1969) – also belong to

WF. The function and usefulness of the productive process of *acronyming* are in fact demonstrated in the present article. They can also be witnessed in the classic use of the linguistic metalanguage in the Prague School of Functional Sentence Perspective (FSP), such as *theme* (T) and *rheme* (R) and *communicative dynamism* (CD); the popularity of acronymy within the framework of transformational-generative grammar (TG) is yet another case in point. In this volume, López Rúa in particular analyses shortening devices such as acronyming in electronic communication.

2.2 Functions of WF

Hohenhaus (1996, 2005) and Lipka (2000b, 2002a) stress the various functions of WF and the importance of *lexicalisation* and *institutionalisation* (L & I), which have long been neglected. Naming, or denotation by means of simple and complex lexemes, was previously considered from the perspective of Lexical Semantics, thus focussing almost exclusively on their nomination function (see Lipka 1999b) but the denotation and the reference of more complex expressions used as *naming units* (NUs) must also be considered from an onomasiological perspective. Other communicative functions, such as purely textual functions, were rarely mentioned in the literature, except in relation to anaphoric and cataphoric lexical cohesion in texts. But much more can be said about such functions. Similarly, proper names and their meanings were often excluded from the domain of linguistics (but see Lipka 2000a) in that they were considered as belonging to general world knowledge, and thus part of encyclopaedias but not of dictionaries. However, given that morphology and semantics are crucial to the selection of multi-lexemic proper names, these are, in effect, relevant to the discussion of lexical creativity.

3. Overview of the volume

The present volume covers a broad spectrum of topics and creative lexical devices and considers a wide range of TTs and registers. Patterns and categories of WF play a central role in most articles, and blends are specifically treated in Lehrer's contribution. Humour, fiction and ASDs are also examined in many of the chapters. Analogy (Veale) – also in translation (Veisbergs) – is treated, as well as allusion, intertextuality, and context-dependent nonce-formations. L & I are occasionally mentioned (esp. in Hohenhaus) and inter-semiotic aspects of creativity, esp. in the use of pictorial and verbal communication in cartoons, is dealt

with by Kuiper. Domains and TTs covered include political correctness, children's literature, science fiction and fantasy narrative, the British media, electronic communication and lexical ontologies (esp. WordNet).

Fischer's chapter is primarily concerned with critical creativity as expressed by politically correct (PC) terms in various speech communities and TTs, viz. style guides for official and academic writing and reference books, including those produced by *The Guardian* and *The Independent*. She identifies the principal strategies in critical creativity as those of unification (truncation, synonymy, neologisms), specification and euphemism. The goal of PC is respectful and equalitarian treatment of people by avoiding social stigmatisation, and thus alternative terms are recommended to replace those which are seen as discriminatory. Linguistic creativity in this case is related to sociocultural phenomena in that new PC terms are created as alternatives to institutionalised, established lexical items considered to be offensive.

Hay's contribution represents a unique approach to lexical creativity in various respects: hers is a study of the phonetic and phonological differences between retrieved words and productively coined words in a corpus of spontaneous New Zealand oral speech, a study which also takes into consideration diachronic and psycholinguistic aspects of creativity. Affixal (*unfortunately*) and non-affixal (*understand*) forms of *un-* are analysed for vowel quality, length and pitch accents and the author concludes, on the basis of her statistical data, that words which are more easily decomposable (such as *unburstable*) tend to have prefixes which are more prominent phonologically than other frequent, highly lexicalised forms (such as *unfortunately* or *unnecessary*). Particular attention is given to the phonetic context of distinctly creative nonce-words, and ultimately Hay suggests that phonetic reflexes of morphological structure may allow us to assess productive WF processes.

Hohenhaus uses allusion (to Austin's 1962 book/lectures) as an ASD in his title to draw attention to functions of nonce WF other than that of naming, and his chapter focusses on fundamentally pragmatic, communicative aspects of nonce formations. Naming has often been seen as *the* central role of novel words so that other functions have typically been neglected in the research on WF, but Hohenhaus topicalises these, including deixis or ASDs. He also discusses nonce formations which appear in particular TTs, including colloquial spoken dialogue and fictional conversation, identifying a variety of meta-communicative functions. The intentions of the text producer or sender are considered (among which humour, objection to forms of address, fictional hypostatisation, stereotypification, etc.). A specific function relating to a new pattern of WF – as investigated in an earlier article by Hohenhaus (2004) – namely Identical Constituent Compounding (ICC), e.g. *friend-friend* (meaning: 'a proper, true friend'), serving to clarify

(potential) misunderstandings, is also discussed, as are several more specialised functions, e.g. in genres of ‘virtual reality’.

Kuiper analyses lexical creativity in an inter-semiotic environment which combines both verbal and visual communication, in the cartoons by Cathy Wilcox published in the *Sydney Morning Herald*. The aspects of creativity studied here are the manipulation of existing phrasal lexical items (PLIs) serving a humorous effect in a non-serious TT in which language play and artistic deformation is crucial; wordplay here is based on the reader’s set of stable expectations. Kuiper’s approach is primarily psycholinguistic in that PLIs are seen as being stored and retrieved from the speakers’ mental lexicons (e.g. *The Leaning Tower of Pisa*) and the subsequent modifications or deformations (in this case phonological: *the leaning piece of towser*) can only be perceived in relation to the familiar stored item. Humour thus derives from the interplay between the modified lexical items and the reader’s expectations.

Lehrer exploits the WF process of blending in the title of her chapter: *Blendalicious*, which we may consider as an ASD. She states that some blends (*workaholic*, *brunch*, *motel*) become completely conventionalised while others remain one-off inventions, and that blends are particularly frequent in newspaper and magazine headings, advertisements and new product names. Her method for collecting empirical data is basically *participant observation* (PO, see Lipka 2002a) but she also uses experiments on the processing of novel blends in context. She states that blends, which are probably used in all languages, are often “cute” or amusing, and that they occur most frequently in jargons such as computer speak, blog speak and on the Internet. They are increasingly common in the written medium, and in advertising they are often accompanied by pictures, photos, graphic devices and other visual material, in general serving as ASDs to invoke humour.

López Rúa investigates a corpus of English and Spanish linguistic material used in electronic communication, namely two types of data: 1. text messages (SMSs) and 2. widely used vocabulary items (not lexemes!) gathered from glossaries, journals, e-magazines, articles and web pages. The empirical data are analysed with regard to three aspects: morphological mechanisms, motivations and implications for everyday language. Creativity and productivity may be illustrated by the recent *cyber thief*, *data mining*, *keylogging*, *jewel case* (plastic box for a CD), which are more or less lexicalised and institutionalised beyond the jargon of expert users, known as *netizens* and *mobile freaks*. Among the motives underlying these new items are economy of time and space, ludic and sociolinguistic needs (in-group marking). Electronic communication is shown as a fertile ground for neologisms and creativity, with text messaging as a particular subcode using a variety of shortening devices viz. clipping, blending, acronyming, initialising and the use of homophone numbers (2, 4) for words as in *Wan2Tlk?*

Munat is specifically interested in the function of novel lexical formations (NLFs) as style markers in texts. She focusses here on two genres, science fiction (SF) narratives and another defined by addressee: children's literature. The data drawn from the two genres are compared with a view to identifying the different types of creative formations that occur, and these are explained in relation to a variety of stylistic, pragmatic, cultural and reader-related factors. Preliminary to the discussion of her data, Munat considers a number of theoretical issues, such as creativity vs. productivity (of categories, types and patterns), context and functional concerns of nonce formations, as well as motivation and pragmatic constraints in naming. She sees the *raison d'être* of NLFs as a momentary naming need related to the pragmatics of the text, in particular that of "world creation" (a pseudo-scientific textual world) in Science Fiction. Marchand's (1969) discussion of "phonetic symbolism" within the scope of WF is certainly relevant here, as is Crystal's (1998: 1ff.) *ludic language*.

Porto Requejo studies naming and lexical categorisation in fantasy fiction, demonstrating the way in which creative formations are disambiguated by the reader through a process of creative reorganisation of lexical categories. Semantic change viz. metaphor and metonymy (M & M) and WF responds to the need to name new objects and concepts that are part of a different (fictional) reality and these new words may effect a change in our perception of that reality. By continually adjusting the information that is stored in the mental lexicon, the reader is able to cope with nonce formations and neosemes pertaining to the fictional world. Porto Requejo argues that our mental lexicon is organised in prototypical categories and that creative categorisation in a fantasy novel can be viewed from both a semasiological and onomasiological perspective.

Renouf's study focusses on the TT of British media as represented by broadsheet journalism from *The Independent* and *The Guardian*. Her data is primarily drawn from a newspaper corpus of over 700 million words, plus additional data from the web. She examines lexical creativity such as metaphorical extension and wordplay, as exemplified by the pun on *Chav* (equivalent to US *white trash*) in the title of her chapter. She employs the method of Observational Linguistics (OL), assisted by software and electronically-held texts, exploring productivity and creativity in a diachronic corpus. Renouf illustrates, through the construction of graphs of individual prefixal and suffixal patterns over time, that it is possible to plot periods and years of peak productivity. Creative playful variants such as *mass distraction/banalisation*, *class destruction* or *chavs* and *chav-nots* are only a few of the numerous examples. The functions of new words and collocations, intertextual allusions (*annus horribilis* from *annus mirabilis*), wordplay and stylistic questions, plus the "life cycles" of neologisms are carefully analysed. The approach happily combines theory, methodology, and a wealth of empirical data.

Veale explores the dynamic creation of analogically motivated terms and the construction of new lexical categories or complex words on the basis of analogies. Research is supported by WordNet, a comprehensive electronic thesaurus of English word meanings containing over 100,000 lexical concepts, and designed to reflect psycholinguistic insights into the structure of the mental lexicon. This study, in examining ad hoc concepts and lexicalisations that arise as by-products of lexical analogy, is an exploration of the ways in which analogy can be used both implicitly and explicitly to drive lexical innovation. Quite specifically, the role of analogy in the creation of new compound terms is addressed. Veale assigns to analogy a central role in linguistic inventiveness and suggests that a thorough understanding of the processes underlying term creation may lead to the construction of computational systems that exhibit genuine linguistic creativity.

Veisbergs' topic is an excellent demonstration of the use of participant observation and Observational Linguistics (OL). The influence of English WF processes on translation and their effects on Latvian are closely examined with a wealth of examples drawn from advertising, newspapers and internet discourse. Highly original and amusing creative formations such as *kleptokratija*, *kokakolonizacija*, *Putinocets*, *pedofilgeitu* and *sexperts* serve as ASDs, and more generally, as cohesive devices, while reflecting the ingenuity of wordplay and the reliance on intertextuality. The role of parody and the ludic factor as vital sociopragmatic aspects of such creative processes are a central focus of this chapter. Distortions (*do your teeth a flavour*, *serial killer*) as well as allusions and other types of creative transformation have influenced Latvian and contributed to an "awakening" of the lexicon following the country's isolation during fifty years of severe Soviet censorship which ostracised jokes, wordplay and, more generally, a colloquial style. Direct loans from English have strengthened the creative processes of linguistic and social transformation in Latvian that can be seen as a case of globalisation and reorientation.

4. Concluding remarks

This brief overview of the wealth of original research in the present volume will serve to illustrate that novel lexical items – formed by a variety of productive or creative processes – and their occurrence and frequency in texts, depend not only on speaker intentions, context, co-text, genre, TT or register, but also on the topic, theme or domain as well as on extralinguistic reality (real or fictional) and audience-related factors. These issues are examined here in greater detail than has hitherto been the case and the research in this volume demonstrates a particularly wide range of different methodologies. This collection is thus a significant step

forward in the exploration of the relationship between lexical creativity, its many functions in discourse and related aspects of morphological theory, lexicology and beyond.

References

- Austin, J. L. 1962. *How to do things with words*. London: OUP.
- Bauer, L. 1983. *English word-formation*. Cambridge: CUP.
- Beaugrande, R. de and Dressler, W. U. 1981. *Introduction to Text Linguistics*. London/NewYork: Longman.
- Biber, D. 1988 *Variation across Speech and Writing*. Cambridge: CUP.
- Brown, G. and Yule, G. 1983. *Discourse analysis*. Cambridge: CUP.
- Clark, E. V. and Clark, H. 1979. "When nouns surface as verbs." *Language* 55: 767–811.
- Crystal, D. 1998. *Language Play*. London: Penguin.
- Downing, P. 1977. "On the creation and use of English compound nouns." *Language* 53 (4): 810–842.
- Esser, J. 1991. "Text-type as a linguistic unit." In *Anglistentag 1990 Marburg*. C. Uhlig and R. Zimmermann (eds), 142–153 Tübingen: Niemeyer.
- Halliday, M. A. K. 1978. *Language as Social Semiotic*. London: Edward Arnold.
- Hohenhaus, P. 1996. *Ad-hoc Wortbildung – Terminologie, Typologie und Theorie kreativer Wortbildung im Englischen*. Frankfurt/M., Berlin, Bern, New York, Paris, Wien: Peter Lang.
- Hohenhaus, P. 2004. "Identical Constituent Compounds – a corpus-based study." *Folia Linguistica* 38: 297–331.
- Hohenhaus, P. 2005. "Lexicalization and Institutionalization." In *Handbook of Word-Formation*. P. Štekauer and R. Lieber (eds), 335–373. Berlin, Dordrecht, Heidelberg, Norwell: Springer.
- Kastovsky, D. 2005. "Hans Marchand and the Marchandians." In *Handbook of Word-Formation*. P. Štekauer and R. Lieber (eds), 99–124. Berlin, Dordrecht, Heidelberg, Norwell: Springer.
- Lipka, L. 1994. "Wortbildung, Metapher und Metonymie – Prozesse, Resultate und ihre Beschreibung." *Münstersches Logbuch zur Linguistik (MLL)* 5: 1–15.
- Lipka, L. 1996. "Words, metaphor and cognition: A bridge between domains." In *Words. Proceedings of an international symposium*. J. Svartvik (ed.), Royal Academy, Stockholm: Kungl. Vitterhets Historie och Antikvitets Akademien. Stockholm. 49–69.
- Lipka, L. 1999a. "The metalanguage of Text Linguistics and the categorization of English texts: Notational terms and the principle of diminishing returns – with particular reference to intertextuality." In *Actas del XXI Congreso Internacional de la A.E.D.E.A.N., Sevilla 1997*. F. Toda Iglesia et al. (eds), 75–99. Sevilla: Universidad de Sevilla, Secretariado de publicaciones.
- Lipka, L. 1999b. "*Blairites, teletubbies, spice girls and wheelie bins* – neologisms, the word of the year, and the nomination-function of 'words'". In *Form, function and variation in English. Studies in honour of Klaus Hansen*. U. Carls and P. Lucko (eds), 40–47. Frankfurt/M.: Peter Lang.

- Lipka, L. 2000a. "Word-formation and (proper) names: A neglected field." In *Words: Structure, Meaning, Function. A Festschrift for Dieter Kastovsky*. C. Dalton-Puffer and N. Ritt (eds), 187–203. Berlin-New York: Mouton de Gruyter.
- Lipka, L. 2000b. "English (and general) word-formation – the state of the art in 1999." In *Anglistentag 1999 Mainz. Proceedings*. B. Reitz and S. Rieuwerts (eds), 5–20. Trier: Wissenschaftlicher Verlag.
- Lipka, L. 2002a. *English Lexicology. Lexical structure, word semantics, and word-formation*. Tübingen: Gunter Narr.
- Lipka, L. 2002b. "Non-serious text-types and German death notices – an unlikely pair." In *Text Types and Corpora. Studies in Honour of Udo Fries*. A. Fischer, G. Tottie and P. Schneider (eds), 59–66. Tübingen: Gunter Narr.
- Lipka, L. 2002c. "Names, onomasiology and semantics. Functional & semiotic lexicology." In *Language: Context and Cognition. Papers in Honour of Wolf-Dietrich Bald's 60th Birthday*. S. Scholz et al. (eds), 217–225. München: Langenscheidt-Longman.
- Lipka, L. 2003. "Observational Linguistics and semiotics." In *Language and function. To the memory of Jan Firbas*. J. Hladký (ed.), 211–222. Amsterdam/Philadelphia: Benjamins.
- Lipka, L. 2005. "Lexicalization and Institutionalization: Revisited and extended." *SKASE Journal of Theoretical Linguistics* 2 (special number), 40–42. <http://www.skase.sk>.
- Lipka, L. 2006. "Naming Units (NUs), Observational Linguistics and reference as a speech act or What's in a name". *SKASE* 3, 30–39. <http://www.skase.sk>.
- Marchand, H. 1969. *The categories and types of present-day English word-formation*. 2nd revised edition. Munich: Beck.
- Searle, J. R. 1969. *Speech acts. An essay in the philosophy of language*. London: CUP.
- Spiegl, F. 1977. *A Small Book of Grave Humour*. London: Pan Books.
- Spiegl, F. 1983. *Dead Funny. A second book of grave humour*. London/Sidney: Pan.
- Strang, B. 1968. *Modern English Structure*. 2nd edition. London: Edward Arnold.
- Tournier, J. 1985. *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Paris-Genève: Champion-Slatkine.

Lexical creativity in discourse

How to do (even more) things with nonce words (other than naming)

Peter Hohenhaus

The functional side of word-formation is often regarded as an understudied area of the field. Indeed only 'naming' is commonly presented as the sole rationale behind forming words. This chapter goes beyond this simple monolithic function (without denying its saliency) and outlines a range of functions quite distinct from naming, as observed in various registers/styles/genres of actual performance (both written texts and spoken conversation), i.e. in nonce word-formation. This chapter may be roughly divided into two halves: general functions of word-formation (e.g. textual deixis, hypostatisation) and more specific metacommunicative functions. Some are fulfilled by evidently creative techniques, others by readily available, more syntax-like means. Thus the distinction between 'creativity' and 'productivity' is also inextricably linked to the functional side of word-formation.

1. Preliminary remarks

This paper takes as its frame of reference my own theory of nonce word-formation as outlined in Hohenhaus (1996, 1998 and 2005b). Since it involves a couple of perhaps not too commonplace concepts together with a set of notational terms, which require clear nominal definitions,¹ a few preliminary remarks are necessary before we can move on to the main discussion of functional considerations.

1.1 The distinction between productivity and creativity – a cline

In contemporary linguistics a distinction between productivity and creativity has by now become fairly well-established. Both specialised terms were used to mean

1. See e.g. Welte (1974: 108, s.v. 'definition') for a typology of kinds of definitions.

roughly the same thing not too long ago, e.g. in (earlier) generative grammar,² namely something between ‘range of potential applicability of a rule of grammar’ (including word-formation) and ‘actual rule-application in performance’ (based on subconscious knowledge of those rules). These days, however, the term ‘creativity’ is adopted as a more narrowly technical notational term with the meaning of ‘a process of (conscious, even ingenious) creation that is not, or not completely, rule-governed’. It is thus to be separated from strictly rule-governed productivity in the narrower sense. (See in particular Bauer 2001 for a fuller discussion of the terminology and concepts involved here.)

Productivity, even in this narrower sense, is, however, still not an absolute value in word-formation (henceforth WF) but a fuzzy concept in which fully unrestricted productivity is the exception rather than the rule,³ and various restrictions on productivity are commonplace. Likewise, creativity: creation *ex nihilo* is a rare occurrence. “Creative coinings” tend to be at least rule-related, up to the point where a distinction from ‘governed by productive rules’ becomes problematic. In my view, it is preferable to consider the difference between creativity and productivity as a cline. And in this chapter I will demonstrate that phenomena at all points on that cline are functionally significant in relation to textuality. The cline between productivity and creativity is also reflected in different degrees of “noteworthiness” of the formations in question: the outputs of rules at the more productive end of the scale tend to pass without much notice, while the more creative “coinages” tend to be more foregrounded – i.e. they draw attention to themselves. Both productive and creative WF are needed in the full range of communicative functions that may be fulfilled by WF.

2. Cf. e.g. Radford (1988: 17). In that vein German *Kreativität* /*kreativ* ‘creativity/creative’ was used in Hohenhaus (1996) as more or less synonymous with ‘productivity’, but this has led to misinterpretations of this usage as implying something *other* than productivity, even as something opposite to the concept of ‘representative of the language system’ – a most unfortunate misunderstanding.

3. This assertion might be contested by proponents of theoretical models that only allow for full productivity (potentially, within the domain of any process), such as Di Sciullo and Williams (1987). But how many derivational patterns are there in English that are truly non-restricted? Not many. (See also Tuggy 2005: 252 for a similar observation from the point of view of Cognitive Grammar). Candidates for more or less “full” productivity may include affixation by *-ness* or adverb-forming *-ly*, which can be affixed to virtually any noun or adjective, respectively. But even if a derivational affix is considered to be fully productive within a carefully-defined morphosyntactic domain, there are still stylistic, semantic/pragmatic or, more substantially, blocking restrictions (sometimes affecting supposedly fully productive affixes, as evidenced by cases like *hardly*) – see e.g. Aronoff (1988).

1.2 General functions of word-formation: Naming and beyond

Naming, or “labelling”, has all too often been seen not only as the prime function of WF but frequently as its sole function, as part of its role of expanding the lexicon (e.g. Motsch 1977; Štekauer 2002, 2005). It cannot be denied that naming is indeed the central function of words as such – especially from the point of view of vocabulary-building (i.e. from a purely lexical point of view). Earlier discussions of the functional aspects of WF, however, have already considered functions that cannot easily be subsumed under ‘naming’: in particular, textual functions (cf. Dederding 1983; Kastovsky 1978; Lipka 1987; Wladowa 1975, etc.); yet other functions such as foregrounding or a general information condensation effect are connected to naming but clearly go beyond that function (some of these will be discussed in detail in Section 2 below).

Kastovsky (1982, 1986) suggests we distinguish between two contrasting sides of the functions of WF, namely a cline between the two poles of “labeling vs. syntactic recategorisation.” The former is largely lexical, namely serving to provide lexical items “to designate segments of extralinguistic reality”; the latter, instead, “is basically syntactic” in that this function merely “nominalises, verbalises, adjectivises, or adverbialises sentences, thus transforming them into parts of sentences” (Kastovsky 1986: 594). These two general functions can of course overlap. There also appears to be a tendency that the more “syntactic” the function of a particular formation, the less likely it is to attract attention to itself, which, as indicated above, suggests a correlation with the creativity/productivity distinction. But this is asymmetric: naming can be done by both creative and productive means, while “syntactic recategorisation” leans heavily towards productivity.

There is still more to the functional side of WF as we will see, but at this point it is sufficient to affirm that naming is only one of the functions of word-formation.

1.3. Lexeme formation vs. nonce word-formation

We must be aware that we are dealing with an area where conceptual and terminological confusion is rife (cf. Bauer 1983: 45; Hohenhaus 1996: 15ff.). Therefore some up-front terminological clarification is required. Most crucially: I distinguish between neologisms and nonce formations. The former I use as the notational term⁴ for words that are “young”, diachronically speaking, but which nevertheless have already entered the language as more or less institutionalised

4. On the need to be aware of the *notational* nature of much of our terminology see e.g. Lipka (2002: 1f., 13) – cf. also his chapter in the present volume.

vocabulary items. As such they are no longer the output of productive or creative processes but are already given, i.e. listed in the lexicon. ‘Nonce formation’ (henceforth NF) on the other hand is the notational term I use to refer to words that are in fact new – in the sense of newly, actively formed in performance, as opposed *to* being retrieved from the lexicon. To avoid common misunderstandings it seems necessary to stress that in this reading ‘nonce’ does *not* necessarily mean “odd” or stylistically marked.⁵ It is simply a cover term for all genuinely new formations, whether perfectly regular outputs of productive rules *or* the result of creative, planned word manufacture.⁶ Whether a particular form belongs to the former or to the latter category is, as already indicated, linked to the functions such forms fulfil.

2. Functions of NFs

In this section, a survey of various communicative functions⁷ will be provided, with an emphasis on functions *other* than naming, as suggested by the title. Obviously, a large proportion of NFs encountered in daily performance *will* have that function. After all, new names for new entities constantly expand the lexicon – and these new words have necessarily started out as NFs. However, if we consider those areas of nonce WF that are *not* meant to be “kept” as parts of longer-term vocabulary, i.e. *ad hoc* formations that do not end up in the lexicon (and in many

5. Nor is *nonce* in any way related to either “nonsense” or “non(-word)”, both (mis)interpretations that can be encountered (the latter systematically in psycholinguistic experimental studies, usually without the notational difference being made overt – cf. Hohenhaus 2005b: 363). Without falling into the trap of the “etymological fallacy” it may still be useful to recall the word’s history: according to the OED, *for the nonce* goes back to ME *for þan anes* ‘for the once/the one occasion only’ (i.e. *nonce* only got its initial <n> through the same “error” as *newt*, from *an ewt*). Hence the German terms *okkasionelle Bildung* ‘occasional formation’ and *Augenblicksbildung* ‘formation for the moment’ are actually closer to this etymology and more transparent.

6. Bauer (2001: 38f.) suggests *coining* as an overarching cover term to include both neologisms and NFs. In Hohenhaus (2005b: 364) this is rejected on the basis of the risk of the connotations of *coining* in the sense of “intended permanence”, which at least for a rather large subset of NFs is not the case.

7. I use the term ‘function(s)’ loosely to mean micro-structural communicative goals – or: the function of a WF is the *point* or *purpose* for which it has been formed in a particular way, where some alternative linguistic means (or none at all) might instead have been chosen, such as a clausal description instead of a naming device, or a regular morphological construction rather than a deliberately deviant one.

cases could not be listed, because they are too dependant on their context of occurrence), we can observe a variety of other functions:

2.1 Textual deixis – discourse functions of nonce formations

The application of the term deixis to word-formation was most prominently introduced by Downing (1977) in connection with her example *apple-juice seat*. She characterised this as a “deictic compound”, since in its situational context it is merely employed to pick out one chair out of a set of chairs at a table, namely the chair in front of which a glass of apple juice had been placed. Thus, the compound serves only as a totally fortuitous (spatial) deictic reference, without implying “the existence of a [name-worthy] subcategory of seats known as *apple-juice seats*” (Downing 1977: 819, 823).

Others have pointed out a “pronominalising” function of WF (Kastovsky 1978/1982; Lipka 1987; Seppänen 1978), including personal deixis – as for *patter* in the following example (given in Kastovsky 1978: 363):

- (1) ... and I patted her shoulder. Patting a shoulder can be anything from an apology to a promise, and only the *patter* can say which.

An extreme case of deictic function is fulfilled by what I have dubbed “dummy-compounds” (DCs) (see Hohenhaus 2000). These consist of a semantically depleted head constituent, usually either *thing* or *business*, plus a first constituent taking up an element from the surrounding, usually preceding context (as nearly all DCs are anaphoric). Their function is simply to refer to a stretch of context; i.e. here the naming function is replaced by that of textual reference alone. A couple of examples shall suffice to illustrate this phenomenon.

- (2) Willie: ‘[...] I want to talk about what we’re going to do on our vacation. [...] We decided that we should all stay right here.’ [... the conversation digresses ...]
Willie: ‘Let’s just go back to this *vacation thing*, right?’
Lynn: ‘Yeah. I think we should talk about it. How come we’re not going anywhere?’

(from “ALF”, Lorimar Home Video, Alien Productions Inc., 1987)

- (3) ‘I’m sorry to interrupt, Professor Baldwin,’ said Lavine. [...] ‘It’s just - Well, I’ve been looking into solstitial winds at Ross, and I think - I think I may have stumbled onto something really big.’ [...] ‘Tell me about this *Ross thing*,’ he [Baldwin] said, [...].

(from *The New Yorker*, 8/5/1989, pages 44 and 47)

Dummy-compounding is so productive, so “syntax-like” that it does not normally attract attention in the way that creativity does.⁸ Moreover, DCs are typical of informal conversation. Thus they are linked to a specific ‘text type’ in the widest sense – including written dialogues e.g. literary or dramatic dialogues where the use of DCs can help create the “natural flavour” of real conversation.

However, their use in a specific register/text type does not exhaust the relevance of dummy-compounds for the issue of textuality. Consider the example *Margaret business* in the following context, from the opening of a novel:

- (4) He’d been drawn into the *Margaret business* by a combination of virtues he hadn’t known he possessed: politeness, friendly interest, [...]. It had seemed only natural for a female lecturer to ask a junior, though older, male colleague up to her place for coffee, and no more than civil to accept. Then suddenly he’d become the man who was ‘going round’ with Margaret [...].

(from Kingsley Amis, *Lucky Jim*, London: Penguin, 1953, p. 10)

One could argue that here the reference is partly anaphoric (the proper name had been mentioned immediately before) and partly cataphoric, since the contextual information necessary to fully decipher the reference is only given later in the chapter/novel, at least from the reader’s point of view. However, we have to distinguish this from the perspective of the characters *in* the novel (and from the writer’s perspective, one may add).⁹ For the characters involved, the reference of the dummy-compound is as easily interpreted as any other anaphoric reference. From their perspective, what is being referred to is part of their contextual knowledge – which, however, is outside the contextual frame shared with the reader. I would argue that it is precisely this use of a DC as informal deixis that is crucial in “dragging” the reader “into” the story. Given that the reader cannot access the reference at the point in which the DC occurs, he/she has to rely on vague assumptions, hoping to be able to verify them by later contextual clues, in hindsight as it were. We can thus see this as a particular literary technique which exploits

8. The fact that compounds can be seen as being in between WF proper and syntax has frequently been pointed out; Di Sciullo and Williams (1987: 18) place compound in between word and phrase in their hierarchical scale of units whose poles are morpheme and sentence respectively. From a conceptual perspective, DC with *thing* or *business* could be interpreted as a particularly productive incarnation of what has been subsumed under the much wider heading of “shell nouns” by Schmid (2000). This concept also includes other nouns of less depleted but still rather general meanings, such as *problem*, *idea*, *guess*, etc.

9. Cf. e.g. Saukkonen (2003: 402f.) for a Systemic-Functional Linguistics framework explicating the distinction between the two frames ‘text world’ and ‘situation frame’ (which together make up the ‘discourse frame’). For our purposes, though, it is not necessary to go into similar theoretical detail

a specific pattern of WF, contributing to textuality. (Moreover, the choice of an informal deictic expression can also be said to help create an illusion of familiarity between reader and text world/author.)

A related but different phenomenon are what Heringer (1984) has called episodic compounds (“episodische Komposita”). Their status is, as it were, in between that of general naming units and strictly context-dependent deictic compounds. They may be established within small-group speech communities in that they relate to a past context, or episode, that is familiar only to the in-group, and hence are not, or not fully, interpretable for “outsiders” and the speech community at large.¹⁰ Heringer’s (1984:9) example is German *Mäusebibel* ‘mice bible’ meaning (for the family members who use this as an episodic compound) ‘the bible we once found in a barn and which had teeth marks from mice who must have nibbled at it’.

Obviously, episodic compounds are not limited to real-world small-group speech communities but also occur in fictional texts where they have the additional function of involving the reader in the story by making him/her “a member” of the small-group community, so that he/she is able to correctly interpret the compound on the basis of preceding episodes in the fictional context. A prime example of this (cf. Hohenhaus forthcoming) is *Dead Duck Day* in Nick Hornby’s novel *About a Boy* (London: Penguin, 1998). It is used several times by the main character, the boy Marcus, to refer to an eventful day – which did include the accidental killing of a duck. But the reference is to the wider situation, the whole episode¹¹ of that day, which also includes his mother’s attempted suicide (which had a much stronger emotional impact than the actual demise of the duck). From the perspective of textuality we can note that the reader shares with the characters of the novel the necessary episodic knowledge and can thus interpret the reference correctly, but an outsider could not do so – nor could someone who hasn’t read the novel (obviously enough), *nor* other characters *in* the novel who were not involved in the *Dead Duck Day* episode(s) or who had not been told of it (and who do not share the reader’s omniscience either).

10. See Hohenhaus (2005b:360–362 and forthcoming; also discussion in Section 4) on some crucially important distinctions regarding different kinds of speech communities.

11. Munat (2003:41) makes a similar observation about her example *body snatch* which she identifies as a NP that serves to “condense an entire episode in the novel”.

2.2 Hypostatisation

This term has been introduced into the terminology of WF in particular by Lipka (1975: 200; 1977: 161) to refer to the power of words in concept-formation, namely that the mere existence of a name implies to a speaker that a corresponding entity must exist in extralinguistic reality (see also Schmid 2005: 78ff. for a fuller discussion from a cognitive perspective). Normally hypostatisation is a side-effect of naming. However, it can also be exploited for purposes of textuality. We can speak of functionalised hypostatisation if a WF names something that does *not* actually exist but is part of the illusion of a fictional context, thus further increasing the overall fictional illusion – through “fiktive Wörter” or fictitious words (Ortner 1985: 255). This has been noted for Science Fiction (SF) in particular (e.g. as early as in Kastovsky 1978: 360), but is of course not restricted to this genre – cf. also the contributions by Porto Requejo and Munat in this volume.

If a hypostatisation-prone genre such as SF is combined with humour or satire, this reinforces the role of hypostatisation. Humour, as well as fiction, are both “non-*bona-fide*-modes of communication” – in the terminology of Raskin (1985). It is thus not surprising in pragmatic terms that a combination of the two should yield a feedback effect on the linguistic means employed. Thus, in Douglas Adams’ popular SF-spoof trilogy (of five parts!) *The Hitch-Hiker’s Guide to the Galaxy* (London: MacMillan, Pan Books, 1979ff.) we not only find all sorts of humorous formations and an even higher than normal proportion of exaggerated “fictitious” words that are often far from transparent (as is typical of such words employed in SF – cf. Kastovsky 1978: 360), while still purporting to have a naming function within their contexts. But we also encounter degrees of hypostatisation where it seems to replace any genuine naming function. For instance, in Douglas Adams’ SF-spoof there is a markedly heightened frequency of certain genre-typical prefixes such as *hyper-*, *mega-* or *ultra-* that do not really fulfil their semantic functions but rather appear to be used solely for the purpose of parody of the genre: e.g. WFs such as *hypermathematics* or *ultramahogany*. These are merely mentioned in passing with no indication whatsoever as to what they might actually mean. The point, of course, is not that they *mean* anything but simply that they *sound* typical of SF. It clearly forms part of the literary technique of SF-*parody* that the hypostatisation function of WF is exploited in an exaggerated manner. Not only are such formations employed more frequently, but also their semantic “pointlessness” and lack of motivation is more conspicuous.

Again it may be useful to distinguish between the a) internal perspectives and b) external perspectives from the reader’s point of view. For a) hypostatisation (as an effect of fictitious naming) is dominant, for b) it is the humorous effect resulting from exaggerated reliance on hypostatisation without any actual basis in real-

ity (in addition to other morphological/stylistic effects that may be involved, such as deliberate formal and/or semantic deviation).

An extreme case of hypostatisation for nothing but hypostatisation's sake is the introduction of fictitious words from "alien" languages. Douglas Adams' SF-parody presents a set of such invented words, for example: what an alien species of swamp-dwelling mattresses (sic!) do, such as *flurble*, *flollop*, *glurry* – where perhaps a remote suggestion of sound symbolism is the only trace of semantic meaning whatsoever.¹² Another famous example is the artificial language of (the extra-terrestrial race of) the Klingons in the *Star Trek* TV series and movies. Originally this must have had a purely hypostatising function, i.e. more or less arbitrarily combined strings of sounds to function as an unintelligible alien language. But this has been taken to exceptionally elaborate heights in that a full-blown artificial (but pseudo-natural) Klingon language, complete with grammar and dictionary, has been constructed (Okrand 1992).¹³

2.3 Attention-seeking/foregrounding

It has frequently been pointed out, in particular by Lipka (1987, 2000), that one important function of WFs is to serve as so-called "attention-seeking devices" (ASDs), which can be understood as a subfunction of more general 'foregrounding'. ASDs are quite frequently found in text types and domains where this is the primary overall function, as in titles of newspaper/magazine articles or in advertising. To achieve this, all sorts of wordplay and deliberate breaking/bending of rules are employed (see Crystal 1998:93ff.). One of the most drastic examples that I have encountered is the use of the single German word *unkaputtbar* (literally 'un-broken-able') to advertise plastic bottles for soft drinks, together with a picture of one such bottle (containing a well-known, even iconic, American soft drink). There is no other text. The formation is, of course, morphologically deviant, as German *-bar*-suffixation, like English *-able*, requires a transitive verb base, but *kaputt* is an adjective (cf. Hohenhaus 2005b:369).

Often wordplay serving as an ASD merely involves graphemic foregrounding (cf. Lipka 2000:8) or bears phonetic similarities to (but significant deviation

12. Such sound symbolic functions have been assumed e.g. for the initial clusters <fl-> and <gl->, but also for final elements such as <-op> and <-urry>; see Marchand (1969:411f., 423, and 425) – and Neuvel and Singh (2001:314) for a more recent attempt at justifying the assumption of such so-called "phonesthemes".

13. A fairly comprehensive sketch of this artificial language is available online at: <http://www.explore-ytelevision.com/star_trek/K/Klingon_language.html> or at <www.kli.org>

from) established expressions which are in some way connected to the new situation. Thus the following example employs wordplay on the theme of *star dust* (cf. Hohenhaus 1996: 344f. and Renouf, this volume):

(5) [title:] *Czar Dust*

[text:] Glasnost took a grave turn last week, when Soviet writer Geli Ryabov claimed he had found the bodies of Czar Nicholas II and his family. [...]

(from *Time* magazine, 24/4/1989, p. 51)

Frequently, ASDs extend beyond single words/WFs, and may even include elements of riddles as in (6), in which the “apparent” compound *ISA Day* together with its indefinite article reads like “have a nice(r) day”:

(6) Have an ISA Day

(advertisement for *Maxim ISA*, spotted in Nottingham in 2004)

All such ASDs are of course closer to the creativity end of the productivity/creativity scale, and since the phenomenon has received a good deal of attention elsewhere, we need not elaborate further on it here. However, it should be remembered that such foregrounding through unusual, strikingly creative means is not limited to titles and advertising but also plays an important role within texts – at least in certain kinds of texts – and is often linked with the overarching function of humour (cf. Hohenhaus 1996: 296–317). One further example:

- (7) Will [tried to come up with] something that a twelve-year-old boy might want to do on a Saturday afternoon [...]. Maybe the coolest thing he could do for Marcus was let him run wild on Saturday – give him some money, take him to Soho and leave him there. He had to admit, though, that while this might score points on the *coolometer*, it didn’t do quite so well on the responsible in loco parentis scale (from Nick Hornby’s *About a Boy* (1998), p. 74f.)

This involves an analogy-based WF on the model of *speedometer* which, due to the rather marginal productivity of that pattern, together with the strangeness of the semantic concept (“coolness” can hardly be measured exactly), is highly foregrounded.

2.4 More specific metacommunicative functions

The functions discussed so far can all be characterised as general communicative functions. In the remainder of this chapter, however, we shall focus on a number of functions that are metacommunicative in nature in that they all constitute

comments on linguistic material and/or speech acts in their given contextual environments.

2.4.1 *Identical constituent compounding/contrastive reduplication*

One rather productive type of metalinguistic nonce formation has been dubbed “Identical Constituent Compounding” (ICC) (Hohenhaus 2004; also 1998:255–257), or “Contrastive Reduplication” (CR) (in Ghomeshi et al., 2004) and can be exemplified by forms such as *job-job*, *friend-friend*, *home-home*, etc. Such “doublings” often occur in a contextual frame in which the reduplication contrasts with either the single constituent or with another construction/compound containing this constituent. Often one of these contrasting pairs is negated or appears in an interrogative construction; as in the following gloss:

‘I don’t mean (Y)X but XX’ /
‘Do you mean XX or (Y)X?’

The ICC/CR serves to clarify the word in question by narrowing down a wider, vague or ambiguous meaning to a “prototype” or “core” reading. So *job-job* is to be understood as ‘a proper, 9-to-5 job’, *friend-friend* as ‘a true friend’ (or as ‘only a friend – no more than that, i.e. not a lover’), and *home-home* is intended as ‘permanent or original, i.e. parental home’.¹⁴

Following are a few more recent attestations in their original contexts.¹⁵

- (8) ... the sunshine of madrid, the coastline and mountains of the canaries,
the warmth of a welcoming people rather than the hung-up, pretentiousness
of surrey! that’s the real *spain spain*, not your tourist spain ...
(from an email communication, hence the lack of capitalisation)

14. Although the studies of this phenomenon by Ghomeshi et al. and myself were carried out independently, each unaware of the other’s research until the results were published almost contemporaneously (in 2004), the empirical results are strikingly similar. The difference lies in the approach: while mine was primarily an empirical corpus study, Ghomeshi et al. is more theoretical/technical, focussing on attempts at describing the phenomenon in relation to two generative models, Minimalism and, in particular, Parallel Architecture (based on Jackendoff 1997, 2002). The theoretical as well as empirical details, however, are beyond the scope of this chapter.

15. For an even fuller range of examples of the phenomenon see the collection of attestations that Ghomeshi et al. (2004) is based on at <www.umanitoba.ca/linguistics/russell/redup-corpus.html> or see the appendix to Hohenhaus (2004: 326–331), which lists the full returns of the corpus searches together with context excerpts and details of sources. I do not call this list “a corpus”, however, because it was drawn from the *results* files of various separate searches of several corpora. These corpora comprised both generally available corpora such as the BNC and specially compiled corpora of my own (in particular, one compiled from online TV-scripts). The examples presented here are new.

- (9) 'She's mad [...] Not *mad* *mad*, but, you know. Out of control.
(from Nick Hornby's *About a Boy*, p. 257)
- (10) My hair was falling out – I mean, not *falling-out falling-out*, but falling out of
the clip [i.e. not in the sense of hair loss/balding]
(from a personal conversation)

As the examples show, ICC/CR goes far beyond the conventions of nominal compounding and may even involve proper names, adjectives or verbs. Ghomeshi et al. (2004: 321) also recorded some occurrences of phrasal CRs such as *considered-it-considered-it*.

But leaving aside all formal considerations and concentrating on our functional perspective, it is clear that we are dealing with a metacommunicative function because it *picks up* an element of the on-going discourse and makes a comment on it, in the attempt to disambiguate, clarify or pre-empt a potential misunderstanding in the given context. What is more, the type of communicative parameters that give rise to the phenomenon appears to be fairly sharply delimited: almost all of the examples are attested in dialogues (whether authentic or scripted), generally in a fairly informal conversational register, as was the case for DC.

Also like DC, ICC/CR lies towards the productivity end of the creativity/productivity scale put forward in Section 1 – but probably not quite as far to the extreme end as DC, because forming ICCs may involve more speaker awareness due to their metalinguistic nature. The prerequisites (and the potential need) for an ICC/CR must first be noticed, and only then reacted to – which suggests a greater consciousness on the part of the speaker than the more “automatic” performance of productive deictic processes such as DC. However, these compounds themselves are not perceived as creative WF, but are rather syntax-like.¹⁶

ICC/CR is a specialised case of the more general metacommunicative function of clarification/disambiguation that can also be fulfilled by non-identical constituent compounds (which in fact often occur in the contrasting co-texts of ICCs/CRs – see above). Consider, for example, the case *angry-mad* (discussed in Hohenhaus, 1996: 346f.) serving to disambiguate *mad*, which in the immediately preceding context was misinterpreted as ‘insane’ rather than the intended meaning ‘angry’. This too is a case of ‘disambiguation’ but in a stricter sense, serving to select one of two readings already available in the lexicon (as opposed to the

16. In fact, Ghomeshi et al. consider CR to be primarily a syntactic phenomenon (see Ghomeshi et al. 2004: 322, footnote 8, where the compounding hypothesis is rejected, and Hohenhaus 2004: 320ff., footnote 4, who argues in its defence), although in their Parallel Architecture framework the difference is less crucial than in theoretical models that involve a strict separation of word-formation and syntax in completely distinct modules.

“narrowing down” of an unspecified range of possible contextual interpretations in the case of ICCs/CRs).

A similar example is the following

- (11) Marcus wasn't daft. Well, OK, he was daft sometimes, like with the singing, but he wasn't *stupid-daft*, just brush-daft.

(from Nick Hornby's *About a Boy*, p. 94)

Arguably, this kind of disambiguation is more effective, more precise, because it uses a synonym (or gloss) for one of the meanings of a polysemous word; however, ICC/CRs despite being inherently less precise, are more flexible as they can resolve a wider range of potentially vague meanings, not just lexical ambiguity in the narrower sense.

2.4.2 *Direct analogical correction*

Without going into the vexed question of the general role of analogy in WF (cf. e.g. Bauer 2001:75ff.), I would like to illustrate one very specific metacommunicative form of direct analogy. By direct I mean that the model for the analogical formation is overtly present in the preceding context and is commented on and corrected by means of the new formation. Consider the following examples from Douglas Adams' SF-parody – the episode is drawn from a part of the novel in which a “useless” third of a planet's population, including advertising executives, management consultants, hairdressers and telephone sanitizers (sic!) are sent off into space to crash-land on a distant prehistoric planet, where they try to resume their former routines despite the complete absence of their previous familiar contexts and creature comforts – the following extract is from a “formal” meeting held by the survivors of the crash:

- (12) ... Ford looked round the crowd. ‘He's mad you know,’ he said, ‘this is a prehistoric planet.’ ‘Address the chair!’ snapped the management consultant. ‘There isn't a chair,’ explained Ford, ‘there's only a rock.’ [...] ‘Well, call it a chair,’ [...] ‘Why not call it a rock?’ asked Ford. ‘You obviously have no conception,’ said the management consultant, [...], ‘of modern business methods.’ [...] A girl with a strident voice leapt to her feet and used it. ‘Shut up, you two,’ she said, ‘I want to table a motion.’ ‘You mean *boulder* a motion,’ tittered a hairdresser. ‘Order, order!’ yapped the management consultant. [...] ‘This is futile,’ he [Ford] exclaimed, [...] ‘If you would care,’ said the girl with the strident voice, ‘to examine the agenda sheet ...’ ‘*Agenda rock*,’ trilled the hairdresser happily. ...

(from Douglas Adams, *The Restaurant at the End of the Universe*, 1980, p. 170ff.)

The role of the NFs (*to*) *boulder* (*a motion*) and *agenda rock* is to act as adaptations of the conventionalised expressions *to table* (*a motion*) and *agenda sheet* in order to make them “fit” the new situation.

The conditions for such analogical adaptations are of course highly context-dependent. This is particularly evident in the following example (the summary of the situation is my own):

- (13) The hotel staff try to conceal from a health inspector’s attention the fact that one of them keeps a rat as a pet. To make matters worse, the pet rat has accidentally been let loose in the hotel. One of the hotel guests sees the rat and attempts to shoot it. The guest is admonished by the hotel owner and told not to reveal the truth, instead claiming that he was shooting at starlings in the garden. Later, in the dining room, the same guest discovers that the health inspector is having lunch there too and shouts ‘So you’re the *rat inspector*!’ To cover up his *faux pas* the guest quickly follows this up with: ‘sorry, *starling inspector*.’ The health inspector merely repeats blankly: ‘*starling inspector*?’
(from the British BBC sitcom “Fawlty Towers”, series 2, episode 6)

This is a succession of increasingly context-dependent analogies: first, the transparent and established compound *health inspector* is adapted to *rat inspector* (already rather context-bound), and this direct analogy is then further “corrected” to yet another direct analogy: *starling inspector*. This final step does not actually make much sense semantically, as it is the result of lexical “evasive action” rather than a genuine act of naming. The fact that (the reason for) this second analogy is almost wholly uninterpretable by anyone who does not share the knowledge of the sequence of contextual information is in fact clearly marked here by the health inspector’s expression of lack of comprehension. (For more on this phenomenon of marking nonce-ness see Hohenhaus 1996: 323–328).¹⁷

2.4.3 “Delocutive conversion”

This term is borrowed from Brekle (1976),¹⁸ who discusses certain verbs such as *to madam*, *to sir*, *to brother* which he calls “delokutiv” (here translated liter-

17. One of the examples given in Munat (2003: 41) of “one-off nonce formations whose referents are identifiable in the immediate textual world” – a broad category that this chapter tries to elaborate on – could in my view be classified as a case of the type of analogy under discussion here, namely *silver policy* (which an insurance salesman’s son is said to be born with in the novel) modelled on *silver spoon*. However, in the absence of the full concrete co-text, it is not possible to see whether it is also an overtly direct analogy.

18. Originally the term has been coined by the French linguist Emile Benveniste – cf. also De Brabanter (2005) for a more comprehensive and recent overview of a range of related phenomena.

ally as ‘delocutive’) as they are speech act-bound formations that “cite” a form of address and transform it into a (zero-derived or converted) verb. There are also delocutive verbs in English that are lexicalised, such as *to okay*, *to welcome*. These are not dependent on an actual verbatim citation of a particular locution – it is possible to *okay* something or *welcome* someone without actually saying “okay” or “welcome” (cf. Welte and Rosemann 1990: 155). On the other hand, the same delocutive mechanism, of citing some linguistic material by converting that material into a verb, is sometimes exploited in nonce word-formation. This is the case, in particular, in communicative situations of conflict, where a speaker rejects the form of something said to him/her (not necessarily a form of address). It typically appears in the frame ‘*don’t (you) __ me!*’ (Hohenhaus 1996: 328ff.), as in the following examples:

- (14) a. People are straining to hear ...
 A: do you mind? ... I can’t hear a word he’s saying.
 B: Don’t you *do-you-mind* me!
 (from the Monty Python movie “The Life of Brian”)
- b. Ruby: “Well, what?”
 Stacey: “Don’t *well-what* me!”
 (from an episode of the British BBC soap “Eastenders”, 25/08/2005)

Interestingly, not only verbs and whole phrases can undergo delocutive conversion, but even extra-phonological (vowel-less) interjections, as in: “don’t ‘*ssh*’ me!” (from the same source as 14 a.). There is also the verbal form *to shush someone* – i.e. a lexicalised verb having the same meaning. Similarly, *tut-tut* is the phonologically lexicalised form of the interjection [ts ts (ts)]. As a nonce mechanism of metacommunicative comment, however, delocutive conversion does not appear to require a similar adaptation to English phonology.

2.4.4 Metalinguistic extraction

A phenomenon that is somewhat similar to the one discussed in the previous section, but less specifically bound to a particular speech act, might be described as metalinguistic extraction. Consider the function of *oid-y* in the following context:

- (15) It’s an *oid-y* world out there. Tabloids run factoids about humanoids on steroids. In a world gone synthetic, why should movies offer something as organic as a hero? Welcome, then, to the age of the heroid.
 (from *Time International*, 24/07/89, page 59)

Again an element from the surrounding co-text is quoted and incorporated into a new formation. However, here it is not really an act of quoting and commenting

on a single speech act, but rather the extraction of a linguistic element, the suffix *-oid*, and a comment on its use in more than one speech act. Note that semantically the resulting formation is perfectly legitimate, based on the underlying productive semantic pattern of *y*-suffixation ‘full of *x*-es’, as in *stony*, *wordy* etc., while on the surface it should be a morphologically impossible word (there are no formations with a suffix as base for another suffix).¹⁹ It is only possible because, through metalinguistic extraction, the base has itself been nominalised and thus brought in line with the morphological subcategorisation for *-y*-suffixation.²⁰

Another, more recent, attestation follows, which is, however, less problematic than *oid-y* from a morphological perspective, as we are now dealing with a proper word that has been extracted from the previous context (but it nonetheless constitutes a similar sort of metacommunicative comment):

- (16) ‘I think we will always be friends,’ Marcus continued. ‘But I don’t know. I think I ought to look for someone less – ’ ‘Less rude and mad? Less violent? Less bloody stupid? There are a number of *lesses* I can think of.’

(from Nick Hornby’s *About a Boy*, p. 265)

This is of course similar to established uses of *ifs* and *buts* and the like – however, as a NF *lesses* is dependent on the co-text from which the extracted element has been taken up verbatim, whereas there is no such direct co-text dependency for the established uses of *ifs* and *buts*. These do not necessarily imply the actual (literal) occurrence of these two words in the text, but rather refer to the (abstract) types of speech acts associated with them. Cf. the same distinction with regard to delocutive conversion in the previous section, e.g. *to welcome*, a speech act that

19. An objection could be raised here on the grounds of forms such as *isms* – cf. Štekauer (2000: 110f.) – which also involve a suffix (*-ism*) as a base without appearing particularly deviant. However, such well-established examples of bound forms used as free form nouns are different in that they do not involve derivational morphology. The apparent counter-examples *anti-ists* and *anti-isms* are also different in that *anti* has already undergone independent conversion into an adjective (as evidenced by usages such as “she’s very anti rock music”) and as a listeme is thus lexically available as a base for WF, similar to other adjectival unbound prefixes like *mega-* or *hyper-*.

20. This description appears to be in conflict with my earlier accounts of this particular example (Hohenhaus 1996, 1998, or 2005b), where I classified this attestation as an example of a morphologically deviant formation. In actual fact, of course, independent of its context it is deviant. It is only by making the tacit assumption, based on the context, that *-oid* must first have undergone nominalisation that the temporary suspension of the morphological deviation is made possible. But because this suspension only works temporarily, in this metacommunicative context, the formation as such remains non-lexicalisable. This shows how questions of deviation can become entangled with context-dependency.

may or may not involve the word *welcome* verbatim, vs. e.g. “don’t *well what* me”, which would not be possible without the actual presence of the converted phrase in the immediately preceding context.²¹

2.4.5 *Ad hoc stereotypification*

This section requires a brief theoretical prelude. Aronoff (1976), in his seminal book on WF in generative grammar, proposed the widely discussed Word-Based Hypothesis. This stipulates that both input and output of word-formation rules have to be words – e.g. in a modular model of language theory the morphology component does not have access to the syntax. This part of the Word-Based Hypothesis also became known as the “No Phrase Constraint” (NPC): WFs cannot contain syntactic constructions. However, some existing phrasal compounds (such as *under-the-counter sales*) appear to pose counter-examples.²² Thus it is clear that the NPC cannot be a global restriction applying to all WF. But neither should it be given up altogether.

Two lines of argument are relevant here: firstly, if the phrasal element in question is itself a listeme, the NPC does not hold. This allows for such well-known idiomatic-phrase compounds as *big bang theory* or *under-the-counter sales* and also lexicalised phrase-derivatives such as *up-to-date-ness*.²³ Such items are only apparent counter-examples that do not affect the real scope of the NPC. This is because the phrases in question are already available in the lexicon and are thus not genuinely syntactic any longer; they are not actually composed online but selected as ready-made wholes (like any other listed lexical items that serve as input into WF processes). Furthermore, for a phrasal unit to become a listeme it does not necessarily have to be a full-blown idiom with idiosyncracies such as non-compositional meaning. As has been demonstrated by Jackendoff (1997) for example, many listed phrases are merely stereotypes, i.e. compositional but formulaic expressions of high frequency and thus high recognisability (which is

21. Cf. also De Brabanter 2005, esp. Section 4.1.2, who uses the term “de-autonym” for such established (lexicalised) cases in order to distinguish them from genuinely metalinguistic “autonyms”.

22. Another one is, rather ironically, the term *no-phrase constraint* itself.

23. The phrases making up the first compound constituent are indeed idiomatic, although not completely demotivated; the NP *big bang* in *Big Bang Theory* denotes not just any big bang (in fact it may not have been a bang at all; it is just a metaphor). Likewise *under-the-counter* is no longer a neutral locative adverbial PP, but has a semantic specialisation roughly equivalent to a combination of ‘illegal’ and ‘stealthy’; and *up-to-date* has acquired a meaning akin to ‘modern’, ‘current’.

the argument for assuming they are listemes too) – see also the contribution by Kuiper in this volume.

Secondly, as I have argued extensively elsewhere (esp. in Hohenhaus 1996: 218ff.), the NPC is best conceived of as a restriction on the lexicon only, not one on word-formation in general. In other words: WFs to be lexicalised *do* have to obey the NPC, but for non-lexicalisable nonce WFs it is suspended. Therefore, violations of the NPC are tolerated in performance (see examples below). However, the products of such nonce processes, when they involve genuine phrasal constructions, are barred from the lexicon (unless of course the phrase in question acquires listeme status itself – through, at least, stereotypicality – in which case, as pointed out above, the NPC does not apply).²⁴

From a functional point of view it is interesting to observe, however, that the influence of the NPC does extend into active (i.e. nonce) WF performance too, namely in the form of what I call ad hoc stereotypification where the use of a phrasal WF can help to create the impression that the phrase involved has the status of (at least) a stereotype, even if it does not actually have that status outside of its ‘text world’. Let me illustrate this by the following example of phrase-to-noun conversion (or zero-derivation):

- (17) [...] on the whole, academic subject groups are self-defining, exclusive entities. Each has its own jargon, pecking order, newsletter, professional association. The members probably meet only once a year – at a conference. Then, what a lot of hellos, howareyous, and *whatareyouworkingons*, over the drinks, over the meals, between lectures.

(from David Lodge’s novel *Small World*, London: Penguin, 1985, p. 233)

Here, the final nonce phrase conversion in a string of conversions, involving a) a single word, and b) an established phrasal listeme, takes on a subtle metacommunicative function in that it c) assigns to the syntactic construction *what are you working on* the same status of a common stereotype as the other two in the string. In fact, in the novel this utterance is nearly as stereotypical as the real idioms since it does indeed appear repeatedly on various comparable occasions. Normally, however, this simple question would probably not be considered to have phrasal listeme status in the real world. It is only through its incorporation into WF that the reader is almost directly prompted to interpret that construction

24. There is of course a good deal more to be said about phrasal word-formation. Not just any syntactic construction can be incorporated into word structure. But for the purposes of this chapter, this rough outline must suffice. For a more detailed discussion see e.g. Scalise (1986: 154–156), Dressler (1986), and the survey of different approaches to the issue in Hohenhaus (1996: 218–228) (see also Veisbergs this volume).

as a stereotype in this particular text world. Here are some further examples to illustrate this function:

- (18) 'Is this the right time to tell you that I'm staying here tonight? [...] I booked a room. Just in case.' I drain my drink and walk out. [...] 'What was that all about?' he asks me the next time I see him [...]. 'Why did you walk out on me at the hotel?' And I make some weak *what-kind-of-girl-do-you-think-I-am* joke [...]
(from Nick Hornby's *How to be good* (2001), page 28)
- (19) David shoots me a look. Not the old-style, *I-hate-you-and-I-wish-you-were-dead* look I would have got, once upon a time; this is the new-style *I'm-sooooo-disappointed* look, and for a moment I am nostalgic for the days when hatred was our common currency.
(from Nick Hornby's *How to be good* (2001), page 96)

The two phrasal compounds in (19) are, in a way, even more metacommunicative. Not only do they stereotypify the phrases involved, they almost quote them metalinguistically, or rather they interpret non-verbal clues in the form of a verbal *paraphrase* according to the pattern 'a look that means "..."' (cf. De Brabanter 2005, Section 4.3).

2.4.6 *Emoting through conversion and reduplication in MUDs*

MUDs (short for "multi-user dungeon/dimension") present a highly distinctive variety of computer-mediated communication (CMC), namely within so-called 'virtual realities'. Crystal's (2001: 180ff.) outline of the genre contains aspects that are relevant here too. MUDs are (at least originally) related to more traditional role-playing games, but use computer networks in ways akin to chatrooms, with some special linguistic features. One important aspect is the distinction between "saying" and "emoting". A player's assumed character can not only "speak" (i.e. the player types in "say: Y", the other players see on their screens: "X says Y") but also perform certain actions, including, in particular, gestures such as nodding, smiling, greeting, etc., and also (due to contextual settings typical of MUDs) kicking, poking, killing. One direct way of putting these into action is using the emote command (i.e. player types "emote Y", other players see "X Ys"). There seems to exist a certain established stock of emotes, but frequently this function also gives rise to "idiosyncratic word-formation" (Crystal 2001: 182) which is of a metalinguistic nature, such as "Anthony *ohboys*" or "Pete *actuallies*". Here the command "saying" is replaced by "emoting", hence what would normally be a verbal gesture (use of an interjection or sentence adverb) is given the status of an action like nodding or grinning.

In some MUDs, emotes are frequently given the form of a reduplication, such as *nodnod* (spelled as a single word) with basically the same function. Further examples (from one of the corpora used in Hohenhaus 2004) of such reduplications for emotes feature in the following lines: “Molly *pokepoke* fur inna eye” and “Molly *kickick* fur inna groin”.²⁵

These phenomena (like various other characteristics discussed in Crystal 2001) are of course genre-markers, an instrumental feature of creating MUD-textuality. On the other hand, being so intimately connected to one such specialised subgenre of CMC (and one whose future may well hang in the balance – cf. Crystal 2001: 194), these are of virtually no relevance to WF outside their strictly delimited domain. (For more on CMC see also López Rúa’s contribution in this volume.)

2.5 Special cases

Occasionally, certain attestations resist any classification and appear to be truly one-off cases (cf. Hohenhaus 1996: 128–131); this is also true from a functional perspective. Consider the curious compound *suitcase sauce* in the following example

- (20) Mr Hamilton: ‘Could you make me a Waldorf salad. [...] it’s celery, apples, walnuts, grapes.’ [...]
In the kitchen. Basil is rummaging frantically in a large cardboard box. [...]
Basil: There’s no celery, there’s no grapes ... walnuts! That’s a laugh, easier to find a packet of sliced hippopotamus in *suitcase sauce* than a walnut in this bloody kitchen.’
(from “Fawltly Towers”, series 2, episode 3)

Here the speaker wants to express with exaggerated emphasis that something is not available (walnuts) – and he does so by putting the word for it in a comparative syntactic structure with a construction that denotes something even less available, in fact one which doesn’t even have any potential real-world referents because it is semantically impossible.²⁶ This is thus a very odd case of naming.

25. Both were drawn from a log at: <www.apocalypse.org/pub/u/lpb/muddex/reality.txt>

26. Formally, there is nothing exceptional here, in fact one could argue that *suitcase sauce* is not a compound at all but a ‘complex nominal’ – cf. Giegerich’s (2004) distinction between proper compounds and noun + noun phrases. For our purposes, however, it is of secondary importance whether this example is or is not a compound in the strict sense as it has only been cited for functional considerations, relating to naming and concept forming. Since that function can also be served by noun phrases, the relevance of the example to our context is not really affected by the question of compound-hood here.

Obviously, the naming of such a (near-)impossible concept is intended to be maximally ludicrous – which is certainly achieved, with an exceptional degree of creativity. I am not saying that it is the non-existence or impossibility of the concept that makes this example so unique. After all, there are scores of names for non-existent and highly unlikely concepts in fictional contexts, especially SF – and some of these, such as *warp speed*, have even become institutionalised (outside their original contexts). The difference is, however, that (see Section 2.2) in such cases functionalised hypostatisation comes to the fore to imply the existence of the named concept (despite technical or other limitations in actual reality). With *suitcase sauce*, instead, we seem to be looking at the opposite of that hypostatisation effect – one could perhaps even call it “negative hypostatisation” (cf. Hohenhaus 1996: 356) – the WF emphasises the *non*-existence of what is being named.

Another highly individual case, which is interesting from both a formal and a functional point of view, is that of *sizenessness*, as reported in Crystal (1998: 31) as an instance of a method of language play that has become a habit between two teenagers. In their shared code they apparently add the *-ness* suffix routinely to all manner of bases (cf. also *upstairsness* and *bookness*) “in order to emphasise a particular abstract notion”, or rather its very abstractness, which apparently worked (self)recursively in echo replies giving further emphasis – the conversation was reported as follows:

(21) A: ‘Look at the sizeness of it!’ B: ‘Cor, the sizenessness!’

Formally, this is obviously a massive deviation from English morphological rules; there are no recursive suffixes in English (only a small handful of prefixes such as *re-* as in *re-read*, *re-re-read*, *re-re-re-read* etc. – cf. Bauer 1983: 67f.). It also violates the condition that *-ness* not only requires an adjectival base but must also be a terminal suffix (disallowing any further derivative suffix). Functionally, the phenomenon is interesting in that it seems to simply detach the component ‘abstract’ from the morphosyntactic-semantic role of *-ness*-nominalisations, making it applicable to *any* notion. (And clearly there is no real conceptual *need* for the formation here – it is neither a case of naming nor syntactic recategorisation).

While *suitcase sauce* was an example of extreme individual context-dependency, the *-ness-* examples seem to work just as well without any supporting context once this particular pattern has been established. Here it is rather the extreme degree of idiosyncrasy in rule-changing creativity that is so striking. Its productive application is probably only possible in such a maximally narrow “micro-dialect” (a setting which allows for much greater creative freedom than would be viable in any larger-scale speech-community – cf. Hohenhaus 2005: 361).

3. Conclusion

In this chapter I hope to have demonstrated that the functional side of WF is significantly more complex than simple ‘naming’ (and/or ‘syntactic recategorisation’ etc.). This is particularly evident once we go beyond the rather stable area of the lexicon and its conventionalised word-stock and begin considering active, nonce WF, the more productive as well as the truly creative aspects, where the very act of forming a complex word serves a specific purpose in the textual and/or situational context. Often rather particular, specialised functions can be discovered, many of which are of a metacommunicative nature.

Needless to say, however, the overview that this chapter has provided is necessarily far from exhaustive.²⁷ But it will hopefully serve as a) an illustration of the range of phenomena in an area that, all too often, is not considered seriously enough in the literature on WF, and b) an encouragement to devote more attention to these intriguing aspects of the study of word-formation.

References

- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Aronoff, M. 1988. “Review of Di Sciullo.” A.-M., Williams, E. 1987 *Language* 64 (4): 766–770.
- Bauer, L. 1983. *English Word-Formation*. Cambridge: CUP.
- Bauer, L. 2001. *Morphological Productivity*. Cambridge: CUP.
- Brekke, H. 1976. “Delokutive Verben: Ein Sprechakttheoretisch fundierter Wortbildungstyp.” In *Grammatik. Akten des 10. Linguistischen Kolloquiums, Tübingen 1975*, vol. 2, K. Braunmüller and W. Kürschner (eds). 69–76. Tübingen: Niemeyer.
- Crystal, D. 1998. *Language Play*. London: Penguin.
- Crystal, D. 2001. *Language and the Internet*. Cambridge: CUP.
- De Brabanter, P. 2005. “Autonymy and its impact on the lexicon”, to appear in *WORD* 56 (2), available online at: <http://jeannicod.ccsd.cnrs.fr/documents/disk0/00/00/06/02/index.html>
- Di Sciullo, A.-M. and Williams, E. 1987. *On the definition of word*. Cambridge, MA.: MIT Press.
- Dederding, H.-M. 1983. “Wortbildung und Text: Zur Textfunktion (TF) von Nominalkomposita (NK.)” *Zeitschrift für Germanistische Linguistik* 11: 49–64.
- Downing, P. 1977. “On the creation and use of English compound nouns.” *Language* 53 (4): 810–842.
- Dressler, W. U. 1988. “Preferences vs. Strict Universals in Morphology: Word-Based Rules.” In *Theoretical Morphology: approaches in modern linguistics*. M. Hammond and M. Noonan (eds). 143–154. San Diego: Academic Press.

27. E.g. metaphor/metonymy is a functional area not touched upon in this chapter; but cf. Hohenhaus (1996: 336ff.); see also Lipka (2002: 138ff.).

- Giegerich, H. 2004. "Compound or phrase? English noun-plus-noun constructions and the stress criterion." *English Language and Linguistics* 8: 1–24.
- Ghomeshi, J., Jackendoff, R., Rosen, N. and Russell, K. 2004. "Contrastive Focus Reduplication in English (the salad-salad paper)." *Natural Language & Linguistic Theory* 22: 307–357.
- Heringer, H. J. 1984. "Wortbildung: Sinn aus dem Chaos." *Deutsche Sprache* 12: 1–13.
- Hohenhaus, P. (1996) *Ad-hoc Wortbildung – Terminologie, Typologie und Theorie kreativer Wortbildung im Englischen*. Frankfurt/M., Berlin, Bern, New York, Paris, Wien: Peter Lang.
- Hohenhaus, P. 1998. "Non-Lexicalizability as a characteristic feature of nonce word-formation in English and German." *Lexicology* 4 (2): 237–280.
- Hohenhaus, P. 2000. "An Overlooked Type of Word-Formation: Dummy-compounds in German and English." In *German Studies towards the Millennium* C. Hall and D. Rock (eds), 241–260. Oxford, Bern, Berlin, Bruxelles, Frankfurt/M., New York, Wien: Peter Lang.
- Hohenhaus, P. 2004. "Identical Constituent Compounds – a Corpus-based Study." *Folia Linguistica* 38 (3–4): 297–331.
- Hohenhaus, P. 2005a. "Elements of traditional and 'reverse' purism in relation to computer-mediated communication." In *Linguistic Purism in the Germanic Languages*. [Studia Linguistica Germanica 75] N. Langer and W. Davies (eds), 204–220. Berlin, New York: De Gruyter.
- Hohenhaus, P. 2005b. "Lexicalization and Institutionalization." In *Handbook of Word-Formation*. [Studies in Natural Language and Linguistic Theory 64] P. Štekauer and R. Lieber (eds), 353–373. Berlin, Dordrecht, Heidelberg, Norwell: Springer.
- Hohenhaus, P. (forthcoming) "Where Word-Formation cannot Extend the Vocabulary – Creativity, Productivity and the Lexicon in Synchronic and Diachronic Morphology." In *Lexical Change and the History of the English Vocabulary*. D. Kastovsky (ed.), Berlin, New York: Mouton de Gruyter.
- Jackendoff, R. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
- Jackendoff, R. 2002. *Foundations of Language*. Oxford: OUP.
- Kastovsky, D. 1978. "Zum gegenwärtigen Stand der Wortbildungslehre des Englischen." *Linguistik und Didaktik* 36: 351–366.
- Kastovsky, D. 1982. "Word-Formation: A Functional View." *Folia Linguistica* 16: 181–198.
- Kastovsky, D. 1986. "The problem of productivity in word formation" *Linguistics* 24: 585–600.
- Lipka, L. 1975. "Re-discovery procedures and the lexicon." *Lingua* 37: 197–224.
- Lipka, L. 1977. "Lexikalisierung, Idiomatisierung und Hypostatisierung als Probleme einer synchronen Wortbildungslehre." In *Perspektiven der Wortbildungsforschung. Beiträge zum Wuppertaler Wortbildungskolloquium vom 9. – 10 Juli 1976. Anlässlich des 70. Geburtstages von Hans Marchand am 1. Oktober 1977*. H. Brekle and D. Kastovsky (eds), 155–163. Bonn: Bouvier.
- Lipka, L. 1987. "Word-Formation and Text in English and German" In *Neuere Forschungen zur Wortbildung und Historiographie der Linguistik*. B. Asbach-Schnittker and J. Roggenhofer (eds), 59–67. Tübingen: Gunter Narr.
- Lipka, L. 2000. "Englisch (and General) Word-Formation – The State of the Art in 1999." *Anglistentag 1999 Mainz Proceedings*. B. Reitz and S. Rieuwerts (eds), 5–20. Trier: Wissenschaftlicher Verlag.
- Lipka, L. 2002. *English Lexicology*. Tübingen: Gunter Narr.
- Marchand, H. 1969. *The Categories and Types of Present-Day English Word-Formation*. 2nd revised and enlarged edition. Munich: Beck.

- Motsch, W. 1977. "Ein Plädoyer für die Beschreibung von Wortbildungen auf der Grundlage des Lexicons." In *Perspektiven der Wortbildungsforschung. Beiträge zum Wuppertaler Wortbildungskolloquium vom 9. – 10. Juli 1976. Anlässlich des 70. Geburtstages von Hans Marchand am 1. Oktober 1977*. H. Brekle and D. Kastovsky (eds), 180–202. Bonn: Bouvier.
- Munat, J. 2003. "When is a noun string a phraseological unit?" *Ranam* 36: 31–47.
- Neuvel, S. and Singh, R. 2001. "Vive la différence! What morphology is about." *Folia Linguistica* 35 (3–4): 313–320.
- Okrand, M. 1992. *The Klingon Dictionary*. New York: Pocket Books.
- Ortner, L. 1985. "Wortbildungs- und Satzbildungsmittel zum Ausdruck von Metaphern und Vergleichen in Science-Fiction-Texten." *Studien zur deutschen Grammatik*. E. Koller and H. Moser (eds), 255–275. Innsbruck: Germanistische Reihe.
- Radford, A. 1988. *Transformational Grammar*. Cambridge: CUP.
- Raskin, V. 1985. *Semantic Mechanisms of Humor*. Dordrecht: Reidel.
- Scalise, S. 1986. *Generative Morphology*. Dordrecht: Foris.
- Schmid, H.-J. 2000. *English abstract nouns as conceptual shells. From corpus to cognition*. Berlin: Mouton de Gruyter.
- Schmid, H.-J. 2005. *Englische Morphologie und Wortbildung*. Berlin: Erich Schmidt Verlag.
- Seppänen, L. 1978. "Zur Ableitbarkeit von Nominalkomposita." *Zeitschrift für Germanistische Linguistik* 6: 133–150.
- Saukkonen, P. 2003. "How to define and describe genres and styles." *Folia Linguistica* 37 (3–4): 399–414.
- Štekauer, P. 2002. "On the Theory of Neologisms and Nonce-formations." *Australian Journal of Linguistics* 22 (1): 97–112.
- Štekauer, P. 2005. "Onomasiological Approach to Word-Formation." In *Handbook of Word-Formation*. [Studies in Natural Language and Linguistic Theory 64] P. Štekauer and R. Lieber (eds), 207–232.
- Tuggy, D. 2005. "Cognitive Approach to Word-Formation." In *Handbook of Word-Formation*. [Studies in Natural Language and Linguistic Theory 64] P. Štekauer and R. Lieber (eds), 233–265. Berlin, Dordrecht, Heidelberg, Norwell: Springer.
- Welte, W. 1974. *Moderne Linguistik: Terminologie/Bibliographie*. Munich: Max Hueber.
- Welte, W. and Rosemann, P. 1990. *Alltagssprachliche Metakommunikation im Englischen und Deutschen*. Frankfurt a.M., Bern, New York, Paris: Peter Lang.
- Wladowa, E. W. 1975. "Okkasionele Wortbildungen mit dem gleichen Stamm als Satz- und Textverflechtungsmittel." *Textlinguistik* 4: 71–87.

The phonetics of ‘*un*’

Jen Hay

This paper analyses words prefixed with *un-* in a large corpus of New Zealand English, in order to test the hypothesis that the difference between retrieved words and productively coined words may be discerned phonetically. The results reveal that factors which tend to facilitate decomposability (boundary-like phonotactics; high frequency of the derived form relative to the base) are associated with phonetically longer prefixes. That is, more decomposable words (e.g. *unburstable*) tend to have longer prefixes than less decomposable words (e.g. *unfortunate*).

In addition, the study reveals an intriguing change over time, in which contemporary New Zealanders seem to be using *un-* less productively than their counterparts born in the 19th and early 20th Century. Contemporary New Zealanders tend to use *un-* slightly less frequently than older New Zealanders, and are more likely to use sequences of *not*+adjective.

1. Introduction

Most data for research on morphological productivity/creativity (and indeed, morphology in general), has come from intuition, elicitation or experimentation. Recent work has also examined patterns in large corpora, which contain either written texts, or transcriptions of speech events. However, very little work has examined ‘spoken morphology’, investigating the way in which morphological structure is actually reflected in the phonetics of speech.

This paper is an exploratory study of morphological creativity in a large corpus of spoken New Zealand English, giving special attention to the phonetic realisation of the prefix *un*.¹ The corpus contains interviews with hundreds of New

1. I’m grateful to Karen Malcolm for her assistance with the phonetic analysis conducted here, Robert Fromont for the development of the software that enabled easy extraction of appropriate examples, and Judith Munat for her helpful comments on the paper. This paper uses data from the Origins of New Zealand English (ONZE) project at the University of Canterbury.

Zealanders born between 1850 and 1971 (see Gordon et al. 2004). The earliest recordings were made in the 1940s. I demonstrate that the length of the prefix is related to the degree to which it is being used as an independent morpheme and that the use of the prefix *un-* has varied over the history of NZE.

The finding that individual words can contain some phonetic evidence regarding the degree to which they are morphologically decomposable has great potential, not only for furthering our understanding of the contexts in which affixed words tend to be creatively coined, but also for the possibility of gaining insight into the extent of productive word-formation in the speech of specific individuals on specific occasions.

2. Background

It may seem surprising for a book on lexical creativity to include a paper examining details of the phonetics of a single, relatively common affix. What, after all, is creative about the use of a straightforward prefix like *un-*? Indeed – sometimes nothing. When a speaker uses words like *unfortunate* or *unnecessary* the chances are that they are being retrieved whole from memory, rather than creatively constructed. The listener, too, is likely to understand the meaning of *unfortunate* without necessarily decomposing the word into its constituent elements or analysing it semantically as *un+fortunate*. In some other words, however, the prefix *un-* may carry an important and distinct part of the meaning: rather than being retrieved whole from memory, the word may be created compositionally, and the listener will have to decompose it in order to acquire its meaning. For example, there is likely to be a greater degree of lexical creativity in producing the word *unburstable* than the word *unfortunate*.

While the focus in this paper is on the affix *un*, this study is intended to stand as a test-case for a more general question. Can the difference between retrieved words (e.g. *unfortunate*) and creatively coined words (such as *unburstable*) be discerned phonetically? If so, this would open up a range of research possibilities, enabling us, for example, to closely study the circumstances under which words tend to be retrieved whole or creatively coined.

The particular focus in this paper is on the length of the affix *un*. Examples (1)–(6) are utterances produced by speakers in the Origins of New Zealand Eng-

The Mobile Unit data was collected by the Mobile Disc recording unit of the NZ Broadcasting Service, the Intermediate Corpus data was collected by Rosemary Goodyear, Lesley Evans and members of the ONZE team. The work done by members of the ONZE team in preparing the data, making transcripts and obtaining background information is gratefully acknowledged.

lish Corpora (described further below). Following each utterance the length of the prefix *un* is given in milliseconds. The utterances are presented in order of increasing length of the prefix. Since these utterances come from different speakers who no doubt have different speaking rates, comparing these raw length measurements may not tell us much at all. However, the variation here is quite large – the prefix in (6) is more than 3.5 times as long as the prefix in (1). This variation suggests that it is worth investigating whether the length of the affix correlates in some sense with the degree to which it is functioning as an independent constituent.

- (1) but I unfortunately . couldn't open the door to look in (.06ms, nasal only – no vowel produced)
- (2) that was very unsatisfactory (.09m)
- (3) and he let me know it in no uncertain terms (.11ms)
- (4) to get an unboring job (.17ms)
- (5) and had many a trying rough trip on the then rough unmetalled tracks (.22ms)
- (6) when you're looking at it from an unbiased point of view (.22ms)

There is some evidence that phonemes which are prone to deletion are less likely to be deleted if they are affixes. For example, studies on final *t*-deletion in English show that /t/ is more likely to be deleted in monomorphemic words (such as *past*), than when the /t/ is itself an affix (e.g. *jumped*) (Guy 1980, 1991; Labov 1989).

If an affix comprises a meaningful constituent of a particular word, it is advantageous and perhaps even necessary for that affix to be phonetically salient rather than reduced or deleted. Indeed, in (1), above, the vowel of the affix was not produced at all; the affix was represented by a syllabic nasal. However, one would be surprised to find the same level of reduction in a word such as *unboring* (in 4). In fact, if a listener simply heard '*boring*' preceded by a syllabic nasal, it is not clear that the intended meaning would be available at all. The same line of reasoning leads to the hypothesis that highly compositional, highly creative words may contain phonetically longer prefixes than less compositional, highly lexicalised words.

3. The storage and processing of affixed words

Before proceeding to investigate whether the degree of reduction of *un* varies in any systematic way, it is first necessary to outline some assumptions about the processing of affixed words. In the previous section I invoked a contrast between

un words which are lexicalised, and those which are creatively coined. However, it is more productive to conceptualise morphological decomposability as a continuum. At one extreme are simplex words, or words that have undergone semantic drift to the extent that they are no longer seen as being composed of separate morphemes, and so, for all intents and purposes, they are processed as morphologically simple. At the other end of the continuum are words which may never have been encountered before (such as *unburstable*), and are therefore created and perceived as compositional. Most affixed words, however, are likely to fall somewhere between these two extremes. Hay and Baayen (2005) outline a wide range of literature which suggests that morphological complexity is gradient. Affixed words seem to be 'affixed' to varying degrees. Those which have been previously encountered are stored in the lexicon – including words with inflectional affixes. Both retrieval and composition play a role (to varying degrees) in lexical access. A useful metaphor for considering the balance of retrieval and composition in lexical access is the dual route model suggested by Baayen and Schreuder (see, e.g. Baayen 1992; Schreuder and Baayen 1995; Baayen and Schreuder 1999). According to this model, lexical access in perception proceeds as a race between direct access of the 'full form' (e.g. *unhappy*), and parsing of the constituents (*un+happy*). While either route may be solely responsible for access on any given occasion, in general the two routes converge interactively on the correct meaning representation. The relative contribution of the different routes is not simply a matter of lexical access on that particular occasion, but rather the access route impacts the storage of the word. Words for which the compositional route dominates tend to be more decomposable than words for which the direct route dominates – that is, they are stored together with stronger associations to their component parts. Words for which the direct route tends to dominate have weaker associations with their component parts. Such words are less 'tied' to the semantics of the constituents, more prone to semantic drift, and are rated in experiments as feeling less morphologically complex (Hay 2003). Thus, we can conceive of a word as highly affixed if the component parts play a strong role in lexical access, in which case the word is stored with strong associations with its parts. A word is less affixed if the component parts play a relatively minor role in lexical access, and the word is stored with weak associations with its parts. Less decomposable words fall toward the less creative end of this continuum and are dominated by whole word storage, while more decomposable words fall toward the more creative end of the continuum, and are characterised by a high degree of compositionality.

Having established the notion of degrees of decomposability, it is worth considering some factors which contribute to the degree of decomposability of affixed words. In Hay (2001) I argue that one important factor determining the degree

of decomposability of a morphological form is the relative frequency of a derived word and its base. Words which are more frequent than their bases (e.g. *illegible* is more frequent than *legible*) tend to be less decomposable than words which are less frequent than their bases (*illiberal*, less frequent than *liberal*). *Unfortunate*, for example, is more frequent than *fortunate*, which would tend to facilitate the role of whole word access rather than access via its component parts.

The majority of affixed words in English, unlike *unfortunate*, are less frequent than their bases. The general intuition behind this effect is that the relative salience of the base word within the derived word facilitates the perception of the component constituents. This, in turn, leads to a stronger association between the affixed word and its component parts. Consequently, words which are more frequent than their bases tend to be rated as less 'complex' than words which are less frequent than their bases (Hay 2001), and affixes which are dominated by the latter type of word tend to be more productive overall (Hay and Baayen 2002, 2004).

There is some phonetic evidence for the role of relative frequency in complex words. Hay (2003) demonstrates that words such as *swiftly* (which is more frequent than *swift*) tend to be pronounced with less of a /t/ than words such as *softly* (which is less frequent than *soft*). Note that *swiftly* and *softly* are approximately equal in frequency, it is the relative frequencies of the derived forms and the bases that differ. Words which are less decomposable tend to be associated with a greater degree of reduction at the morphological boundary. In addition, Hay (2003) found that words which are less frequent than their base are more likely to attract a contrastive pitch accent on the prefix – suggesting that the prefix is a transparent and meaning-bearing constituent.

Another factor that seems to facilitate decomposability in English is the phonotactics across the boundary between an affix and a base. It is well established that listeners (even infants) are sensitive to the co-occurrence probabilities of phonemes, and use low probability phoneme transitions to spot the boundaries between words (Saffran et al. 1996a, 1996b). This sensitivity carries over to some degree to affixed words, such that low probability phoneme transitions across morpheme boundaries facilitate parsing and reinforce strong associations between derived forms and their component parts. For example *bowlfu* tends to be rated as less decomposed than *pipefu*, because the l][f transition is attested inside monomorphemic words in English (e.g. *dolphin*), whereas the p][f transition is not. Affixes which tend to create many low-probability phoneme transitions tend to be more productive (Hay and Baayen 2004). Prefixed words containing illegal phoneme transitions tend to be more semantically transparent, less polysemous, and are more likely to be less frequent than their base (Hay 2003). Thus, we might expect a word like *uncork*, all other things being equal, to be less decomposable, and stored with weaker associations with its parts than a word like *unhinge* be-

cause the n][k transition is fairly well attested inside English monomorphemic words, but the n][h transition is not.

The general hypothesis that *un* may be longer when it is more meaning-bearing is consistent with work on the length of words, which shows that the informativeness of a word in its local context can predict its duration. More informative words are longer than words which are highly predictable in the context in which they occur (Gregory et al. 1999; Jurafsky et al. 2000).

If we entertain the hypothesis that *un* may be longer in words which have strong associations with their component parts, then two more specific hypotheses will follow. One is that *un* may be shorter in words which are more frequent than their bases, and the other is that *un* may be shorter in words which contain legal phonotactics across the morpheme boundary. If it is the case that the length of *un* correlates with the decomposability of the derived word, then this would provide evidence that gradient morphological structure can be reflected in the fine phonetic detail of spontaneous speech. This would open up a variety of possibilities for future research, including the study of the speech of individuals where a particular word may be formed ‘creatively’ by one speaker (or on one occasion), while being retrieved whole by another speaker (or by the same speaker on a different occasion). Phonetic evidence on this point would enable us to conduct considerably more sophisticated and detailed work on what exactly constitutes ‘lexical creativity’ both within individuals and across different groups of speakers.

4. Methodology

I searched for the orthographic string *un* in the recordings of a subset of speakers from the Origins of New Zealand English (ONZE) Corpora of New Zealand English at the University of Canterbury. The earliest born ONZE speakers were interviewed for a radio broadcast in the 1940s, and the most recent speakers were interviewed by undergraduate students as part of their course requirements. The corpus also includes some interviews which were originally conducted for Oral History projects. A full description of the recordings is available in Gordon et al. (in press), and extensive phonetic analysis of the early recordings has been reported in Gordon et al. (2004). The full set of recordings comprises over 1000 hours of audio. We are in the process of migrating these recordings to a digitally interactive system (Fromont and Hay 2004; Gordon et al. 2005), and the recordings we analysed were those available in our interactive database as of October 2005. From this data, we extracted tokens of *un* in the speech of 244 speakers born in New Zealand between 1857 and 1982. This data-set includes 109 females and 135 males.

We analysed all tokens in which *un* was an affix, and also (for comparison) tokens of *un* in the words *unless* and *until*. These items were chosen for comparison as, unlike other potential candidates such as *uncle* and *under*, they more directly parallel the most common stress pattern for the affixed words (i.e. stress on the base rather than on the prefix). A total of 359 affixed forms and 310 non-affixed forms were analysed.

The length of *un* was measured in milliseconds. Whether the vowel in *un* was full, absent, or reduced to schwa was also recorded. We first compare the phonetics of the affixes and the non-affixes to investigate whether an affixal *un* tends to be longer than a non-affixal *un*, and whether it is less prone to reduction.

5. Results

5.1 Affixes versus non-affixes

In considering the difference between affixed and non-affixed forms, we first examine the likelihood of the vowel being reduced or completely absent. Table 1 shows the distribution, for both affixed and non-affixed forms. Not surprisingly, we find that the *un* in the non-affixed forms is considerably more likely to be reduced. Very few prefixes actually contain a reduced or absent vowel. One of the phonetic qualities that is associated with the prefix *un*, then, is a relatively robustly present full vowel. If the *un*- carries a meaning component, then there is a disincentive to reduce the vowel.

We then considered the relative length of *un*.

Table 1. Comparison of degree of reduction of affixed and non-affixed forms

	Not-affixed	Affixed
Vowel Absent	95	4
Vowel Reduced	122	7
Full Vowel	93	348

Because there are considerable differences in individual speech rate, the comparison of raw length measurements may be misleading. Instead, we attempted to normalise for speech rate to some degree by calculating the length of *un* relative to the length of the following syllable. In defining the following syllable, we assume the maximal onset principle (such that, e.g. the /l/ in *unbalanced* belongs to the following syllable, but the /l/ in *unhealthy* does not). We then fit a linear model designed to predict the relative length of *un*.

Using the relative length of *un* normalises for speech rate to some degree. However, this also introduces other complicating factors related to inherent differences in the lengths of following syllables. For example, the following syllable may vary in length depending on whether it is stressed, and whether the word occurs utterance finally. In order to inspect reasonably comparable items, we restricted this analysis to cases where *un* did not carry a pitch accent, and where it was produced with a full vowel. Affixed words which did not have lexical stress on the first syllable of the base were also excluded from the analysis. Whether or not the word was utterance final was included as a factor in the statistical model.

Probably the most important factor affecting the length of the following syllable is the weight of that syllable. We therefore coded the following syllable as “light”, “heavy” or “superheavy”. Light syllables contained a short vowel (i.e. one mora). Heavy syllables contained a long vowel, or a short vowel and a coda (two moras). Superheavy syllables contained at least a long vowel and a coda, or a short vowel and two coda consonants (i.e. three moras). We collapsed *unless* and *until* together into a separate “unaffixed” category in the same factor group.

The details of the linear regression are shown in Tables 2 and 3. Whether or not the word was utterance final had a significant effect. This is because words which are utterance final are prone to phrase-final lengthening (Wightman et al. 1992). This would lengthen the base, and make the relative length of *un* somewhat shorter.

The nature of the following syllable also had a significant effect. The model predictions (i.e. holding utterance-finality constant) are shown in Figure 1. As

Table 2. Wald statistics for model predicting the relative length of *un-*, over both affixed and unaffixed forms

Factor	d.f.	Partial SS	MS	F	p
Syllable Type	3	2.999	0.999	22.2	< .0001
Utterance Final	1	0.745	0.746	16.55	< .0001
REGRESSION	4	3.828	0.957	21.25	< .0001
ERROR	332	14.953	0.045		

Table 3. Coefficients for model predicting relative length of *un-* over both affixed and unaffixed forms

Intercept	0.754
Syllable type=light	0.056
Syllable type=super	-0.223
Syllable type=unaffixed	-0.124
Utterance Final=yes	-0.115

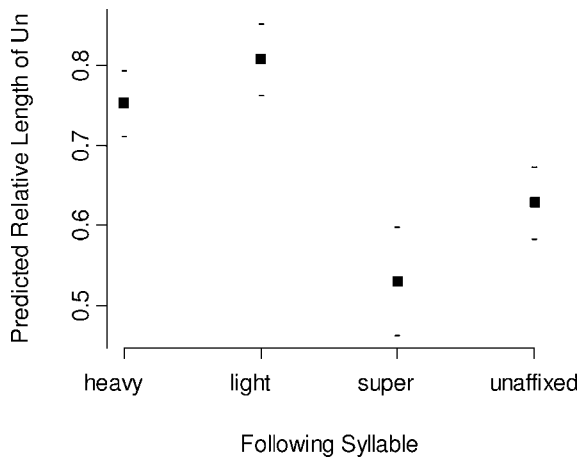


Figure 1. Model predictions showing the effect of the following syllable

expected, the heavier the following syllable, the shorter “un” is relative to that syllable. *Unless* and *until* both contain following syllables coded as heavy. However, as can be seen from Figure 1, the *un* in these words is actually considerably (and significantly) shorter than the *un* in the comparable (heavy) affixed forms.

This provides some baseline evidence that when *un* is an affix it tends to be relatively longer than when it is simply a non-meaning-bearing syllable. It will be noted that the words we have used as controls are both function words – an inevitable choice in order to avoid tokens in which the *un* was stressed. Some readers may consider this an unfair comparison, as function words tend to be more reduced than content words (see e.g. Jurafsky et al. 2001; Lavoie 2002; Shi et al. 2005). This argument could, indeed, be rallied against the results regarding the phonetic reduction of the vowel, and so these results should be regarded with some caution. However, in terms of the relative length of *un*, we believe that this should not affect the results. This is because in the function word, we expect the whole word to be prone to shortening so, while the absolute length of *un* may be shorter, the following syllable should also be shorter.

This comparison of affixed forms with *until* and *unless* was intended to establish a simple baseline comparison for our more detailed investigation of affixed forms. When *un* is an affix (as opposed to part of a simple word) there is some evidence that it is less prone to phonetic reduction (either in vowel quality, or through vowel deletion), and it also tends to be longer.

In terms of our wider hypothesis about the phonetics of affixed forms, then, we expect highly creative, highly decompositional words with *un* to have relatively longer *uns* than words which are highly lexicalised. Following Shi’s (2005: 359) ar-

gument that the reason function words tend to be reduced “can be attributed to their low semantic load and high predictability”, we can assume that affixes which carry a smaller semantic load (by virtue of being in highly lexicalised words) should also show more phonetic reduction than affixes which appear in productively coined words and which carry a greater semantic load.

This possibility is explored in the following section.

5.2 The phonetics of affixal *un*

Only 11 affixes were, in fact, reduced. Obviously this is too small a number to assess what factors may facilitate reduction. It is, however, worth pointing out that the tendency regarding relative frequency goes in the expected direction. As shown in Table 4, while 2.5% of forms which are less frequent than their bases contain reduction of *un*, this is true for 4.3% of the forms which are more frequent than their bases. While by no means statistically conclusive, this is at least consistent with the notion that words which are more frequent than their bases are less likely to be highly decompositional. In such words, the *un* may not function as a distinct meaning-bearing unit, and is thus available for reduction.

As outlined above, in order to assess the effect of length, we measured the length of the affix in milliseconds, relative to the length of the following syllable (assuming maximal onsets). This calculation of the relative length of *un* goes some way toward eliminating speech rate effects. It does, however, introduce other potential effects – namely independent variables influencing the length of the following syllable, including its weight, whether it carries lexical stress, whether it is word final, and whether the entire word is phrase final. In an attempt to remove any significant effects of this kind, these were considered in the modelling process. We also considered the part of speech (i.e. distinguishing between verbal and adjectival *un*), whether the base had been recently mentioned in the discourse, and the age and gender of the speaker. The factors of particular interest were whether the base or the full word was more frequent, and the phonotactics of the transition.

An ordinary least squares model of the relative length of *un* was fit by hand, starting with a relatively saturated model, and removing factors which did not

Table 4. Comparison of degree of reduction of vowel in affixed forms, by relative frequency of the derived form and the base

	Reduced or Absent	Full Vowel
Base more frequent	6 (2.5%)	234 (97.5%)
Word more frequent	5 (4.3%)	111 (95.7%)

reach significance. The model is shown in Tables 5 and 6. Syllable weight (light, heavy, super-heavy), whether the syllable carried lexical stress, and utterance finality all played a significant role. These effects (from the model – i.e. holding all other significant effects constant) are shown in Figures 2 and 3. As would be expected, factors which would lead to the increased length of the following syllable (a heavier syllable, a lexically stressed syllable, and an utterance final word) all decrease the length of *un* relative to the length of the following syllable. It is hoped that the inclusion of such factors in the model effectively subtracts any differences

Table 5. Wald statistics for model predicting the relative length of *un-* in affixed words

Factor	d.f.	Partial SS	MS	F	P
Most frequent (Factor+Higher Order Factors)	2	0.648	0.324	5	0.007
All Interactions	1	0.647	0.647	9.98	0.002
Period (Factor+Higher Order Factors)	2	0.681	0.341	5.26	0.006
All Interactions	1	0.647	0.647	9.98	0.002
Syllable type	2	4.504	2.252	34.77	< .0001
Strong-following	1	1.412	1.412	21.8	< .0001
Phonotactics	3	1.586	0.529	8.16	< .0001
Utterance Final	1	0.324	0.324	5	0.026
Most frequent * Period (Factor+Higher Order Factors)	1	0.647	0.647	9.98	0.002
REGRESSION	10	15.073	1.507	23.27	< .0001
ERROR	334	21.637	0.064		

Table 6. Coefficients for model predicting the relative length of *un-* in affixed words

Intercept	1.216
Most frequent =word	-0.128
Period=late	-0.082
Syllable type=light	0.129
Syllable type=super	-0.249
Strong-following=yes	-0.249
Phonotactics=good	-0.199
Phonotactics=same	-0.099
Phonotactics=vowel	-0.0001
Utterance Final=yes	-0.071
Most frequent=word * Period=late	0.194

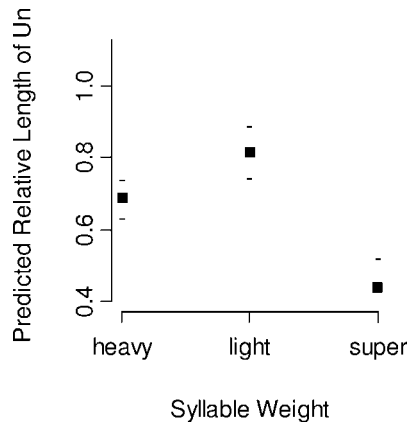


Figure 2. The effect of weight of the following syllable on the relative length of *un-*

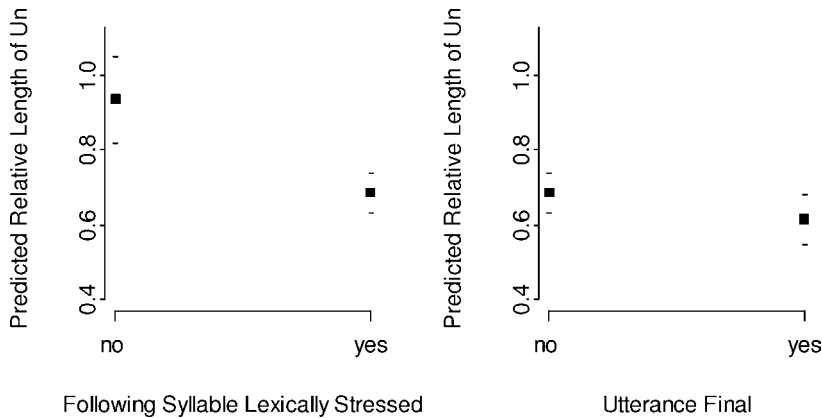
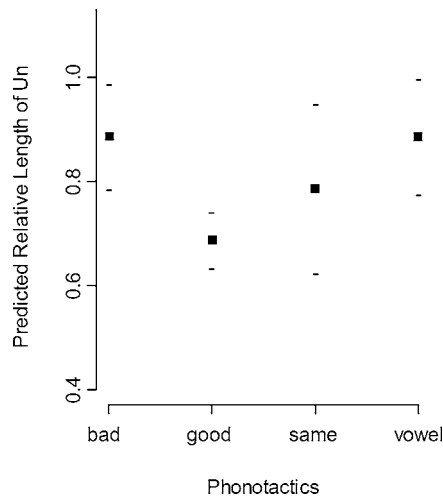


Figure 3. The effect on relative length of *un-* when the following syllable carries lexical stress and when the word is utterance final

in relative length which are attributable solely due to variation in the length of the following syllable. We hypothesised that any further variation in relative length may be attributable to effects carried by variation in the length of the affix itself. A result which is interesting in this regard is the effect of the phonotactics across the word boundary. This factor contains four categories – the base begins with a vowel (e.g. *uneasy*), with an alveolar nasal (e.g. *unnatural*), with an onset leading to a phonotactically legal transition (e.g. *untidy*, *unsafe*), and with an onset leading to a phonotactically illegal transition (e.g. *unwell*, *unpleasant*). Legality was established by checking whether the transition occurs in any monomorphemic words in English, by using a corpus of monomorphemes from the CELEX lexical database (see Baayen et al. 1995 for a description of CELEX and Hay et al. 1999 for a description of the subset of monomorphemes). For example the transition



bad=base begins with an onset creating an illegal transition (e.g. unpleasant);
 good=base begins with an onset creating a legal transition (e.g. untidy);
 same= base begins with alveolar nasal;
 vowel= base begins with vowel.

Figure 4. The effect of the phonotactics across the morpheme boundary on the predicted relative length of *un-*

n][t occurs inside monomorphemic words (e.g. *winter*), whereas the transition n][w does not.

This factor proved to be significant, and the effects are shown in Figure 4. Interestingly, the three categories 'bad' (i.e. containing illegal consonantal transition), 'vowel' and 'same' (containing an alveolar nasal) are not significantly different from one another. This is perhaps not surprising as there is a sense in which all of these are illegal. We kept the sequence of nasal-vowel and nasal-n separate as, while these are technically illegal, they do allow for the possibility of resyllabification over the morpheme boundary. That these patterned with the 'bad' consonant-initial bases suggests that this didn't happen. Intriguingly, in words in which the base begins with a consonant that forms a legal transition, the length of the *un* is considerably reduced. This reinforces Hay's (2003) claim that such words are more likely to be accessed whole, and so are more prone to semantic drift. The individual parts are less likely to play a dominant role in such words, relative to the derived form, and so the length of the *un* can be reduced. Note that it is particularly revealing that this result is significant, given that the 'bad' cases are more likely to begin with consonant clusters (e.g. *unpleasant*, *unscramble*). If anything, this should lengthen the following syllable, thereby reducing the relative length of

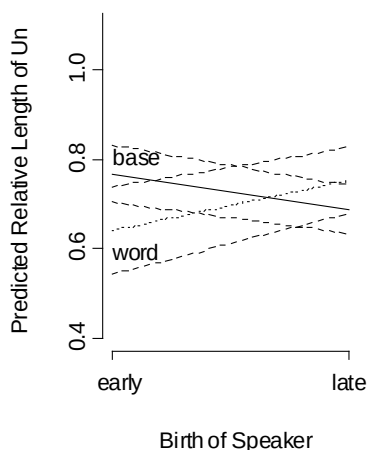


Figure 5. Interaction between a speaker's period of birth and the frequency of word or base

un. However, *un* in such cases is still significantly longer, despite this possible bias in the opposite direction.

This result, then, provides evidence in favour of the hypothesis that the relative decomposability of words prefixed in *un-* may be reflected in the length of the prefix.

In order to test for any potential changes over time, the speakers were divided into two groups: 'early' speakers, born before 1920, and 'late' speakers, born in 1920 or later. Recall that I had hypothesised that words where the base is more frequent than the full form are likely to contain a longer *un* than words where the full form is more frequent. This hypothesis was supported, but only for the early speakers. This interaction is shown in Figure 5. For early speakers, as predicted, words which are more frequent than their bases tend to have a relatively shorter (i.e. less affix-like) *un*. For late speakers, this distinction does not hold up (and, if anything, goes slightly in the opposite direction).

This change over time was certainly unexpected – why should only one group show a difference between greater and lesser degrees of lexicalisation?

One possible explanation is that there has been a change in the way that the *un-* prefix is being used in New Zealand English. Affixes certainly come and go in terms of fashion at different times, and can be influenced by a variety of social and contextual factors (see, e.g. Keune et al. 2005; Plag et al. 1999). Perhaps the contemporary New Zealanders are using the prefix less productively. This may explain why there is no distinction between words which are more frequent than their bases and those which are less frequent – perhaps the young New Zealanders are retrieving a greater proportion of their *un-* words whole.

The contemporary New Zealanders do seem to be using a slightly smaller range of *un* words. The type/token ratio for the early speakers is .49 (68/138), whereas for the later speakers it is .41 (89/218). This is a small difference, but it may suggest that the proportion of creatively-coined *un*- words being used by the later speakers is somewhat smaller.

Given that the overall amount of speech analysed differed for the early and late speakers, comparing the raw number of *un*- tokens does not help establish whether *un*- words are actually being used more frequently by the earlier speakers. In an attempt to crudely isolate environments where individual speakers could have used an *un*-word but did not, I counted all instances of the word 'not' occurring in the same data-set. The results of this comparison are shown in Table 7, from which we discover that the earlier speakers are using significantly more *un*- forms (relative to the word 'not') than the later speakers.

Table 7. Comparison of words containing *un*- and the lexical item *not* for early and late speakers

	<i>Un</i>	<i>Not</i>
Early	138 (16%)	713 (84%)
Late	218 (13%)	1512 (87%)

While part of speech was not a significant predictor of prefix length, it is worth separating the two types of *un* here. The data-set collapses two different affixes (verbal, as in *unwind*, and adjectival as in *unwell*). If there has been a change over time, it is worth considering whether this change is carried by one or both of these affixes. The comparison between early and late speakers in Table 7 reaches significance if considered over the verbs alone (3% vs. 1% 'un', $p < .003$), but not for adjectives alone, (14% vs. 12% 'un', $p < .13$). However, for both groups of words, the later speakers do seem to be producing a lower *un/not* ratio than the early speakers are.

A more direct investigation, in which we analyse the instances in which speakers could have produced an *un*- form but didn't, is difficult to conduct for the verbal items. However, it is possible in the case of adjectives. In order to more directly compare the relevant frequencies for adjectives, I took all occurrences of *not* followed by an item that could be an adjective – that is, 'adjective' was listed as a potential part of speech for that lemma in the CELEX lexical database. I then went through each of these by hand, identifying cases where the meaning could potentially have been conveyed by a prefixed form (e.g. *not pleasant*, *not healthy*, *not sure...*). There were 50 such examples amongst the early speakers, and 135 examples produced by the late speakers. The late speakers, therefore, have a sig-

nificantly lower ratio of *Un+adj* versus *Not+Adj* than the early speakers do. This distribution is shown in Table 8.

Table 8. Comparison of *un-* prefixed adjectival forms and sequences of *not* + adj. which could be expressed with a prefixed form

	Un+adj	Not+Adj
Early	114 (70%)	50 (30%)
Late	199 (60%)	135 (40%)

This analysis presents further (and more direct) evidence that the later speakers may be using adjectival *un* less productively than earlier speakers. They have a lower type/token ratio, at least for adjectives, and adjectives also have a lower “*un/not*” ratio.

While it is harder to find direct evidence with the verbs, it does seem that the use of *un* with verbs may also be decreasing. The earlier speakers actually use more verbs relative to adjectives than the later speakers (17% of the early speakers’ tokens are verbs, and only 9% of the late speakers’, $X^2 = 5.2003$, $df = 1$, $p < .03$). However, the overall number of verbal tokens is small and, unlike the adjectival tokens, there is no straightforward alternate phrasing that one could easily search for.

6. Discussion and conclusion

This study has provided evidence that the gradient decompositionality of affixed forms is reflected in phonetic detail of spontaneous speech. The particular focus of this paper was the length of *un*, and we found that factors which tend to facilitate decomposability are also associated with longer prefixes. Highly creative, productively coined *un-* words, then, are likely to reveal this productivity in the relative length of the prefix.

In addition, the study also reveals an intriguing change over time, in which contemporary New Zealanders seem to be using *un-* less productively than their counterparts born in the 19th and early 20th Century. This is interesting, particularly as we first discovered this shift by observing a subtle change over time with regard to the importance of relative frequency in predicting the length of *un*. More detailed probing revealed that contemporary New Zealanders tend to use *un-* slightly less frequently than older New Zealanders, and are more likely to use sequences of *not*+adjective.

This suggests that attention to the phonetic detail of affixed words may be a valuable source of information in the study of morphological creativity, and may

potentially enable us to assess varying degrees of compositionality in productive word-formation in the speech of specific individuals on specific occasions.

Hohenhaus (this volume) argues that there is a cline between lexical creativity, on the one hand, and productivity on the other hand (where productivity is highly rule-governed, and lexical creativity is not). This certainly seems to be a continuum, and one could also argue that the cline extends further to include monomorphemic words. There is a continuum between words which contain no identifiable subparts, through phonaesthemes (such as the *fl* in words relating to liquid such as *flow* and *float* – see Bergen 2004), through affixed words (showing different degrees of productivity), right through to highly creative nonce words which have been consciously coined.

If we conceptualise this continuum from a cognitive perspective – reflecting the varying degrees of compositionality that a speaker engages in to produce a word – then merely observing the existence or use of a word can not help us position that word with any accuracy on this continuum. This is because the degree of composition involved for a given speaker may reflect a number of factors that vary across speakers, across social contexts, and over time. The degree of composition involved may reflect whether (or how often) the relevant word or its component parts have previously been encountered, if at all. It may also reflect whether the component parts were uttered in the preceding discourse. Finally, it may depend on whether the particular word formation strategy being used is a highly productive one for that era, and/or that speaker. For example, blends such as those discussed by Lehrer (this volume), may reflect high levels of creativity by the speaker who first coins them. But for a listener who hears the blend and repeats it on a later occasion, this second usage is already somewhat less creative and less compositional than the first occurrence. In time, blends such as *motel* or *email* may function as effectively monomorphemic for many speakers. For speakers of early New Zealand English, a word like *unbiased* may once have been highly compositional whereas for a speaker of contemporary New Zealand English, the word would be positioned closer to the monomorphemic end of the continuum.

The case study presented in this paper has revealed that degrees of decompositionality may be reflected in phonetic detail. Thus, phonetics may prove to provide an insight into the nature of the continuum between stored words and highly creative or nonce words. Specific words may be differently positioned on this continuum for different speakers, or on different occasions. By examining the phonetic detail of morphologically complex words, we may be able to make some progress in understanding the highly dynamic and changing nature of lexical creativity, and the ways in which it varies across speakers, contexts and time.

References

- Baayen, H., Piepenbrock, R. and Gulikers, L. 1995. The CELEX Lexical Database (release 2) cd-rom. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania (Distributor).
- Baayen, H. 1992. "Quantitative aspects of Morphological Productivity." In *Yearbook of Morphology 1991*. G. Booij and J. van Marle (eds), 109–150. Dordrecht: Kluwer Academic Publishers.
- Baayen, H. and Schreuder, R. 1999. "War and Peace: Morphemes and Full Forms in a Non-interactive activation Parallel Dual-Route Model." *Brain and Language* 68: 27–32.
- Bergen, B. J. 2004. "The psychological reality of phonaesthemes." *Language* 80: 290–311.
- Fromont, R. and Hay, J. 2004. "Development of an Interactive Digital Corpus of New Zealand English." Paper presented at New Zealand Language and Society Conference, Palmerston North, September 2004.
- Gordon, E., Campbell, L., Hay, J., MacLagan, M., Sudbury, A. and Trudgill, P. 2004. *New Zealand English: Its Origins and Evolution*. Cambridge University Press, Cambridge.
- Gordon, E., MacLagan, M. and Hay, J. (in press). "The ONZE Corpus." In *Using Unconventional Digital Language Corpora Volume 2: Diachronic Corpora*. J. C. Beal, K. P. Corrigan, H. Moisl (eds). Palgrave.
- Gregory, M., Raymond, W., Bell, A., Fosler-Lussier, E. and Jurafsky, D. (eds) 1999. "The effects of collocational strength and contextual predictability in lexical production." *Proceedings of the Chicago Linguistics Society* 35: 151–166.
- Guy, G. 1980. "Variation in the Group and the Individual: The Case of Final Stop Deletion." In *Locating Language in Time and Space*. W. Labov (ed.), 1–36. New York: Academic Press.
- Guy, G. 1991. "Explanation in Variable Phonology: An Exponential Model of Morphological Constraints." *Language Variation and Change* 3: 1–22.
- Hay, J. 2003. *Causes and Consequences of Word Structure*. New York and London: Routledge.
- Hay, J. 2001. "Lexical Frequency in Morphology: Is Everything Relative?" *Linguistics* 39(6): 1041–1070.
- Hay, J. and Baayen, H. 2005. "Shifting paradigms: gradient structure in morphology." *Trends in Cognitive Sciences* 9(7): 342–348.
- Hay, J. and Baayen, H. 2004. "Phonotactics, Parsing and Productivity." *Italian Journal of Linguistics* 15(1): 99–130.
- Hay, J. and Baayen, H. 2002. "Parsing and Productivity." In *Yearbook of Morphology 2001*. G. Booij and J. van Marle (eds), 203–235. Kluwer Academic Publishers.
- Hay, J., Pierrehumbert, J. and Beckman, M. 2004. "Speech Perception, Well-formedness and the Statistics of the Lexicon." In *Phonetic Interpretation: Papers in Laboratory Phonology VI*. J. Local, R. Ogden and R. Temple (eds), 58–74. Cambridge University Press: Cambridge.
- Jurafsky, D., Bell, A., Gregory, M. and Raymond, W. 2001. "Probabilistic relations between words: Evidence from reduction in lexical production." In *Frequency and the Emergence of Linguistic Structure*. J. Bybee and P. Hopper (eds), 229–254. Amsterdam: John Benjamins.
- Keune, K., Ernestus, M., Van Hout, R. and Baayen, H. 2005. "Social, Geographical, and Register Variation in Dutch: From Written MOGELIJK to Spoken MOK." *Corpus Linguistics and Linguistic Theory*. (In press.)
- Labov, W. 1989. "The Child as Linguistic Historian." *Language Variation and Change* 1: 85–97.

- Lavoie, L. 2002. "Some influences on the realization of for and four in American English." *JIPA* 32: 175–202.
- Plag, I., Dalton-Puffer, C. and Baayen, H. 1999. "Productivity and register." *Journal of English Language and Linguistics* 3: 209–228.
- Saffran, J., Aslin, R. and Newport, E. 1996a. "Statistical learning by 8-month old infants." *Science* 274: 1926–1928.
- Saffran, J., Newport, E. and Aslin, R. 1996b. "Word Segmentation: The Role of Distributional Cues." *Journal of Memory and Language* 35: 606–621.
- Shi, R., Gick, B., Kanwischer, D. and Wilson, I. 2005. Frequency and Category Factors in the Reduction and Assimilation of Function Words: EPG and Acoustic Measures." *Journal of Psycholinguistics Research* 34, 4: 341–364.
- Wightman, C. W., Shattuck-Hufnagel, S. and Ostendorf, M. 1992. "Segment durations in the vicinity of prosodic phrase boundaries." *JASA* 91(3): 1707–1717.

Lexical creativity in texts

The press

Tracing lexical productivity and creativity in the British Media

‘The Chavs and the Chav-Nots’

Antoinette Renouf

In this chapter, ‘the British media’ refers to broadsheet journalism in a corpus of over 700 million words. By ‘lexical productivity’ (Renouf & Baayen 1996), we mean the inflectional spread of a word or phrase, or part thereof, over a period of time, for reasons such as fashion and topicality. By ‘lexical creativity’, we mean the range of creative processes that are wrought upon words and phrases, as new items become assimilated into the language (Renouf 2004), and existing items newly become the object of word play. The study thus assumes a diachronic perspective (Renouf 1993, 2002; Renouf & Baayen 1998). Within the context of our particular methodology, the paper will report on some lexical formation and creativity in recent data, such as the pun on the word ‘Chav’ above (roughly equivalent to US ‘white/trailer trash’), and discuss the motivation and underlying processes.

1. Introduction

The aim of this chapter is to examine the processes of lexical productivity and creativity as they are revealed in the empirical observation of journalistic texts over time. This observation is assisted by software systems designed to identify and monitor lexical innovation and change in chronological, electronically-held texts, and to extract language resources in the form of analysed databases of lexical ‘knowledge’ which is not accessible or retrievable manually.

The chapter will look at some recent lexical coinages and revived uses which journalists have either produced themselves or which they cite, and the key concepts of productivity and creativity will be defined within the terms of modern diachronic corpus study.

With the diachronic perspective on lexis, the degrees of frequency, productivity and creativity indicate how active and important a word is in the language at a given point in time, and they provide the means whereby its 'life-cycle' may be charted across the years.

2. Data

In order to face the challenge of tracking and tracing lexical change over time (Renouf 1994, 1997), we have accessed open-ended stretches of electronically held texts which are amenable to chronological analysis. The data underpinning this chapter thus comprise over 700 million running words, or tokens, of UK 'broadsheet' newspapers covering the period from 1989 to the close of 2005 in an unbroken flow.¹ Though lexical productivity and creativity vary according to the sub-domain of journalism in which they occur (sports reporting, for example, being surprisingly rich in word play and metaphor), we shall deal with these processes in 'undifferentiated' text across entire newspaper issues.

Data extracted from the World Wide Web by our 'WebCorp' tool will also be consulted.² Web-based text supplements the linguistic information found in even the most-up-to-date tailored text corpora by furnishing more recent coinages and usage (Kehoe & Renouf 2002; Renouf 2003; Renouf et al. 2003, 2005 & forthcoming; Morley 2006), although its lack of dating information inhibits the reliable diachronic study which a designed diachronic corpus permits (Kehoe 2006).

3. Methodology and tools

The research on which this chapter is based has followed a coherent intellectual path since 1990.³ Hypotheses about the relationship between meaning and surface textual patterning have proved to be sound, allowing us cumulatively to develop algorithms for automated systems capable of identifying a number of lexical and lexico-semantic phenomena in text across time. These phenomena include

1. Broadsheet newspapers comprising: *the Times* (timesonline.co.uk), *The Telegraph* (telegraph.co.uk), *the Guardian* (guardian.co.uk), *the Observer* (observer.co.uk), and/or *the Independent* (independent.co.uk)

2. <http://www.webcorp.org.uk/>

3. In a series of major research projects funded by the UK Engineering and Physical Sciences Research Council, Department of Trade and Industry, commercial companies and successive universities; see <http://rdues.uce.ac.uk/>.

neologisms (Renouf 1993b; Baayen & Renouf 1996), new word senses⁴ (Renouf 1993 c, d, e), sense relations and changing sense relations within text⁵ (Renouf 1996; Pacey et al. 1998); and the kinds of productivity and creativity of new words entering news text.⁶ The methodology for each of these procedures varies, but basically involves ‘feeding’ a specific time chunk of chronologically sequenced, fresh textual data through a set of software filters which detect novel words as well as new collocational environments of existing words.⁷ The illustrative material for this chapter has been produced via these various word filters, as well as by regular concordancing and word-processing software.

4. Productivity

Productivity is the term used to refer to the word formation processes wrought upon a lexeme. If a word is ‘productive’, it means that associated grammatical and derivational variants are being produced.⁸

4. New word senses were researched in the AVIATOR (Analysis of Verbal Interaction and Text Retrieval) project <http://rdues.uce.ac.uk/aviator.shtml>

5. Sense relations in text were researched in the ACRONYM project (Automatic Collocation Retrieval of ‘Nyms’) <http://rdues.uce.ac.uk/acronym.shtml>

6. New words entering text were researched in the APRIL project (Automatic Prediction of Innovation in Language) project <http://rdues.uce.ac.uk/april.shtml>

7. *Neologisms* – As the new text flows through the filters, a word is deemed to be new if it has not appeared before in an accumulating master word list. It is then recorded as new and appended to the master list with accompanying records of the circumstances of its occurrence (and subsequent recurrence).

New Word Senses – The meaning of an existing word is deemed to be new if it begins to be surrounded by collocates which do not match those in the established ‘collocational profile’ or wordprint which is held for each word.

Sense Relations – Two (or more) words are deemed to be sense-related (e.g. synonymous) by establishing that they typically appear in text within very similar ‘collocational profiles’ or wordprints.

Productivity and Creativity – Words encountered for the first time are analysed and classified as candidate compounds, inflections, derivations and so on by software filters based on dictionaries and by character-level morphological analysers. This gives a picture of the pattern of productivity and creativity in text as a whole.

8. The corpus linguist who works diachronically tends to view productivity as an observed fact, and is not concerned with it as a probabilistic phenomenon in the way that theoretical linguists are (notably Aronoff 1976). However, productivity is a probabilistic matter for corpus linguists working with sparse text such as early documents.

Table 1. Prefixes with growth or decline in corpus frequency ranking across 16 years

	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05
Cyber	8	9	8	9	8	5	3	2	2	2	3	2	3	4	4	5	5
Euro	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	4
Mid	2	3	3	3	4	4	3	3	3	4	4	4	5	5	5	5	5
Techno	6	7	6	6	5	5	3	3	4	4	5	4	5	5	7	6	6
Under	2	2	3	3	3	3	3	3	3	3	4	4	4	4	5	4	5
Dis	3	4	5	4	5	5	5	5	6	6	6	5	6	6	6	7	6
Inter	2	3	3	4	4	4	4	4	4	5	5	5	5	5	5	5	6
Mis	2	3	4	4	5	5	5	5	5	5	6	5	5	6	5	5	6
Mock	4	4	5	5	5	5	5	4	5	5	5	5	5	6	6	6	7
Faux	9	9	8	9	7	8	6	5	6	6	5	6	7	6	5	5	5
Trans	4	6	5	6	5	5	6	5	5	7	7	6	6	7	6	6	8
Poly	4	5	5	5	6	7	6	6	6	6	7	7	9	7	7	7	7
Uber	10	10	10	9	9	9	7	7	6	6	5	7	7	7	6	6	5
Fore	6	7	7	7	8	8	9	7	8	7	8	8	8	8	9	7	9
Vice	6	7	8	8	9	7	9	9	9	8	9	9	9	10	8	9	9

Productive items may be words or sub-word morphemes. Our analytical systems reveal that the most common word formations to appear in chronological text include compounds, derivations, proper names, acronyms and words containing numerals; less common entrants include loans, blends and conversions.

We have not yet ranked the individual roots or lemmata across time for productivity, though we will at a later stage. The analysis involved is only partially automatable, given morpheme-boundary ambiguity. But we do have a chronological record of the ranked inventories of morphological or, more precisely, affixal productivity for each year from 1989–2005.⁹ This is a rather pragmatic definition of affix, which more accurately equates to ‘combining element, hyphenated or unhyphenated’, and includes affixes, neo-classical combining forms and even words which seem to be grammaticalising – e.g. *cod(-)*, *faux(-)*, *(-)*gate (Renouf & Baayen 1998). Within this framework, the annual ranking for each affix is interesting for the insight it gives into contemporary social and media preoccupations. It confirms that the level of productivity for each affix is remarkably stable, a stability sustained below the commonest, most established band of affixes – of the ‘*un(-)*’ and ‘*in(-)*’ variety – into the lower ranks. It also reveals some slightly more productive exceptions. Tables 1 and 3 isolate the few affixes showing an upward or

9. This is not 100% accurate for the reasons of morpheme-boundary ambiguity mentioned above, but we have used a unique and sophisticated backward-forward character-level chart parser to achieve the lowest-cost and thus most likely analysis in each case.

Table 2. New prefixal formations in years of peak productivity for *cyber(-)*, *faux(-)*, *techno(-)* and *uber(-)*

cyber 2000	uber 2005
cybertainment	uberpundits
cyber-squatting	uber-waif
cyber-geeks	uber-hyped
cyber-potential	uber-talkshow
cyberdream	ubermogul
cyberpicketed	uber-media-saturated
faux 2005	techno 1996
faux-ruefully	techno-nerdish
faux-handwritten	techno-glam
faux-autism	techno-ambient
faux-clanger	technofreak
faux-cringes	techno-guru
	technoboffins

downward trend in our data, their rankings represented in terms of the frequency band into which they fall each year. There are ten bands (or ‘deciles’), with Band 1 ranking for affixes with the highest frequency score, and so on.

In Table 1, the ranking of prefixes reveals a noticeable rise for the vogue items *cyber(-)*, *faux(-)* and *uber(-)*. In addition, the prefix *techno(-)* shows definite but temporary peaking in the mid 90s.

In Table 2, which itemises the derived forms built upon these prefixes, we see that their rise is due to their function both as semantic components and as discourse signals of modernity and humour, and that they are ready to attach to most grammatical word-classes (an argument for classifying modifiers like *faux(-)* not as adjectives but as grammaticalising affixes (Renouf & Baayen, op. cit.)).

Meanwhile, for the suffixes in Table 3, we see an even more consistent profile across time, with a slight rise in rank only for vogue suffixes *(-)fest* and *(-)athon*.

Both *(-)fest* and *(-)athon* designate or allude to large-scale public events, and are thus ideal for use in news reporting. As seen in Table 4, it is clear that at the end of 2005 there is a semantic set of words combining with *(-)fest*, of the smarm, swank, schmalz, slush, gush variety. The combining form *(-)athon* is meanwhile seen to attach to a semantic group connoting triviality: *froth*, *bitch*, *cringe*, *grope*, *yawn*. Both suffixes also attach to topical proper names, such as *Sven* (Sven-Göran Eriksson – English football team coach involved in scandals) and *Bard* (metonymic reference to a Shakespeare festival), demonstrating that they are used to add glamour to the reporting of publicised events.

Table 3. Suffixes with growth/decline in frequency across 16 years

	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05
Able	2	3	3	3	3	3	3	3	3	3	4	3	4	4	5	5	5
Ation	2	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5
Fest	7	7	7	7	7	6	6	5	5	4	5	4	5	5	5	5	5
Ment	4	5	6	5	7	6	6	7	7	7	7	5	7	7	7	7	7
Ville	3	4	4	4	5	4	5	4	5	5	5	5	6	5	5	5	6
Ful	4	5	5	5	5	6	5	6	5	7	7	6	7	7	7	6	7
Ship	4	6	5	6	6	6	7	5	5	6	7	6	8	7	7	7	7
Athon	9	8	9	8	9	7	8	7	7	8	8	7	8	7	7	7	7
Ize	5	6	7	7	7	7	6	7	5	6	7	7	6	6	7	7	8
Able	2	3	3	3	3	3	3	3	3	3	4	3	4	4	5	5	5

Table 4. Some new suffixal formations in year of peak productivity for (-)fest and (-)athon

	fest 2005	athon 2005
	miaowfest	frothathon
	backslapfest	Svenathon
	smarmfest	Bardathon
	Potterfest	bitchathon
	kinderchucklefest	cringe-athon
	swankfest	improvathon

Morphologically, once a neologism begins to take hold, it typically starts to spawn inflections, derivations and even base forms. Such productivity may occur almost at once, especially if the word is in the public eye.

Alternatively, the productive process may take longer, or not occur at all. Reduced or delayed productivity may be due to one of a number of inhibiting factors. Possible ‘inhibitors’ are tentatively suggested below.

Inhibitor 1: a term may be introduced into general English which people find difficult to integrate into regular syntax; common examples are terms borrowed metaphorically from scientific domains, such as *symbiosis*. This first took on a non-biological sense in the 1600s, with *symbiosis* first recorded in 1622.¹⁰ The foreignness of this word and its variants is perhaps what has kept them stably rare in our text over the years. The first occurrence of *symbiosis* in our modern-day news

10. For the first instance of *symbiosis* in a transferred sense, OED online cites: Misselden, Edward (1622) in *Free trade; or the meanes to make trade flourish*: “To study and inuent things profitable for the publique Symbiosis”. OED does not cite *symbiotic* in a figurative sense before 1970: *Nature* 6 June 905/1: “Throughout its auspicious history the Botanical Society of Edinburgh has had a symbiotic relationship with the Royal Botanic Garden”.

Table 5. *Symbiosis* and variants with frequencies in *Independent/Guardian* 1989–2005

symbiosis 339	symbiotic 564
symbioses 5	non-symbiotic 1
symbiosis-travel 2	quasi-symbiotic 1
sapper-symbiosis 1	symbiotically 27

corpus is in 1987, and of *symbiotic* in May 1986, where the latter is glossed as follows, possibly indicating that its comprehensibility is still not taken for granted:

86.05 Part of the reason lies in the symbiotic relationship between hardware and software – **one cannot work without the other.**

Table 5 shows that there is little productivity over the years for *symbiosis*. Of the few formations, *symbiosis-travel* is part of a web address (www.symbiosis-travel.co.uk) rather than a genuine language item.

Inhibitor 2: Other examples of words which may integrate with difficulty are foreign or classical terms, such as *synergy*.¹¹ Having its origin, like *symbiosis*, in the 17th century, this term featured in our modern text at a low level of frequency from the outset. However, its derived adjective *synergistic* first appeared in a non-scientific context in our news data only in March 1989, and then only once, in the following context:

89.03 Turner's broadcasting networks and MGM film archives are more '**synergistic**'.

This is an instance of an established word form being presented in inverted commas, perhaps indicating that the writer feels it needs to be signalled as new because of its rarity and enduring strangeness.¹² The overall productivity of *synergy*

11. For the first instance of *synergy* in a transferred sense, OED online cites Heylin, Peter (1660), in *Historia quinqu-articularis; or a declaration of the judgement of the Western churches, particularly of the church of England, in the five controverted points reproached by the name of Arminianism* (1.9): "They speak only of such a **Synergie**, or cooperation, as makes men differ from a senseless stock, or liveless statua, in reference to the great work of his own conversion". Meanwhile, the variant *synergistic* is cited in OED only in 1965 for Ansoff, Harry Igor, in *Corporate strategy: an analytic approach to business policy for growth and expansion* (v. 76): "This step certainly provides for some of the **synergistic** interactions".

12. As we know, inverted commas can be applied at any point, not just to signal unfamiliarity or assumed unfamiliarity on the part of the reader, but also authorial distance, etc. The enduring strangeness might also be due to the multiple technical terms deriving from this one etymology, introducing ambiguity which could further inhibit productivity.

Table 6. Productivity of lexeme *synergy*, 1989–2005

non-synergistic 1	synergised 4	synergy 846
synergic 1	synergises 1	synergy's 1
synergical 1	synergising 2	synergy-driven 1
synergie 3	synergizing 1	synergy-less 1
synergies 906	career-synergy 1	synergy-unlocking 1
synergise 6	dis-synergy 1	

in terms of the number of each variant form has remained fairly low over 16 years, as shown in Table 6. The range of variants is also limited: there is no *synergistically* and there are few compounds.

Inhibitor 3: A lexeme which is not in widespread use may also be limited to one or two inflectional variants. The figures for *synergy* and *symbiosis* in Tables 5 and 6 indicate that the nominal form seems to be preferred for *synergy*, but both the nominal and adjectival forms appear for *symbiosis*. Some nouns may be retained in particular texts in their nominal form for rhetorical purposes, e.g. as thematic markers. Nouns which mark themes in news text include abstract nouns such as *inequality* or *unemployment* which have a sociological bias and are used in news reports on social issues. In these cases, the usual stylistic prohibition on exact repetition in written text may be obviated by inserting them into 'nominal frameworks' of the kind '*the NP1 of NP2*'.¹³

Indeed, while '*the NP1 of NP2*' is commonly found within most textual domains and types, a web search shows that '*the NP1 of unemployment*' is particularly frequent in news text, where its NP1 comprises nouns clustering in various semantic areas, as follows:

- *problem, topic, issue, question*
- *effect, consequence*
- *level, rate*
- *risk, pressure*
- *spectre, curse, blight, tragedy, ghost, humiliation, misery, scourge*

The type of noun ranges from multi-purpose semi-lexical discourse organisers such as *question* and *issue*, to more specifically lexical, evaluative items like *spectre* and *scourge*. The nouns *problem* and *risk* occur most frequently in this lexico-grammatical pattern, especially when NP2 is actually a sequence of NPs. We find for example:

13. See also 'discontinuous collocational frameworks', in Renouf A. and Sinclair J. (1991).

Table 7. Productivity for lexeme *NIMBY/nimby*, 1989–2005

NIMBY 7	Nimbyists 2	nimbyism 101
NIMBYS 1	Nimbyness 1	nimbyist 5
NIMBYism 2	anti-Nimby 1	nimbyists 2
NIMBYs 5	non-Nimby 1	nimbys 42
Nimby 124	Nimbys 62	nimbies 12
Nimby's 7	nimby 71	nimbys-come-lately 1
Nimby-ism 2	nimby-ish 1	not-in-my-back-yard 8
Nimbyish 5	nimby-ism 1	not-in-my-back-yarders 1
Nimbyism 83	nimby-ites 1	not-in-my-backyard 19
Nimbyist 1	nimbyish 2	

- 05.07 a summit dominated by **the problems of high inflation and growth**.
 04.11 the social contract underwrites **the risk of unemployment, ill-health and old age**
 6.06 a government to tackle **the problems of unemployment, low pay, and education and health services in decline**

Corpus frequency figures indicate that the nominal forms for *unemployment* and *inequality* predominate numerically – see below:

Unemployment	38,610	inequality	4,421
unemployed	18,542	inequalities	2,552
		unequal	1,983
		inequal	2

Inhibitor 4: A word may only be fleetingly topical at the outset, and subsequently only intermittently applicable to events or thinking, and thus not lend itself to productive word formation. The acronym *NIMBY*, standing for ‘Not In My Back Yard’, had a limited impact after its first appearance in 1989, and displays little variation over 17 years, as shown in Table 7.¹⁴

Although *NIMBY* very soon adopted a lower-case orthographic variant, *nimby*, this was not the sign of growing currency that such an event usually represents (Renouf 1993a). The sparse use of a limited repertoire of standard inflectional variants indicates that this acronym has not set the media world on fire over 17 years, and that, on the whole, what has been said about it has been factual rather than playful. Of course, it represents an entire phrase, but as a quasi-word it has

14. *NIMBY*: a 1980s coinage brought to general attention in 1988 by a Conservative minister, Nicholas Ridley in reference to hypocritical colleagues supporting undesirable developments while opposing their implementation in their own constituencies.

not entered into any compound formations or, with only one exception in our data – *nimbys-come-lately* – any creative use (see discussion of ‘Creativity’ in the next section).

In the context of productivity, it is pertinent to speculate as to whether anything in the nature of a generalisable principle can be deduced about the order in which these different inflections and variants are generated in language. As yet, we have not deduced any general pattern of evolution, and expect that it is likely either to be serendipitous as to which form occurs first, or determined by a particular text in which several references to the same concept arbitrarily follow one another, according to the writer’s style and the flow of the argument.

5. Creativity

Lexical creativity in journalism, as in literary prose or poetry, is employed for the purposes of achieving certain stylistic effects, such as humour or irony (for a full discussion see Munat, this volume). It is also a device used to convey to the reader a sense of authorial learnedness, sophistication, distancing and so on. It manifests itself in punning and other word play, metaphorical extension, willful error and duplication or usurpation of the role of an existing formation. These devices have been classified more systematically and exhaustively by Kuiper (this volume), and other functions proposed by Hohenhaus (also this volume).

Creativity is typically thought of as the act or quality of an unpredictable departure from the rules of regular word formation, and in poetic texts it can take surprising and thrillingly unexpected turns. To the long-term observer of newspaper text, however, the repertoire of journalistic creative devices is rather more familiar and restricted, probably due to the pressures on journalists to produce large amounts of eye-catching copy to deadlines (Renouf 2005a, b). However, one can still marvel at the ability of the writer to invent quite novel and allusive language, and of the reader to interpret it.

Different policies operate in the different newspaper publishers concerning the level of readers’ real-world knowledge and regularity of readership that should be assumed by the reporting and editorial staff, and these doubtless also have a consequence for the degree and type of lexical creativity employed, which would be a subject for further study.

Whether the full creative potential of a word or phrase is actually realised is another matter. Each word has its own frequency profile across time. Like *NIM-BY*, mentioned earlier, some remain stably rare. Other obsolescent or dormant words may suddenly find themselves revived and pressed into service (e.g. *poll-tax*, *cod-*; and *tsar*, which we shall come to later). Some new words emerge and

display a surge in popularity on the crest of a new media topic or vogue, receding from view once media interest has shifted.

In journalism, the time constraints on writing copy may not only formalise linguistic experimentation but also induce the writer to prefer certain words and avoid others. At one extreme, with a consistently low-ranking term like the acronym *NIMBY*, we find just one creative treatment in our data, in July, 2001 – the phrase *nimbys-come-lately* – in:

01.07 Mitchell of the Alliance strongly denies representing bourgeois *nimbys-come-lately* who don't want the hoi polloi lowering the tone¹⁵

referring to two Fulham football clubs, by analogy with the phrase 'Johnny-come-lately'. This is thin pickings for a lexical item so long in existence, considering that it did have some time in the sun.

It is clear that vogue terms, foregrounded by real-world events for whatever reason, will be in the front line for exploitation, as will words and phrases with inherent potential for allusion and punning. A highly topical or fashionable word or phrase is typically accompanied not just by a flurry of productivity but also of creativity.

6. Life-cycle

In order to demonstrate the evolution in the fortunes of a word or lexical unit, we take the noun phrase '*Weapons of Mass Destruction*'.¹⁶ This originally emerged in our data in 1989, in the context of the first Gulf War. However, it only really caught the media imagination in 2003, with the threat of a second Iraq war, as is shown by the 'time series plot' of its frequency pattern from 1989–2005, in Figure 1. The phrase is not specified for case, allowing upper and lower case counts to be conflated.

Figure 1 indicates frequency bulges in the early and late 90s, (represented as a dotted line for simple 'frequency per million words', and as a solid line for the more interpretable 'smoothed frequency'), as well as an enormous upsurge in the frequency line between 2002 and mid 2003, when the phrase was (re)introduced by the Bush administration. This pattern is followed by a slight downward trend, as the rate of occurrence slowed down again by the close of 2005.

15. *Nimbys-come-lately* is a pun on *Johnny-come-lately*, meaning 'upstart', first cited in the OED in 1839.

16. Some data for *Weapons of Mass Destruction* are also displayed on our web-site

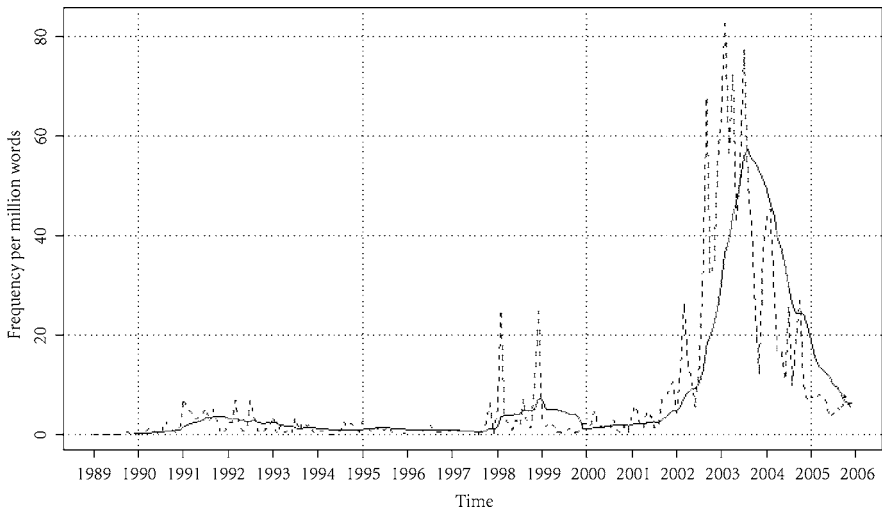


Figure 1. Time series plot: *weapons of mass destruction*, 1989–2005

The surge in frequency of the noun phrase ‘*Weapons of Mass Destruction*’ was accompanied by a surge in creative play with its lexical content. In order to discover more precisely what shape this new creativity took, we targeted the primary areas of creative potential within the phrase, namely the slots containing the words *mass* and *destruction*, and foregrounded these by the selection of 3 lexico-grammatical search frameworks, one [*weapons of *destruction*] another [*weapons of mass **], and the third [*weapons of * **]. A requirement for creativity was imposed by replacing the terms *mass* and *destruction* by wild-card open slots, in order to suppress the retrieval of the conventional phrase. The results obtained for each of these search patterns are presented in Table 8.

As the results in Table 8 show, on the same day that the fateful term *weapons of mass destruction* was newly uttered, Feb. 20, 2003, we already find quite a number of creative variants emerging. Some of these will have come from texts posted prior to the date, in fact, but it is still a remarkable crop.

Meanwhile, the noun phrase *Mass Destruction* is replaced by phrases which emphasise the seriousness of the issue through devices such as punning and bathos.

However, these results are relatively few when compared with the tremendous surge of additional creativity which occurred the very next day, as shown in Table 9.

Typically, in the life-cycle of a word or phrase in text, a rise in lexical creativity is followed by a fall. Accordingly, for *weapons of mass destruction*, in the time

Table 8. Variants of *weapons of mass destruction*, Feb. 20, 2003

weapons of mass * (no destruction)	weapons of * destruction (no mass)
weapons of mass distraction	weapons of modest destruction
weapons of mass obstruction	weapons of scant destruction
weapons of mass disruption	weapons of ass destruction
weapons of mass deception	weapons of indiscriminate destruction
weapons of mass discomfort	weapons of ministerial destruction
weapons of mass terror	weapons of market destruction
weapons of mass casualties	weapons of most destruction
weapons of mass entertainment	
	weapons of * (no mass, no destruction)
	weapons of bluff, brinkmanship & bloody-mindedness
	weapons of bumps and bruises
	weapons of nasty scratches

Table 9. Creativity for *weapons of mass destruction*, Feb. 21, 2003

weapons of mass * (no destruction)	weapons of * destruction (no mass)
weapons of mass banalization	weapons of catastrophic destruction
weapons of mass construction	weapons of class destruction
weapons of mass consumption	weapons of complete destruction
weapons of mass deconstruction	weapons of further destruction
weapons of mass distortion	weapons of global destruction
weapons of mass distractions	weapons of godlike destruction
weapons of mass effect (WME)	weapons of great destruction
weapons of mass hysteria	weapons of less destruction
weapons of mass instruction	weapons of local destruction
weapons of mass irritation	weapons of mad destruction
weapons of mass media (WMM)	weapons of massive destruction
weapons of mass partying	weapons of media destruction
weapons of mass protection	weapons of mischievous destruction
weapons of mass repression	weapons of murderous destruction
	weapons of nuclear destruction
	weapons of pattern destruction
	weapons of terrible destruction
	weapons of wealth destruction
	weapons of moose destruction

period following the initial linguistic and historic event, we see vastly reduced creative activity in web text (retrieved from the same number of web pages), as shown in Table 10.

Table 10. Decreased creativity of *weapons of mass destruction*, Feb. 28, 2002–May 29, 2004

weapons of mass * (no destruction)
weapons of mass photography
weapons of mass media (WMM)
weapons of * destruction (no mass)
weapons of 'grass destruction'
weapons of singing destruction
weapons of * (no mass, no destruction)
weapons of math instruction

7. Rules of creativity

Creativity, as seen in Figure 1, and Tables 9, 10, is shown not to be random but to follow a clear set of conventions, involving substitution on the basis of phonological, morphological, semantic and other types of similarity as well as allusion.

Destruction is replaced by words that have a

- difference of one phoneme – *distraction*
- same initial letter – *distraction, disruption, deception, discomfort*
- same prefix – *deception*
- same suffix – *distraction, obstruction, disruption, deception*
- shared semantics – *disruption, discomfort, terror, casualties*
- other – (for effect) e.g. *entertainment*

Mass is replaced by words that have a

- difference of one phoneme – *ass*
- same initial letter – *modest, ministerial, market, most*
- semantic similarity/contrast – *indiscriminate, most (vs. modest, scant)*
- other – (also collocate of *destruction*) – *indiscriminate*

The change in the meaning of a word, or the adoption of an additional meaning or metaphorical extension, is arguably a type of creativity. As we outlined in our initial section on Methodology, a change or extension in meaning is suggested to us by our automated software which signals a change in the collocational profile of a word or phrase. Novelty is in part a linguistic response to a real-world event, and the new use will be exploited in the media in creative play. For example, we would expect to find creativity in the wake of events such as the shift of primary reference of the acronym ASBO from organisations such as *Association of School Business Officials* to the newly-introduced *Anti-social Behaviour Order*.

8. Case studies of some recently popular neologisms in the UK media

Words which most clearly demonstrate such change patterns are those which are coined or revived in response to world events and topics that capture the public (and media) imagination for a period. One area of lexis of recurrent interest in British culture and reflected in UK journalism is the predilection for characterising sub-sets of UK society, through their clothing, behaviour or other attributes and associations. This preoccupation is peculiarly but not exclusively British; it is clear that this phenomenon occurs in other media and other societies, particularly in the US, whence many formations later make their way into UK text.

Some of these lexico-social attributions have been chosen from two currently topical and thus linguistically-active areas, namely:

- words for young people of low social status and life-style, e.g. *chav*, *neet*, and *hoodie*
- words for highly-paid people appointed to trouble-shoot in industry, government and other institutions, e.g. *tsar*

These words are new and possibly just at the start of a longer life-cycle. We shall focus on the salient productive and creative activity that each has produced in its relatively short life.

8.1 Neologism 1 – CHAV

The lexeme *chav* first emerges in our text in a series of dialectal uses in the nineties: *chavo* (1995), *chaveys* (1997), *chava* (1999) – where, like these, it is a Romany (and possibly also Essex rural) term for a young child. However, it is only in 2004 that we find the term revived in our data, this time in reference to a young British person of low education, having insufficient means to live away from home though sufficient to indulge in the purchase and wearing of hitherto socially-prestigious items of clothing, such as Burberry caps and other fashion accessories. The name gained currency swiftly through its publicity and incorporation into the name of the website, *chavscum.co.uk.*, a site pandering to, while ridiculing, the eponymous membership.

It is clear from our data that the base term *chav* itself emerged in 2004, at once rose precipitously to the top frequency ranks of voguish media vocabulary, and has remained there ever since. Table 11 reveals that the frequency was accompanied by quite a healthy level of productivity.

The range of variation in productivity based on the word *chav* in Table 11 is high, representing a considerable range of the canon of available suffixes and

Table 11. Inflections of *chav* emerging during 2004–2005

0403 chav 170	0502 chav-rap 1	0506 chav-nots 1
0403 chavs 95	0502 chavdom 1	0506 chavsta 1
0403 chavspotter 1	0502 chavalanche 4	0506 <i>chavtowns.co.uk</i> 2
0403 chavtastic 4	0502 chav's 3	0507 chavification 1
0412 chavvie 4	0503 chav-tastic 2	0508 chav-goth 1
0410 chav-spotting 1	0503 chav-in-the-street	0508 chav-infested 1
0410 <i>chavscum.co.uk</i> 1	0503 chavistocrats 1	0509 chavocrat 1
0412 chav-fave 1	0503 chavish 4	0510 chav-associated 1
0501 chavved-up 1	0503 chavvy 4	0510 chav-style 1
0502 chavocracy 1	0505 chav-central 1	0511 chavy 1
0502 chavier 1	0505 chavettes 1	
0502 chaviest 1	0506 chavette 1	

other post-positional combining elements. There is a striking predominance of *chav* compounding with colloquial or ‘youth’ terms like *spotter*, *spotting*, *fave* (‘favourite’), *rap* and *central*; we also find *chav* in apposition to *goth*; and *chav*-based derivations achieved through combination with less conventional affixes and affix-like elements.

The creativity associated with *chav* in its new sense is considerable, given its short life as compared with the 17-year-old acronym *NIMBY*. We find the following word play in *Independent* text from 2004:

Table 12. Creative variants based on *chav*, 2004–2005

1.	04.02 Chavs and chav-nots : the eternal divide. Charvers, skangers, scallies – Britain’s burgeoning peasant underclass may be revolting. But a threat to civilisation?
2.	05.02 A bickering chavocracy that marries in register offices should suit advocates of a bicycling monarchy
3.	05.03 the tulip-shaped skirt becomes de rigueur at Pangaea or wherever it is soap stars and chavistocrats hang out these days
4.	05.05 you would think that our society is on the brink of collapse from a chavalanche of scary teenagers.
5.	05.05 Vogue may be hoping to broaden their brand and attract a whole new readership but it seems pretty unlikely that the average chav-in-the-street is going to pick up a copy of Vogue at the local newsagent with their packet of fags
6.	05.06 After the Selfish Revolution in the Eighties, better-off people just shut themselves away behind their garden walls and moaned about the chav-nots .
7.	05.09 the polo-playing chavocrat who wears Nazi gear to a party and gets into a drunken brawl outside a nightclub
8.	05.10 It’s not simply a case of comedy for the chavs and chav-nots

It is plain from the word-play in Table 12 that this fertility results from *chav* being used to refer not just to the British social sub-class of poor youth, but also to the Royal Family and other rich and aristocratic members of British society, with the clear implication that the lives of the privileged are similarly tasteless and undistinguished. The neologistic derivations *chavocracy* (2), *chavistocrat* (3) and *chavocrat* (7) above actually refer respectively to Charles and Camilla, to rich people and to Prince Harry.

One of these puns, namely *chavs and chav-nots*, appears to have caught on particularly, by analogy with ‘the haves and the have-nots’, meaning ‘the rich and the poor’. Further instances of this phrase appeared in other broadsheet newspapers for the period, e.g:

- Times 05.02 **Chavs and chav-nots.** Julie Burchill says there is nothing wrong with chav culture
- Times 05.04 Sadly, in **the battle between the chavs and the chav-nots**, the ones with the bling-bling make the most noise.

8.2 Neologism 2 – *HOODIE*

We find a total of 159 occurrences of ‘*hoodie*’ in our news data between 1989 and end 2005,¹⁷ and the evolution in its referential and connotative meaning is exemplified in Table 13. Its creativity lies primarily in the use of metonymic extension and allusive reference. It first occurs in 1992 as a colloquial term for a ‘hooded crow’ (1). By Dec. 2001, however, *hoodie* was adopted to signify metonymically the hooded jacket or top made of soft material primarily associated with particular groups of society. Initially, a *hoodie* was just the preferred clothing for boys, young men and their peer groups (2), before it moved on to refer to an identification with fashion and youth icons in June 2002 (3).

Thereafter it was used to refer to adult fashion, as in (4) above (Oct. 2002); subsequently restricted to young female adult fashion but in the context of the need for warmth (see (5), Oct. 2003). It was not until Jan. 2004 (6) that the term *hoodie* was associated with the social underclass, even though it also continued to be used in reference to fashion items associated with sport, musical performers, ghetto stars, boy singers and music festivals. In this early period, only one signifi-

17. Available in the US folk lexicon since the 1870s, semantically derivable from ‘*hoodlum*’ (etymology unknown), shortened form ‘hood’, specifically associated with organised crime, the term waited until a century later to appear in our UK data, and even then, it is first used in reference to the US (see *hoodies*). Some theories also associate it with *neighbourhood*.

Table 13. Evolution of *hoodie* references in *Independent/Guardian* data 1989–2005

1.	92.09 Here live only deer and eagles, the odd fox and hoodie crows.’
2.	01.12 He buys a black ‘ hoodie ’ with ‘Alien Ant Farm’ on the back
3.	02.06 his progress is tracked by an army of hoodie-wearing , baggy-jeaned pre-teens. ‘It’s Matt Hoffman’, they whisper reverently.
4.	02.10 you can’t buy a decent hoodie from the Issey Misake or Jean Paul Gaultier sites.
5.	03.10 Small bags. Impossible in the winter to tote round a zip-up hoodie , extra vest...
6.	04.01 But Waynetta Slob [TV character] wore matching jogging pants and hoodie , if you catch my drift.
7.	04.11 How would anyone breaking a hoodie Asbo be caught? The hood turns any gentleman into a hoodlum.

Table 14. Collocational pattern for *hoodie* in Nov. 2004

wearing	4470	5+0	13.707
cap	1712	1+9	8.922
teenager	2109	5+2	5.930
ban	5974	1+5	4.829
baseball	774	0+5	3.981
wore	1298	5+0	3.968
black	13345	4+1	3.679
Bluewater	127	2+2	2.997
jeans	1078	1+3	2.978
debate	8248	0+4	2.832

cant collocat for hoodie was generated from its few occurrences by our analytical software, namely crows.

From June 2004, however, *hoodie* increasingly functioned as a negative evaluation of youth fashion sense, and in Nov. 2004, it adopted its first association with anti-social behaviour (7). From Nov. 2004, the negative sociological connotation of the term can be calculated by our analytical software. Its collocational pattern significantly increased and changed – no longer *crows*, but a list of collocates redolent of its new meaning, as shown in Table 14. The accompanying figures represent respectively ‘corpus frequency’, ‘plus/minus 1 span frequency’ and ‘collocate z-score’, three statistical measures which combine to rank the collocates for significance.

Similarly, until Feb, 2005, the plural form *hoodies* was too infrequent to have any statistically significant collocates. But thereafter, there is a stretch of increased frequency of occurrence as follows:

05.05	>	48
05.06	>	32
05.07	>	13
05.08	>	8
05.09	>	24
05.10	>	6
05.11	<	26

which is accompanied by an upsurge in significant collocates, revealing a clear change in connotation, as can be seen in Table 15, with new items such as *ban*, *banned* and *yobs*.

The first outing for the doubly metonymic use of the term *hoodie* was in May 2005. The meaning of the term was transferred from the hood to the jacket and now to the wearer of the jacket. It was by then also imbued with the sense of the wearer being socially undesirable. In the example in question, we furthermore see the term used allusively, applied not to the usual hooded skateboarder or footballer, but to a tennis player.

05.05 As a result, Mancini is now officially as welcome in Genoa as a hoodie in the home counties

The neologisms *chav* and *hoodie* each evaluate and to some extent caricature members of the social underclass by reference to their choice of clothing, assigning respectively the vices of poor taste and fecklessness, and lack of social responsibility and petty criminality. They conjure up amusing images, have great metaphorical application, and allow one section of society to make jokes about and feel superior to another. They are thus ready candidates for use by the public, and not simply by the media, and probably for this reason have become so productive.

Table 15. Collocational profile for *hoodies* in Feb. 2005

wearing	3319	24+0	22.439
caps	799	1+8	7.945
ban	4090	8+1	7.724
baseball	586	2+6	6.964
boys	4367	7+1	6.735
wear	2700	6+1	5.853
banned	2370	4+2	4.887
young	19320	4+2	4.149
yobs	168	3+2	3.993
youths	606	4+1	3.975
teenagers	1773	5+0	3.927
shopping	2241	3+2	3.908
people	79115	6+0	2.262

8.3 Neologism 3 – *NEET*

By contrast, the acronym *NEET*, (standing for ‘not in education, employment or training’) was coined not by the media but by the government Dept. of Health to classify still another social sub-set of British youth. Constant changes in the educational and social service provision in the UK lead to such a plethora of terms that acronyms are widely used to condense information. *NEET* is a bureaucratic term, neither particularly amusing nor readily exploitable by the public or even the media, so it is less productive than *chav* or *hoodie*. Until 2004, the word *neet* was simply a transliteration of the regional dialectal variant for ‘night’. We find in our news data six such instances, e.g (1) in Table 16. Towards the end of 2004, the new acronym *NEET* emerged (2–4). As in the earlier cited case of *NIMBY*, *NEET* immediately moved to lower-case orthographic status, *neet*, though again this was not the consequence of widespread currency (Renouf 1993a). The only instances of productivity (5–6) show a minor syntactic shift, from noun to noun modifier.

With the meagre opportunity for creative use afforded by *NEET*, there has been just one single wordplay, as shown in Table 17, a pun trading on the somewhat marginal phonological similarity between *neet* and the word ‘*nit*’, meaning ‘louse’.

Table 16. Instances of *neet* since late 2004

1.	02.01	I’ve taken part in many a merry neet with hunting folk
2.	04.10	The health ministry estimates over 500,000 so-called neets
3.	04.11	Japanese researchers know their nation has neets too.
4.	04.11	The no. of young Japanese neets is an estimated 600,000
5.	04.11	The neet group in Britain tends to be from the lower socio- economic groups
6.	04.11	The neet phenomenon may be fundamentally different in both countries

Table 17. Creativity of acronym *neet* since its inception in late 2004

04.11	For most people, the suggestion that a sizeable number of British teenagers has a neet problem will prompt images of strong hair lotions and fine-tooth combs
-------	--

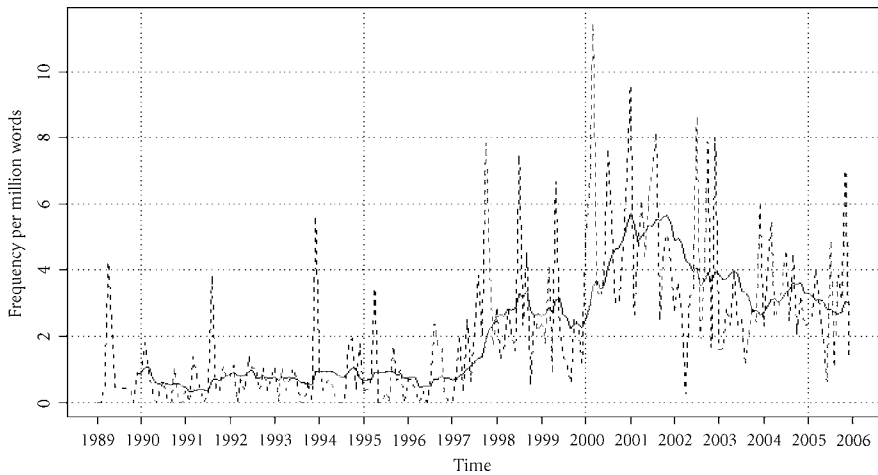


Figure 2. Time series frequency plot for *tsar*, 1989–2005

8.4 Existing term with new meaning – *TSAR*

Many long-established words can lie fallow until a new use is found for them. The word *tsar* is an example.¹⁸ This used to have a fairly low-key but consistent life, bumping along in the bottom reaches of the lexicon, chiefly in relation to events in the Soviet Union.

It has become one of many similar terms – *guru*, *supremo*, *pundit*, *troubleshooter* and so on – which have in recent years come back into fashion, used to characterise the New Labour government cronies, media personalities and other personages who oversee the functioning of a public institution or commercial body, or who are experts in some field. The terms often even appear in juxtaposition:

99.01 it is impossible not to trip over a **Government-appointed tsar, task force, regulator, or watchdog** surveying our lives

00.07 the arts council **supremo / guru / tsar** Gerry Robinson

Among these, *tsar* and *pundit* are the most productive and creative, probably due to the linguistic accident that they are more easily manipulated and lend themselves to affixation more readily than their synonyms. Since *tsar* is the more in vogue of the two, we shall select this as our object of investigation.

Our 'time series' graph of word frequency in Figure 2 shows that in recent history, specifically from 1989, this term has hovered in the lower reaches of the

18. What is said about the creativity of *tsar* does not hold for its spelling variant *czar*.

Table 18. *Tsar* and its inflections, 1989–2005

anti-tsarist	2	tsar's	94	tsarina's	3
anti-tsarists	1	tsar-gazing	1	tsarinas	10
co-tsar	1	tsar-hating	1	tsarism	17
e-tsar	1	tsar-in-waiting	1	tsarism's	2
ex-tsar	1	tsar-like	3	tsarist	234
mini-tsar	2	tsar-wars	1	tsarist-era	8
neo-tsarist	1	tsardom	9	tsarists	4
pro-tsarist	1	tsarevich	7	tsaritsa	1
super-tsar	1	tsarevna	1	tsaritsas	1
tsar	1702	tsarina	41	tsars	308

lexicon, initially restricted in reference to the former Soviet Union, to its imperial heritage and its past and present political leaders. As in Figure 1 for *Weapons of Mass Destruction*, the frequency pattern is represented as a dotted line for simple ‘frequency per million words’, and as a solid line for the more interpretable ‘smoothed frequency’.

The long, steady life-cycle of the root form *tsar* has been shared with long-established inflections such as *tsarist*, *tsaritsa/s*, mostly transliterations of Russian titles used with literal reference. It can be seen in Table 18 that the bulk of productivity has been rather conventional derivation, with a few more recent exceptions such as *e-tsar*, and a rash of creative wordplay, chiefly punning on *tsar/star*, which we shall consider in more detail later.

The frequency graph in Figure 2 shows that the stably low frequency of the word *tsar* suddenly rose steeply in our text from 1996, peaking in 2000 and 2001. The peaks corresponded to periods when it was the fashion for the Labour Party and other institutions to appoint public figures to oversee most aspects of national life. This historical context has had a particular effect on the productivity pattern for *tsar*.

Statistically significant collocates positioned immediately to the left of *tsar* in our data from 1996–2005 are ranked as in Table 19. What this tells us is that, from about 2000, the productivity of *tsar* primarily takes the form of a multitude of compounds of the ‘government’s transport *tsar*’ kind, with *tsar* as an all-purpose noun-head carrying a range of noun-modifiers. Together, these form designated roles and titles specifying a brand of advisor, usually in the public eye, highly-placed and highly paid, but exceptionally honorary, whose role is to monitor areas of practice and expenditure and solve problems to try to improve performance in different social arenas, from government to sports. The term *tsar* seems to impart a certain glamour and power to these roles.

In contrast, in Table 20, the collocational partners of members of the same semantic class – *guru*, *supremo* and *pundit* – whilst also designating assorted ex-

Table 19. Significant left-hand collocates for *tsar*, 1996–2005

drugs	yob	sports	anti-corruption
drug	intelligence	economic	Drug
health	behaviour	car	disability
cancer	patient	anti-drugs	Cancer
homelessness	digital	ethnicity	Children's
Drugs	housing	equality	Crime
Russian	heart	economics	healthcare
care	maths	design	countryside
children's	IT	competition	Food
counter-terrorism	transport	policy	poverty
food	homelessness	Asbo	emergency
people's	architecture	film	arts
crime	rural	ex-drugs	TV
security	rail	anti-drug	education

Table 20. Significant left-hand collocates for *guru*, *supremo* and *pundit*, 1996–2005

guru	supremo	pundit
PR	PR	TV
management	policy	television
style	security	football
lifestyle	election	media
advertising	transport	political
design	marketing	racing
investment	media	soccer
fashion	press	radio
marketing	arts	Sky
Indian	communications	sports

perts, bosses and top dogs, do not imply the same degree of national or governmental status or authority as do the *tsar* compounds.

Many compounds of *tsar* are longer multi-word units than those suggested by the collocational profile. Examples are shown in Table 21.

Moreover, the de-contextualised listing of significant collocates in Table 21 largely obscures the many humorous and parodic coinages. A sample of the effects achieved, both consciously and sub-consciously, is catalogued under the following headings:

– **compounds which humorously juxtapose colloquial words and basic concepts with the elevated *tsar***

- 97.10 South Boston has a '**trash tsar**' – in charge of street cleaning
- 97.10 **Bug tsar** appointed to stop millennium computer crash
- 97.10 **Rat tsar** needed to crack down on the booming population of rodents

Table 21. Multi-word compounds based on *tsar*, 2000–2005

	arms control tsar
	EU foreign policy tsar
	the treasury’s banking competition tsar
	the government’s workplace equality tsar
	the European Commission’s competition tsar
	hospital design tsar
	digital television tsar
	learning disability tsar
	The NHS information technology tsar
	the government’s university admissions tsar
	the first criminal justice IT tsar

– compounds which label ludicrous official roles:

(passim) ‘NHS **waiting lists tsar**’; ‘**waiting list tsar**’

– compounds which apply ludicrous epithets to official roles

- 01.07 Jean Glavany, the French agriculture minister and **tree tsar**
- 03.11 Every NHS trust is to appoint a **bug-buster tsar**

– compounds which humorously paraphrase official terms

homelessness tsar

- 98.07 the ‘**streets tsar**’
- 00.12 The government, **roofless tsar** and all
- 01.04 **Rough sleepers tsar** sweeps ahead with targets

older people’s tsar

- 98.12 Call for ‘**grey tsar**’ to champion the elderly
- 05.11 The **old-age tsar**, Professor Ian Philp

de-regulation tsar

- 00.12 The government’s **red tape tsar**, Lord Haskins
- 03.03 Lord Haskins is the **multi-purpose regulation/farming/euro tsar**

food tsar

- 00.09 Loyd Grossman to be hospital ‘**foodie tsar**’

– compounds which parody the official use of *tsar* in miscellaneous journalistic ways

These generally involve making fun of public figures by characterising them ironically as *tsars* – distinguished and powerful – while simultaneously highlighting

their insignificance, triviality or other shortcomings, particularly arrogance. See for example:

- 99.03 Milton Hershey had become America's **caramel tsar**
- 00.08 a friend of lipless, **rightwing morals tsar** William Bennett
- 00.12 Labour's watchdog; a kind of **anti-William Hague tsar**
- 01.03 the **after-dinner tsar** Kenneth Clarke

– **compounds as part of everyday journalese**

- 99.01 The Foreign Office security consultant, or '**kidnap tsar**'
- 01.10 sources close to Radio 5 Live's **phone-in tsar**
- 01.10 In April, he appointed a gay activist as **Aids tsar**

Our selection of the term *tsar* from within its semantic group is not accidental; the term is fertile ground for the study of pun creation. Some words lend themselves more readily to wordplay than others: *guru*, for example, is less ambiguous, phonetically or otherwise, and so inherently less exploitable. For *tsar*, we find that the phonetic similarity between *tsar* and *star* is heavily traded upon. Indeed, on some occasions it is almost impossible to avoid punning, or to know whether it is a deliberate device. See, for instance, the marginal case:

- 11.05 Learning disability tsar Rob Greig revealed that ministers were to appoint a learning disabled **co-tsar**

Table 22 (below) displays some of the deliberate creativity found in our corpus.

This shows that over the years, virtually every phrase and cliché involving the word *star* has been exploited, from *star-gazing*, to *star quality*, *rock star*, *rising star*, *porn star*, *a galaxy of stars*, *pop star*, *super-star*, *wandering star*, *star attraction*, *stars in one's eyes* and *star turn*. In addition, many of the puns have been wrung from song lyrics such as '(Catch) a Falling Star' and 'When you wish upon a star', and film titles such as 'Star Wars' and 'a Star is born'. Some allusions are more loosely made; it is not in fact 'a star' which *risks in the east*, but 'the sun'. It is clear that these puns have been strategically created to add a humorous or allusive quality to the writing, and are manipulated to fit the new contexts semantically. That the majority function as titles or sentence openers indicates their purpose as awakeners of reader interest.

Table 22. Puns on the vogue word *tsar*, 1996–2005

03.89	History has consigned previous US tsar-wars to oblivion
03.91	exhibition ‘Twilight of the Tsars’ no chance for tsar-gazing
05.92	he seems to have seen tsar quality in him
03. 98	Nowadays he’d be a rock tsar
05.99	The falling tsar . Yeltsin brooks no rivals at court.
05.99	In Cannes, a tsar is born : The Russians are coming to the film festival
09.99	A tsar rises in the east : Helen Wallace on classical music
10.99	Spoken like a rising tsar
03.00	Porn tsar to save Utah from sex
08.00	Managed by New Labour pop tsar Alan McGee
10.00	Putin piles on super-tsar pretensions
04.01	When you wish upon a tsar . Lithuanian magnate opens theme park.
01.02	A wandering tsar blesses union between human and alien
10.02	an army of unelected policy advisers - Blair’s galaxy of tsars .
02.03	New Tsar is born as artist gives facelift to regal past.
10.03	A sprinkling of tsar quality . Sarah Eldridge meets the man
06.04	A tsar is born : Prince Michael of Kent a cult hero in Russia.
03.05	rise and fall of a pop tsar – pop’s most colourful character.
06.05	Bulgaria’s falling tsar Simeon Saxe-Coburg faces bleak exit.
08.05	Tsar attractions . From couture to high street, Moscow style everywhere
11.05	Tsar turn : our rag trade king rocks Russia.
11.05	Government has tsars in its eyes . Learning disabled tsar to be appointed

9. Conclusion

In this chapter we have sought to make explicit some of the insights about the nature of lexical productivity and creativity in text by employing a research methodology designed to trace lexical activity over time as it is used in UK journalistic text. We have focussed on key concepts of ‘productivity’ and ‘creativity’, which we have defined within the terms of modern diachronic corpus study. We have conducted a modest case study of selected neologisms, and within the limitations of their relatively short life-spans and correspondingly limited occurrences, we have tried to illustrate their patterns of use.

Stepping back from this catalogue of minutiae about the individual words, one is able to make a few more general and hopefully thought-provoking observations, which may encourage the reader to revisit the data here and elsewhere in a fresh light. In tracing productivity and creativity chronologically over 16 years of journalistic text, one is struck by a number of findings which emerge from this methodological approach.

Probably the central revelation is that words have a life-cycle consisting, in the most general terms, of birth or re-birth, followed by gentle or steeper upward trajectories in frequency of use and leading to brief or lengthier moments at the zenith of popularity, after which they take faster or slower downward paths, until they reach a stable level of use. During this life-cycle, words which make a sufficient impression on the public imagination will also spawn a number of productive and creative variants. Indulging in hyperbole for an instant, this is evidence of the marvellous self-generating and self-renewing power of language, words like stars soaring into the linguistic firmament in a blaze of popularity, spinning other words off hither and thither, and at length tumbling back down into the calm(er) reaches of servitude.

A related observation is that, for various reasons, not all words are equally destined for success. Some inhibitors have been suggested; another factor is that words do not all have universal franchise. Some stem from grass-roots movements in society, some from developments at more rarified social or scientific levels; and some are media coinages. The ability of a word to establish itself as an icon which furthers the media cause seems to be the most crucial criterion for its securing a place in the lexicon of the press. The acronym *NIMBY* is a political term for a somewhat drab concept, and little used by the public or journalists. In contrast, a phrase like '*annus horribilis*', coined by the Queen's speech writer in 1992, whilst opaque to the majority of the British public, has fared better. Though foreign to many in meaning, etymology, provenance (from *annus mirabilis*¹⁹), and stylistic intention, its meaning is still guessable and so of service to journalists. Thus, following a brief if heady ascent and a fairly swift decline, it has settled at a stably low but resilient level of productivity, where it seems to be come in handy for journalists in their characterisation of key events at the close of each year. Our statistics on its occurrence indicate around double its average number of monthly appearances between November and January, assigning an overall 'seasonality' score of '.04', which means that it is less seasonal than the terms *Wimbledon* or *Christmas*, each with top scores of 0, but still quite significant.

19. Dryden (1667) wrote a poem entitled '*Annus Mirabilis: the year of wonders 1666*', in which he opines that the 1665 plague and 1666 fire of London could have been worse, in part no doubt because he himself survived both.

A final observation is that our diachronic, empirical study of the data reveals and traces parallels between language and contemporary world events, and thus acts as a window on contemporary culture; in particular on recent world events and on aspects of youth culture as reflected in the media. These data and these findings raise issues which tip over into neighbouring disciplines: sociolinguistic questions, about the circumstances of language use and the fortunes of words; psycholinguistic questions, about the processes underlying the subtleties of mutual comprehension of allusion and wordplay; stylistic questions, about the preference for certain choices and not others, and so on. It is hoped that these concluding thoughts may inspire the reader to investigate such mysteries further.²⁰

References

- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Baayen, R. H. and Renouf, A. J. 1996. 'Chronicling the Times: Productive Lexical Innovations in an English Newspapers', in *Language*, 72.1: 69–96.
- Kehoe, A. & Renouf, A. J. 2002. *WebCorp: Applying the Web to Linguistics and Linguistics to the Web*. World Wide Web 2002 Conference, Honolulu, Hawaii, 7–11 May 2002.
- Kehoe, A. 2006. 'Diachronic Linguistic Analysis on the Web using WebCorp', in *The Changing Face of Corpus Linguistics*, A. J. Renouf and A. Kehoe (eds), 297–308. Amsterdam & New York: Rodopi.
- Morley, B. 2006. 'WebCorp: A Tool for Online Linguistic Information Retrieval and Analysis', in A. J. Renouf and A. Kehoe, 283–296.
- Morley, B. and Renouf, A. J. 2003. 'Linguistic Research with the XML/RDF aware WebCorp Tool', in *Proceedings of WWW2003, Budapest*. <http://www2003.org/cdrom/papers/poster/p005/p5-morley.html>
- Pacey, M., Collier, A. and Renouf, A. J. 1998. 'Refining the Automatic Identification of Conceptual Relations in Large-scale Corpora', in *Proceedings of Sixth Workshop on Very Large Corpora, Montreal, 15–16 August 1998*, COLING-ACL, E. Charniak (ed.), 76–84.
- Renouf, A. and Sinclair, J. 1991. 'Collocational frameworks in English', in *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, K. Aijmer and B. Altenberg (eds), 128–143. Harlow: Longman.
- Renouf, A. J. 1993a. 'H.E.R.E. Today and 'here' tomorrow', unpublished paper delivered at 14th ICAME Conference, University of Zurich, Switzerland.
- Renouf, A. J. 1993b. 'A Word in Time: first findings from dynamic corpus investigation', in *English Language Corpora: Design, Analysis and Exploitation*, J. Aarts, P. de Haan and N. Oostdijk (eds), 279–288. Amsterdam and New York: Rodopi.
- Renouf, A. J. 1993c. 'Sticking to the text: a corpus linguist's view of language', in *ASLIB Proceedings, Volume 45*, 131–136.

20. Researchers interested in working with our data are welcome to access, occasionally via password, any of our systems at <http://rduces.uce.ac.uk/>

- Renouf, A. J. 1993d. 'What the Linguist has to say to the Information Scientist', in *The Journal of Document and Text Management*, vol. 1/2, F. Gibb (ed.), 173–190.
- Renouf, A. J. 1993e. 'Making Sense of Text: Automated Approaches to Meaning Extraction', in *Proceedings of 17th International Online Information Meeting*, 7–9 December 1993, 77–86.
- Renouf, A. J. 1994. 'Corpora and Historical Dictionaries', in *Early Dictionary Databases*, I. Lancashire and T. Wooldridge Russon (eds), 219–223. University of Toronto, 1–8 Oct. 1993.
- Renouf, A. J. 1996. 'The ACRONYM Project: Discovering the Textual Thesaurus', in *Papers from English Language Research on Computerized Corpora (ICAME 16)*, I. Lancashire, C. Meyer and C. Percy (eds), 171–187. Amsterdam & New York: Rodopi.
- Renouf, A. J. 1997. 'Tools for the Diachronic Study of Historical Corpora', in *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*, 'Mémoires de la Société Néophilologique de Helsinki', vol. 52, T. Nevalainen (ed.), 185–199.
- Renouf, A. J. and Baayen, R. H. 1998. 'Aviating among the hapax legomena: morphological grammaticalisation in current British Newspaper English', in *Explorations of Corpus Linguistics: Proceedings of the 18th ICAME Conference*, A. Renouf (ed.), 181–189. University of Liverpool, 21–25 May 1997.
- Renouf, A. J. 2003. 'WebCorp: providing a renewable data source for corpus linguists', in *Extending the scope of corpus-based research: new applications, new challenges*, S. Granger and S. Petch-Tyson (eds), 39–58. Amsterdam & New York: Rodopi.
- Renouf, A. J., Kehoe, A. and Mezquiriz, D. 2003. 'The Accidental Corpus: issues involved in extracting linguistic information from the Web', in *Proceedings of 21st ICAME Conference, University of Gothenburg, May 22–26 2002*, K. Aijmer and B. Altenberg (eds), 404–419. Amsterdam & New York: Rodopi.
- Renouf, A. J. 2005. 'Phrasal creativity viewed from an IT perspective', in *RANAM (Recherches Anglaises et Nord Américaines) n°38: Language chunks and linguistic units*, A. Hamm (ed.), 113–122. Université Marc Bloch, Strasbourg.
- Renouf, A. J., Kehoe, A. and Banerjee J. 2006. 'The WebCorp Search Engine: A holistic approach to web text search', in *Proceedings from the Corpus Linguistics Conference Series*, Vol. 1, no. 1, www.corpus.bham.ac.uk/PCLC
- Renouf, A. J. and Kehoe, A. (eds) 2006. *The Changing Face of Corpus Linguistics*. Amsterdam & New York: Rodopi.
- Renouf, A. J., Kehoe, A. and Banerjee, J. 2007. 'WebCorp: an integrated system for web text search', in *Corpus Linguistics and the Web*, C. Nesselhauf and M. Hundt (eds), 47–68. Amsterdam & New York: Rodopi.

Lexical creativity in texts

Cartoon art

Cathy Wilcox meets the phrasal lexicon

Creative deformation of phrasal lexical items for humorous effect*

Koenraad Kuiper

Phrasal lexical items, since they are phrases, may exhibit a range of idiosyncrasies which single word lexical items cannot (Burger 2003; Moon 1998; Wray 2002). These idiosyncrasies will be outlined in this chapter because they provide resources for artistic deformation (Gläser 1995; Naciscione 2001). Such artistic deformation will be illustrated by reference to the work of the gifted cartoonist, Cathy Wilcox of the *Sydney Morning Herald*. Humour, it may be supposed, can result from the cognitive dissonance arising from a disjunction between the expected conventional properties which a PLI has in its lexical entry and which are accessed when the item is recognised, and the unexpected twist given to one or more of such properties by the cartoonist.

‘Conventional, formulaic, and memorised expressions of all kinds play a large role in language play and linguistic creativity. The innumerable phrases known to fellow speakers are alluded to, toyed with, and varied just a little for emphasis and effect. Such phrases arise within families, organizations and groups of all kinds. Humorists and cartoonists draw on them constantly and journalists rely on them as “eye catchers” in their story titles.’
(Van Lancker 1987: 100)

1. Introduction

Phrasal lexical items (PLIs) are lexical items which have phrase structure, as contrasted with single word lexical items which have word structure. For example,

* I am grateful to Cathy Wilcox for permission to reproduce her work in this chapter and for the pleasure that her work has provided. I am also grateful to Andrew Biddington with whom I have discussed these cartoons and to John Paolillo for his paper on Gary Larson’s cartoons (Paolillo 1998) which takes a different tack but with similar material.

dog-eared is a compound adjective and thus a word while *a dog's breakfast* is a lexicalised phrase. Like all lexical items, PLIs are coined at a particular point in time and may come to be lexicalised when some use is found for them. It might be that they express a complex predicate such as, for example, *make allowances for*, or that they perform a useful social role, such as, for example, *Make sure you ...*. Unlike single word coinages, however, potential PLI coinages occur almost every time someone constructs a new phrase or clause since any phrase which is uttered is potentially a new PLI.¹ Creativity in terms of the coinage of new PLIs is, therefore, hard to study. As Noam Chomsky has tirelessly pointed out, speakers are being creative every time they put a new phrase or sentence together. However, that creativity is not in itself intentional.

If we make a distinction between intentional and unintentional creativity then there are certainly PLI coinages which are intended to be creative. In domains such as new product onomastics or slogans and pop song lyrics it may be possible to discern an intent to be creative. Frequently, this is manifest in the coined word or phrase(s) being repeated so many times by its creator(s) that it becomes lexicalised, sometimes reposing in our mental lexica 'like a hectic'. John Lennon's *Give peace a chance* and McDonalds' *You deserve a break today* or *I'm lovin' it* come, regrettably, to mind.

In this chapter another form of lexical creativity will be explored, namely the creativity which takes an existing lexical item and performs an operation on it which is not conventionally permitted. Lexical items, since they are stored and retrieved from the mental lexicon, generate stable expectations on the part of their users as to what may conventionally be done with them. *Finnegans Wake* shows that these stable expectations may be tampered with in a variety of ways. Such word play involves a form of intentional lexical creativity.

The approach to PLIs adopted here is thus primarily psycholinguistic because PLIs are stored and retrieved from speakers' mental lexica. Whatever properties PLIs have, they have on that account. Whatever properties are revealed by their use in texts result from the interplay between the items accessed from the mental lexicon and speech production or perception processes in context. The mental representation of PLIs thus has primacy. In production and perception, where the standard form PLI is used as a template for modification, that template is the set of stable expectations against which lexical play is to be seen.

1. I ignore, for the purposes of this chapter, the possibility that PLIs are subject to constraints on their structural representations (Kuiper and Everaert 2000).

2. Potential locations for word play in PLIs

PLIs have potential idiosyncrasy in a number of areas of their representations. We could see these as rigidities of various kinds and thus potentially a source of 'artistic deformation' (Melčuk 1995: 213–4).² Such deformation is local in space and time. Therefore PLI deformation has a unique creative aspect. Being local it is normally context dependent. For example, subeditors who create headlines based on PLIs normally use the particular changes once only.

The approach to be taken here is to enumerate the potentially idiosyncratic properties of PLIs and then to see which of them are further available for humorous manipulation, after which we will see how one cartoonist actually does manipulate them. This is thus a theory-driven approach as compared with studies which provide taxonomies of humour based on the study of examples (Alexander 1997; Chiaro 1992).

In order to drive our study from its origins in the linguistic resources of the phrasal lexicon, it is necessary first to understand what the relevant properties of PLIs are. It needs to be supposed, in the following discussion, that each PLI is a lexical item with its own entry in the mental lexicon of a speaker who knows it. Such knowledge is potentially complex. Furthermore, not all native speakers of a language who know a particular PLI will necessarily know it in exactly the same way, i.e. some will know it as having properties which other speakers will not (Fraser 1970: 23; Melčuk 1995: 171); but for every speaker there must be the mental equivalent of what Naciscione (2001: 20) terms, the PLI's *base form*. Without such a base form a number of kinds of creative lexical play could not be recognised. That they are, corroborates the claim that PLIs have base forms.

This leads to two more observations. The first is that a speaker must know the PLI in order to be able to perceive artistic deformation of it. That seems obvious but it needs saying. All intentional deformation has as its intent that a perceiver perceive the difference(s) between the base form and its variant. That can only happen if the perceiver knows the PLI and its base form and can therefore access it from their mental lexicon. For this reason many second language learners cannot perceive this kind of lexical play since they do not know the PLIs on which it is based. I shall term this *the accessibility condition*.

The second observation is that, if the base form of a PLI is to be recognisable perceptually, then any artistic deformation of it must be such that a perceiver who knows the base form can retrieve it from his or her mental lexicon. To put it dif-

2. See Naciscione (2001) for a discussion of various approaches to such manipulation.

ferently, after artistic deformation there must remain sufficient perceptual cues to allow the PLI to be accessed. I shall term this *the recoverability condition*.

2.1 Idiosyncratic properties of PLIs

To see what those perceptual cues might be requires a listing of the idiosyncratic properties by which PLIs are recognised and which form part of their mental representation.

PLIs are phrases of which speakers know at least two of the words. We can term these words **lexicalised constituents** and define such a constituent as one where the lexical content of the constituent is given in the lexical entry of the PLI (Verstraten 1992). For example, in the *let alone* construction discussed by Fillmore, Kay, & O'Connor (1988) the words *let alone* are lexicalised constituents of the construction which I take to be a PLI. Testing for this property is often done by substituting for a lexicalised constituent (Gläser 1986:20, 1995; Verstraten 1992:27). If the substitution results in the phrase no longer being recognisable as a PLI then this shows the original phrase was a PLI.

PLIs may also contain a **bound word** (Aronoff 1976). These are single words which occur only within a PLI. For example *take umbrage at* contains the word *umbrage* which cannot occur freely, and occurs in no other PLI. In German these are termed *unikale Komponenten* (Soehn 2003).

Since PLIs are phrases where not every word must needs be specified, PLIs can contain **slots** (Koopman and Sportiche 1991; Williams 1994). A slot in the syntactic representation of a PLI is a position in the structure which requires to be 'filled' with other lexical items but which is not filled in the representation of the item in the lexicon. For example, in the PLI *take NP to task*, the NP is an obligatory complement of the verb which must be filled for the phrase to be used grammatically but the lexical content of the NP is not given in the lexical entry of the PLI.

Sometimes slots also have an additional **slot restriction**. While the syntactic category of a slot constrains what it may contain syntactically, there are frequently other constraints of an arbitrary or semantic kind. For example, some slots must be filled with animate or human NPs when that is not an inherent requirement of the verb of which the NP is a complement, i.e. not the result of the s selection properties of the verb (Chomsky 1996:54). Again examples of slot restrictions are given in Fillmore et al. (1988) for the *let alone* construction.

Some PLIs have **optional constituents** which may or may not be used. They are part of what the speaker knows when (s)he knows the PLI but their use is optional. For example, in the English PLI *breathe one's last breath* the final noun

is optional; speakers can and do just say *breathe one's last*. Note that optional constituents are not just adjuncts which may be added freely. The form of words is particular and is part of what native speakers know of the particular PLI.

In some PLIs there appears to be more than one lexical item functioning in the same position. *To be in a bad mood* is equivalent to *being in a bad temper*. It seems that *mood* and *temper* function as alternative last nouns in this PLI. But there are no other possible nouns here that are 'known' as part of knowing the PLI. These two thus constitute a **selection set**. Selection sets only occur where the PLI is semantically and pragmatically equivalent regardless of which member of the set is used.

Some PLIs will take freely inserted adjunct constituents. Others will not. This can be termed the PLI's **modifiability** (Nicolas 1995). For example, one can *get annoyed* or *get very annoyed* but one cannot modify the dismissive PLI *Get lost!* to *Get very lost!*

PLIs have greater or lesser degrees of syntactic **flexibility** under movement, supposing a theory of syntax which allows movement. Classically the PLI *kick the bucket* will not passivise (Nunberg, Sag, & Wasow 1994).

Restricted collocations occur (Mackin 1978; Melčuk 1998). For example, if one wishes to use a bus as a means of public transport, one is said to *catch the bus* and then *get on the bus*. One does not trap the bus or get in the bus. Restricted collocations involve preferential selection of word combinations where such combinations are arbitrary. They may also be idiomatic, i.e. not semantically compositional. Catching the bus is, in some sense, idiomatic but getting on the bus could be seen quite literally to be placing one's feet on the floor of the bus or oneself on its seats. Wine is (classically) either white or red; it is not ever purple or light green regardless of the truth conditions of these colours in collocation with *wine*. *To the best of one's abilities* is what English speakers say rather than *at the best of one's abilities*. In terms of their semantic properties neither preposition is preferable. Both create semantically well-formed and appropriate compositional meanings in this construction. Yet one is lexicalised as a restricted collocation. The other is not.

If the meaning of the whole PLI is a compositional function of the meaning of its constituent parts then it is fully compositional. Thus PLIs with this property will have all the possible meanings available from the semantic interpretation of the senses of their constituents. For example, the checkout farewell, *have a nice day* is fully compositional but is a PLI.

A lexical item which is non-compositional in its meaning, i.e. in which the meaning of the whole is not a predictable semantic function of its constituents words is **idiomatic**.

It is possible for a PLI to be compositional in that the meaning of the whole is a compositional function of the meaning of its constituents, but for not all the possible readings to be available in the lexicon. For example, a political party could be a social occasion which is political, but in its lexicalised form it is an organisation which functions to select and have elected members of a legislature. This is one of the possible compositional meanings of *political party*, but only one, given that *party* is polysemous. It is thus **selectively compositional**.

In some PLIs only one of the words has an idiomatic sense, i.e. a sense that it does not have when on its own. This sense only exists in combination with the other words in a particular PLI. Such PLIs are **unilaterally idiomatic**.

In some PLIs more than one word has a sense that it has only in the PLI. For example a *red herring* is neither red nor a herring, i.e. both words have special senses that they have nowhere else but in construction within this PLI (Weinreich 1969). Such PLIs are bilaterally idiomatic.

It seems important in the discussion of the semantic properties of PLIs to clearly differentiate these from the syntactic properties of the same PLI. The work of Melčuk, as exemplified in work such as Melčuk (1995), makes it clear that PLIs can, in many cases, be seen as mapping semantic predicates idiosyncratically onto verbs for specific arguments. So, for example, the weather is forecast, rather than predicted.

The properties above relate to the formal properties of a PLI in terms of its syntax and semantics but many PLIs also have conventional conditions of use. A **formula** can be regarded as a PLI with contextually restricted conditions of use. For example, *I'm sorry* is a PLI which is used to offer an apology. Speech act theory provides examples of formulae and subclassifications of types of usage conditions. However this is just a beginning. Every small scale ritual tends to be accompanied by formulae: cabin crew on aeroplanes use them, 'What would you like to drink, Sir/Madam?' Flight crew use them: 'This is your captain speaking.'

2.2 Creative artistic deformation of PLIs

It is now possible to survey which of the above properties is available for artistic deformation.³ Since PLIs have syntactic structure, that may be deformed, for instance by adding a modifier where conventionally the PLI is not permitted to take such modifiers. For example the PLI *as scarce as hen's teeth* cannot have the modifier *red* added before *hen* or the word *white* before *teeth* however appropri-

3. We will only survey deformation of those PLI properties which result from their being PLIs.

ate these adjectives might otherwise be. The semantic properties of a PLI may be independently deformed. For example, a lexicalised constituent of a PLI may be semantically ambiguous in the general case but have only one reading in the PLI. The PLI may be re-contextualised in such a way that the other reading becomes accessible. In many instances, however, both kinds of deformation may occur together. A pun created by the substitution of a homophone for a lexicalised constituent of a PLI has both structural and semantic effects. As Chiaro (1992) suggests, many of these kinds of deformation are closely paralleled by kinds of slips of the tongue.⁴ The leading hypothesis of what follows is that, in general, artistic deformation will be of this kind.

2.2.1 Possible phonological deformations

It is possible to create phonological deformation through substitution and exchanges in the same way as these are found in slips of the tongue. Spoonerisms are a classic example in which the onsets of stressed syllables are exchanged.

2.2.2 Range of structural deformations

a. Lexicalised constituent substitution. Lexicalised constituents might be substituted for. For example, the clause *It takes a thief to catch a thief* might have the word *thief* substituted for by *cop*. Recall that a crucial aspect of such deformation, as with all the other processes to be outlined below, is that the original PLI must still be identifiable or else the deformation is not identifiable as such. For example, if one substituted the word *bus* for the word *cake* in the PLI *to take the cake*, the original PLI would not be accessed by any hearer. Instead, if the word *biscuit* were substituted for *cake* and the PLI was used in a context where it was clearly the speaker's intent that the phrase should be taken as an expression of astonishment, the substitution could then be seen as a kind of lexical play. That rests on the fact that *biscuit* and *cake* are co-hyponyms, whereas *bus* has no relationship with the PLI's lexicalised constituent *cake* for which it has been substituted.

In this regard, puns can be seen as being of two kinds; either a lexical substitution of a homonym (homophone or homograph) has taken place, or a non-conventional sense of a polysemous lexicalised constituent within a PLI is accessed. I will suppose that, in every case, some preferred sense or lexical item has been substituted for the one in the PLI. There are clearly complexities here that have to

4. See Kuiper (2004) for some examples of PLI slips which parallel cases of the artistic deformation of PLIs. Note that not all slip types are plausible conscious artistic deformations and so not all slip types will be used in this study. For example, anticipation and perseveration slips arise as the result of activation patterns which are unlikely to be replicated as a conscious process, while metatheses and exchanges such as Spoonerisms can be.

do with the accessing of lexical items and word senses in the mental lexicon. Since this chapter deals with the way in which the properties of PLIs are utilised, such complexities will be ignored.

Substituting for a bound word would normally yield a freely formed phrase since the rest of the phrase would give no clue that the word substituted for was a bound word. For example, if one substituted the word *offence* for *umbrage*, *take umbrage at* would not be recoverable. Hence such a form of lexical substitution is unlikely in the kind of lexical creativity under discussion here.

The same holds for addition to a lexical selection set. Such a creative addition can essentially be seen as a lexical substitution for all the members of the set. Here again the recoverability of the PLI may be hampered. For example, if one substituted *equilibrium* for *cool/rag/temper* in *to lose one's cool/rag/temper* the PLI would not be recoverable. But if one substituted *cloth*, there is potential recoverability on the basis of the fact that *rag* is unilaterally idiomatic in this PLI and in its literal sense it is semantically closely related to *cloth*. That allows for recoverability of *rag* in the appropriate context.

Optional lexicalised constituents provide little opportunity for artistic deformation since any lexical substitution in an optional constituent would likely lead to it being read as a normal adjunct.

Lexical substitution may be further examined in terms of the characteristics of the lexicalised constituent and what is substituted for it. A number of characteristics of the word might be similar or related. For example, they might rhyme or be synonyms and so forth.

Since restricted collocation is a kind of minimal idiosyncrasy for a PLI, substituting for either of the words which are involved in such a collocation is likely to be perceived as non native-like selection (Pawley & Syder 1983). For example, the substitution of *accept* for *take* in the PLI *take offence* would probably not be taken as a form of artistic deformation. A native speaker would generally know that the conventionally prescribed verb is *take* and so recoverability can be maintained.

b. Lexical exchanges. Some PLIs allow for words of the same syntactic category to be exchanged. For example the proverb *A bird in the hand is worth two in the bush*, might have its nouns exchanged to read *A bird in the bush is worth two in the hand*.

c. Violating slot restrictions. While filling in, say, a noun phrase in an open slot in a PLI would not yield anything remarkable, breaching arbitrary slot restrictions in a PLI would. Take, for example, the PLI *take NP to task*. The noun phrase slot has the conventional restriction that it can only be a human NP. If one inserted

an inanimate NP, say, *the rock*, this would be akin to the violation of a selectional restriction. In other cases, slots are restricted in terms of co-reference. Many possessive PLIs have arbitrary restrictions on their antecedent co-reference. For example, in *get NP's goat* the goat which is being got must not be co-referential with the subject of the verb *get*. So the sentence *Andrea got her goat* cannot be read as involving the subject of the PLI in question.

d. *Modifying unmodifiable PLIs.* If a PLI is conventionally unmodifiable then it can be modified as a form of artistic deformation. For example, indicating that someone is *a moderately bad egg* could be seen as a form of artistic deformation.

e. *Transforming inflexible PLIs.* One could also transform an inflexible PLI, provided the recoverability condition was met. For example, in a context where it was clear that the death of a person was involved, one could indicate that *the bucket was kicked by old uncle Maurice* which involves passivising *kick the bucket*, a PLI that conventionally does not allow passivisation. But if one clefts *take for granted* as in *Granted for he certainly was not taken* then the resulting construction is barely recognisable as a form of *take for granted* and thus on the border of irrecoverability.

f. *Blending two PLIs.* Sometimes two PLIs are semantically related and this allows for them to be blended. For example, *cross that bridge when NP come(s) to it* and *burn NP's bridges* can be blended to *burn that bridge when NP comes to it*.

g. *Foregrounding structural ambiguity.* PLIs can be syntactically ambiguous (Alexander 1997: 42ff.). However, conventionally only one of the possible readings is normally available. In such a case the ambiguity may be foregrounded by artistic deformation. For example, the PLI *hail-fellow-well-met* is conventionally used as a pre-nominal modifier. But suppose it were written *Hail, fellow. Well met.* Then an alternative parsing of the phrase has been foregrounded without the form of words being altered.

Not all of these options are mutually exclusive. It is possible to perform both a lexical substitution and transform an inflexible PLI at the same time. Some of the humour illustrated below shows multiple structural deformations.

2.2.3 Possible semantic deformations

a. *Literalising of idiomatic readings.* PLIs may be idiomatic in whole or in part as we saw earlier. Various means are available to bring to mind the literal sense of either one of the lexicalised constituents which are conventionally metaphorical, or of the whole PLI idiom (Alexander 1997). For example, the PLI *raining cats and dogs* can have attention drawn to its literal meaning by lexical substitution as

in *raining moggies and hounds*. Speaking of his intimations of the whereabouts of Brünhilde, Siegfried could say that *a little bird told me* which, in Wagner's opera, is literally the case. Such literalising is a function of substitution within the analysis conducted here. Usually if a lexicalised constituent is a metaphor that has been literalised, the whole PLI is then also reparsed in a literal sense.

b. Pun on a constituent lexeme. PLIs often contain lexemes which are polysemous but which in the PLI have only one reading. By providing a targeted context the unconventional sense can be foregrounded. Conventionally, *being left speechless* involves being bereft of speech, but a speaker who has left his or her notes behind is also left speechless since the word *speech* is ambiguous as between the power to speak and a public address.

Puns, as suggested above, can be forced through the use of homonyms in lexical substitutions. For instance, by substituting the word *soul* for the word *sole* one can indicate that an elderly jazz singer is the soul survivor of a group. Such puns are common in newspaper word play (Alexander 1997; Cowie 1991, 1998).

The relationships between structural and semantic deformation are not mutually exclusive. While a structural deformation may not have a semantic effect, it often will, particularly in cases where the intent is to create humour.

3. Word play in a corpus of cartoons by Cathy Wilcox

While all of the above options appear to be available, it is not necessarily the case that all are equally likely to lead to humour or that, even when they are, a humorist will be equally drawn to all of them. Particular humorists may prefer some kinds of deformation to others. What follows aims at two goals. The first is to illustrate how an artist might use artistic deformation of PLIs for humorous effect. In place of an inductive approach this will be done deductively by allowing the analysis to proceed in a resource-based way from the above possibilities for artistic deformation. The second goal is to provide a model of the way in which a particular humorist exploits the resources of the phrasal lexicon. The leading hypothesis is that such exploitation differs from one humorist to another and from one medium to another. We would suppose that Cathy Wilcox's cartoon art differs in this respect from that of Giles or Gary Larson and that the humour of Wallace and Gromit differs from that of the Muppet Show.

Cathy Wilcox is an award-winning Australian cartoonist whose work is to be found in the *Sydney Morning Herald* and the *The Age*. The data for this study is the complete set of cartoons from her first collection (Wilcox 1991). Only those cartoons were selected where there was clear reference to one or more PLIs; those

Table 1. Number of PLIs per cartoon

Number of PLIs	Instances
1 per cartoon	73
2 per cartoon	14
3 per cartoon	3
Total number of PLIs	110

where such reference was evident were then coded for the various artistic deformations that they had undergone.

Cartooning is a medium which has its own potential for PLI humour in that the PLI must normally be in a caption, in the balloon text of the characters or both, while the drawing can shed a contextual light on the (very brief) text(s). There is a substantial tradition of word play in cartoons. *Punch* and the *New Yorker* both publish cartoons of this kind. In one *New Yorker* cartoon two men are sitting behind identical desks facing one another on opposite sides of an otherwise featureless office. One says to the other, '*One day you will be sitting where I am now sitting.*' Conventionally this PLI is used by a senior member of an institutional hierarchy to a junior member of the same hierarchy as a way of indicating that the junior is on his way up the institutional ladder. However, the featureless office invites the statement to be taken literally and, since one cannot see any difference between the situation of the two men, one is likely to examine the nature of promotion. On the front cover of *Punch* a doctor is writing a script for a Japanese patient who is dressing behind a screen. The doctor says '*It's all right, Mr Hayakawa. I'll just give you something to open your bowels.*' Mr Hayakawa has just produced a samurai sword and appears to be going to commit harakiri. In both cases a PLI is being literalised.

The Wilcox corpus consists of 240 cartoons of which 90 contained reference to one or more PLIs. The rest also involve humour based on varieties of word play but these lie outside the scope of this chapter. (See Table 1.)

3.1 Phonological deformation

There are four cases of phonological deformation in the data. Three involve exchanges. In one a dog is leaning on a crutch and the caption reads *The leaning piece of towser* where the exchange source is *the leaning Tower of Pisa*. In a second, one cow is saying to the other *I am just a tense simmental fool* where the exchange source is *I'm just a sentimental fool*. In a third an approaching ocean liner named HMS Pinafore is hailed from a small boat with the words, '*Your ship is slowing.*' This last exchange is a Spoonerism.

Table 2. Lexical substitutions

Number of substitutions/cartoon	Total	%
single substitution	39	73.6
double substitution	10	18.9
triple substitution	3	5.7
quadruple substitution	1	1.9
Number of substitutions	72	
Number of cartoons containing substitutions	53	

Phonological addition occurs in a cartoon showing a cow on a hillside with a caption reading *A ruminant with a view*. This latter case can also be seen as a lexical substitution, in which case it is a malapropism.

The next set involve structural deformation (and possible attendant semantic effects).

3.2 Lexical substitutions

Of the 90 cartoons containing a PLI, 53 contained one or more substitutions for a total of 72 substitutions. (See Table 2.)

One of the more ingenious cases of substitution is a cartoon showing a mother, father and small child around the dinner table. The caption is *Nuclear family*. The child asks ‘*What’s for DNA?*’. Here the substitution is the acronym *DNA* for the word *dinner*. In Australian English these are close to homophonous. The mother replies, ‘*You’re just like your father.*’ picking up the *DNA* substitution and getting an ambiguity as to whether the child has the same *DNA* as his father or whether that is what the father always asks at dinner time. I take it that the word *nuclear* in the PLI *nuclear family* has the conventional sense ‘parents plus children’ but has been substituted for by the word *nuclear* with the sense ‘pertaining to the cell nucleus’.⁵

Lexical substitutions also allow one to examine the relationship between the original word and its substitute. Here the relationship was usually homonymous. Where the substitute was also a homophone one has a clear case of punning. Where the phonological relationship is close, i.e. when the words are paronyms (Attardo 1994), the closeness allows for the recoverability condition to be met, thus producing a near pun. (See Table 3.)

A typical example of lexical substitution involving a phonological relationship in a PLI is a cartoon in which the male protagonist says, ‘*What say we burn some*

5. Whether these are two senses of one word or two different words, i.e. homonyms, is not pertinent to the argument.

Table 3. Relationship between lexical items

Type of relationship	Number of instances	% of insertions
homonymy	39	54.2%
homophony	39	54.2%
homography & homophony	26	36.1%
homograph only	0	0%
paronymy	21	29.2%
co-hyponymy	2	2.8%
antonymy	1	1.4%
same semantic/pragmatic field	1	1.4%
syntactic	2	2.8%
none	6	8.3%

hedges?' The substitution here is that *bridges* has been substituted for by *hedges*. Both words are disyllabic and they share the coda of their first syllable and all of their second syllable. The female protagonist then says, '*Sterling idea.*' alerting us to the ambiguity of *hedge* as between a physical entity and a financial institution.

Homophony without homography is found in cases such as a cartoon where two men are on board a speeding boat. One says, '*I'm afraid I've thrown away the bits that slow us down.*' The other replies, '*Them's the brakes.*'

Semantic relationships are created when a substitute is, for example, a converse as in a cartoon showing the start of a race at a dog track where one hound says to the other, '*This is where the buck starts.*'

Syntactic relationships occur in these cartoons when one pronoun is substituted for another, as in a case where *your* is substituted for by *my* in the Muir and Nordenesque Mae West misquotation, '*Is that a gun in my socket or are you just pleased to seize me?*' which shows a soldier placing the muzzle of a rifle in the speaker's eye.

In some cases, for example where there are PLI blends, the substitutions have no individual relationship. Here the relationship often arises from the context. For example a dog is scratching itself and the caption reads *Scratch and smell*. Here the PLI target is *Scratch and win*, which is a game of chance.

3.3 Lexical exchanges

Five of the 90 cartoons contain exchanges. (See Table 4.) Two of these are lexical. In one, one sheep says to the other, '*Do you know "Click go the shears"?*' The other replies, '*No, but if you baa a few hums ...*' where the exchange source is *hum a few bars*. In another, two dogs are bemoaning the state of urban living. One says to

Table 4. Lexical exchanges

Type	Instances
Spoonerism	1
Syllabic exchanges	2
Lexical exchanges	2
Total exchanges	5

the other, ‘*This dog’s going to the country.*’ Here the exchange source is *This country is going to the dogs.*

3.4 Violating slot restrictions

There are three cases which can be interpreted this way. The phrasal verb *let NP out*, when its complement is a domestic pet means that the pet is to be allowed/ made to leave the house; when its slot is a garment it means that the garment is to be made larger. In the cartoon the husband, in bed with his wife with a large cat sitting on the bed, asks, ‘*Didn’t you let that cat out?*’ The wife replies, ‘*No dear, I just bought a bigger one.*’

The phrasal verb *NP go off* has two sets of slot restrictions; if the NP is food then it means the food has become unfit to eat and smells; if it is an explosive then it means that the explosive has exploded. One cartoon shows a woman bending over a fish. She sniffs and says, ‘*I think this fish has gone off.*’ In the next frame the fish has exploded and her face and hands are covered with black soot.

3.5 Adding a modifier where none is conventionally permitted

No cases were found.

3.6 Transforming a frozen PLI

One case only occurs where the one hound says to the other, ‘*This is where the buck starts.*’ The canonical form of the source PLI is *The buck stops here.*

3.7 PLI blends

There were four PLI blends. In one *The Lone Ranger* was blended with *loan shark* to make *The lone shark*. In another *The Bermuda triangle* was blended with *Bermuda shorts* to give *The Bermuda shorts.*

3.8 Structural re-analysis

There are nine cases where a non conventional structural representation of a PLI is foregrounded. *The Queen's English* comes up for re-analysis in a cartoon in which the lady says, 'I was taught the Queen's English.' Her disreputable-looking male interlocutor replies, 'I already knew she was.'

The following cases involve semantic deformation (much of which arises from structural deformation)

3.9 Literalising the sense of a metaphor within a PLI

There are seven cases of punning using this device. At a bar a gentleman asks a lady, 'What do you do for a crust?' The lady replies, 'Eat out the soft bit.' Note that literalising one word in a PLI usually involves reparsing the meaning of the whole PLI.

3.10 Literalising the sense of a figurative PLI

There are eighteen such cases. In many cases where an individual substitution has taken place the reading of the whole PLI has become literal. A brain surgeon holding up the whole of a brain says to an assistant, 'This is the stuff that dreams are made of.' A man delivering a box of underpants to a prison says to the guard at the gate, '200 pairs of Y fronts for maximum security, Sir.'

3.11 Calques and other paraphrases

There is one ingenious case of a loan translated PLI. The lady asks an academic-looking gentleman, 'Do you speak Latin?' The gentleman replies, 'Pontifex catholicus est?' The latter is a loan translation of *Is the pope a catholic?*

In another cartoon, the PLI *rose-tinted glasses/spectacles* is paraphrased as *pink contact lenses*.

3.12 Pragmatic incongruity

Many PLIs have conventional conditions of use which can be foregrounded by placing the PLI in an unconventional context. For example, under the caption *Humane lobster cooking* a chef is looking into a boiling pan saying 'I hear what you're saying.' Many of the cartoons also show unconventional contexts alongside other verbal deformations. In one a hunter in pith helmet holding a gun asks a

lion, 'Why won't you talk to me?' The lion replies, 'You've hurt my pride.' The pun on *pride* exists alongside the pragmatic incongruity of a hunter asking a lion why the lion won't talk to him. The pragmatic incongruity is partly occasioned by the pun to come.

3.13 Inferred PLIs

In some measure any PLI which has been artistically deformed must be inferred and accessed from the mental lexicon on the basis of what remains after the deformation. However, in the majority of cartoons, a PLI is directly quoted in whole or in part. Occasionally, however, the PLI must be largely inferred in that it is not stated but just hinted at. There are four such cases among the PLIs found in the cartoons. In one ingenious case one snake is saying to another which has a large bulge in its middle and a dummy in its mouth, 'Spit it out, Bill, that's only the garnish.' The only word of the PLI *spit the dummy* which is overt is *spit*. The dummy can be seen in the picture and it must be inferred that the baby that once sucked the dummy is the bulge in Bill's middle, therefore the main course that has the garnish.



Figure 1. Cathy Wilcox cartoon

In the Bermuda shorts cartoon (Figure 1) above, the word *triangle* in the PLI *the Bermuda triangle* must be inferred from the fact that the wife has lost her husband in his Bermuda shorts. The contextual real world knowledge that things disappear without trace in the Bermuda triangle must be brought to bear in order to make this inference.⁶

6. It remains a largely unresearched question as to when hearers have enough information to look up a PLI.

Table 5. The PLI deformation signature of Cathy Wilcox

Type of deformation	Total	Type/PLI
substitutions	72	.655
literalisation of a PLI	18	.163
structural ambiguity	9	.081
literalisation of a lexical sense	7	.063
Phonological	4	.036
PLI blends	4	.036
slot restriction violations	3	.027
lexical exchanges	2	.018
calques and paraphrases	2	.018
frozenness violation	1	.009

4. Conclusion

It is clear that Cathy Wilcox has a preference for puns and near puns over other forms of artistic deformation. Since we have suggested that puns are a function of substitutions, then this is borne out by the fact that the total number of substitutions (given that there can be more than one per cartoon) is almost as large as the total number of cartoons under analysis, and that more than half of the cartoons contain substitutions. That is not to say that Wilcox is unaware of the other avenues for deformation. This is evidenced by the fact that the only deformation device of which there is no instance in the data is that of an additional modifier where none is conventionally permitted. Given the long list of potential sources of deformation, it is noteworthy that so many of them can be used for humorously intended deformation.

If we look at Table 5 which summarises Wilcox's selective preferences in the dataset of 90 cartoons, then we have a comparison against which other cartoonists who work with PLIs might be assessed. Note that, since any given PLI can undergo more than one of these processes, the index is calculated in terms of the number of cases that appear per PLI in the data.

Lexical creativity can therefore be seen to manifest itself not only on each occasion that a cartoonist produces a new cartoon involving lexical play but also by the personal set of preferences for certain kinds of creativity.

5. Artistry and artistic deformation

The conclusions above are fine as far as they go, but that is not far enough. Artistic deformation, i.e. intentional changes to the stable expectations arising from the



Figure 2. Cathy Wilcox cartoon

idiosyncratic properties of PLIs in the mental lexicon, do not give us quite enough sense of the artistry which a fine cartoonist like Cathy Wilcox exercises. Here we are drawing on the kind of distinction between *performance* used in the Chomskyan sense and *performance* in the sense of Bauman (1975). A sense of artistry is essentially qualitative and not all artistic deformation is qualitatively alike. Such quality is not easy to analyse in linguistic terms and, furthermore, to do so would often mean doing violence to the intuitive sense one has with many of Wilcox's cartoons that the cartoonist is doing something which is really clever. Analysis can also be seen as 'brushing the bloom off the butterfly'. However, another instance will perhaps suffice to show what is meant here.

In a hold up with a hostage (Figure 2) the PLI *hand over the ... or the ... gets it* is conventionally understood to mean that the addressee is to hand over the money, diamonds etc to the speaker, the man with the gun, or else the hostage will be shot. In the cartoon, a mother is the addressee and her small son is the hostage. By uttering a second PLI to her son the mother uncovers a multiple ambiguity in the PLI which it does not conventionally have. *Hand something over* is conventionally taken to mean "hand it over to the robber" and that is what the phrase appears to mean here. But there are three participants: robber, hostage and person being held up. The phrase could therefore be taken as an instruction to hand the cash over to the hostage. Conventionally the phrase *so and so gets it* means that the hostage will be shot or killed. But if the cash is to be handed over, then *the kid gets it* could be a structural alternative meaning to hand it over to the robber since the two clauses are connected by *or*. Since the mother would, naturally, prefer that her son get her money rather than the robber, she exhorts her lad not to spend all her money at once. Here the second PLI, *Don't spend it all at once*, is used to expose the structural ambiguity of the first PLI. This is, clearly, a clever and creative use of the two PLIs, but it is also totally unexpected. One expects a mother to want to protect her son who has a pistol aimed at his head. One does not expect him to be exhorted to fiscal responsibility at such a time. It is this element of surprise

and delight which comes from 'making it new', in the sense of seeing a hackneyed scenario and its attendant language anew through its creative deformation, and this is what gives a sense of real artistry.

References

- Alexander, R. J. 1997. *Aspects of Verbal Humour in English*. Tübingen: Gunther Narr.
- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.
- Attardo, S. 1994. *Linguistic Theories of Humor*. Berlin: Mouton de Gruyter.
- Bauman, R. 1975. "Verbal art as performance." *American Anthropologist* 77: 290–311.
- Burger, H. 2003. *Phraseologie: Eine Einführung am Beispiel des Deutschen* (2nd ed.). Berlin: Erich Schmidt Verlag.
- Chiario, D. 1992. *The Language of Jokes: Analyzing Verbal Play*. London; New York: Routledge.
- Chomsky, N. 1996. *Minimalism*. Cambridge, MA: MIT Press.
- Cowie, A. P. 1991. "Multiword units in newspaper language." *Cahiers de l'Institut de Linguistique de Louvain* 17(1–3): 101–116.
- Cowie, A. P. 1998. "Creativity and formulaic language." *Linguistica e Philologia* 8: 159–170.
- Fillmore, C. J., Kay, P., and O'Connor, M. C. 1988. "Regularity and idiomaticity in grammatical constructions." *Language* 64: 501–538.
- Fraser, B. 1970. "Idioms within a transformational grammar." *Foundations of Language* 6: 22–42.
- Gläser, R. 1986. *Phraseologie der Englischen Sprache*. Tübingen: Max Niemeyer Verlag.
- Gläser, R. 1995. "The stylistic potential of phraseological units in the light of genre analysis." In *Phraseology: Theory, Analysis, and Applications*. A. P. Cowie (ed.), 125–143. Oxford: Clarendon Press.
- Koopman, H. and Sportiche, D. 1991. "The position of subjects." *Lingua* 85: 211–258.
- Kuiper, K. 2004. "From intent to malutterance." *Landfall* 207: 90–96.
- Kuiper, K. and Everaert, M. 2000. "Constraints on the phrase structural properties of English phrasal lexical items." In *PASE Papers in Language Studies: Proceedings of the 8th annual conference of the Polish Association for the Study of English*. B. Rozwandowska (ed.), 151–170. Wrocław: Aksel.
- Mackin, R. 1978. On collocations: "Words shall be known by the company they keep". In *In Honour of A. S. Hornby*. P. Strevens (ed.), 149–165. Oxford: Oxford University Press.
- Melčuk, I. 1995. "Phrasemes in language and phraseology in linguistics." In *Idioms: Structural and Psychological Perspectives*. M. Everaert, E.-J. van der Linden, A. Schenk and R. Schroeder (eds), 167–232. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Melčuk, I. 1998. "Collocations and lexical functions." In *Phraseology: Theory, Analysis, and Applications*. A. P. Cowie (ed.), 23–53. Oxford: Clarendon Press.
- Moon, R. 1998. *Fixed Expressions and Idioms in English: A Corpus based Approach*. Oxford: Clarendon Press.
- Naciscione, A. 2001. *Phraseological Units in Discourse: Towards Applied Stylistics*. Riga: Latvian Academy of Culture.

- Nicolas, T. 1995. "Semantics of idiom modification." In *Idioms: Structural and Psychological Perspectives*. M. Everaert, E.-J. van der Linden, A. Schenk and R. Schroeder (eds), 233–252. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Nunberg, G., Sag, I. and Wasow, T. 1994. "Idioms." *Language* 70 (3): 491–538.
- Paolillo, J. C. 1998. "Garry Larson's Far Side: Nonsense? Nonsense!" *Humor* 11 (3): 261–290.
- Pawley, A. and Syder, F. 1983. "Two puzzles for linguistic theory: Nativelike selection and nativelike fluency." In *Language and Communication*. J. Richards and R. Schmidt (eds), 191–226. London: Longman.
- Soehn, J.-P. 2003. *Von Geisterhand zu Potte gekommen: Eine HPSG-Analyse von PPs mit unika-ler Komponente*. Unpublished MA, Universität Tübingen, Tübingen.
- Van Lancker, D. 1987. "Non propositional speech: Neurolinguistic studies." In *Progress in the Psychology of Language*. A. W. Ellis (ed.), Vol. 3, 49–118. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Verstraten, L. 1992. "Fixed phrases in learners' dictionaries." In *Vocabulary in Applied Linguistics*. P. J. L. Arnaud and H. Béjoint (eds), 28–40. Basingstoke: Macmillan.
- Weinreich, U. 1969. "Problems in the analysis of idioms." In *Substance and the Structure of Language*. J. Puhvel (ed.), 23–81. Berkeley: U. of California Press.
- Wilcox, C. 1991. *Throw Away Lines: Cartoons by Cathy Wilcox*. North Ryde: Collins/Angus and Robertson.
- Williams, E. 1994. "Remarks on lexical knowledge." *Lingua* 92 (1): 7–34.
- Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Lexical creativity in texts

Advertising and the media

Blendalicious*

Adrienne Lehrer

Although blends have been marginalised as a word-forming process, they are now common. Some are completely conventionalised (e.g. *workaholic*, *telethon*, *cityscape*), while most are nonce forms.

Blends are compounds consisting of a whole word and a splinter (part of a morpheme) or two splinters. A frequent splinter may become a combining form. Final combining forms like *-thon* and *-holic* and initial ones like *Mc-* (*McMansions*) and *e-* (*e-tail*) have achieved this status.

What and how much can be omitted is limited by the need for the hearer or reader to recover what is missing. Word frequency, neighbourhood effects, and semantic plausibility are some of these constraints. I summarise several psycholinguistic experiments on understanding novel blends.

Blends occur mostly in media, advertisements, and product names, functioning to gain our attention.

1. Introduction

Neologisms in English and probably all languages use the whole spectrum of word-formation devices. The type of neologism I will discuss in this article, blends, has long been considered marginal. However, in the last few decades it has become increasingly common, so much so that it is losing its marginal status and more linguists are examining the properties of blends. The interesting question for me is why blends and other types of neologisms have become increasingly popular. They generally make comprehension more difficult because the hearer or reader has to figure out their meaning, as they are typically presented without glosses or explanations. Like other neologisms, such as rimes and allusions, blends are often cute and amusing. They work as a form of word play, which Kelly describes as “lexical teases” (1998:586). My hypothesis about their increasing frequency –

* I wish to thank Judith Munat for her many valuable suggestions.

perhaps it is only a speculation at this time – is that we are surrounded by stimuli vying for our attention (newspapers, magazines, radio, TV) and thousands of advertisements in all of the above. Therefore, using a novel clever word is likely to catch our attention and get us to read or listen to what is being presented. Then when a word-formation device like blending becomes common, other speakers and writers create similar forms by analogy simply because it is fashionable to do so; they want to show that they, too, are trendy, creative, and cool.

In this article I will first provide a characterisation of blends, discuss their structure, and examine the development of their constituent parts (splinters) into new morphemes. I will consider how novel blends are processed, and finally, the kinds of contexts in which they tend to occur.

2. What are blends?

Blends are underlying compounds which are composed of one word and part of another, or parts of two (and occasionally three) other words (Algeo 1977:48).¹ The word part is called a **splinter**. Although a splinter is a clipping, it cannot occur alone as a word. For example, in *dramedy* < *drama* + *comedy*, neither *dram-* nor *-edy* may be used independently. One cannot say **I saw a really gripping dram on TV last night*. In *infotainment* < *information* + *entertainment*, *info* is a clipping and can stand as an independent word, but *-tainment* must be attached to something else, like *info* or *winter* in *wintertainment*, the name of a winter festival.

A few blends were recorded as early as the 15th century,² but they are now obsolete. However, some current blends originated in the 19th century,³ such as *brunch* < *breakfast* + *lunch* (1886, OED, MW); *slanguage* < *slang* + *language* (1879) and *soliloquacity* < *soliloquy* + *loquacity* (1895). Blends became more common in the 20th century: *motel* < *motor* + *hotel* (1925), *transceiver* < *transmitter* + *receiver* (1934), and *permafrost* (1943). Many of these have entered the lexicon as conventional words and some speakers are no longer aware of their underlying complex sources.

New blends are sometimes introduced with a full account of their **source words**, that is, the underlying compound and its meaning, especially if this is not

1. Some scholars (e.g. Plag 2003; Kubozono 1990) use a more restricted definition, dealing only with blends in which both elements are abbreviated.

2. Adams (2001:141) lists *foolosopher* (OED, 1592), *niniversity* (OED 1590), and *knavigation* (OED, 1613).

3. Dates are taken from Merriam-Webster Online (MW) or the Oxford English Dictionary.

obvious. The blend *vog*, < *volcano* + *smog*, the emission of gases from smoking volcanoes which produce air pollution, is often glossed as ‘volcanic smog’. When I entered this word on Google, the examples that appeared on the first screen generally provided definitions. Another example is *Craisins*[®] < *cranberries* + *raisins*, a brand name for dried cranberries, defined on the package. Sometimes the source words are recognisable, but the interpretation is not. A university newsletter described some people as *negaholics*, which is easy to identify as coming from *negative* + *-holic* (the latter being a splinter from *alcoholic*). However, since the meaning is unclear, a definition was provided: ‘Negaholics are people who always think the worst of themselves, and others, and of outcomes.’ Another example is a blend coined spontaneously by a news commentator who was reporting the weather. She said that it would snow or maybe drizzle and then laughingly said to her colleagues at the news desk that it would *snizzle*.

More often, however, new blends are introduced in a context where the hearer or reader is left to figure out the underlying compound, and a plausible interpretation must be found for that context. The first step is to parse the blend so that the two source words may be identified. If one part consists of a whole word such as *oildraulic* or *deskercise*, the parse is fairly easy, and the next task is to identify the source of the splinter (*-draulic* < *hydraulic*, and *-ercise* < *exercise*). If the blend consists of two splinters, the parse may be more difficult, depending in part on the phonology of the blend. For example, two blends which subjects in my experiments (described below) found difficult to understand are *snizzle* < *snow* + *drizzle* and *swacket* < *sweater* + *jacket*. The problem for the hearer or reader is that the new word may be divided in more than one place. In *burbclave* < *suburb* + *enclave*, however, the cut will most naturally be made between *b* and *c* because *bc* is not a possible syllable internal combination in English.⁴ (This example is particularly unusual because *burb* is a clipping which deletes the first syllable.)

After the source words have been identified, a plausible meaning must still be found, but this is the same problem that exists in interpreting novel compounds, since blends are shortened forms of compounds.

3. The structure of blends

The commonest type of blend in my corpus is a full word followed by a splinter: *wintertainment* < *winter* + *entertainment*; *chatire* < *chat* + *satire*; *vodkatini* < *vodka* + *martini*. Blends can also begin with a splinter, followed by a full word: *narcoma*

4. I am indebted to Jen Hay, personal communication, for pointing this out.

< *narcotic* + *coma*, *cinemenace* < *cinema* + *menace*; *squangle* < *square* + *angle* (a type of tool); *Amerindian* < *American Indian*; *administrivia* < *administration* + *trivia*.

Also common are blends consisting of two splinters. Two types are possible: (1) the beginning of one word is followed by the end of another: *psychergy* < *psychic* + *energy*; *hurricoon* < *hurricane* + *typhoon* or *monsoon*; *cheriodical* < *cheery* + *periodical* or (2) both splinters are the beginning of words: *biopic* < *biographical* + *picture*; *sitcom* < *situation* + *comedy*, and *cabsat* < *cable* + *satellite*. If the splinter precedes a full word or another splinter, it must be the first part of a word. One cannot have a blend like **glyson* < *ugly* + *person* or **ictionblem* < *addition* + *problem*. (*Blog* < *web* + *log* is a major exception.)

A fourth type of blend involves complete overlap of one or more phonemes, often of whole syllables. Some part of the word has to be counted twice, as belonging to both source words. (English spelling may require deleting some letters, like the silent *e*, and other minor spelling changes may occur which do not affect pronunciation.) This blend type has become increasingly common. The examples I found some 15 years ago were *sexploitation* > *sex* + *exploitation*; *sexpert* < *sex* + *expert*; *palimony* *pal* + *alimony* ‘money paid to an unmarried partner after separation’ and *cocacolonization* < *Coca Cola* + *colonization*. (Lehrer 1996). More recently I have found: *clandestiny* < *clandestine* + *destiny*; *cashmiracle* < *cashmere* + *miracle*; *netiquette* < (*Inter*)*net* + *etiquette*; *airobics* < *air* + *aerobics* (exercises during air travel); *Wheatables*[®] < *wheat* + *eatable* (a cracker); *snapppetizer* < *snap* + *appetizer* ‘easy-to prepare appetizers’, *faddition* < *fad* + *addition*; *cattitude* < *cat* + *attitude*; *cinnamincredible* < *cinnamon* + *incredible*; *telelectorate* < *tele* (British clipping for *television*) + *electorate*; *blobjects* < *blob* + *object* (art exhibit featuring functional objects with amorphous shapes);⁵ *cutensils* < *cute* + *utensils*; (items from the *blobjects* exhibition); *Yahooligans!* < *Yahoo* + *hooligans*, a Web site guide for kids, and *Freakonomics*, a book title.

A less common type of blend involves a discontinuous element. A word or clipping is embedded in part of another source word as an infix. Examples are *entrepornneur* < *entrepreneur* + *porn(ography)*; *chortle* < *chuckle* + *snort*, created by Lewis Carroll; *delinquancy* < *delinquency* + *lingual* ‘misusing language’; and *enshocklopedia* < *encyclopedia* + *shock* ‘knowledge of scary films’

In addition to complete overlap, blends often present partial overlap when letters or phonemes occur in both source words. For example, in *wintertainment*, the letters in bold belong to both source words: *winter* and *entertainment*. In *cinemenace*, the *m* belongs to both words and in *communostolga* < *communism* +

5. This blend also has a philosophical meaning involving the totality of being.

nostalgia the *n* belongs to both words. Sometimes the overlap is not simply of contiguous letters or sounds but occurs in various places in both parts. Examples are *astrocity* < *astronaut* + *atroc*ity, *flustratred* < *flustered* + *frustrated* and *woofle* < *woof* + *waffle* ‘a waffle-shaped dog biscuit’, which all distribute overlapping letters (and sounds) noncontiguously.

A semantic distinction that cuts across the above classifications is that between syntactic compounds (also called abbreviated or exocentric compounds) and coordinate compounds (also called copulative compounds). In syntactic compounds the last source word is the semantic head and the first is a modifier, as in *sci-fi* and *motel*. Semantically a motel is a kind of hotel (Plag 2003: 123). (Qualifications are discussed below). In these blends the second splinter can be from either the first or last source word. Coordinate compounds such as *beefalo* denote semantic elements from both entities. With relatively few exceptions, the two splinters consist of an initial splinter from the first source word and the final splinter from the last part of the second source word.

4. The phonology of blends

Kubozono (1990), Kelly (1998) Plag (2003), Gries (2004a), López Rúa (2004) and others have examined some of the phonological and prosodic characteristics of blends. One problem for the hearer or reader is where to divide the blend in order to identify the source word(s), while from the creator’s perspective, the issue is what to omit from the source word(s).

Plag (2003), who deals only with blends consisting of splinters, analyses the syllable (as do others) into the onset, nucleus (or peak) and coda, the last two forming the rime. In the case of monosyllabic blends, the onset of the first word combines with the rimes of the second. E.g., *goat* + *sheep* yields *geep* and *smoke* + *fog* yields *smog*.⁶

Polysyllabic blends can follow the same patterns, but in addition, syllables can be deleted or combined (Plag 2003: 123–124). In Plag’s examples, the number of syllables in the blend tends to be the same as the number of syllables in each source word. If there is a difference, the blend has the same number of syllables as the second word. Gries (2004a), using a more inclusive corpus, found that if

6. Others have found the same patterns. Kubozono (1990) found that in speech error blends, the division is between the onset and rime in about 75% of the corpus, and in word formation blends in 71% of the items. In Gries’ corpora (2004a) the division between the onset and the rime is 4 times more common than between the peak (nucleus) and the coda. See also Kelly (1998).

the number of syllables in the first and second source word differs, the number of syllables in the blend equals that of the second source word in only 55.7% of the items.

5. Orthographic blends

One important distinction in classifying blends is whether they are spoken or written. Although most blends are identical in speech and writing, some can be processed as blends only when seen in written form. Examples are *shampagne* < *sham* + *champagne*; *buyography* < *buy* + *biography*; *pursonality* < *purse* + *personality*; *awkword* < *awkward* + *word*; *fantesea* < *fantasy* + *sea*; *boisterous* < *boy* + *boisterous*. Conversely, in a blend like *eracism* < *erase* + *racism* the pronunciation takes precedence over the orthography.

Sometimes typographic devices are used to call attention to a part of the blend, which also solves the problem of division, as in *ARTstravaganza* < *art* + *extravaganza* and *Eggs-quisite* < *eggs* + *exquisite* (headline for an article on a restaurant specialising in omelets). Occasionally symbols other than letters are used: *3-peat* < *three* + *repeat* ‘repeat three times’.

6. Blends with more than two splinters

Although most blends consist of two source words, as in the examples above, I have found a few with three source words: *Japornimation* < *Japan* + *porn* + *animation*; *skafrocuban* < *ska* + *Afro* + *Cuban* (music); *Intelevisiary* < *Intel* + *television* + *visionary* (headline from an article in *The Economist*, where complete overlap is used twice).

7. More on splinters

Once a blend is created, the splinter may be reused. *Frankfurter* gave rise to the blend *turkeyfurter* (something like a frankfurter made with turkey) by resegmenting the etymon *Frankfort* + *er* (in the style of Frankfurt). Subsequent blends include *chickenfurter* and *shrimpfurter*. In addition to *deskercize* we can find *jazzercise* < *jazz* + *exercise*. In fact, it may be impossible to tell when a blend is used by analogy with another blend or created anew.

Splinters often occur in groups. Shortly after Bill Clinton was elected president of the United States and brought in staff and advisors from Arkansas, a series

of blends based on his former state were created: *Arkanauts*, *Arkeology*, *Arkansiana*. Blends created by analogy with *Spanglish* < *Spanish* + *English* are *Japlish*, *Czechlish*, *Yidlish* for code-switching styles based on Japanese, Czech, or Yiddish plus English, and recently *globlish* or, more frequently, *globish* from < *global* + *English*. Coffee shops sell *mochaccino* < *mocha* + *cappuccino* and *frappuccino* < *frappé* + *cappuccino*. This last word also gives rise to *cappuchillo* < *cappuccino* + *chill* and *macchichillo* < *macchiato* + *chill* (cold coffee drinks). Examples from Adams (2001: 139) are *-ar* < *star* in *pulsar* ‘pulsating star’, *spinnar* ‘spinning star’, and *collapsar* ‘collapsing star’.

8. Splinters as bound morphemes

When a splinter becomes so common that people start using it frequently, it may lose its connection with the source word and can be considered as a morpheme in its own right. Of course, since there is a scale from a completely novel splinter to a completely conventional morpheme, the transition from splinter to independent morphemehood is a diachronic process. However, some clear candidates for morpheme status are *gate* < *Watergate*, (a)*holic* < *alcoholic*, and *thon* < *marathon*. These elements always appear in final position.

Gate < *Watergate* is used for political scandals involving a cover-up. *Gate* can be added to names of individuals involved (*Monicagate*), a place (*Irangate*), or a common noun *ratagate* ‘an attack on a major telephone company’. The etymology, based on the scandal during the Nixon administration, springs from the Republican party spying on the Democratic offices at the Watergate apartment complex in Washington, DC in 1972. Hundreds of *-gate* neologisms have since been coined, often referring to local scandals, and the words are forgotten as soon as the event blows over. A few examples are *Irangate*, *Contragate* (scandals under Reagan), *Whitewatergate*, *Monicagate* (scandals during the Clinton administration), and some more recent ones, *Fajitagate* and *Memogate*. *Fajitagate* refers to an episode when some undercover policemen beat up two men outside a bar, demanding their take-out fajitas. Criminal charges were brought against several high ranking San Francisco police officers for allegations of a conspiracy to obstruct justice. *Memogate* involved a story on the TV program *60 Minutes* questioning George Bush’s service records. The memos on which the program were based were apparently fake. Other languages have borrowed *gate* for their political scandals:

Pascuagate in French for an incident involving Charles Pascua and *Donaugate* in Hungary for a scandal concerning the Danube River.

(A)*holic*,⁷ a resegmentation of *alcohol-ic* can be added to a noun or verb to refer to people who overindulge in a substance or in the activity named to the extent that they can be considered addicts.⁸ Early blends include *workaholic* and *chocoholic*, which are listed in unabridged dictionaries. Other examples I have collected include *foodaholic*, *drugaholic*, *readaholic*, *shopaholic*, *drivaholic*, *spendaholic*, *milkaholic*, *writaholic*, *reggaeholic* and *typeaholic*. Patrick Hanks from the OED kindly supplied me with many more examples, a few of which are: *aquaholic*, *controlaholic*, *faxaholic*, *golfaholic*, *helpaholic*, *hockeyholic*, *holidayholic*, *huntaholic*, *leisureholic*, *operaholic*, *plasticsurgic-holic*, *rageaholic*, *sexaholic*, *smokaholic*, *soapaholic*, *social-holic*, *sportsaholic*, *videoholic*, *vodkaholic*, *votaholic*, *waste-aholic* and *wordaholic*.

Thon is used to refer to events that last a long time, imply some difficulty, and may also involve fund-raising for a good cause. The etymon is drawn from *Marathon*, the place in Greece from where a messenger ran to Athens with news that the Greeks had defeated the Persians. However, the most salient meaning today is for a foot race of 42 kilometers (26 miles). The earliest use of *thon* was *telethon*, recorded by Merriam-Webster and the OED in 1949. Other examples are *walkathon*, *bikathon*, *jogathon*, *thrashathon*, *roackathon*, *dancethon*, *paint-a-thon*, *performancethon*, *pianothon*, *poolathon*, *rockerthon* (jobless contestants compete for cash prizes that will be awarded to the person who can keep a rocker going the longest), and *swimathon*. As with *holic*, an *a* is sometimes inserted for prosodic reasons and hyphens may be added to clarify the segmentation. This form has also spread to other languages. In France a four-hour long TV program on AIDS (the French acronym is *SIDA*) was called a *SIDathon*.

Some other examples of splinters that reappear in several blends follow, but it is debatable whether these deserve morpheme status (and in fact it may not be theoretically important).

-*Licious* < *delicious*: *applicious*; *bubblicious* (from a bubble gum ad); *piglicious* (item on a menu at a restaurant called Hog Heaven); *gobblicious* (a cat food flavor from Meow Mix®); *berry-licious* (on a fresh strawberry package); *Hoge Day-licious* (an exhibit by artist Hoge Day). In *cerealitious* spelling predicts < *cereal* + *nutritious*, but the pronunciation evokes *delicious* as well.

7. The OED lists both *-holic* and *-aholic*, while Merriam-Webster lists only *-aholic*.

8. An extended use of *addict* is used in glosses. The X-*holic* is not physically dependent on the substance or activity, but he or she devotes an unhealthy amount of time, energy, and importance to X, sometimes to the point of ruining or endangering things that other people would consider more important in life. Often the term is used humorously or as an exaggeration.

-*Wich* < sandwich: *bagelwich*, *croissanwich*; *Kempswich*[®] (Kemps ice cream sandwich).

-*Umentary* < *documentary*: *mockumentary* (satirical play or film in the form of a documentary); *rockumentary* (documentary about rock music); *soapumentary* (soap opera documentary).

-*Tel* occurs in two different sets of blends: 1) < *hotel* and 2) < *telephone*. In the first set *hotel* gave rise to *motel*, which for many younger speakers is not even perceived as a blend. Other *tel* blends are *boatel* (a hotel reached by boat or a hotel on a boat), *Hometel* (name of a hotel with home-like suites, and *skytel* (a hotel near the airport). The splinter often occurs in proper names like *Westel*, *Astrotel*, and *Sofitel*. Members of the second set have provided us with *Nextel*[®], *Nortel*[®] and another *skytel* (telephone via satellite). The Home Telephone Company is listed on the Internet as *Home Tel*.

-*Opoly*. When the patent on the popular board game Monopoly expired, the company started making versions of the game for universities, with the name of the Institution or the name of the sports team as the first element: *Harvardopoly*, *Yaleopoly*, *MITopoly*, *UCLAopoly*, *Arizonaopoly*, *Beaveropoly* (Penn State), *Soon-eropoly* (Oklahoma), *Irishopoly* (Notre Dame).

Two initial splinters that have acquired bound morpheme status are *Mc-* and *e-*. *Mc* < *McDonald's* was used in many of the corporation's own products, such as *McMuffin*, *McNuggets*, and *McParking* (for a McDonald's parking lot). But it is also used to indicate "inexpensive, convenient, or easy but usually low quality or commercialized version of something" (MW Online). Most neologisms with *Mc-* have unfavorable connotations. *McMansions* are large assembly-line houses in new developments; *McJobs* refers to low-paying employment, no or few benefits, and little hope of advancement. Environmentalists Helen Steele and Dave Morris created *McGreedy*, *McDanger*, and *McGarbage* and were then sued by the McDonald Corporation for libel. An article describing that lawsuit was titled *Mclibel* (Guttenplan 1996), which became the title of a movie about the event. Other recorded words are *McDoctor*, *McBook*, and *McTherapy*.

Another very popular initial splinter is *e-* < *electronic* (Safire 2003), occurring in dozens of items: *e-commerce*, *email*, *e-tail* (*electronic* + *retail*), *e-slammer* 'prison system with a web site' (*American Speech* 75.4:435), *e-bucks* 'electronic money', *e-Bay*[®], and *e-celebrity* 'a famous person promoting an Internet company', (*American Speech* 76.1:86).

An initial splinter that occurs in more than one blend is *docu-*, which appears in *docudrama* and *docu-opera* < *documentary* + *soap opera* (a variant of *soapumentary*, cited above). However, this does not yet appear to be productive enough to count as a morpheme. A variation is *jockudrama* in the title of a review of Spike Lee's film 'He Got Game.'

9. What kind of morphemes are these splinters?

There is some disagreement in labeling splinters like *Mc-* and *-thon*. The OED-Online classifies *-gate*, *-holic*, and *-thon* as suffixes, but *Mc-* and *e-* as combining forms. The Merriam Webster Online labels *e-* as a prefix and the other four as combining forms. There are compelling reasons for considering them to be combining forms and not affixes. Since this issue has been discussed in depth elsewhere (Bauer 1983, 1998; Marchand 1969; Adams 1973; Stein 1977; Warren 1990; and Lehrer 1998), I will not go into the criteria for distinguishing affixes from combining forms. What is critical for the discussion here is that combining forms are bound bases that combine with full words or with other combining forms. The most common and typical combining forms are those appearing in neo-classical compounds like *psychology* or *sociopath* which consist of two combining forms. Bauer (1983) classifies combining forms as either first or last elements, but some can function in both positions. For example, *path* occurs initially in *pathology* and finally in *psychopath*; *anthrop* is initial in *anthropology* and final in *misanthrope*. A few neoclassical combining forms have become clipped in certain contexts, such as university jargon (*psych*, *bio*, *antho*).

Some combining forms are derived from the Germanic part of the English vocabulary, such as *scape*, from *landscape*. Examples include specific kinds of landscapes such as *mountainscape*, *forestscape*, *moonscape*, *bridgescape*, *cityscape*, and *streetscape*. Many *-scape* words are names of paintings or kinds of paintings, such as *abscap* < *abstract* + *landscape*.⁹ There are further semantic extensions beyond visual phenomena which have produced neologisms such as *mindscape* and *soundscape*.

Besides the combining forms that occur only in combination with words or other combining forms, there are some that are homonyms of independent words and are semantically and etymologically related to them, but which have a special meaning. Some examples are *speak*, *fare*, *cast*, and *quake*.

Speak refers to a kind of jargon, derived from *newspeak*, coined by George Orwell in 1984. Neologisms that have been formed include *doublespeak*, *geek-speak*, *nounspeak*, *sportspeak*, *blandspeak*, *computerspeak*, and *blogspeak* (*blog* < *web* + *log*). A couple of dated items are *Haigspeak* (based on Alexander Haig's language) and *Valspeak* < *Valley Girl* (a song whose slang e.g., *grody to the max*, has also disappeared).

Fare, meaning 'public assistance to the poor', comes from *welfare* but has generated *workfare* and *learnfare* (recipients must work or be enrolled in learn-

9. Keith Lehrer uses this term. See www.u.arizona.edu/~lehrer/ga.htm.

ing programs to receive benefits). *Cast* < *broadcast* has provided us with *telecast*, *sportscast*, *colorcast*, *simulcast*, *narrowcast* (broadcast to a small audience), *dogcast* (broadcast information about dogs) and recently *podcast*, < *i-pod* + *broadcast*. *Quake* < *earthquake* has led to *moonquake*, *seaquake*, and *Timequake* (a novel by Kurt Vonnegaut).

10. Burgers

Burger is a resegmentation of *hamburger*, named after the German city of Hamburg, just as *frankfurter* is named after Frankfurt and *wiener* after Wien (Vienna). It has given rise to dozens of blends: *cheeseburger*, *chiliburger*, *baconburger*, *Californiaburger* (all of which are hamburger plus something else served on a bun) and *fishburger*, *chickenburger*, or *veggieburger*, where the meat is replaced by other ingredients. *Burger* is unusual in that it has become an independent word, not just a combining form like the other items cited above. Since *-thon*, *-holic*, and *-gate* have definite meanings, they could in principle become free forms, too, but this has not happened.¹⁰ It appears that most splinters tend to remain bound morphemes.

11. A note on the semantics of combining forms

Plag (2003:122) notes that in the type of blends he classifies as shortened or abbreviated compounds, the first element modifies the second. Thus a *mocamp* < *motor* + *camp* is a kind of camp, and a *breathalyzer* is a kind of analyzer. In many cases the original source word for the second element of a blend serves as the hyponym for new blends with the same head (Lehrer 1998). Therefore subjects accept definitions like:¹¹

- (1) a. A mountainscape is a kind of landscape.
- b. Carjacking is a kind of hijacking.
- c. A croissanwich is a kind of sandwich.
- d. A swimathon is a kind of marathon.

10. Cynthia Allen (personal communication) reported that her child's school has a fundraising event every year, which is referred to as a *thon*. One could ask, "What kind of thon should we have this year?" A plant nursery near San Francisco is called *Scapes*.

11. Subjects were undergraduates in introductory linguistics classes. They were given extra course credit for their (voluntary) participation.

However, they reject definitions like

- (2) a. A skyscape is a kind of landscape
- b. A meatitarian is a kind of vegetarian.
- c. Warnography (< war + pornography) is a kind of pornography.

The difference is that in (1a) the meaning of *mountain* is included in that of *landscape*, while in exmple (2a), *sky* is in contrast with *land*. In (1b–d) the first element (*hi*, *sand*, *mara*) has either no semantic content or has a meaning that is totally unrelated. In (2b) and (2c) the contrast between the meanings of the first elements in the new blend (*meat* and *war*) are semantically incompatible with the meanings of *vege* and *porn*.

12. Experimental data on the processing of novel blends

My past research on blends and combining forms has focussed on the reader's or speaker's processing speed and accuracy of identifying and interpreting novel blends. The details of this research are reported in Lehrer (1996) and Lehrer (2003). The following conditions appear to contribute to the identification of the source words making up the blend and to facilitate an interpretation:

- 1. context
- 2. the number and percentage of letters (or phonemes) of the source word present in the splinter
- 3. the frequency of the source words of the splinter
- 4. the number of neighbours of the source words (explained below)
- 5. the semantics of the blend

That context should facilitate identification is no surprise and, in fact, novel blends often occur in contexts which involve graphic as well as linguistic clues. In Lehrer (1996) two examples of contextual clues were taken from advertisements: *swacket* < *sweater* + *jacket* and *swingle* < *swinging* *single*. Both blends were hard to identify in isolation, but when presented in context, subjects identified the sources more easily.

SWACKET: The way to a Golfer's Heart. Our all-wool double knit...tailored swacket

SWINGLE: Hilton Swingles Week. We created a week for people like you.

The second factor – the amount of material measured in letters or phonemes – seems intuitively obvious. The more elements that are present, the easier identifi-

cation should be. I found a weak correlation between the number and percentage of letters of the source word present in the splinter and the identification of the source word, but this depends on the kind of blend. The positive effect of a greater number of letters is more evident with blends composed of two splinters.¹² In the case of complete overlap, e.g. *palimony*, there is instead a negative correlation. In fact, nothing has been deleted, but part of the blend has to be counted twice. In my earlier experiments, relatively few blends fit this pattern and subjects apparently did not consider the repetition of letters or syllables. However, since blends with complete overlap have increased in frequency, perhaps results would be different now.

A third factor that contributes to identification is the frequency of the source words, but this interacts with the fourth factor, which is the number of neighbors that the source word has. In the literature on lexical retrieval, a neighbor is defined as a word that differs from the target word by one letter or phoneme. For example, in the word *bland*, neighboring words are *gland*, *brand*, *blend*, and *blank*. Words with no or few neighbors are identified more quickly than those with many neighbors. I use *neighbor* in a slightly different way. A *neighbor* is any other word that could be a possible source for a splinter because it contains the same letters (or phonemes) as the source. In a blend like *applicious*, any word ending with *-cious* would be a neighbor. Blends are easiest to identify when the frequency of the source word is greater than any of its neighbors.

Finally, the semantic connection of the source words plays a significant role in identification. In *applicious*, *vicious* is a more frequent neighbor than *delicious*, but *vicious apple* is much less plausible than *delicious apple*. *Swacket* was difficult to identify, but subjects who got *jacket* could correctly guess *sweater* because the two words belong to the same semantic field.

Other timed experiments, reported in Lehrer and Veres (2000), were carried out in order to determine whether blends could be identified and interpreted as quickly and automatically as other kinds of compounds. In the first set of experiments the subject had to respond as quickly as possible when a blend was presented on a computer screen. Some of the same correlations were found as in the untimed experiments.

The second set of experiments was a lexical decision task, using a masked prime. A masked prime is a prime that is presented for a very short time (usually 50 ms) before the target to which the subject must respond, the target being a string of letters. The subject is asked to push the YES button if the stimulus is a

12. See Gries (2004a and 2004b) for quantitative data on much larger corpora.

word and the NO button if it is not. Subjects were presented with the following on the computer screen:

a string of hatch marks	#####	500ms
a masked prime	fruitopia	100ms
a target word	##FRUIT##	1000ms

The hatch marks call attention to the screen. A masked prime appears for such a short time that most subjects do not even see it and are unable to read it; however, its appearance has a demonstrated effect (Forster 1985). Subjects who are presented with a prime respond significantly faster to a target word that is either identical, morphologically related (*work-worker*), or semantically related (*doctor-nurse*). There were three conditions for words. In the first condition, the masked prime was a blend word (as above). In the second condition the masked prime and the target were identical. In the third condition the masked prime was unrelated, such as *chair* for the target *fruit*. Subjects saw 126 items, half of which were words and half non-words.

The prediction was that the targets with an identical prime would be identified correctly faster than the others, while those with blend primes would be next, and recognition time for targets with unrelated primes would be the slowest. The order was as predicted, but an analysis of variance showed that the time differences were not significant at the .05 level.

13. In which domains do blends tend to occur?

The commonest places for blends to occur are in product names, advertisements, newspaper and magazine headlines and titles. Blends are found primarily in print, but they can easily be produced in speech, and people do sometimes coin blends when speaking. What these genres and texts in the printed domain have in common is that they are all attempts to catch the attention of the reader. Since those of us in literate societies are surrounded by an excess of stimuli competing for our attention, the media and commercial sectors try to get the reader's attention through novel and catchy means, one of which is the creation of neologisms that utilise recognisable parts of words, often with intentional spelling errors. When encountering a new blend, we may 'get it' immediately or it may require some effort to identify the source words and supply a plausible interpretation in the given context. The creator of such blends wishes to elicit a favorable response in the reader, but ultimately the goal is to call his attention to the product, news item, etc. so that he will respond in the desired way. In the case of a product name or an advertisement for a product, we are supposed to be motivated to buy the

product or at least remember the product name. In the case of news articles, we are expected to read the piece, but there is no expectation that we will remember the blend, and many of these novel creations are (and are intended to be) nonce forms – items produced for a specific context and occasion.

Below I have listed some major categories where blends occur most frequently. This is not intended as a scientific taxonomy. The classes are not mutually incompatible, and many items can be placed in more than one category. Some blends do not fit into any of these categories, and it does not seem productive to establish new ones for a single unclassifiable item.

Hybrids

Many of these are classificatory terms for plants and animals that have been bred or cultivated. They include the items classified by Plag as copulative (2003: 123). In general, the referent denotes something which combines properties of both sources. For example, a broccoflower has some characteristics of broccoli and some of cauliflower.

beefalo < *beef* + *buffalo*

broccoflower < *broccoli* + *cauliflower*

tangelo < *tangerine* + *pomelo*

gasahol < *gas* + *alcohol*

Texmati[®] < *Texas* + *Basmati* a hybrid of American long-grain and Indian Basmati rice, mostly grown in Texas

dramedy < *drama* + *comedy*

infotainment < *information* + *entertainment*

infomercial < *information* + *commercial*

edutainment < *education* + *entertainment*

walleyball < *wall* + *volleyball* ‘a game like volleyball that uses walls

pictionary < *picture* + *dictionary* ‘a dictionary with emphasis on pictures’

Kripgenrein < *Kripke* + *Wittgenstein* ‘Wittgenstein’s philosophy as interpreted by Sol Kripke’

Blends related to places

Mexicali < *Mexico* + *California*

Ameurope < *America* + *Europe*

Afropean < *Africa* + *European*

Floribbean < *Florida* + *Caribbean* (cuisine)

Vansterdam < *Vancouver* + *Amsterdam* (refers to tolerance for drug use)

Proper names of companies, stores, organisations, etc.

Successories[®] designs, makes, and sells self-improvement products
Preventronics[®] sells security systems
Shirtique < *shirt* + *boutique* store for shirts
Petcetera < *pet* + *et cetera* a pet store
Odditorium < *odd* + *auditorium* Ripley's Believe It Not Museums
Avant-Card < *avant-garde* + *card* a greeting card shop
cosmodrome < *cosmonaut* + *hippodrome* Russian space center
Massport < *Massachusetts* + *Transport*
Exersaucer[®] < *exercise* + *saucer* suspended seat and tray for tots
Felaway[®] < *feline* + *away* a product to stop cats from marking things with urine

Restaurants and brand names of food and drink

Kabob-Que < *kabob* + *barbecue* 'a restaurant featuring kabobs
veggie-Q < *veggie* + *barbecue* barbecued vegetables
Frutopia[®] < *fruit* + *utopia* a beverage
Wheatables[®] < *wheat* + *eatables* crackers
Conecept < *cone* + *concept* an ice-cream parlor
Count Chocula[®] < *chocolate* + *Dracula* a cereal
Nutrageous[®] < *nut* + *outrageous* 'a candy bar
Kempswich[®] < *Kemps*[®] (ice cream) + *sandwich*
Cinnebon[®] < *cinnamon* + *bon* (French for 'good'), but pronounced like *bun* a cinnamon bun
smoke-tacular, *smoke-stonishing* < *smoke* + *spectacular*, *astonishing* flyer for a Smokehouse Combo from Round Table Pizza[®]
grandscapes < *Grand Marnier*[®] + *landscape* an advertisement with a picture of a bottle of Grand Marnier on an artificial landscape
Coffee drinks based on cappuccino: *frappuccino* < *frappe* + *cappuccino* --> *cappuchillo* 'chilled cappuccino', *macchichillo* < *macchiato* + *chill*; *mochaccino* < *mocha* 'cappuccino with chocolate'
Mixed drinks with Red Bull[®] beer: *Bullgarita* < *margarita* (Red Bull + tequila), *chambull* < *champagne*, *Bullmeister* < *Red Bull* + *Jagermeister*
vodkatini, *crantini* < *vodka*, *cranberry* + *martini* cocktails

Titles, headings, captions and other names

Genethics© < *gene* or (*genetics*) + *ethics*: *Technical Intervention in Human Reproduction as a Philosophical Problem* by Kurt Bayertz
Volumetrics© < *volume* + *metrics* a book on dieting by Barbara Rolls

- Adventures of a Verbivore*© < *verb* + *carnivore* a book about words by Richard Lederer
- Affluenza (The Book)*© < *affluence* + *influenza* by J. de Graff, D. Wann, T. H. Naylor, & D. Horsey
- Pornocrates* < *pornography* + *Socrates* the title of a painting of a blindfolded nude woman; the painting has Greek letters
- Ghettoriginal* < *ghetto* + *original* the name of a dance group. (When pronounced, it is homophonous with *get original*.)
- Basebrawl* < *baseball* + *brawl* heading of a review of the film *Up for Grabs* directed by Spike Lee
- McLibel* < *Mc* (now a combining form) + *libel* film about McDonald's libel suit against Steele and Morris

Blends inspired by the Internet

The Internet and World Wide Web have given rise to numerous blends, some of which appear below.¹³

- blog* < *web* + *log* 'personal website full of commentaries'. This blend has become a verb, which has generated *blogger* and *blogging*.
- webliography* < *web* + *bibliography*
- netiquette* < (*inter*)*net* + *etiquette*
- netizens* < *net* + *citizens* or *denizens*

Miscellaneous blends

- dogbella* < *dog* + *umbrella* umbrella for a dog
- monokini* < *bikini* swimwear with only a bikini bottom
- tankkini* < *tanktop* + *bikini* swimwear with tank top
- snivilization* < *snivel* + *civilization* a story about Jerry Springer in *Rolling Stone* (14 May, 1998, p. 43)
- zombedy* < *zombie* + *comedy* a movie genre
- communostalgia* < *communism* + *nostalgia*
- adminstrivia* < *administrative* + *trivia*
- legaldygook* < *legal* + *gobbledygook* legal jargon, sometimes called *legalese*
- synthespian* < *synthetic* + *thespian* "an electronic actor ...a 3-D computer animator term for a human form that is used in a virtual reality world" (www.netlingo.com) or *vactor* < *virtual* + *actor*
- frenemies* < *friends* + *enemies* "people you're sometimes forced to spend time with but heartily dislike", coined by Jessica Mittford 1977

13. The Internet has given rise to many neologisms. A regular column in *American Speech*, *Among the new words*, devoted two issues to computer-related neologisms and the Internet in 1999 (73, 298–323 and 74, 403–425). (See López Rúa's paper in this volume.)

shrimply delicious < *shrimp* + *simply*

nocebo < *nocerer* (Latin 'to harm') + *placebo* a placebo that induces a bad feeling when the person taking it [eg. a drug] is told it may have negative side-effects

Art-O-batics < *art* + *acrobatics* performance by the Circus Center of San Francisco

14. Summary and conclusions

Novel blends have become increasingly common, so common that they should no longer be considered as a marginal word-forming device. Although most of the examples presented above are from the written medium, I have also picked up several from conversations, ad hoc creations for the occasion. These nonce forms usually disappear. In the case of many brand and product names, the original etymology underlying the blend may be forgotten or ignored, just as the metaphorical origin of a word or expression becomes a dead metaphor over time and is treated as a 'new' entry in the lexicon, losing its underlying metaphorical sense. However, speakers can usually retrieve the etymological source words of a blend, just as they can figure out the connection between the literal and metaphorical meanings of words. When a splinter becomes common, it can take on the status of a morpheme, and it behaves linguistically like a combining form. However, these morphemes tend to remain bound and rarely become free morphemes.

Some of the data I collected a decade ago involved less common kinds of blends such as those with complete overlap and embedded elements, but these have increased in frequency more recently. It would be interesting to replicate the earlier experiments to see if younger speakers, who are more accustomed to encountering blends, respond to novel blends with greater speed and accuracy.

Although blends can be found in many genres, they are most frequently used where the creator wants to call attention to something by using a novel word or a misspelling, as in advertisements, titles, newspaper and magazine captions and headlines, or in product and company names. In advertisements blends are often accompanied by pictures, photographs, or other graphic devices that facilitate identification of the source words, thus allowing the reader to provide a meaning.

References

- Adams, V. 1973. *An Introduction to Modern English Word-formation*. London: Longmans.
- Adams, V. 2001. *Complex Words in English*. Harlow: Pearson Education.
- Algeo, J. 1977. "Blends, a structural and systemic view". *American Speech* 52: 47–64.
- American Speech*. Journal of the American Dialect Society.
- Bauer, L. 1983. *English Word-formation*. Cambridge: Cambridge University Press.
- Bauer, L. 1998. "Is there a class of neoclassical compounds, and if so, is it productive?" *Linguistics* 36: 402–422.
- Forster, K. 1976. "Accessing the mental lexicon." In *New Approaches to Language Mechanisms*, R. J. Wales and E. W. Walker (eds), 257–287. Amsterdam: North Holland.
- Forster, K. 1985. "Lexical acquisition and the modular lexicaon." *Language and Cognitive Processes* 1: 87–108.
- Gries, S. T. 2004a. "Shouldn't it be breakfrunch? A quantitative analysis of blend structure in English." *Linguistics* 42.3: 639–667.
- Gries, S. T. 2004b. "Some characteristics of English morphological blends." *Papers from the 38th Regional Meeting of the Chicago Linguistic Society: Vol. II. The Panels*. M. Andronis, E. Debernport, A. Pycha and K. Yoshimure (eds), 201–216. Chicago, IL.
- Guttenplan, D. 1996. "McLibel." *Columbia Journalism Review*, July–August 1996, 13–15.
- Kelly, M. H. 1998. "To 'brunch' or to 'brench': Some aspects of blend structure." *Linguistics* 36.3: 579–590.
- Kubozono, H. 1990. "Phonological constraints on blending in English as a case for phonology-morphology interface." *Yearbook of Morphology* 3. G. Booij and J. van Marle (eds), 1–20. Dordrecht: Foris.
- Lehrer, A. 1996. "Identifying and interpreting blends: an experimental approach." *Cognitive Linguistics* 7.4: 359–390.
- Lehrer, A. 1998. "Scapes, holics, and thons: The semantics of English combining forms." *American Speech* 73.1: 3–28.
- Lehrer, A. 2003 "Understanding trendy neologisms." *Italian Journal of Linguistics/Revista di Linguistica* 15.2: 369–382.
- Lehrer, A and Veres, C. 2000. "Processing blends online." Unpublished manuscript, Tucson, Arizona.
- López Rúa, P. L. 2004. "The categorial continuum of English blends." *English Studies* 85: 163–171.
- Marchand, H. 1969. *The Categories and Types of Present-English Word-formation*. Munich: C. H. Beck.
- Merriam-Webster OnLine Dictionary <<http://www.m-w.com>>
- Oxford English Dictionary OnLine <<http://dictionary.oed.com>>
- Plag, I. 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.
- Safire, W. 2003. "The e-lancer eats a bagelwich." In *No Uncertain Terms*, 78–80. NY: Simon & Schuster. [First appeared in the *New York Times Magazine* Nov. 8, 1998.]
- Stein, G. 1977. "English combining forms." *Linguistics* 9: 140–148.
- Warren, B. 1990. "The importance of combining forms." In *Contemporary Morphology, Trends in Linguistics, Studies and Monographs* 49. W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer, & J. R. Rennison (eds), 111–132. Berlin: de Gruyter.

Lexical creativity in texts

Electronic communication

Keeping up with the times

Lexical creativity in electronic communication

Paula López Rúa

In the last decades mobile phones and the World Wide Web have revolutionised communication, providing fresh evidence that language is a dynamic instrument which can evolve with the times and adjust to new contexts. The purpose of this chapter is to investigate the extent of that capability by examining a corpus of linguistic material used in electronic communication, from ready-made text messages (*AFAIK*) to lexical creations like *WiFi* or *blog*. The analysis shows that electronic communication is a fertile ground for new vocabulary, since it uses productive word-formation processes which already exist in the language (such as compounding, derivation, conversion, or different types of shortening) and contributes to the popularisation of new items which may eventually enter the lexicon (for example, *to text*: “to send a text message to”). As regards the reasons behind the use of the above mechanisms, system-derived and user-derived motives are identified and discussed. The former refer to the need to speed up communication by saving time and space, and relate to the features and requirements of the message, the channel and the context. The latter combine ludic and sociolinguistic needs, and are in turn connected with the characteristics of the primary users of the system, in-groups that may be defined by age, occupation and interests.

1. Introduction

Technology has opened up new channels of communication which have become popular among young and old in record time. Having a mobile phone or a laptop is no longer the privilege of businessmen or scientists but simply a sign that one is keeping up with the times. The natural outcome of the universalisation of *digital jewellery* (high-tech handheld devices like pagers and cell phones) is that users approach electronic communication with different aims and degrees of expertise. Consequently, there are occasional users who benefit from the main asset of

the new “medium”¹ (i.e. speed), although they do not change the code: in other words, their e-mails and SMS messages do not greatly differ from conventional letters or notes. Others (amateurs or average users) may modify the code slightly so as to adjust it to their needs, thus creating subcodes to be used by in-groups with the aim of saving time and space while creating or reinforcing bonds. Lastly, there are expert users (adepts and specialists) sharing computer-related interests or jobs, for instance programmers, technicians, people in e-business, designers and regular chatters. These users develop their own jargon, their main motive being exclusivity, since by means of the jargon they strengthen their bonds and exclude outsiders.

There are many previously existing lexical items whose “sense-spectra” (Cruse 1986: 71) are enriched with added senses to refer to new realities developed in the world of electronic communication. Some of those derived senses have spread into the general language: even someone who is new to the Internet or to computers in general knows the meaning of items like *mouse*, *refresh* or *surf* in the context of computers, and novices in mobile phone communication will have little difficulty in deciphering abbreviated formulas like *txtme ASAP* or *RUOK?* Besides, there is a great number of pre-existing shortened items (particularly items constructed from initials, clippings and abbreviations) which have been simply adopted by the electronic code. Along the same lines, average Internet users may be familiar with the computer-specific meanings of *server*, *bug*, *cookie*, *bookmark*, or *hard copy*; however, they may not be so well acquainted with the meaning of *thread* (a chain of postings on a topic) or *flame* (an insulting or controversial message sent by e-mail or posted on a newsgroup) unless they are subscribers to newsgroups or mailing lists. Likewise, they will probably be at a loss as to the meaning of *hamster* (tailless mouse), *handle* (on-line nickname), *bit bucket* (the place where lost e-mails and documents end up), *skyscraper* (an on-line ad which is vertically taller), or *vanilla flavour* (standard hardware or software), since those items belong to the jargons of more specialised users, and they will be unable to make sense of items belonging to the jargon of ‘subcommunities’, for example people working for research laboratories of a certain company or university: thus *double DECKers* are married couples where both partners work for *DEC* (‘Digital Equipment Corporation’), and *to be hosed* means to be down (i.e. not operating) in the jargon of Unix programmers.

1. Computer networks are, according to Herring (2001: 614), a new “medium”, in between speech and writing but different from both. In this respect, see also Collot and Belmore (1996: 14) or Crystal (2001: 28). As used here, *medium* comprises all types of electronic devices that allow text-based exchanges.

Linguistic specificity and geographical variation are other aspects which should be mentioned when discussing electronic communication. Although Internet exchanges have certain conventions and formulas which are respected across languages (the rules of netiquette), each language exhibits its own particularities, especially when it comes to shortening mechanisms. Therefore, as described in López Rúa (2005), English text messaging favours initialisation and tends to abbreviate words by using letter and number homophones, whereas Spanish, a language with different phonotactic constraints and a tendency towards polysyllabism, often resorts to abbreviation by letter reduction (particularly consonant elision). Local varieties must also be taken into account: the distinction between British and American English or between American and Commonwealth English also extends into electronic channels. Thus, for example, American and British Hackish exhibit lexical differences: *hamster* is a tailless mouse, but in British hackish it is also any item of hardware made by Amstrad; similarly, the word *proglet* is only used by British programmers to refer to a short program written to meet an immediate but temporary need.

The aim of this chapter is to examine lexical creativity in text-based electronic communication via mobile phones and the Internet, since, as Crystal (2004: vii) points out, terms initially belonging to these fields “have begun to make their appearance in the language as a whole” (though usually as part of slang or informal language). One example is the verb *click*, i.e. ‘to press a key on a computer mouse’: outside computing, this term is used as a noun meaning ‘very short distance’ (as in ‘it is just a click away’).

Although there is obviously more lexical innovation in the jargons of specific groups of Internet users (hackers, web designers, computer scientists and technicians, programmers or people involved in e-commerce), it was felt that a thorough description of particular jargons would be both overly lengthy and impractical, as these are restricted codes with numerous terms which do not come into general use. The chapter thus offers a survey of common methods and sources of lexical creativity and illustrates them with examples taken from different jargons, particularly focussing on those items which have achieved mainstream currency (*blog*, *adware*, *e-comm*, etc.).

The analysis of mobile phone communication via text messaging has been chosen as the starting point in the description of lexical creativity in electronic contexts for two reasons: first, because the code is less sophisticated, as it is fundamentally devised for informal exchanges on a small scale; second, because lexical innovation is far more restricted than in on-line communication, since it is basically a system of shorthand. Besides, it must be noted that the abbreviations and shortening mechanisms resorted to are also used on the Web, particularly in e-

mails and different types of on-line chat (Internet Relay Chat or *IRC*, newsgroups, mailing lists, bulletin boards, etc.).

2. *LokinIn2Txtin*: Devices and trends

Text messaging – also called texting or *p2p* (person-to-person) messaging – is the activity of sending and receiving short electronic messages via handheld devices, basically cell phones but also pagers or PDAs ('Personal Digital Assistants', i.e. handheld portable computers). Text messages can also be exchanged on-line in real time by means of IM ('Instant Messaging'). Although SMS ('Short Message/Messaging Service') was originally created to allow operators to inform their customers about the network, over the last few years it has become the main channel of communication for a whole generation of mobile phone users. Typing on a mobile keypad is convenient in many ways: it is fast, cheap and discreet, it can help people express things that they would not dare say face to face, and it can even contribute to the reinforcement of group bonds. All that is required is to stay within the limits of around 160 characters per message; for this reason words are multifariously abbreviated and combined with symbols intended to express feelings or tones, the so-called emoticons.

2.1 Shortening devices in text messaging

Since mobile phone texting is severely constrained by space requirements, shortening devices are the main sources of lexical innovation. However, not all shortened items can achieve the status of new words because very often they are just convenient written abbreviations. In contrast with what happens in on-line communication, productivity at the level of major word-formation devices is quite limited. Only a few cases of affixation (for instance *polytone*)² and compounding (*faceplate*, *handsfree*, *voice mailbox*), together with occasional instances of conversion (as in *to text* meaning 'to send a text message'), seem to have entered the lexicon, but most of these refer to technical specifications and cell phone accessories.

Following is a summary of the analysis of a corpus of abbreviated words, phrases and complete text messages in English, compiled from specialised books

2. Following Quirk et al. (1985:991) and Adams (2001:41, 116), *poly-* is regarded here as a "number prefix", although it could also be classified as a borderline case between the category of affixes and that of combining forms.

and articles, guidebooks, dictionaries and on-line glossaries (see References) and classified according to the shortening devices used. Out of a common core of abbreviated words and expressions in both mobile phone communication and the Internet, a selection was made so as to focus on those items most commonly used in the former,³ although in the case of initialised expressions there is a high degree of overlap between the two contexts. The Internet has undoubtedly furnished a number of shortened items used in texting (basically items built out of initials and combinations with letter and number homophones like *AFAIK*, *BTW*, *RUOK?*, *F2F*, or *CU l8r*, that is, ‘as far as I know’, ‘by the way’, ‘are you OK?’, ‘face to face’ and ‘see you later’) as well as conventions such as the avoidance of capital letters except for established items and for rules occasionally applied. As Crystal (2001:229) points out, this overlap may partly be due to the fact that the users – a “generally youthful population” – form quite a homogeneous group.

As regards the abbreviating methods used in mobile phone texting, the analysis shows that all the shortening mechanisms which are already established in the language are extensively used. This implies that, besides resorting to conventional abbreviations like *hr* (‘hour’), *fwd* (‘forward’) or *wk* (‘week’), conventional clippings such as *hols* (‘holidays’) or *sec* (‘second’), and existing initialisms (i.e. expressions reduced to initials) like *ASAP* (‘as soon as possible’) or *SWALK* (‘sealed with a loving kiss’), texting productively uses abbreviation, clipping and initialisation for the creation of new items. Some of them are rapidly attaining the status of accepted conventions judging by their frequency of occurrence in the sources consulted. A few examples of new but already popular items are abbreviations like *pls* (‘please’), *spk* (‘speak’) or *hv* (‘have’); clippings like *cell* (cellular phone) and *handy* (‘handheld device, i.e. cell phone, actually a combination of clipping and suffixation); and initialisms like *HAND* (‘have a nice day’), *PAL* (‘parents are listening’) or *TTUL* (‘talk to you later’). Initialisation is by far the most productive of these devices, which is not surprising since English has always employed the reduction to initials and is in fact one of the languages (or rather the language) bearing the highest number of initialised expressions. The result of the process may be an acronym (an initialism read in accordance with the rules of orthoepy such as *KISS*: ‘keep it simple, stupid’), an alphabetism (an initialism read letter by letter, like *TTFN*: ‘ta ta for now’), or an abbreviation composed of initials and read by retrieving the original expression it replaces, for example *AFAIK*). In texting, the abbreviations which are not composed of initials often result from eliding certain mid-letters (usually vowels and double consonants).

3. As the analysis was intended to focus on linguistic signs, emoticons were likewise ignored.

Apart from the shortening methods mentioned above, texting also resorts to phonetic respellings, homophones and a number of minor devices such as symbols or onomatopoeic renderings, all of them used either on their own or in combination with one another. Phonetic respellings are shortened homophones of genuine words (*wot* for 'what'), clipped words (*ne* /ni/ comes from respelling only the second syllable of 'any' /'eni/), colloquial contractions (*yer* for 'you're') or longer expressions (*wenja...?*: 'when do you...?'). Some of these are already existing items (*cos/ coz, luv*) while others are new creations to save space in writing (*afra*: 'after', *wen*: 'when'). They may even combine with other items to form anomalous compounds (*ne1*: 'anyone', *yaslf*: 'yourself'). The use of letter and number homophones is another popular shortening method in English text messaging; thus entire words or parts of words are replaced with numbers or letters which share the same pronunciation. Consider the following frequent correspondences and examples: 0 ('nothing'); 1 ('want', 'one', 'won-'), as in *I12CU* ('I want to see you'), *no1*, *1daful*; 2 ('to', 'to-', 'too'), as in *2gthr*, *up2U*, *2U2* ('to you too'); 4 ('for', 'for-', '-fore'), as in *4U*, *4(e)vr*, *B4*; and 8 ('-ate', '-eat'), as in *L8r*, *Gr8*, *m8*. The fact that English is a Germanic language undoubtedly favours the extensive use of this device, since many words of very frequent use in everyday language are actually monosyllables which happen to be easily replaceable with letters or numbers. Lastly, texting also resorts to symbols and onomatopoeic expressions such as *zzz* ('boring, tired') or *grrr* ('angry'). Symbols keep their established uses or acquire new ones, although the rules of equivalence are not consistently applied, for example: @ ('at', as in *@wrk*: 'at work', *w@*: 'what'), \$ (which sometimes replaces double *s*, as in *M\$U*: 'miss you'), or % (which replaces '-oo-' or '-oul-': *C%l*, *W%d*).

As a final remark, it should be added that a source of difficulty for encoding and decoding text messages is the possibility of applying different shortening methods to the same word or expression with different results, for instance, *Ya*, *U*, *Y*: 'you'; *som1*, *sum1*, *sme1*, *some1*: 'someone'; *wu?*, *sup?*, *wassup?*, *wuzup?*: 'What's up?'; or *4e*, *4evr*, *4vr*, *4eva*: 'forever'.

2.2 Trends in text messaging: 'conservative' and 'advanced' texting

The last point to be made with respect to text messaging is the existence of two parallel trends as regards shortening rules, each of them resulting in a different type of text talk. The one that could be described as 'conservative' resorts to the basic principles already mentioned, although the few existing rules of shortening are not strictly applied. In fact, users are allowed a certain degree of freedom to create their own rules or to abbreviate words on the spot by using their intuition and creativity, as long as the message remains intelligible for both the sender and the receiver.

Capital letters are usually restricted to initialised items and letter homophones, and spacing between words is generally maintained except in ready-made expressions of frequent use (*Wan2tlk?*, *CU2moro*, *Jstcllme*). In contrast, 'advanced' texting is a more recent trend popularised by guidebooks and compilations apparently intended to reflect general use, although it actually looks more like an attempt to turn a particular code (that of a subgroup) into the norm.⁴ The most remarkable features of advanced texting are the omission of spacing, the expanded use of strokes with positive and negative meanings (*D/*: 'do', *D*: 'don't'), and the use of capital letters not only for initialised expressions and letter homophones, but also to indicate the beginning of a word (that is, to replace spacing), the existence of a double consonant, a double vowel or a long vowel sound or diphthong, as shown in these examples: *BeTa*: 'better'; *FrE*: 'free'; *GOd*: 'good'; *MAk*: 'make'. Capitalisation thus becomes a particularly complex shortening rule which, together with the absence of spacing and the fact that letters and numbers may coincide in shape (for example, capital *I* and small *l*, or capital *O* and zero), can considerably slow down the process of encoding and decoding. Consider the following examples of shortened sentences in advanced texting: *2DAsYaLkEDA*: 'today's your lucky day'; *NETImNEPlAc*: 'any time, any place'; or *ThAAntCEn0Yt*: 'they ain't seen nothing yet'.

As can be seen, advanced texting takes the principle of space saving to extremes, with the result that messages are actually shortened but their intelligibility is sometimes put at risk. On the other hand, the prescriptivism embodied in advanced texting might be facing the users' rejection because, when SMS started to catch on, it was precisely the absence of rules that enhanced the appeal of texting in the eyes of its main users (i.e. teenagers), a group traditionally reluctant to follow any kind of imposed rules. Text messaging succeeded because, starting from a few general guidelines, users were allowed to develop their own private subcodes for in-group communication, a kind of personal language which would remain opaque to outsiders. It is, as Crystal (2001: 230) puts it, "a desirable safeguard against unwelcome surveillance" which reinforces the feeling of group belonging, the same feeling that underlies the rise and expansion of institutionalised initialisms and abbreviations in specialised fields such as science and technology, the military or computing. If rules are universalised, the sense of belonging to the in-group may disappear.

4. The issues of 'prescriptivism' and 'proscriptivism' on the Net are extensively discussed in Crystal (2001) and Hohenhaus (2005), their conclusions being also applicable to texting. Crystal accurately remarks that guidelines and regulations in on-line communication may be community-biased, representing a "partial or prejudiced view" (2001: 68). This also seems to apply to text messaging. Moreover, in relation to style-guides on on-line usage, both Crystal and Hohenhaus observe that users seem to be more influenced by the way other users interact than by recommendations or rules, which is also true for SMS communication.

In any case, whatever shape it may adopt, the advantages of text messaging make it unlikely to go out of fashion, at least in the short term. Although, as already mentioned at the beginning of the section, lexical innovation is quite limited in this mode of electronic communication, texting arises as a valuable testing ground for the productivity of the shortening devices of the language. Moreover, its role in the popularisation of new or existing abbreviated items cannot be ignored, since those items may eventually replace their expanded equivalents in everyday language, thus qualifying as cases of lexical creativity. For example, the Spanish clipping *finde* (from the noun phrase ‘fin de semana’: ‘weekend’) is nowadays popular outside electronic channels due to its extensive use in texting.

3. The world of *Netspeak*:⁵ *hackers*, *crackers*, *bloggers*, *Usenetters*, *websters* and other *netizens*

The vocabulary, word-formation devices and written conventions which will be described in the following sections comprise the contributions of various communities of computer users, from ordinary e-mail writers or on-line chatters via IRC⁶ to the elitist world of hackers. Leaving aside those professionals engaged in computer technology (language programmers, engineers, system operators, designers of Web pages or e-businesspeople), Internet users can be divided into as many communities as there are technologies for on-line communication.⁷ Thus, for example, *MUDders* are regular users of *MUD* (the ‘Multi-User Dimension or Dungeon’, a simulation environment with a programming language which allows the textual construction of characters and fantasy worlds), whereas *Usenetters* are basically contributors to discussion boards or newsgroups. It is also necessary to distinguish between *hackers* and *crackers*, two groups defined by their interests and often confused. Hackers are computer experts who enjoy exploring the capabilities and limitations of programs, whereas crackers devote themselves to breaking security on systems. Besides, each jargon has distinctive features; in the case of crackers, for example, there is wide use of the *k-* prefix or terms borrowed from skateboarding and underground rock slang.

5. This is a general term for ‘Internet language’. Following Crystal (2001), it is used as an alternative to ‘electronic language’ or ‘CMD’ (‘computer-mediated discourse’).

6. IRC is a synchronous mode of CMC or computer-mediated communication (see Werry 1996).

7. For a thorough account of technical specifications, discursive features and layout conventions in different “Internet situations”, see Crystal (2001). Alternative classifications of “genres” or “modes” of CMC are offered in Herring (2001: 615) and Wood and Smith (2001: 9ff.).

The term *jargon* will be consistently used throughout this chapter to refer to the languages of all these communities, since the single varieties do not necessarily present a distinctive style, even though each contains a great number of informal terms. In accordance with the “narrow” definition of register provided by Holmes (1992), these are different registers rather than slang, inasmuch as they are varieties used by groups of people with common interests or occupations:

Registers, on the other hand, when they are distinguished from styles, tend to be associated with particular groups of people or sometimes specific situations of use. Journalese, baby-talk, legalese, the language of auctioneers, race-callers, and sports commentators, the language of airline pilots, criminals, financiers, politicians and disc jockeys (...) could all be considered examples of different registers. (Holmes 1992: 276).

The items to be examined have been drawn from the same sources as those used in mobile phone communication (see Section 2 and References). The origins of many lexical items in computer jargons can be traced back to a miscellaneous collection of sources, all of them related to the users’ specific culture, interests, values and experiences. The shared background which becomes apparent through language strongly contributes to reinforcing the feeling of belonging to a community. Some of the sources from which terms are drawn or developed are listed below:

1. Commands and instructions in programming languages, for example, *cdr* (/’kudr/: ‘Contents of Decrement Part of Register’) is a verb taken from the programming language *LISP* which means ‘to skip past the first item from a list of things’.
2. Computer games, for instance, *ogg* is ‘to execute kamikaze attacks against enemy ships’ in the game *Netrek*. It comes from *OG*, or ‘Orion ship G’, the initials of the ship of the player who first used the trick. Outside gaming, it means ‘to do anything forcefully’.
3. TV programmes such as quizzes or variety shows. For example, a *cookie monster* (from the TV programme ‘Sesame Street’) is a program which locks up the user’s terminal or console asking for a ‘cookie’ until the right answer is provided. Another example is the term *spam* (see below).
4. Cult films and series. For instance, the famous phrase ‘Beam me up, Scottie!’ in *Star Trek* gave rise to *beam* meaning ‘to transfer a soft copy electronically’; similarly, *Star Wars* is the source of *UTSL* (‘Use The Source, Luke’, meaning ‘consult the source code’ in order to solve a problem).
5. Advertising. This includes trade names and slogans or words from ads, comic strips and cartoons: for example, *for the rest of us* (from the Macintosh slogan ‘The computer for the rest of us’) sarcastically describes a limited program

designed so as not to confuse naive users. Similarly, *Spam* was the trademark of a type of canned meat launched in the 1930s. According to the manufacturers, its original name ('Hormel Spiced Ham') was replaced by *Spam* (a blend of 'spiced' and 'ham') as a result of a naming contest. Due to the lack of fresh products, the public was reluctantly compelled to consume it very frequently in World War II and also in the postwar period. This situation provided the British comedy group Monty Python with the context for their famous *spam* sketch on a TV series first broadcast in 1969. This sketch, set in a café in which every item in the menu included *spam*, is commonly cited as the origin of the internet sense of the word, i.e. 'unwanted e-mails sent to many recipients'.

6. Books, particularly programming manuals and fantasy literature. *The Great Worm*, for example, comes from the dragons featuring in Tolkien's books (a worm is a program which propagates and reproduces itself over a network). Similarly, the terms *jack in* and *jack out* (to enter or leave a virtual reality simulation) come from cyberpunk (a type of science fiction).⁸
7. Some terms are coined by programmers, hackers, computer game designers and other net personalities, while others are borrowed from the slang of radio hams (*NIL*: 'no'), and scientists or engineers (the computer use of *hairy* meaning 'difficult' derives from the same meaning in the term 'long-haired', which is technical slang). Further sources of loanwords are specialised fields like the army (*newbie*, i.e. 'new boy' is a pejorative term for a new – and clumsy – Internet user), or skateboarding (for instance, *lamer* is a common term among crackers and phreakers – phone crackers – meaning 'would-be cracker' or 'slapdash phreaker'. It is the equivalent of *luser*: 'loser + user' among hackers, having the same elitist connotations).
8. Lastly, some terms even originated from typographical errors: for example, *scrool* comes from *scroll* and refers to logging old messages so as to catch up with an on-line conversation.

3.1 Word-formation devices in on-line communication

As Algeo (1991: 14) points out, the need for new words arises when "there are new things to talk about" or when "we want to talk about old things in a new way." That is the pragmatic motivation of neology. But in addition to being an instrument for communication and conceptualisation, language is also "a field for play", and new words function as an outlet for human wit and humour. Moreover, shared words (whether old or new) contribute to creating a sense of identity and, as long as they

8. Munat (this volume) provides a detailed treatment of lexical creativity in science fiction.

enjoy a limited currency, serve to identify the members of a group while excluding outsiders, thus reinforcing group bonds.⁹ As will be shown in the following paragraphs, pragmatic, ludic and sociolinguistic needs permeate lexical creativity in all manifestations of electronic communication. An excellent overview of the productivity of “lexical processes” in Netspeak can be found in Crystal (2001:82ff.).

3.1.1 *Affixation (lossage)*

When dealing with affixation, three processes must be distinguished:

1. The productive use of existing affixes, for example *-able*, *-let* or *-oid*.
2. The creative use of existing affixes. This process comprises meaning extensions such as *-age* in *lossage* (see below), and the attachment of existing affixes to unusual or unorthodox bases, such as *re-* in *rehi* (‘hello again’, i.e. *re-* + interjection) or *dis-* in *disusered*.
3. New affixes proper, for instance, *net*. (read ‘net-dot’) in *net.party*.

Hacker jargon exhibits a tendency towards deviating productivity as regards derivational morphemes, particularly *-ity* and *-itude*. Therefore, many adjectives ending in *-ous* (*mysterious*, *obvious*, *dubious*) whose canonical abstract nouns would add *-ness* (*mysteriousness*) choose *-ity* instead (*mysteriosity*, *obviousity*, *dubiosity*). In agreement with Aronoff (1976), Adams (2001:32) points out that adjectives ending in *-ous* “take *-ness* more readily than *-ity*.” In this case, deviation arises when *-ity* is made to replace its more productive ‘rival’ affix *-ness* (for an account of the ‘rivalry’ between *-ness* and *-ity*, see Adams 2001:32). The suffix *-itude* is similarly used to form abstract nouns out of adjectives, thus *lameitude* (instead of *lameeness*) from *lame*. This ending comes from the Latin suffix *-tudo* / *-tude*, which was also used to form deadjectival nouns: for example, Latin *longus* (‘long’) was the base for *longitūdō* (‘longitude’). Its canonical productivity is, however, quite restricted since it mainly occurs in words derived from Latin or French whose roots are often dependent formatives in English (*hebetude*, *fortitude*, *plenitude*). Another instance of deviating productivity is the possibility of obtaining deverbal nouns by adding *-age* or *-itude* to a large number of verbs. Therefore, *winnitude* (instead of *winning*) or *winnage* are derived from *win*, and *disgustitude* (instead of *disgustedness*) from *disgust*.¹⁰ As regards prefixation, among MUDders and Use-

9. A subtype of sociolinguistic motivation related to individual prestige before the community underlies the use of terms from computer technology in everyday conversation. Crystal (2001:19) provides examples like “She’s multitasking” (she is doing “two things at once”) or “He started flaming me” (“shouting at me”).

10. Cf. the emphatic overgeneralisation of *-ness* described in Hohenhaus (this volume).

Table 1. Prefixation in on-line jargon

Prefix	example of productive use	example of creative use	example of derived word with new prefix
<i>de-</i>	<i>deselect*</i>		
<i>dis-</i>		<i>disusered*</i>	
<i>hyper-</i>	<i>hyperlink, hyperspeed</i>	<i>hyper-win</i> (emphatic meaning)	
<i>inter-</i>	<i>interface</i>		
<i>mis-</i>		<i>misfeature*</i>	
<i>net.</i>			meaning ‘related to the net’: <i>net.party</i> , <i>net.lurker*</i>
<i>non-</i>		<i>nonlinear</i> , <i>non-optimal*</i>	
<i>over-</i>		<i>over-Mac*</i>	
<i>re-</i>	<i>reboot, reformat</i>	<i>rehi</i>	
<i>Super-</i>	<i>superhighway</i>	<i>super-win</i> (emphatic meaning)	

* *deselect*: to click a box or button after selecting it so as to cancel the selection; *disusered*: an adjective applying to a person whose account on a computer has been removed (the verb *disuser* is rarely used); *misfeature*: a bad property of a program; *net.party*, *net.lurker*: these items can be alternatively interpreted as compounds including the clipped element (*Inter*)*net*; *non-linear*: nervous, erratic, unstable; *non-optimal*: bad; to *over-Mac*: to oversimplify a product so that consumers view it as a toy.

netters it is common to use the exclamation mark as a negative prefix, for example *!clue* means ‘clueless’ or ‘no clue’. Similarly, crackers use the *k-* prefix (from the SI symbol *k* for ‘kilo’) with emphatic meaning, as in *k-cool* or *k-awesome*. On these premises, Tables 1 and 2 display some prefixes and suffixes (the list is by no means exhaustive) used in on-line jargons, together with a few illustrative examples.

With respect to inflectional morphology, there are also a few devices worth mentioning, such as the overgeneralisation of foreign plural morphemes in hacker jargon, particularly the Anglo-Saxon *-en*, the Hebrew *-im* and the Latin *-a*, as in *boxen*, *frobbozim* (from *frobbozz*, an unspecified physical object or electronic black box), or *fora* (instead of the more common *forums*). Also, in *warez*, *gamez* or *tunez* *-z* combines the meaning of ‘pirated version’ with that of the plural morpheme. Finally, turning a word into a question by adding the letter *-p* is another Hackish convention derived from *LISP*, thus *foodp?* or *foodp* stands for ‘are you hungry?’ or ‘would you like something to eat?’, and *HELLOP* appears as an opening formula in on-line chat. Although in *LISP* the letter is appended to denote a predicate, the added letter can be morphologically interpreted as an inflectional affix with an interrogative meaning.

Table 2. Suffixation in on-line jargon

Suffix	example of productive use	example of creative use
-able	<i>clickable, crackable</i>	
-age		Augmentative meaning: <i>lossage, winnage, screwage</i>
-er	<i>digitizer, flamer, gamer, netter, sniffer*</i>	
-ful	<i>clueful</i>	<i>flavourful</i>
-y /-ie	<i>jaggies, clippy, techie*</i>	
-ify	<i>webify*</i>	
-itude	<i>winnitude, wedgitude*</i>	
-let	<i>applet*</i>	
-oid	meaning 'poor imitation': <i>modemoid, computeroid</i>	
-ster	<i>webster*</i>	
-o		meaning 'error': <i>scanno, mouso*</i>

* *digitizer*: someone who turns printed material into a digital form; *flamer*: someone who sends flames; *gamer*: a player of electronic games, thus a particular type of player; *netter*: someone with a network address; *sniffer*: a program that steals and spreads passwords; *jaggies*: images with jagged edges; *clippy*: the paper clip assistant in Windows software; *techie*: a technician (here the suffix is attached to a clipped form); *webify*: to transform any kind of information in order to display it on the Web; *winnitude*: the quality of winning; *wedgitude*: the quality of being wedged or stuck; *applet*: a small application (clipping plus suffixation); *webster*: someone who is new to the Web. Note in this case that an already existing word (which is a proper noun and in principle unanalysable) is playfully reanalysed into a new base (*web-*) plus the suffix *-ster* meaning 'person engaged in or associated with an activity' – Quirk et al. (1985: 994) – as in *gangster, prankster, gamester* or *songster*, and thus the suffix becomes productive; *scanno*: error due to a scanning problem; *mouso*: error due to a malfunction in the mouse. In the last two examples, the last letter from the clipping *typo* (typographical error) is used as a suffix with the pejorative meaning 'error'. It could be interpreted as a meaning extension of the existing low-frequency suffix *-o*, which parallels the suffix *-y /-ie* (see Gramley 2001).

3.1.2 Shortening

3.1.2.a. *Simple shortenings: written abbreviations and related items (letter and number homophones, phonetic respellings and symbols: 4eva, phishing.* The reduction of phrases to initials with expanded or unexpanded reading is a handy resource in on-line communication. As mentioned in Section 2, there is an ample core of simple shortenings (i.e. used only in writing) which are shared by all modes of electronic communication. This common core comprises abbreviations involving letter reduction (*sth*: 'something', *ppl*: 'people'), initialisation (*BTW*: 'by the way'), phonetic respelling (*wuzup*, *4eva*) and also by resorting to letter and number homophones (*I h8 it, r u going*). However, some items are almost exclusively used on-line, such as those listed below:

1. Initialised abbreviations: *bcc* ('blind carbon copy', in e-mails), *BSOD* ('Blue Screen of Death'), *GA* ('go ahead', used when two people try to type simultaneously), *HHOJ* ('Ha ha only joking'), *JAM* ('just a minute'), *O* ('over to you'), *OO* ('over and out'), *ROFL* ('rolling on the floor laughing'), *SorG* ('Straight or Gay?').
2. Letter reductions: *ENQ?* ('are you busy?', it expects *ACK* or *NAK* in return), *HD* ('hold'), *KBD* ('keyboard').
3. Letter and number homophones: *B4Y* ('Before You'), *N-A-Y-L* ('in a while'), *P-ZA* ('pizza').
4. Phonetic respellings: *NUFF* ('enough said'). Some lexical items are created out of respelling, for example *phishing* and *pharming*, or *dot-com*. *Phishing* is the practice of acquiring confidential financial information by sending e-mails which direct victims to fake websites. It is a deliberate misspelling which plays on 'fishing', since the e-mail is the metaphorical bait which tricks the victim. *Pharming* is phishing on a large scale, and a *dot-com* (company) is a company with a .com Web site. Crackers often resort to misspellings based on phonetic similarity. However, in contrast with mobile phone communication and on-line chat, where respellings save space in writing, crackers misspell simply out of playfulness. Thus they write *dewd* or *d00d* instead of *dude*, or *troo* for *true*. Along the same lines, they tend to replace the letter *o* with the number *0* (*l0zer*). The playful use of abbreviated forms, which does not necessarily save space, is not exclusive to crackers: for instance, *I-D-L* ('ideal') is a common abbreviation on-line.
5. Onomatopoeic expressions are also used on the Net, for example *::poof::* ('I'm gone'), or *shhh* ('quiet').
6. As regards symbols, Usenetters and MUDders are particularly fond of using logical and relational operators. Other symbols commonly seen on-line are the caret ('^', in *w's ^*, or *wut^*, i.e. 'what's up?') and the tilde ('~', meaning 'approximately').
7. Lastly, it should be noted that on-line communication basically takes place through the written medium, whose limitations call for certain punctuation and emphasis conventions (see Crystal 2001: 87ff.). For instance, it is a well-known rule of netiquette to avoid writing a text all in uppercase, since it means shouting and it is considered rude (although crackers are actually identified for doing so).

3.1.2.b. *Letter-based complex shortenings: initialisms (acronyms: MUD and alphabetisms: BBS)*. Common on-line acronyms are, for example, *GIF* /*gif* ('graphics interchange format'), *plotka* ('press lots of keys to abort', i.e. to press random keys trying to get a response from the system), *SNERT* ('snotty nosed egotistical rotten

teenager', usually a teenage male in a chat room who offends or harasses other users), *cuspy* (an excellent program, from 'commonly used system program'), or *VAX* ('Virtual Address Extension', the name of a minicomputer whose plural is *VAXen*). Illustrative examples of alphabetisms are *BBS* ('Bulletin Board System / Service', a message database where users can log in and leave messages for others), *BRB* ('Be Right Back', at the keyboard), *CTS* ('carpal tunnel syndrome', a computer-related health problem affecting the hand and wrist), or *T* ('true', that is, 'yes'. Among hackers, it is the opposite of *NIL* or 'no'). *FAQ* ('Frequently Asked Questions') belongs to those items which can be read out as both acronyms and alphabetisms, and *JPG* /*jpg* ('joint photographic experts group', a type of image format, read out /'dʒeɪpeg/ exemplifies the items which combine letter-by-letter and orthoepic readings. Lastly, the combination of initials plus the number homophone 2 ('to') is a productive pattern in e-commerce, thus *B2B* ('Business to Business'), *D2D* ('Developer to Developer'), or *M2M* ('Manufacturer to Manufacturer').

Initialisation is a suitable mechanism to express playfulness, irony and political correctness. As a result, previously existing initialisms are often subject to humorous or witty reinterpretations, for example, *IBM* ('inferior but marketable', 'it's better manually', 'it's been malfunctioning'), *BFI* ('Brute Force and Ignorance', a play on *FBI* which describes primitive or inelegant programming style), *SOHO* ('small office, home office', which describes a small business headquartered at home), or *AIDS*, i.e. 'Apple/ Amiga infected disk syndrome', the result of unsafe *SEX* ('software exchange'). Ludic motivations also underlie new coinages patterned after popular acronyms: thus *ALAP* ('as late as possible') and *AFAP* ('as far as possible') are variations of *ASAP*; *WYSIWYG* gave way to *WYSIAYG* ('what you see is all you get') and *YAFIYGI* ('you asked for it, you got it'), whereas *yetties* ('young, entrepreneurial, tech-based twenty-somethings' or 'young entrepreneurial technocrats') are the modern version of *yuppies* ('young urban professionals'). Lastly, euphemistic motivations can be traced in *ID10T* ('idiot') and *I/O error* ('Ignorant Operator'), two terms used by technicians to refer to inexperienced users, or the military borrowing *PONA* ('person of no account', i.e. someone who is not, or has never been on-line).

3.1.2.c. *Syllable- and splinter-based complex shortenings: clippings (puter), blends (netiquette) and clipped compounds (blog)*. Clipping or truncation may apply to the beginning, the middle or the end of a word, and even higher units can be clipped (for example Spanish *finde*, see Section 2 above). Although some clippings eventually replace their sources, they are usually shorter alternatives used in informal register. The following are a few examples of the three main types of clippings registered on-line, i.e. hind, fore and mid-clipping (the first two denominations have been taken from Algeo 1991):

1. Hind clippings: *ACK* (*acknowledge*), *app*(*lication*), *bats* (*battery-powered transmitters*), *char*(*acter*), *gen*(*erate*), *nic*(*kname*), *param*(*eter*), *peeps* (*people*), *pix* (*pictures*), UK *vom*(*it*), US *parens* (*parentheses*)
2. Fore clippings: (*pseudo*)*nym*, (*com*)*puter*, (*ro*)*bot* (a program)
3. Mid-clippings: *prolly* ('probably')

Blending is a very productive word-formation device in on-line communication¹¹ although, as Bauer (1983: 236) points out, the limits of the category are not easy to draw, since the process "tends to shade off into compounding, neo-classical compounding, affixation, clipping and (...) acronyming." The first difficulty in setting borderlines arises when trying to discriminate between blends and clipped compounds, also referred to as "compounds of clipped elements" (Adams 1973) or shortened forms "subject to (...) compounding" (Adams 2001). For some authors (Algeo 1991) clipped compounds fall within the range of blends, although the following three defining parameters make it possible to talk about two continuous categories with fuzzy boundaries:

1. Syntactic-semantic structure: typical blends are copulative, for example *loser*: 'loser + user', or *automagically*: 'automatically + magically'; Adams (1973) calls them "appositional" or "coordinative", whereas Plag (2003) regards these two subcategories as essentially different. However, blends may also be 'headed' (their semantic head being inside or outside the compound), like *bracket*: 'broken bracket' or US 'angle bracket'. According to Plag (2003), headed blends are "abbreviated compounds" and not "proper blends". By contrast, typical clipped compounds tend to be headed (*webcam*, *navbar* – see below), and only a few display a coordinative structure, for instance *Alcan*: 'Alaska + Canada'.
2. Shortening: blends are typically constructed from 'the beginning of one word and the end of another' (Bauer 2004: 22), whereas clipped compounds exhibit other combinations, especially the beginnings of two or more words.
3. Constituent integration, i.e. sound overlap: there is overlap in typical blends, but not in clipped compounds.¹²

In view of the foregoing parameters, the following could be categorised as examples of blends used on-line: *anticipointment* ('anticipation + disappointment'), *boobiecon* ('boobie + emoticon'), *clickly* ('clicking quickly'), *crapplet* ('crap + applet'), *digiterati* ('digital + literati', i.e. knowledgeable about computers and the

11. An extensive treatment of blends is offered in Lehrer (this volume).

12. For a thorough description of different types of overlap, see Lehrer (this volume).

Internet), *emoticon* ('emotion + icon'), *fram* ('friend + spam', i.e. spam sent by friends), *hacktivist* ('hacker + activist'), *mouselexia* ('mouse + dyslexia', i.e. the inability to use a mouse correctly), *netiquette* ('net + etiquette'), *netizen* ('net + citizen'), *nouse* ('nose + mouse', a nose-steered mouse for disabled people), and *pagejacking* ('page + hijacking'). Combining forms may also be parts of blends, as in *featurectomy*, i.e. to remove a feature from a program (from 'feature + -ectomy') or *Internaut* ('Internet + -naut').

Clipped compounds display the following formation patterns:

1. Hind clipping + hind clipping: *HOHA* ('Hollywood hacker'), *sysadmin* ('systems administrator'), *telco* ('telecommunication company'), *Wi-Fi* ('wireless fidelity', coined after *Hi-Fi*).
2. Fore clipping + word: *blog* ('web log'). A blog is a Web site where users can post a chronological e-journal of their thoughts and opinions. Note the productivity of this pattern in *vlog* /*v-log* /*vidlog* /*vblog* ('video blog'), *moblog* ('mobile blog'), *sideblog*, *MP3blogger*, etc.
3. Hind clipping + word: *digicash* ('digital cash'), *favicon* ('favorite icon'), *navbar* ('navigation bar'), *p-book* ('printed book', as opposed to *e-book*), *v-Net* ('voice net', i.e. the phone system), *techspeak* or 'technical speak'. In the last example, *-speak* is rather a combining form (see Lehrer, this volume) or even a "suffix-like" element (Algeo 1991).
4. Word + fore clipping: *chatfly* ('chat barfly'), *geekosphere* ('geek biosphere'), *textitecture* ('text architecture').
5. Word + hind clipping: *newsadmin* ('news administrator'), *webcam* ('web camera').

As can be observed, clipped compounds exhibit different formation patterns, degrees of shortening – cf. *educommerce* ('educational commerce') and *d commerce* ('digital commerce') – and degrees of phonic integration, since the constituents may belong to different syllables or to the same syllable in the resulting form (cf. *v-log* and *vlog*, or *p-book* and *datasphere*: data + biosphere). The items exhibiting greater degrees of constituent integration are closer to blends, whereas those composed of initial + noun are described as "letter compounds" in Algeo (1991).

Bauer (2004:95–96) defines a splinter as "a fraction of a word arising in a blend, then used as an affix to create more words, as -nomics in Reaganomics, Clintonomics, Thatchernomics, and so on." The definition could be stretched to comprise not only parts of blends but also shortened forms of words whose productivity and stable meaning suggest a certain affixal quality. However, since the 'meaning' resides in the word they replace, they could alternatively be described

as affix-like clippings.¹³ Productive ‘splinters’ registered on-line are, for example, *e(lectronic)*, *cyber(netic)*, (*soft*)*ware* and (*ro*)*bot*. It must be noted that *ware* and *bot* exist independently as clippings. *Cyber-* occurs, for example, in *cybercide* (the killing of someone’s virtual persona in *cyberspace*), *cyberslack* or *cyberloaf* (that is, using the Net during working hours for activities not related to work), and *cybertourist* (a weekend Net user). It is also part of blends like *cybrarian* (‘cyber + librarian’) or *cytizen* (‘cyber + citizen’). *E-* is sometimes regarded as a prefix (Crystal 2001: 83) due to its shape, productivity and predictable meaning, although it is actually an initialised constituent in a number of compounds, like *e-comm(erce)*, *e-book*, *e-(maga)zine*, *e-money* or *generation-e* (coined after *generation X* and *generation Y* or *millennials*).¹⁴ The splinter *-ware* occurs in items referring to different types of software, such as *adware*, *careware* / *charityware* (a form of *shareware* for which payment should be made to a charity), *freeware*, *forestware* / *treeware* (printed material), *shovelware* (extra software to fill up space) and *spyware* (software that gathers information about a user’s surfing habits in order to build marketing profiles). Lastly, *-bot*, a shortened form of *robot* (a type of computer program) appears in *mailbot* / *answerbot* (an automatic e-mail responder), *autobot* (a program performing ordinary tasks automatically), or *meatbot* (a human being).

As shown in Table 3 below, the boundary between blends and clipped compounds can be established through examples of marginal cases which combine features of both categories.

3.1.3 Compounding (*mouse potato*)

On-line communication has bred numerous compounds (particularly noun compounds) with various “degrees of opacity” (Cruse 1986: 39), for instance *angry garden salad* (a poorly designed Web site); *copywronged* (a play on *copyright* describing a copy with the copy-protection disabled); to have a *drunk mouse* (when the pointer moves around irregularly), *mouse potato* (the on-line version of a couch potato), *silver surfer* (a Net user aged 50 or over), *snail mail* (postal mail), or *wire jam* (a situation of heavy traffic on the Net). Some compounds include combining forms, such as *-logy* (*fontology*: the knowledge related to the construction and use of fonts), *meta-* (*metafile*, *metaflaming*), *tech(no)-* (*technobabble*), *tele-* (*telecommuting*, *tele-medicine* / *telehealth*) or *-ectomy* (*VAXectomy*: to remove a VAX). Bases can be used productively, as in *toolsmith* and *websmith* (from ‘smith’: craftsman), *chiphead* (a science engineer) and *modemhead* (someone who

13. The productivity of splinters and their categorisation as morphemes or quasi-morphemes are extensively dealt with by Lehrer (this volume).

14. For an account of alternative views as regards the morphological status of *e-*, see Lehrer (this volume).

Table 3. Blends and clipped compounds: borderline cases

Examples of borderline cases*	Typical blends: automagically			Typical clipped compounds: Wi-Fi		
	syntactic structure: copulative (headed)	shortening structure: beginning + end	constituent integration: YES	syntactic structure: headed	shortening structure: beginning + beginning	constituent integration: NO
<i>infotainment</i>	X (cop)	X				X
<i>privacation</i>	X (head)	X		X		X
<i>Balun</i>	X (cop)				X	X
<i>modem</i>	X (cop)		X		X	
<i>Satrack</i>	X (head)		X	X	X	

* *infotainment*: 'information + entertainment'; *privacation*: 'private + publication'; *Balun*: 'balanced unbalanced transformer'; *modem*: 'modulator demodulator'; *Satrack*: 'satellite tracking'.

is always on-line), or *voicenet* / *voice-net* (the phone system) and *paper-net* (the postal service).

3.1.4 Conversion (to mouse) and other minor devices: reduplicatives (*flame, flame*), borrowings (*samizdat*), nonce words (*google*), and eponyms (*to be Gatesed*)

Conversion,¹⁵ particularly from nouns to verbs, is another productive device in on-line jargons. Thus, for example, to *mouse*, to *clipboard* or to *carbon copy* (*cc*), i.e. to send a duplicate copy of an e-mail. In American Hackish, doubling a verb is used as an intensifier or to end a conversation (*bang, bang!*; *lose, lose*).¹⁶

Loanwords are also occasionally resorted to, for example, *samizdat* (to spread documents via underground channels, from Russian 'self publishing'). In relation to borrowings, it should be mentioned that the great influence that English exerts on other languages is shown in the high number of loanwords borrowed from English, either unchanged or adapted to the particularities of the host language, for example the Italian verb *clicare* (from English *click*) or Spanish *chatear* (from English *chat*).

Although an account of nonce formations would fall outside the scope of this chapter (see Munat and Hohenhaus, this volume), it should be noted that, among the endless words and expressions created on the spur of the moment to satisfy a momentary need, a few items (such as *Google*) have actually succeeded in entering the on-line lexicon. Lastly, turning proper nouns and trademarks into common

15. For the use of conversion by MUDders, see Hohenhaus (this volume).

16. For the different functions of reduplicatives, see Crystal (2001:91) and Hohenhaus (this volume).

nouns, adjectives or verbs is another mechanism used to obtain new lexical items. Thus *to be amazonned / amazonized* (from Amazon, a Net company) means that a business is being dominated by a web-based company, and *to be Gatesed* is to have problems with Microsoft technology

3.2 Idiomatic expressions (film at 11) and wordplay (dot's all folks)

Idiomatic expressions are the most evident manifestation of exclusiveness, since they cannot be interpreted unless one belongs to the in-group. A few on-line examples are *a New York minute* ('immediately'); *film at 11* (a parody of TV news-readers, meaning 'there is more to come', or 'the same old story'); *lots of MIPS* ('millions of instructions per second') *but no I / O* ('input /output'), used to describe a computer whiz who is unable to socialise; *to little 'r'* (in e-mails, to click the 'Reply' option and not the 'Reply to All'); or *to bite the bag* ('to fail', thus a *bagbiter*). The origin of this expression is uncertain, but the *On-line Hacker Jargon File* suggests that it might have had obscene connotations, "possibly referring to the scrotum").

All on-line communities are fond of wordplay, as they are of showing witicism through creativity.¹⁷ In fact, it is sheer playfulness that gives rise to many linguistic distortions and creations. On-line wordplay is frequently based on phonetic similarity or structural parallelism, thus *a pain in the net* (an annoying flamer); a *cobweb site* (a dead Web page with broken links), *dot's all folks* (in e-business, it refers to the disappearance of a dot-com company); *Fortrash* (from *FORTRAN*, a programming language); *netlag* (a situation of heavy traffic on the Web, so that messages are delivered in bursts and with delays); or *word-of-mouse* (information spread via e-mail). One feature of hacker jargon is the playful alteration of common words or phrases, for instance *New York Slime* (*New York Times*), or *histerical reasons /raisins* (*historical reasons*).

4. Conclusion

The shared characteristics and advantages of on-line and mobile phone communication explain the similarities observed in lexical innovation, as well as the overlap between the different jargons of computer-related communities. It was this common ground that served as the starting point for the overview offered in this chapter. The productive use of current shortening methods like initialisation,

17. Some reasons for this tendency towards "verbal play" can be found in Werry (1996: 59).

letter reduction, letter and number homophones, clipping or blending illustrates how language adjusts to the particular constraints of these new media where speed and conciseness are of prime importance. As can be seen, the mechanisms exploited are productive word formation processes, but part of the adaptability of language to new situations consists in the economic recycling of these devices in order to obtain new products. Consequently, lexical creativity arises when old devices are used in new ways. This is what happens, for example, in innovative affixation, unconventional reductions and combinations, or in the expansion of “sense-spectra” by metaphorical extension. As regards the reasons behind lexical innovation, the same set of motivations were found to underlie all the devices employed in electronic communication, irrespective of the community of users involved. These include pragmatic motivations (the need to name new realities and condense information), ludic motivations (the desire to express humour, wit and irony), and sociolinguistic motivations (the need to create and maintain group bonds).

In just a few decades the users of electronic communication have built an alternative reality where one can live and die, work, socialise and have fun. Common places, people and activities in *RL* (‘Real Life’) have been straightforwardly transferred to this new dimension. Therefore, the Internet allows one to be an *e-collar worker* who *telecommutes*, or to enhance one’s knowledge by *telestudying* or *e-learning*. It is possible to devote oneself to *webcasting*, *e-commerce*, *on-line banking* or *telemedicine*, bid for products at *on-line auctions*, read *e-books*, buy goods at *e-malls*, have virtual reality pets and key pals, see films featuring *vactors* (virtual actors) or *technopuppets*, be a *cybercop*, a *cyberslacker*, or a victim of *cyberfraud* or *textual harassment*. *VL* (‘Virtual Life’) is a new context, a parallel world which quite naturally calls for a manipulated code, whereas language is a living instrument which evolves with the times while catering for ever-present human needs, such as laughing at life and at ourselves, belonging to a community, or making one’s own rules. The logical outcome of this supply-and-demand relationship is a code that is resourceful enough to adapt to the requirements of a new context and a new medium.

References

The list that follows also includes the sources of the corpus data:

- Adams, V. 1973. *An Introduction to Modern English Word Formation*. London/New York: Longman.
- Adams, V. 2001. *Complex Words in English*. Harlow: Pearson Education Ltd.
- Algeo, J. (ed). 1991. *Fifty Years Among the New Words*. Cambridge: Cambridge U.P.

- All about Texting, SMS and MMS*. <http://www.textually.org> [Consulted April 2005]
- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge MA: The MIT Press.
- Bauer, L. 1983. *English Word-Formation*. Cambridge: Cambridge U.P.
- Bauer, L. 2004. *A Glossary of Morphology*. Edinburgh: Edinburgh U.P.
- Carter, M. 2003. "Why a Kennel Means Trouble." *MediaGuardian.co.uk*. February 3, 2003. <http://www.media.guardian.co.uk/mediaguardian/story/0,7558,887510,00.html> [Consulted April 2005]
- Collot, M. and Belmore, N. 1996. "Electronic language: A new variety of English." In *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives*, S. C. Herring (ed), 13–28. Amsterdam /Philadelphia: John Benjamins.
- Cruse, D. A. 1986. *Lexical Semantics*. Cambridge: Cambridge U.P.
- Crystal, D. 2001. *Language and the Internet*. Cambridge: Cambridge U.P.
- Crystal, D. 2004. *A Glossary of Netspeak and Textspeak*. Edinburgh: Edinburgh U.P.
- Dictionary of Internet Terms*. <http://www.netlingo.com> [Consulted July, 2005]
- Ellison, S. 2001. "UTLKIN2ME? Can U Speak Cell Phone?" *The Wall Street Journal On line*. May 2, 2001. <http://zdnet.com.com/2100-11-529548.html> [Consulted April 2005]
- Gramley, S. 2001. *The Vocabulary of World English*. London: Edward Arnold.
- Herring, S. C. 2001. "Computer-mediated discourse." In *The Handbook of Discourse Analysis*. D. Schiffrin, D. Tannen and H. Hamilton (eds), 613–634. Oxford: Blackwell.
- Hohenhaus, P. 2005. "Elements of traditional and 'reverse' purism in relation to computer-mediated communication." In *Linguistic Purism in the Germanic Languages*. N. Langer and V. Winifred (eds), 204–220. Berlin /NY: Walter de Gruyter.
- Holmes, J. 1992. *An Introduction to Sociolinguistics*. London /New York: Longman
- Karutury, S. <ksm@funSMSnet>. "SMS Dictionary." http://www.funsms.net/sms_dictionary.htm [Consulted April 2005]
- López Rúa, P. 2005. "Shortening Devices in Text Messaging: A Multilingual Approach." *Neuphilologische Mitteilungen* 106 (2): 139–155.
- MacLean, S. 2001. *UWOT? What Those Text Messages Really Mean*. West Sussex: Summersdale Publishers Limited.
- MacMillan Publishers Ltd. 2002. "Terms and Abbreviations Used in Texting." In *MacMillan Essential Dictionary* [on line]. <http://www.macmillandictionary.com/essential/resources/resourcetextingterms.htm> [Consulted April 2005]
- Mander, G. 2003 (2000). *Wan2Tlk?: Ltl Bk of Txt Msgs*. London: Michael O'Mara Books Limited.
- Mobile Phone Text Messaging Glossary*. http://www.top-of-the-tones.com/sms_glossary1.htm [Consulted April 2005]
- Pearsall, J. (ed). 2002. *Concise Oxford English Dictionary*. 10th edition. Oxford: Oxford U.P.
- Plag, I. 2003. *Word Formation in English*. Cambridge: Cambridge University Press.
- Quinion, M. *International English*. <http://www.worldwidewords.org/index.htm> [Consulted June 15, 2005]
- Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. London: Longman
- Raymond, E. (ed). *The On-line Hacker Jargon File*. Version 4.4.7. <http://www.catb.org/~esr/jargon> [Consulted June, 2005]
- Shorthand Dictionary*. <http://www.t-mobile.com/mytmobile/communication/messaging/shorthand.asp> [Consulted April 2005]

- Werry, C. C. 1996. "Linguistic and Interactional Features of Internet Relay Chat." In *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives*. S. C. Herring (ed), 47–63. Amsterdam/Philadelphia: John Benjamins.
- SMS Smiley and Abbreviation Dictionary* <http://www.smartertext.com/mobile/help/smsdictionary.html> [Consulted April 2005]
- Text Talk Dictionary* http://www.funkupyourfone.com/mobile/smsgettexttalkdictionary/text_dictionary_letter_a.html [Consulted April 2005]
- Wingate, P. (ed) 2001. *The Wicked Book of Txt Tlk*. London: Michael O'Mara Books Limited.
- Wood, A. F. and Smith, M. J. 2001. *Online Communication. Linking Technology, Identity and Culture*. NJ /London: Lawrence Erlbaum.
- Wrolstad, J. 2001. "GR8 News for Wireless Users – Oxford Dictionary Decodes Shorthand :)". *Wireless NewsFactor*. July 13, 2001. <http://wireless.newsfactor.com/perl/story/11979.html> [Consulted April 2005]

Lexical creativity in texts

Fictional genres

Lexical creativity as a marker of style in science fiction and children's literature

Judith Munat

This is a descriptive study of novel word formations in science fiction and children's literature, seeking to identify the underlying functional motivations behind creative wordplay, in particular that determined by genre, by characterisation and other reader-related factors. While it is evident that productive word formation rules explain the majority of novel complex words at the level of competence, other explanations must be sought at the level of performance. Stylistic effects and individual authorial style are among the motivations identified in examining the word types in the two textual genres. In SF these are predominantly pseudo-scientific complex words serving to evoke verisimilitude and an effect of estrangement on the reader, while in children's literature the majority of creative formations are phonologically motivated, aiming at a ludic effect.

1. Introduction

The present study originated with my interest in the creative lexical formations or nonce words so often employed in certain literary genres and which appear to be functioning as style markers, especially in science fiction (SF). As the study developed, it opened up more complex questions regarding the nature of lexical creativity as well as related functional concerns. I originally collected approximately 100 novel lexical formations, selected at random from a variety of SF narratives, and later supplemented these with the creative formations drawn from a single SF text (Philip Dick's *The Simulacra*). In order to make a contrastive comparison, I chose Roald Dahl's *The BFG*, a work of juvenile fiction which is particularly fertile terrain for lexical invention, and later augmented my corpus with a selection of novel words from a variety of other works of juvenile literature, so that I might verify my conclusions regarding stylistic motivations in two distinct fictional sub-genres.

In comparing the types of creative formations in my corpus, I was struck by their diversity, which is immediately evident from only a few examples: *hyperspa-*

tialist, *teleportation*, *worldlets* and *plexoneuronics*, typical inventions appearing in the SF texts, in contrast with *babblement*, *scrumplet*, *murderful* and *wigglish*, some of the more analysable formations in juvenile fiction. I consequently posed the following questions in my attempt to explain these differences: in what way does the context condition lexical creativity? Are different Word Formation (WF) rules at work in different textual genres? What relationship is there between genre and specific types of nonce formations? Finally, are these nonce words formed on the basis of the grammar or are they, rather, to be seen as erratic, extra-grammatical formations falling outside the system? And, perhaps most interesting from the stylistician's point of view, what textual functions do these nonce words fulfill, other than that of referring to something in the immediate fictional environment?

The answers to some of these questions inevitably involve issues of productivity, creativity and motivation, and thus I begin with the definition of these notational terms.¹ Ultimately, this study can be seen as an attempt to bridge the divide between functional issues of creative word formation and the more theoretical concerns of morphological productivity, the first being a question of *parole* (or performance), and the second of *langue* (or competence). More specifically, I aim to illuminate the nature of individual (as opposed to societal) creativity by considering novel lexical formations in relation to the text-oriented concerns of stylistics. This approach is comforted by Kastovsky's (1986:593) observation that questions relating to the *application* (my emphasis) of productive WF rules are to be dealt with by text linguistics and stylistics, disciplines that deal with performance phenomena.

After addressing the terminological distinctions between productivity and creativity, those between neologism and nonce formation, and reflecting on the more pragmatic concerns of lexical creativity, I proceed to examine my data, considering both the *types* of novel formations that predominate in these two textual genres, as well as individual *tokens*, in order to identify the stylistic effects achieved in relation to considerations of genre. I further consider reader-related factors underlying these lexical choices and, in closing, draw some general conclusions regarding the functional roles of nonce formations in literature.

2. Productivity/creativity

While creativity in the domain of WF may appear to be a non-equivocal term, it soon becomes apparent that, as with many notational terms, it is used with a

1. For a definition of 'notational term' see Lipka (1994, 2002).

variety of different meanings (see also the discussion in Hohenhaus, this volume). Bauer (1983:63) defines creativity as “the native speaker’s ability to extend the language system in a motivated, but unpredictable (non-rule-governed) way,” in contrast with productivity, which is, instead, defined as “rule-governed innovation.”² Bauer affirms that the formation of new words that do not become established can be explained by the rules in a synchronic grammar just as those words which later enter the lexicon.³ In his later study of morphological productivity (see Bauer 2001:63ff.) the distinction between creativity and productivity is reconsidered; here the author states that the latter term should be reserved for the description of complex words, while simplexes are formed by some means other than productivity, “perhaps by creativity.”⁴ But he suggests two additional meanings for creativity: figurative extension of existing words and individual (as opposed to societal) or non-productive coinages. These three types of creativity are all “non-rule-governed.” In order to emerge from the terminological quandary, Bauer ultimately proposes that creativity and productivity should be taken as hyponyms of innovation, whereby creativity is identified as “non-productive innovation.” The concepts of rule-changing creativity and rule-governed productivity are drawn from Botha (1968:135), but Bauer adopts a slightly different formulation, defining creative coinages as those that *change* the rules while rule-governed coinages *exploit* the rules. Creativity is thus non-productive, rule-changing innovation which presents a lack of generality and predictability. Bauer returns to the suggestion, previously made by Schultink (1961), that words which are consciously created should be excluded from productivity. However, I would object that “consciously created” is a psycholinguistic criterion which is not easily determined given that we are not in the head of the speaker. Rather than speculate about the definition of “consciously created,” I consider it unhelpful as a distinguishing characteristic of creativity.

Bauer concludes that creativity and productivity are not distinct categories but, rather, prototypes which present some more central and other more periph-

2. Bauer (2001:12) raises the question as to whether productivity applies to particular affixes, to morphological processes, to rules or to words, or whether it is a feature of the language system as a whole. Some scholars (Di Sciullo and Williams 1987; Langacker 1987), as Bauer points out, deny that productivity is a matter of grammar at all, seeing it rather as a feature of performance.

3. New words, Bauer states (2001:38), are difficult to identify; in other words, we are unable to tell which words are actually in use until they are listed, and by that time they are no longer new. Those words which ultimately become part of the norm of the community will be classified as neologisms, while those that fail to become part of the norm will be ephemeral nonce words.

4. These would presumably fall under the heading of “word manufacture” which does “not involve any morphological process” (Bauer 2001:206).

eral members, and ultimately hedges the question by stating that “a precise definition of the difference proves elusive.” (Bauer 2006).

Plag (2003:47ff.) instead, sees productivity, somewhat more generically, as the “possibility of creating a new word” and he then focusses on the productivity of derivation, examining the greater or lesser productivity of certain affixes. He contemplates so-called “idiosyncratic complex words” which, he says, “must be stored in the lexicon” since they are non-rule-governed and therefore unpredictable. *Hapaxes* (lexemes which occur only once in a given corpus), according to Plag, may represent a “weird ad-hoc invention by an imaginative speaker,” which is a fitting description for many of the words in my data, though a considerable number of the novel formations in SF do reoccur within a given text and thus are not, strictly speaking, hapaxes. It is, however, irrelevant to my present purposes if these coinages reoccur in the same text since, in my view, the criterion determining their status as novel creations is that they have not been institutionalised in the wider speech community, and are therefore confined to a given genre or to a specific textual world.⁵ The point made by Plag is that *possible* or *potential* words (i.e., those which a speaker might conceivably form) always respect the WF rules, while only *actual* words may be idiosyncratic, i.e. not derived on the basis of rules and thus necessarily stored in the lexicon. I suggest that actual words, such as those appearing in my data, *may* be idiosyncratic (or non rule-governed), but more typically they respect the WF rules, their only difference from established lexemes being their referential oddity, which means that their existence is justified by and limited to the immediate textual world in which they appear, and they consequently have no wider applicability in the speech community. These fly-by-night constructions find their *raison d'être* exclusively in the text for which they have been created and will be stored in the mental lexicon only for the duration of the reading experience.

Bauer (2002:57) expresses “suspicion” toward words which appear to be playful formations, since they are, he states, formed on the basis of morphological processes that are not societally productive, used either for their phonetic qualities or “simply as a means of gaining attention.” I shall return later (see Section 6 below) to the question of attention seeking devices (ASDs), which I believe is a somewhat generic umbrella term for a much wider range of functions, but as

5. Institutionalisation is defined by Hohenhaus (2005: 359) as the stage in the life of a word at the transitional point when a neologism becomes a listeme, and is therefore a generally available vocabulary item. Lipka (2002:6) specifies that institutionalisation relates specific items with their form and meaning to certain social groups or speech communities (e.g., American English). Hohenhaus likewise specifies that a word may be institutionalised for a specific group of speakers but not for others.

we shall see, many of the so-called playful formations in juvenile literature are analysable on the basis of productive WF rules.

Non-rule-governed morphological items are classified by Dressler (2000: 2ff.) as part of extra-grammatical morphology and defined as conscious or sophisticated coinages (in contrast with unconscious productive new formations), a definition which, once again, relies on the criterion of intentionality. He also states that these coinages may apply to any word class, but are most often represented by nouns. This is not a restriction, but rather a result of the pragmatic motivations underlying the creation of what he calls “audacious poetic occasionalisms”. This type of novel formation partially overlaps with Zwicky & Pullum’s (1987: 335) category of “expressive morphology”, which includes words serving an “expressive, playful, poetic or simply ostentatious effect of some kind.” Dressler identifies three subcategories of such lawless morphological formations: 1) *metamorphology*, comprising punning, linguistic games and nonce formations; 2) *premorphology*, or a child’s inventions prior to development of the module of morphological grammar; and 3) *paramorphology*, including echo-words, blends, back-formations and onomastic creations, affirming that these non-productive extragrammatical formations exhibit “varying combinations of violations of basic properties of morphological grammar.” In the present study, only the first and last of Dressler’s categories will be contemplated, and as I have already stated, the resulting novel formations do not always violate the morphological grammar.

There may, in fact, be no real qualitative or measurable distinction between productive (i.e., ‘regular’ or rule-governed) and non-productive (or ‘irregular’, extragrammatical) word formations, as they are all created from “the connections made among related stored items” (Bybee 1995: 452). Though productivity typically refers to rule-governed word formation processes (including analogy), distinct from more creative or unpredictable processes, the two terms, as Fischer (1998: 17) notes, cannot be strictly kept apart. Ultimately, a process will be productive if it is useful (Adams 2001: 148) or when it serves the speakers of a language to form words to name the entities, properties and situations that are part of a specific environment.⁶

6. In the present case, these entities, properties and situations are part of the text world or the context evoked by the author’s lexical creations (see Semino 1997 and Stockwell 2002 for a thorough discussion of text world theory). This is a case of what Lipka (2002) calls hypostatization, which can be seen as an intentional function of nonce formations in literature, particularly in SF (see Hohenhaus 1996: 319ff.). Hohenhaus suggests (personal communication) that this function is even more marked in the less transparent simplex creations and opaque blends and clippings.

In his study of constructional productivity, Boas (2003:80) also casts doubt on the usefulness of existing definitions of productivity, affirming that it “can only be determined with respect to particular constructions.” He considers the traditional concept of “morphological productivity” to be inadequate, since productive word formation phenomena are characterised by “strikingly different combinations of morpho-syntactic, semantic, pragmatic and phonological properties and conditions.” In Boas’ view, which is borne out by the present study, such theoretical distinctions fail to capture the linguistic intuitions of native speakers.

Indeed, I believe there is good cause for perplexity regarding the theoretical distinctions that are often claimed to exist between productivity and creativity, and the analysis of my data casts further doubt on these definitions. We will see that creative formations are less rule-breaking than might initially be thought, often a matter of degree, and, especially when viewed in relation to their textual environments, are neither weird nor eccentric.

3. Neologisms/nonce formations

Nonce formations by their very nature are heavily context-dependent and their interpretation thus requires support from the co-text. They may or may not be deviant with respect to productive WF rules, but as they are not coined for permanence, they are not stored in the lexicon. Hohenhaus (2005) sees them as (prototypically) unlexicalisable.⁷ Their eventual reoccurrence (within the same context) will suggest that they have been *temporarily* stored in short term memory, and even their reappearance in a new or different context might simply mean that they have been reconstructed *ex novo*. Only with repeated occurrences in different contexts in a relatively brief span of time may we begin to consider them as neologisms.

Bauer (1983:45) defines nonce formations as new *complex* words coined by a speaker/writer on the spur of the moment to cover some immediate need, including totally regular new words that may later be accepted in the community. However, he later (2001:38) specifies that nonce words are spontaneous creations coined for the occasion that fail to become part of the norm.

While the distinction between nonce formations and neologisms is often determined on the basis of lexicographical criteria (see Note 3), this is clearly inadequate given that the word in question must have appeared repeatedly in different

7. Lexicalisation, according to Hohenhaus (2005) is the (diachronic) evolution of a complex word or expression which has undergone a shift from a compositional and transparent syntagma to a semantically opaque sign (see also Lipka 2002).

contexts before it can be listed, and only at that stage will it have “lost its status as nonce-formation” (Fischer 1998: 3/4).

From a functional perspective, I define nonce formations as words which cannot enter the lexicon, not because of their structural properties, but due to their heavy context dependence and lack of referential utility in the world at large.

To maintain my neutrality in the nonce/neologism definitional quandary, I have chosen to label the one-off coinages that I will be examining as *novel formations*. These are unlisted, non-institutionalised words, highly context dependent, whose primary characteristic is their ephemeral nature, as they serve only a passing naming need (and here I intend *naming*, as opposed to any of a variety of other functional roles identified by Hohenhaus in his chapter in this volume) in a restricted context. Beyond their referential function we may also identify a pragmatic or stylistic role in relation to the context as well as other, reader-related, functions. In some rare cases these ad hoc formations do become a stable entry in the lexicon (such is the case, for example, of *cyberspace*, and *psychohistory*⁸). But while the words that are eventually institutionalised may often be ‘regular’ with respect to systemic WF rules, structural properties alone are not an adequate predictor of permanence of any given nonce formation.

4. Motivation and the pragmatics of WF

In affirming that an act of naming “responds to the immediate naming needs of a speech community,” Štekauer (2002: 101ff.) places the output of a WF rule at the level of an “actual naming unit” by contrast with a potential word (see discussion in Section 2 above), coined to satisfy a linguistic demand, regardless of whether the demand is that of an entire speech community or an individual member of the community on a single occasion. “What we are considering is the capacity of the system to generate a new, well-formed linguistic sign by means of its productive WF rules whenever the need arises.” In other words, Štekauer focusses on the process of WF rather than the product. He considers context-dependency to be an inadequate defining criterion for nonce formations, since this can only be determined by quantitative measures such as reader’s/listener’s background knowledge which is clearly unquantifiable. In considering possible intrinsic dif-

8. The science of psychohistory in Asimov’s *Foundation* series, Stockwell says (2000: 73), is an amalgamation of socio biology, sociology, cultural psychology and technological progress that predicts the social trends of the future. This has now become a familiar term of reference, even though it is immediately associated with the SF context, and thus to be considered as a marker of genre.

ferences between nonce formations and other coinages, Stekauer affirms that this lies only in “deviation from our pragmatic knowledge and experience,” together with a second, psycholinguistic factor: an instinct to prefer words that are generated on productive and frequently-used patterns (at the system level).⁹

Motivation is generally applied to complex words which are decomposable or analysable (see Kastovsky, 1982). Thus a motivated word stands in opposition to an arbitrary simplex word. However, as Marchand (1969) has pointed out, motivation is also a matter of degree. A word is morphologically motivated if it can be interpreted on the basis of the rules of the system or when its meaning can be derived from the meaning of its constituents (e.g. *tablecloth*). Fischer (1998: 14) adds that motivation corresponds to “transparency on the cognitive level.” Among the various types of motivation that have been contemplated we may enumerate graphological, morphological, semantic and stylistic, in addition to Lipka’s (1994) mixed motivation.

Stylistic motivation, occurring at the level of performance, underlies many of the novel formations created for specific literary contexts. Given that the functional characterisation of a given word formation is found in its “actual textual occurrence” (Kastovsky 1986: 597), and not in the language system as a whole, it is only by focussing our attention on actual words in real contexts of use that we can understand the motivation behind a coinage. If the productivity of a word formation rule refers to both “competence and performance” (Kastovsky 1986: 593), this means that there is an area in which systemic competence overlaps with the more pragmatic concerns of performance. This, I believe, is especially the case with nonce words in fictional contexts.

5. Novel formations in context

5.1 Science fiction

Science fiction is particularly rich in the number and variety of neologisms which appear¹⁰ (Stockwell 2000: 115ff.), including neosemes or new meanings attributed

9. Štekauer (2002) hypothesises three levels of productivity: non-existent but potential naming units (predictable from the productive rules of the system), actually-occurring units (novel formations, nonce words, one-off naming units, neologisms, etc.) and institutionalised naming units.

10. Stockwell (2000: 82) sees the use of neologisms in pulp science fiction as serving to “disguise gaps in scientific knowledge by patching a technically-sounding invented word over a phenomenon.” The root constituents of these new formations almost always have their sources

to already existing words).¹¹ We may even go so far as to consider novel formations as one of the defining features of the SF genre. But Stockwell makes no distinction between neologisms and nonce formations, treating novel lexical creations in SF that have later been institutionalised (e.g., *superman*, *android*, *dystopia*) as having identical status with non-institutionalised novel creations (such as *pantropy* or *psionics*). He identifies the broader general purpose behind this frequent coining of new words in SF as that of world creation through the identification of a system of facts and knowledge (see Note 6), a pseudo-scientific textual world which lends credibility to the fiction. But he also speaks of new coinages in SF that are motivated by the author's desire to create an alienating effect for the reader.

5.1.1 Philip Dick's *The Simulacra*

I have chosen a novel by one of the most prolific and representative SF writers of the 20th century, Philip Dick, identifying all the novel formations in the text, from which I have created a taxonomy that provides an overview of all the types of lexical creativity in a single work. This will allow me to draw some provisional conclusions regarding the functional roles and stylistic effects achieved in relation to the themes in the novel, and to permit a comparison with novel formations in other SF texts.

Following is the taxonomy based on Dick's novel, with representative tokens of each word type:

Blends + clippings – *sim-con* (meaning simulacra construction), *relpol* (religio-political), *pubtrans*, *precogs*, *conapt*, *govpol*, *offtrans*

Combining forms + free morph – *autoshovels*, *mid-space*, *psych-chemist*, *psychopomp*

Compounds – *skypilot* (referring to a father figure, who mediates between the here and now and the eternal); *mechanical reporting machine*, *poop-sheets*, *deepthink* (adj.), *homeostatic beams*, *illness-process*, *instruction-well* (of the cab), *reaction-formation*, *electroconvulsive shock*, *short range audio receiver*, *tapedrums*.

(Pseudo) derivations – *confessionator*, *hallucinosi*, *squashily*, *chuppers* (descendants of Neandertal man – not to be confused with *chuppies*, a more recent

in the sciences of physics, chemistry, biology, and materials engineering. Their referents tend to be material and technical: physical objects as opposed to abstract concepts.

11. Stockwell (2000: 199) lists the various ways in which neosemy occurs: through broadening or narrowing of an existing meaning, through metaphor and metonymy, synecdoche, hyperbole, recontextualisation and litotes. He also distinguishes between neologisms as one-off occurrences and those that eventually enter the lexicon. See also Porto Requejo's discussion of neosemy in this volume.

blend for Chinese yuppies), *psionic*, *psionically*, *wormy* (n.) (some type of recording equipment?),

Neosemes – *jalopy* (a one-man-driven space ship), the *wheel* (a transport vehicle), *simulacrum* (a clone or reproduction of a human being that is taken for the real thing), *fracas of 1980* (a traumatic historical event, not further explained), *ersatz-tobacco* (non tobacco or tobacco substitute), *sports* (radiation survivors),

analogical creations – *unmarsed* (by analogy with unearthed)

foreign or loan words (not commonly used in English) – *der Alte*, *Geheimis*, *Geheimnis* (state secret) or *Geheimnisträger* (bearers of the secret), *geheimlich*, *Weltanschauung*, *Augenblick* (all German), *ukases* (Russian), *quondam* (Latin adv.), *Magna Mater* (primordial mother)

metaphor and metonymy – *wheel* (a vehicle)

phrase and clause reduction – *famnexdo* (family next door), *kinlhelpya*, *al-lost* (all's lost), *buttinski* (from butt in, a meddler).

sound imitation (phonologically motivated?) – *taaaaaanggg* ('spoken' by a television set), *exxxxxcologically*.

semantic or grammatical recategorisation – the following nouns (all referring to spoken or written messages) have been shifted to a new semantic category, taking on animistic or human qualities: *the text inquired* (in the sense of actually speaking), *the memo whined*, *the commercial harangued him*, *the commercial slipped in* (i.e., a dead commercial found under the vehicle seat).

Miscellaneous and simplex formations: *psi* (the 23rd letter of the Greek alphabet, here referring to psychokinesis and ESP) which is inflected (*psis*) and also serves as a base for derivations (see above). Other new simplex coinages are *spalled* (v), *runnels*, *papoola* (a 'totem animal', described only on page 126 as a flat, platter-shaped, animal, orange with black spots and antennae; later the reader is told that it is a mechanical device operated by remote control).

A few brief citations concerning the nature of Dick's fiction will serve as support to my subsequent analysis of these formations. First a quote from the author (see website): "Things are never what they seem" and secondly, a blurb on the back cover of *The Simulacra*: "there is always another layer of conspiracy beneath the one we see." Finally, Roberts (2002: 148) notes that in Dick's fiction machines "can acquire an alarming life of their own."

These themes are all directly reflected in the types of creative lexical formations employed by Dick. Words, like things, are never quite what they seem: *jalopy*, the *wheel*, appear deceptively familiar, but in Dick's textual world they take on new meanings. Derivations are either constructed with non-existent bases (*hallucinos-*) or display incompatibility between base and suffix (*confession-ator*). Transparently descriptive labels such as 'the family next door' become opaque foreign-sounding complex words, *famnexdo*; familiar inert objects take on ani-

mistic qualities as semantic categories are shifted;¹² transparent NPs are transformed into opaque blends (*rempol*, *precogs*, *simcon*) and compounds (the most frequent WF type in SF¹³), ostensibly serving to create verisimilitude, also evoke a sense of estrangement. Familiar items from the lexicon are recombined into uninterpretable opaque nominals. What exactly is *illness-process* or *tapedrums*? And can we visualise a *reporting machine*? The reader encounters familiar lexical items, recombined to form new labels. This sense of displacement is reinforced by the frequent use of foreign, particularly German and Russian, loanwords. Dick's fictional world is a distorted and menacing one in which the American government appears to have obscure ties with a Nazi past and the President (actually a simulacrum) is called *der Alte*; the creative lexical formations contribute to the reader's sense of alienation, and illustrate the way in which the manipulation of word forms serves the author's stylistic purpose.

5.1.2 Lexical data from other SF texts

I shall now look briefly at other novel formations which I have selected from a variety of different SF texts and authors (see Appendix A).¹⁴ A rapid glance at the list confirms that the dominant types of lexical innovation in SF are, in accord with Stockwell's observation, borrowing, derivation, compounding, shortening (back-formations, acronyms, abbreviations). Perhaps more significant is the fact that the majority of so-called 'new' lexemes (Stockwell 2000: 129) are, on closer examination, not as new as they might at first seem. Many of the novel formations, whether compounds or poly-word NPs, employ a familiar headword from scientific register, along with implausible modifiers, or are coupled with classical (Latin or Greek) combining forms (CFs): *arco-*, *carbo-*, *mono-*, *astro-*, *photo-*, *pseudo-*, *orni-*, *hyper-*, *holo-*, *mono-*, *micro-*, *psycho-*, *peri-*, *tele-*, *fuseo-*, *exo-*; among these *pru*, a non-canonical English CF. Many of the scientific terms exist, *in some relat-*

12. In his analysis of Dick's fictional world, Roberts (2000: 150) speaks of things beginning to acquire a power over people, imbuing objects with a kind of fetishism that he identifies as "Marxian reification."

13. Multi-word lexical items are among the most common novel creations in science fiction texts as Stockwell notes (2000: 128), due to their descriptive power and the "flavour" of a scientific register; many include legitimate scientific terminology to describe non-existing realities.

14. In all cases I have chosen to exclude proper names from my data, as these typically allow greater deviation (at the morpho-syntactic level) than is permissible in words of the common vocabulary. In addition, the assigning of proper names appears to be regulated by unique semantic and pragmatic considerations (see Munat 2007).

ed form, but when combined with unfamiliar bases or transformed into unlikely derivations or compounds, their meaning ultimately eludes us.¹⁵

The majority of words in this list are complex lexemes, either composed of two free roots fused into a single orthographic unit, compounds such as *stillsuit*, *voidhawk*, *crysknife*, *godlight*, or formed of an initial (classical) CF and a familiar base (*holoscreen*, *hyperspace*, *microfood*). Other prefixes, such as *trans-*, *super-*, *anti-* and *sub-*, belonging to a category somewhere between CFs and derivational prefixes (like canonical derivational affixes, they freely combine with numerous English bases but, similar to CFs, they carry greater semantic content than most derivational prefixes). Typical derivational suffixes (such as *-er*, *-ist*, *-ic*, *-let*, *-tion*) are attached to non-existent bases (*thopters*, *teleportation*, *nanonics*), representing perfectly regular, though semantically anomalous formations.

Polyword nominals, whether viewed as compounds or NPs, composed of a noun head plus one or more pre-modifiers, appear to be more descriptively transparent than the fused compounds cited above. Examples such as *palm lock*, *blast pistol*, *student-agent*, *poison-snooper*, (*photonic*) *dead zone*, *slow-pellet stunners* function as single naming units and refer to identifiable entities or objects (in the fictional world). Even if they remain unfamiliar and unlexicalised, they are nonetheless analysable and display greater semantic transparency than that of the orthographically fused compounds cited earlier: *stillsuit*, *voidhawk*, *crysknife*, *godlight*. In the latter examples the two base components do not share any evident semantic features and the newly-formed compound remains opaque, despite its structural compositionality. The meanings of these words must be learned, just like those of any arbitrary or symbolic simplex lexeme.

I have found only one instance of conversion or zero-derivation: ‘computer’ used as a transitive verb in the expression ‘computering you’; and one other novel construction ‘eystack panels’, formed by reducing a compound adjectival pre-modifier (easy-stack) to morphemic status in a new complex word by a clever exploitation of orthographic, phonological and morphological rules. An anomalous (and recurring) type of formation is that represented by compounds formed with the pseudo-suffix ‘side’, attached to a prepositional or adverbial base: *upperside*, *dayside*, *farside*, *upside*, (a device which I have found in the fictions of at least two authors) and which have presumably been formed by analogy with existing complex lexemes such as *inside* or *outside*.

Truly novel creations, in the sense of new simplex referring units (thus not involving any identifiable WF process) are relatively rare. Among the few in my data

15. These were checked in the OED online and no entry appeared for any of them, with the exception of *exoskeleton* and *photonic*, which are, however, here combined in poly-word strings with other lexemes that alter (or even contradict) the base meanings.

are: *simps*, *biot*, *bitek*, *baliset*, *kroclion*, *plexons*, *blivet*. None of these violate the phonological rules of English, though the orthographic form of *kroclion* might suggest a word of Greek origin (as opposed to the more typical ‘*cr*’ of English spelling).

In Štekauer’s onomasiological model (2002), nonce formations are defined as regular, transparent, predictable and productively coined naming units, whose ‘oddity’ lies in the interplay between the language and the surrounding reality in that they relate to a non-existing world. This definition is valid for many of the novel formations we have examined, except for the criterion of transparency. Since definitions are rarely provided in the SF context, and transparency is not a quality that is to be expected in a strange or futuristic world, the reader will either attempt to disambiguate these words with the support of context or by association with other familiar entries in the lexicon, but it may also be the case that our normal expectations of comprehensibility are temporarily suspended in the world of SF.

One final observation regards the lexical preferences of the single authors: Asimov’s lexicon is heavily weighted with scientific-sounding CFs, while Herbert’s and Hamilton’s fictions abound in compounds, Hamilton having a predilection for multiword constructions. None of these authors, however, uses the types of blends and clippings that are so frequent in Dick’s fiction, nor is there a similar abundance of foreign loan words. This suggests that, within the wider SF genre, single authors exploit different morphological processes which may be seen as markers of a given authorial style or stylistic markers reflecting particular themes, as we have seen in Dick’s *The Simulacra*. A more attentive study of individual authors and a comparison of different novels by the same author might well reveal unique lexical ‘styles’ serving particular functions in a specific text world.

5.2 Children’s literature

5.2.1 *The BFG*

The nonce formations drawn from Roald Dahl’s *The BFG* (see Appendix B) reveal a variety of different lexical typologies. Even though compounding and derivational processes account for many of the words on the list, the resulting word forms are distinctly different from those identified in SF, in particular for the absence of classical CFs. The primary characteristic of these formations is that many (most?) of them seem to be phonologically motivated, including rhyming compounds (*moocheling* and *footcheling*), assonance (*wispy-misty*), reverse rhyme (*squiff-squiddled*), alliteration (*wopsey whiffing*), and numerous onomatopoeic words (*whooshey*, *swizzfiggling*, *uskyshlush*). Two other pervasive phonological features are: 1) the abundance of consonant clusters typical of sound symbolism, clusters which appear in initial, medial and final positions (*bl/br/cr/fl/fr/gl/sl/scr/*

sq/st/sw/tw) - their presence being far more frequent than their random distribution in the lexicon; and 2) the use of phonetically motivated spellings (e.g., *human bean* for human being, and *cannybull* for cannibal), also to be explained by a child's tendency to seek sound associations with more familiar concrete words.

Other creative formations are represented by a playful combination of phonological and semantic substitutions (or confusions), presumably explainable by the giant's linguistic muddlement: *fridging with cold* is easily recognised as a parallel expression for "freezing cold"; and *frumpkin pie* as a rhymed substitute for pumpkin pie, or *jipping and skumping* in which the initial consonants have been inverted, commonly referred to as a Spoonerism.

Some of the morphologically motivated formations present infixes, a rarely used and unproductive process in English, aside from occasional playful formations such as *absobloominglutely* cited by Bauer (1983: 18) or *unbefuckinglylievable* cited by Stockwell (2000: 127).¹⁶ These exceptions prove the (absence of a) rule: *scrumpdiddlyumptious*, *crockadowndillies* (where *crocadillies* is a childish variant for crocodiles).

Compounds are typically formed of one recognisable free root combined with a nonsense word (*bumplehammers*, *dogswoggler*) or two semantically unrelated words: *bugswallop*, *gobblefunk*, *fizzwiggler*, resulting in vaguely familiar, but ultimately uninterpretable nonce words.

Perhaps the most transparent formations are those formed by semantic analogy, such as 'kidsnatched', recognisable as a synonym for the more canonical 'kidnapped', or derivations formed with existing bases such as *babblement* or *murderful*, which are analysable, though semantically anomalous. Since the majority of these novel formations are uttered by the BFG (or Big Friendly Giant – even the acronym in the title is a novel onymic unit), they might be attributed to the giant's somewhat unique idiolect and seen as partly motivated distortion rather than word creation.

The reader has the impression that many of these formations are only slightly 'off', but just when we think we have identified a pattern, new and uninterpretable formations appear (*maggotwise*, presumably a derived adjective, or 'rotsome?'). Among the more transparent of the complex words is *earbursting*, a reduction of the VP (to burst one's ears or eardrums), and *sizzlepan*, a semantic and possibly onomatopoeic alternative for 'frying pan'. The transparent *brain-boggling* is easily interpreted as a synonym for the more familiar *mind-boggling*, which can either be seen as an analogical formation, or a confusion reflecting the Giant's linguistic befuddlement. The only novel simplex word in the list is *slutch*, which respects

16. See also the examples of infixing reported in Lehrer's paper (this volume).

the phonological constraints of English but lacks a recognisable referent, just as the majority of complex formations created by Dahl. Sound associations seem to be of greater importance here than semantic motivation, and so this simplex creation fits into the pattern that the author has established, allowing the reader to abandon herself to suggestions of sound symbolism.

As Stockwell (2000: 124) observes, the reader intuitively tends to interpret nonsense words in terms of familiar lexical items, thus automatically assuming that they are “derivations or compounds of other meaningful chunks.” The BFG, in effect, speaks a language which is quasi familiar, especially in view of the many canonical derivational suffixes (-*er*, -*ment*, -*ing*, -*ly*, -*wise*, -*ful*, -*some*, -*y*, -*tious*-, -*let*-, -*ish*), even when they are attached to unfamiliar or unlikely bases. So we can make a reasoned guess, even when the resulting derivation is anomalous; such is the case of *glamourly*, though *glummy* and *rotsome* may create greater difficulty. Consider, for example, Dahl’s *scrumplet* by comparison with Clark’s *worldlet*, the first totally opaque because the base is non-existent, while the second is transparent as well as plausible in relation to the SF context in which there may well be a series of smaller worlds populating the horizon.

The majority of Dahl’s lexical creations are ear-catching for a child or “phonologically funny” (intuitively, certain phonemes seem to inspire amusement, a kind of emotional sound symbolism that has an instinctive appeal to a child). The author’s lexical inventions might thus be seen as audience-oriented, his ‘ideal reader’ being a child who will presumably respond, not only to the story, but to the individual sounds and word associations which aim to amuse, markedly different from the more rational, pseudo-scientific lexis of the SF world.

It might also be the case that children make fewer demands as regards comprehensibility; accustomed to a world in which meanings are frequently elusive, they are more willing to be carried by suggestion, closer to the fantasy world of infancy.

Equally important, however, in explaining the choice of nonce formations is the characterisation of a somewhat befuddled protagonist. The BFG’s communicative competence is definitely shaky, verging on the linguistically challenged: “Meanings is not important”, said the BFG. “I cannot be right all the time. Quite often I am left instead of right” (p. 27).

5.2.2 *Novel formations in other juvenile literature*

Appendix C presents a selection of novel formations drawn at random from other juvenile works of fiction, poetry and one song. Along with the variety of novel formations, it is curious to note that a couple of rather similar items appear in different authors: *thingummies* in Milne and *thingummybobs* (perhaps echoing the more familiar *thingamabob*?) in Child; *glumping* in Seuss which vaguely resembles Dahl’s *glummy*, and it is impossible to know whether these have been

'borrowed' with the intention of creating intertextual reference, or are simply casual re-creations. Milne creates his own WF pattern with the *-eezles* formations: first consonant or consonant cluster plus the vowel phoneme of an existing word, followed by his invented suffix (?): thus *wheezing* is *wheezles*, *sneezing* *sneezle*, *please* *pleazle*, etc. Among the more analysable (though semantically opaque) formations in this data are derivations such as *glumping* (adj.), *biggered* (v.), a non-canonical inflection formed from *bigger* (comp.) + *ed* (p.tense), *once-ler* (n.), *gruvvulous*, *memmily*, *crummies*, *explectionary* or *quackery*, exploiting canonical derivational suffixes (*-ing*, *-er*, *-ous*, *-ly*, *-ie/y*, *-tion*, *-ary*, *-ery*), but attached to non-existent bases (except for the anomalous *once-ler*, hyphenated in the original and referring to a person who uses up the environment). There are a few relatively interpretable creative compounds – *smogulous smoke*, *spaceburger*, *zooglobber* or *knothead* – *spaceburger* seeming quasi transparent (until we discover that it refers to a pirate ship of the cosmos) and *knothead* being at least guessable.

Some of the simplex lexemes created (thus falling under word manufacture, in that they are not based on productive processes) violate the phonological rules of English (*thneed* and *thwerll*) while others appear to be onomatopoeic or, in any case, phonologically motivated, formations like *g-r-r-zapp*, *g-r-r-zibb*, etc. which resemble the category of sound imitation in Dick's fiction (see Section 5.1.1). Other formations, such as *grackle-grass*, *screebees*, *dinko blam*, and perhaps *plinkee*, reminiscent of Dahl's phonologically-motivated formations, seem to be based on sound play. *Zitzkis* has a definitely foreign ring (and look) to it and a few others seem unlikely as English words (*floog*, *vengo glop*, *upaloo*, *inyews*, *kibblezy*), but what is significant in terms of the present discussion is that, with only minor variations, these formations are not appreciably different from the types encountered in *The BFG*, while they differ substantially from the pseudo-scientific creations in SF. Thus we may provisionally conclude that juvenile literature prefers phonologically motivated novel creations and the bases for derivation and compounding are more likely to be nonsense words, substantially different from those identified in SF. The semantic, morphological and phonological distortion of the BFG's idiolect is clearly not the only explanation for Dahl's sound play and pseudo derivations, as similar characteristics are present in the second set of data as well. This suggests that the underlying stylistic motivations are ultimately related to a distinct fictional world and its intended audience.

6. Functional motivations and ASDs

I shall briefly return to the question of novel formations as attention seeking devices (ASDs) originally raised in Section 2 above. This function is attributed to a

variety of different WF processes – novel creations as well as metaphors, blends, puns, etc., frequently those employed in media headlines and advertising (see Lehrer and Veisberg, this volume). I believe, however, that there are finer distinctions to be made, depending on the specific type of lexical play and the precise context in which it occurs. Lehrer (2003:380) suggests that novel blends and other neologisms serve a perlocutionary intent, that of slowing down the reader's response with the use of a clever or puzzling new word, given that new words generally require greater processing effort. Thus a novel formation demands greater attention, in addition to providing pleasure, amusement and entertainment.

Among the textual functions of nonce words mentioned by Lipka (2002: 189) are those of creating coherence and cohesion, information condensation (as in the case of acronyms), and various, not better defined, "stylistic effects" of meta-textual discourse markers serving as "attention getting devices", particularly in journalistic texts. Among other functions, Hohenhaus (this volume) considers the "foregrounding effect" of certain creative lexical devices and in an earlier study (Munat 2003), the compression of entire sections of narrative text was identified as a functional motivation underlying the creation of certain ad hoc compound formations.

However, limiting my self to the data examined here, it appears that creative formations fulfill a variety of stylistic or textual functions that go well beyond attention-seeking. Among these I would identify: the construction of textual worlds (by creating an illusion of verisimilitude and estrangement for the reader in SF), characterisation (the unique idiolect of the BFG) and, more generally, a purely ludic function (phonological play for purposes of amusement). I believe that the very incomprehensibility of these incongruous or semantically opaque nonce words is instrumental to the specific textual goals.

While non-sense words such as those in the children's fiction are unlikely candidates for institutionalisation, the more rational lexical creations in the SF world do on occasion make their way into the lexicon, perhaps proving that reality is stranger than fiction.

One further reflection regards the question of simplex words vs. motivated complex formations: the recombination or manipulation of existing morphemes and/or lexical units (e.g. unfamiliar bases combined with familiar affixes or CFs and the juxtaposition of existing words to form new compounds) constitutes the most frequent 'creative' patterns in SF. This preference for familiar components might be explained by psycholinguistic factors, both that of facilitating memory and that of suggesting analogical associations (in SF) with other existing technical or scientific terms. Juvenile literature instead displays greater WF anarchy, paying mere lip-service to canonical WF rules and creating a greater number of unanalysable simplex lexemes.

7. Closing considerations

From the examination of my data it is possible to affirm that complex novel formations are mostly rule-governed, giving no evidence of a net demarcation between productivity and creativity. This suggests that the “audacious nonce formations” described by Dressler, as well as Plag’s “weird ad hoc inventions” or Bauer’s “playful formations” are mostly explainable on the basis of productive rules. Creative formations do not change the rules, with the exception of ad hoc patterns such as Milne’s *-eezle* formations, but these are no more permanent (and ultimately no more rule-changing) than any of the one-off playful creations encountered. When we have difficulty in assigning meaning, complex constructions appearing in SF such as *astrogration review*, *carbotanium fuselage* and *datavised proximiting*, are ultimately analysable on the basis of WF rules, and the bases have some vague referential value. What escapes us is their precise meaning in relation to our familiar world. Similarly, formations such as *glam-ourly* (*flavour*), *glummy* and *hipswitch* in juvenile fiction, even though semantically uninterpretable, are nonetheless analysable in terms of the grammar. The simplex creations in juvenile literature are mostly suggestive of sound symbolism and other phonological play intended to amuse the reader, therefore appropriate to the context and intended audience.

We have considered novel creative formations both from the perspective of the writer and that of the reader. The writer’s role in lexical creativity clearly involves the exploitation of WF rules at the level of *langue*, but these original creations represent individual *performance* as solutions serving the specific needs of the fiction. The reader, as she attempts to interpret these novel formations, will be only marginally concerned (if at all) with systemic WFRs, focussing primarily on the more text-oriented concerns of performance in trying to disambiguate the word formations with reference to the context.

Writers draw on the entire range of WF rules offered by the system, rarely extending the limits of the grammar. For the present, at least, we must conclude that stylistically motivated nonce formations are not inherently extragrammatical or rule-changing. They exploit the potential of the system in unusual ways and may ultimately have an echo effect in serving as models for further analogical creations.

The distinction made between actual vs. potential words by Plag (2003) and by Štekauer (2002) seems to bear little significance as regards the formations we have been examining. The actual words that appear in my data are not appreciably different from any potential (or productively-formed) word, except for their referential ambiguity.

As we have seen, different types of lexical creations emerge in different genres, determined by context and motivated by a variety of phonological, morphological, semantic, pragmatic and stylistic factors. Different literary genres give rise to different types of novel formations, thus confirming that it is the context which determines the nature of the novel word. This is also borne out by Renouf's study of the media and López Rúa's investigation of electronic communication, as well as the examples that Lehrer and Veisberg (all in this volume) have drawn from advertising. But even beyond the major textual varieties, we may identify finer (sub)generic distinctions reflected in the types of lexical creativity exploited. This creative variation bears witness to the flexibility of a system that allows infinite original solutions in response to the demands of context.

References

- Adams, V. 2001. *Complex Words in English*. Harlow: Pearson Education.
- Bauer, L. 1983. *English Word Formation*. Cambridge: Cambridge University Press.
- Bauer, L. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Boas, H. U. 2003. "Constructional Productivity in English". *Ranam* 36/3: 79–91.
- Botha, R. 1968. *The Function of the Lexicon in Transformational Generative Grammar*. The Hague and Paris: Mouton.
- Bybee, J. 1995. "Regular morphology and the lexicon". *Language and Cognitive Processes* 10.5: 425–445.
- Dahl, R. 1982/2001. *The BFG*. London: Penguin Books.
- Dick, P. 1977/2004. *The Simulacra*. London: Orion Publishing Group.
- Di Sciullo, A. M. & Williams, E. 1987. *On the Definition of Word*. Cambridge, Mass.: MIT Press.
- Dressler, W. U. 2000. "Extragrammatical vs. marginal morphology". In *Extragrammatical and Marginal Morphology*, U. Doleschal and A. Thornton (eds), 1–10. Munich: Lincom Europa.
- Fischer, R. 1998. *Lexical Change in Present-Day English: A Corpus-Based Study of the Motivation, Institutionalization and Productivity of Creative Neologisms*. Tübingen: Gunter Narr Verlag.
- Hohenhaus, P. 1996. *Ad-hoc-Wortbildung – Terminologie, Typologie und Theorie kreativer Wortbildung im Englischen*, Frankfurt/M etc.: Peter Lang.
- Hohenhaus, P. 2005. "Lexicalization and Institutionalization". In *Handbook of Word-Formation*, P. Stekauer and R. Lieber (eds), 353–373. Dordrecht: Springer.
- Kastovsky, D. 1982. "Word-formation: A functional view". *Folia Linguistica* 16: 181–198.
- Kastovsky, D. 1986. "The Problem of productivity in word formation". *Linguistics* 24: 585–600.
- Langacker, R. 1987. *Foundations of Cognitive Grammar Vol. 1: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lehrer, A. 2003. "Understanding trendy neologisms". *Rivista di Linguistica* 15.2: 369–382.
- Lipka, L. 1994. "Lexicalization and Institutionalization". In *The Encyclopedia of Language & Linguistics*, Vol. 4, R. E. Asher (ed.), 2164–2167. Oxford.

- Lipka, L. 2002. *English Lexicology: Lexical structure, word semantics and word formation*. Tübingen: Gunter Narr.
- Mackey, W. F. 1965. *Language Teaching Analysis*. London: Longman.
- Marchand, H. 1969. *The Categories and Types of Present-day English Word-formation*, 2nd ed. Munich.
- Munat, J. 2003. "When is a noun string a phraseological unit?" *Ranam*. 36/2003 (2), 31–47.
- Munat, J. 2007. "Onymic Noun Phrases in English: The importance of fields and frames in as-signing meaning." In *I nomi nel tempo e nello spazio. Atti del XXII Congresso Internazionale di Scienze Onomastiche, 2005, Pisa*. M. G. Arcamone, D. Bremer, D. DeCamilli, B. Porcelli (eds) Pisa: ETS (forthcoming).
- Plag, I. 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.
- Roberts, A. 2002. *Science Fiction*. London: Routledge.
- Schultink, H. 1961. "Produktiviteit als morfologisch fenomeen." *Forum der Letteren* 2: 110–125.
- Semino, E. 1997. *Language and World Creation in Poems and Other Texts*. London: Longman.
- Štekauer, P. 2002. "On the Theory of Neologisms and Nonce-formations". *Australian Journal of Linguistics* 22.1: 97–112.
- Stockwell, P. 2000. *The Poetics of Science Fiction*. London: Longman.
- Stockwell, P. 2002. *Cognitive Poetics: An introduction*. London: Routledge.
- Zwicky, A. and Pullum, G. 1987. "Plain morphology and expressive morphology". In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistic Society*, J. Aske, N. Beery, L. Michaelis, H. Filip (eds), 330–340. Berkeley: Berkeley Linguistic Society.
- http://www.philipkdick.com/media_articles.html (Consulted Nov. 2005)

Appendix A

Nonce formations in science fiction

Asimov *Prelude to Foundation*

computering you
gravitic lift
microfood
student-agent
upperside

Asimov *Nemesis*

holoscreen
hyper-assistance
hyperatomic motor
hypercommunication
hypermessages
hyperspace
hyperspatialists
photon-complex characteristic
plexoneuronics
plexons
spaceship acromegaly
superluminal flight

As. *Foundation and Empire*

blast pistol
hyperwave
 news broadcast
speedster bird-chases

As. *Nightfall*

godlight

Herbert *Dune*

baliset
chromoplastic
crysknife
dew collectors
hypno-ligation
kindjal-type
opafires
ornithopters
palm lock
phosphortubes
poison-snooper
projecto-stylus

pru-barriers
pru-shields
slow-pellet stunners
solido tri-D projection
stand-mounted
stillsuit
thopter (n)

Clark *Rend. with Rama*

biots
camshafts
teleportation
worldlets
simps

Hamilton *The Reality Disfunction*

anti-matter-confine
 chambers
arcology's refuge
astrogration review

bitek
carbotanium fuselage
datavised proximiting
dayside/farside/upside
ezystak panels
fuseodollars
geneering
kroclion
monobonded-carbon
 exoskeleton
neural nanonics
photonic dead zone
picoseconds
pseudofabric structure
scaploose way
submunitions
thermal resistant boron-
 beryllium
voidhawk
xenoc researchers
zero-tau pod

Appendix B

Nonce formations in Roald Dahl's *The BFG*

babblement	muckfrumping
Bloodbottler	murderful
brain-boggling	rotrasper (n)
buggles	rotsome
Bugswallop!	rubbsquash
bumplehammers (n)	scrotty
bunkumhouse (n)	scrumdiddlyumptious
buzzburgers	scrumplet (n)
cannybull	scuddle (v)
cannybully	scuddling (v)
chittering	scumscrew
crockadowndillies	sizzlepan (n)
crodscollop (nb)	slime-wanglers
crotching mood	slutch
crumpscoddle	snitching (adj)
diddly	squiddlerpibbling little...
dillions	squiff-squiddled
dispunge v	squinkers
dogswoggler	squoggle v
earbursting	strawbunkles and
fizzwiggler (n)	cream
flungaway (adj)	swallow
flushbunking	swiggling (v)
fridging with cold	swinebuggler
frisby north (adj)	swishfiggler
froboscottle party	switchfiddling
frumpkin pie	swizzfiggling
glamourly flavour	telly-telly bunkum box
glimp (n)	titchy (adj)
glummy	twitch-tickling
gobblefunk (v)	uskyshlush (adj)
gogglers (b)	whiffswiddle (n)
hipswitch (adv)	whiffsy time-twiddlers
humplecrimp	whizzpopping
jipping and	whoopsey-splunkers!
skumping (v)	whooshey
jumbly (adj)	whoppsy big
kidsnatched (v)	wigglish
maggotwise	winksquiffler
mince my maggots	wispy-misty buggles
moocheling	wondercrump
and footcheling (v)	wopsey whiffing
	(human beans)

Appendix C

Novel formations from other juvenile literature

A. A. Milne

wheezles
sneezles
teazle
pleazle, etc.
thingummies

Seuss

biggered
bustard
chug
crummies
flupp
flustard
g – r – r – zapp
g – r – r – zibb
g – r – r – zopp
glumping
grackle-grass
grinch
gruvvulous
gusset
joats
lerkin
mulligatawny
once-ler
screebees
shuvv
slupp
smogulous smoke
snergelly
thneed
thwerll

Siracusa

dinko blam

fleenzil
floog
plinkee
upaloo
vengo glop

Pinkwater

fleegix
snarking out
spacebirger
zitzkis

Boynton (song)

Beezifies (v.)
deen/voo/nizee
explectionary
ezzeleboo dornut
febbin (v.)
glorble snop (n.)
inyews (n.)
kibblezy
memmily blitt
quackery
snoofly
ud (conj.?)
yuddle (v.)
zooglobble (n. + v.?)

Earley

knothead
poot

Child

punctuationy thingummybobs
exceptionordinarily

Creative concept formation

Dynamic creation of analogically-motivated terms and categories in lexical ontologies

Tony Veale

As creative processes go, it is uncontroversial to suggest that metaphor and analogy are extremely knowledge-hungry processes. Their appetite for structured knowledge ranges from the taxonomic to the image-schematic to the purely propositional. In this paper, I demonstrate that new lexical analogies often rely, implicitly, on the meaningfulness of a corresponding, and equally creative, lexical combination. I therefore argue that the interpretation and generation of novel lexical analogies has much in common with the interpretation and generation of creative noun compounds, so much so that a mechanism for the latter can be formulated, in a computationally precise manner, in terms of a mechanism for the former.

1. Introduction

Ontologies are, for the most part, static organisations of categories and relations that attempt to model some aspect of the world. But such organisations can be strained to the breaking point when creative actions in this sub-world necessitate a dynamic response in the corresponding ontological structures. An important case in point is when, through some purposeful activity, an agent dynamically creates a new category of entity on the fly that must then be accommodated within the ontology (e.g., see Way 1991; Veale 2003a, 2005). This contrasts with the creation of new senses for existing ontological terms (as, for example, described in Pustejovsky 1991). Such ad hoc categories, as described by Barsalou (1983), are typically created in response to a specific goal or task, and may thus be considered task-specific or goal-specific constructs. Examples of ad hoc categories include “things to take on a camping trip”, “useful wedding presents”, “substances from which sugar can be extracted”, “objects that would make unusual murder weapons”, and so on. Ad hoc categories do not correspond to the existing hierarchical

categories of an ontology, and their members are rarely clustered in the same localised area of the ontology. Rather, the members of a given ad hoc category may be drawn from many different established categories. Ad hoc categories thus constitute a horizontal rather than vertical slice of an ontology, cutting across conventional hierarchical structures, and as such, the creation of ad hoc categories may intuitively be seen as a form of “lateral thinking” (de Bono 1994).

The cross-cutting nature of ad hoc categories means that they cannot easily be lexicalised with any of the labels associated with existing hierarchical categories; in fact, as seen above, the lexical label given to an ad hoc category can be quite a mouthful, as a new multi-word expression must be constructed to capture the functional rather than taxonomic nature of the category. Nonetheless, some ad hoc categories can be given compact labels that may subsequently find use as meaningful collocations in their own right and which merit their own individual listings in the lexicon. The creativity inherent in the construction of ad hoc categories can thus apply at two different levels of representation, involving the creation of not just new ontological categories but of new lexical entries as well.

Of course, these levels of representation become one and the same when the ontology in question is a lexical ontology (see Veale et al. 2004, Hayes and Veale 2005), that is, an ontology that concerns itself only with those units of meaning, called lexical concepts, that correspond to specific words or compound terms in a language. In this case, the creation of a new ontological category will correspond to the creation of a new lexical concept, forcing all new categories to assume a lexical label that serves a useful indexing role in the lexicon. For example, a new category like French-Food will serve to cluster together the various foodstuffs, dishes and wines that can be considered French into a single ontological category. At present, a number of large-scale lexical ontologies are available to support research in this area, such as WordNet, a comprehensive electronic thesaurus of English (Miller 1995) whose design reflects psycholinguistic insights into the structure of the mental lexicon. Thus nouns, verbs and adjectives are represented by WordNet in different ways, and each lexical concept partakes in one or more relations to other concepts. HowNet (see Dong and Dong 2006; Carpuat et al. 2002) is an equivalent Chinese ontology (with English translations) in which each lexical concept is associated with a propositional semantic structure. Both WordNet and HowNet are more properly described as weak ontologies since they exhibit neither the relational richness nor formal precision of the structures normally called ontologies by philosophers. In contrast, the Cyc ontology of Lenat and Guha (1991) is a rich axiomatic system explicitly designed to support logical inference and knowledge-based problem-solving, but which also maps word-forms onto a selection of underlying logical forms, both atomic and formulaic. However, though Cyc deserves

to be called a strong ontology, it is not a lexical ontology, since natural language is not a motivating factor in its design; consequently, Cyc contains a high proportion of unlexicalised concepts and new concepts are not required to have corresponding lexical forms. Philosophers of language often contrast the role of the dictionary and the encyclopaedia when considering the knowledge demands of language; a lexical ontology aims to capture key aspects of both the dictionary and the encyclopaedia, and so constitutes the ideal framework in which to explore the mechanics of term creation.

Since ad hoc categories are goal-specific, different goal contexts might give rise to different kinds of ad hoc categories, suggesting that these are best studied from the perspective of a specific cognitive goal or task. One such important task to which lexical ontologies have been directed is the construction and interpretation of lexical analogies. Analogy has been identified as a reasoning mechanism at several different levels of linguistic operation, from native speaker intuitions about pronunciation (see Baron 1977) to intuitions about morphological inflection (see Trask 1996) to intuitions about semantic relatedness (see Rumelhart and Abrahamson 1973). At a lexico-conceptual level, analogies such as Fructose is to Fruit as Lactose is to Milk (e.g., see Veale 2003a and Veale 2003b, 2004, 2005) exhibit creativity not only in their production, since they constitute novel linguistic artefacts, but also in their interpretation, which frequently requires the dynamic construction of new ad hoc categories. Consider the joke given in Freud (1905), and analysed as an analogy in Attardo et al. (2002):

“A wife is like an umbrella. Sometimes one takes a cab”

Attardo et al. (2002) provide the missing concept, prostitute, to complete the analogy:

wife : prostitute :: umbrella : cab

To understand the analogy (and thus the joke) the listener must recognise that wives are personal lovers, while prostitutes are hired lovers; and that umbrellas are personal resources, while cabs are hired resources. This recognition necessitates the creation of the ad hoc categories Personal-Lover and Personal-Resource (both sub-types of Personal-Belonging, to take a Victorian view of marriage and wives), as well as Personal-Resource and Hired-Resource. The burden of creativity is not borne solely by the creator of the joke, as the listener must also carry much of this burden through the creation of new categories that mirror the mind-set of the joker. Now, categories like Hired-Lover are highly goal-specific, and may not persist beyond the immediate context of the analogy that gives rise to them. However, if an ad hoc category demonstrates some long-lasting value, its lexical label may also

persist, to the point that it becomes a unit of common currency in the language and a permanent entry in a speaker's lexical ontology.

In this paper I consider how a particular kind of lexicalised ad hoc category is created in a lexical ontology when the motivating task is that of analogical reasoning. We argue that the interpretation of lexical analogies often necessitates the creation of new conceptual categories that in turn necessitate the creation of new lexical items. These lexical terms may be as short-lived as the analogies themselves, but a corpus analysis can be used to reveal those terms that have sufficient durability to merit a place in the lexical ontology. For our current purposes, we ground our investigation in the context of WordNet, and explore a variety of ways in which analogy can be used to drive the creation of lexical innovations that do not already exist within the WordNet lexicon.

I am careful to note that, in the context of a lexical ontology, the term “analogical categorisation” is an ambiguous one. It can mean either the creation of new categories, like *Hired-Lover* and *Personal-Resource*, to resolve a particular analogy, or it can denote the use of analogy as an explicit term-creation mechanism, in much the same way that analogy can be used to suggest spelling, morphology and pronunciation. Since each reading denotes a process that is meaningfully performed within a lexical ontology, the ambiguity is a benign one that reflects the multipurpose nature of lexical analogy. So for the sake of completeness, I consider both of the foregoing uses in this paper. In Sections 2 and 3 I explore the use of analogy as an explicit and quite deliberate mechanism of term creation, while in Section 4 I consider how new ad hoc concepts and their corresponding lexicalisations can arise as by-products from the interpretation of lexical analogies. More specifically, Section 2 considers how analogy can be used to produce new words, while Sections 3 and 4 explore the role of analogy in the creation of new compound terms. In Section 5 I then consider the application of these ideas to the WordNet lexical ontology, which allows me to empirically evaluate their effectiveness in the context of a real analogical retrieval task. I then conclude with some remarks on the limitations of these ideas in Section 6.

2. Analogy as a mechanism of word creation

To begin, one might well consider how analogy is implicated in the determination of morphological inflections, even in the face of stronger and more accepted linguistic principles. Indeed, as noted by Trask (1996), analogies can be so persuasive that any fallacious conclusions that can be drawn may seem even more natural than those of a first principles analysis. For instance, the lexical analogy in (1)

is compelling even though the +us→+i pluralisation rule is valid only for words of Latin origin, as exemplified by *radius/radii* and *succubus/succubi*.

- (1) *cactus : cacti :: octopus : octopi*

The correct inflection, following the Greek origins of the word *octopus*, is “octopodes”, yet this is far less favoured by English speakers, to the extent that the automatic spelling corrector provided by Microsoft Word deems “octopi” to be valid and “octopodes” to be a misspelling. This usefulness of analogy in dealing with irregular plurals even extends to the treatment of regular verbs, where a compelling analogy can make even a regular verb like “dive” seem irregular. The analogy *drive:drove:dive:dove* will rightly strike some readers as invalid, yet many Eastern American speakers strongly prefer “dove” to “dived” (Trask 1996). In contrast, the analogy *teach:taught::catch:caught* seems a valid one, though as Trask notes, “*catched*” is actually the historically favoured past tense of *catch*. Again, however, analogy prevails to the extent that most spelling checkers will flag “*catched*” as a misspelling (no doubt due to the fact that spelling checkers are based on a corpus analysis of how language is actually used, rather than how it should be used).

Morpheme-level analogies can do more than suggest inflection patterns, and can even be used to derive new words and meanings that frequently exhibit a high degree of lexical creativity. Consider an example of analogy-based derivational morphology:

- (2) *astronomy : astronomer :: gastronomy : gastronomer*

Neither WordNet nor the spelling checker for Microsoft Word recognise “*gastronomer*” as a valid word, though a web-search reveals that it is a real word with the semantics that one would expect from the analogy (i.e., a specialist in gastronomy). The analogy in (2) is semantically sound since *Astronomy* and *Astronomer* are strongly related concepts, but as one allows the analogy to veer towards the speculative, and to rely as much on sound similarity as semantic similarity, one can achieve even more innovative results, as in

- (3) *astronomy : astronaut :: gastronomy : gastronaut*

The relation between *astronomy* and *astronaut* is a good deal more tenuous than that between *astronomy* and *astronomer*, but a relation does exist (one observes the stars, the other explores the stars, in name at least). Indeed, one can argue that *Astronaut* is itself analogically derived from *Argonaut*. A *gastronaut* might thus denote anything from an adventurous *gastronomer* to a food tourist; at the very least, we know that a *gastronaut* is a person, with some of the signal characteris-

tics of an astronaut (bravery, perhaps), who takes his directions from the field of gastronomy.

If it seems that the process of lexical creativity can be as strongly influenced by phonetic concerns as semantic concerns, this should not be too surprising a conclusion. New words survive and thrive for a whole host of reasons, but an important factor in their survival is euphony: natural sounding words are more likely to secure a lasting place in the lexicon than those that are difficult to pronounce. Analogies with existing words can transplant the euphony of an original form onto a newly minted neologism only if phonetic similarity is also allowed to influence the mapping. Indeed, the most innovative creations may give so much prominence to phonetic similarity that an analogy may lack a credible semantic basis. Nonetheless, as demonstrated in (4), terms predicated on a false analogy can still be seen as lexically innovative:

- (4) astronomy : astrodome :: gastronomy : gastrodome

Of course, there is no real semantic connection between astronomy and astrodome. Nonetheless, the word “gastrodome” can be seen as a deliberate malapropism that amply suggests a place where gastronomy is performed, and perhaps even celebrated. The analogy works, despite its lack of semantic grounding, because “dome” is itself a word denoting a large enclosed space where people congregate, like an arena or stadium. A second, implied analogy can be used to tease out its precise meaning:

- (5) astrodome : stadium :: gastrodome : restaurant

That is, just as an astrodome is a large, impressive stadium, a gastrodome is a large impressive restaurant (where restaurant is itself implied by the morphological conjunction of gastronomy and place in “gastrodome”). Ultimately, “Gastrodome” is preferable to the neologisms “Gastroarena” and “Gastrostadium” in part because the largely phonetic analogy ensures that it is a euphony combination of morphemes, each of which can be considered in isolation to provide a compositional meaning to the neologism as a whole.

3. Analogy as a mechanism of term creation

Moving from the level of morphemes to that of words, analogy again reveals itself as a powerful force in the creation of compound terms. Consider the analogy of (6):

- (6) Greek-Alphabet : Hebrew-Alphabet :: Greek-Deity : Hebrew-Deity

The analogy captures a basic symmetry both in the way concepts can be differentiated and how such differentiations are lexically expressed as compound terms. Both “Greek” and “Hebrew” denote a cultural amalgam of people, language and belief, so it makes sense to conclude that if “Greek” can be used to culturally differentiate a particular concept, then so can “Hebrew”. In fact, WordNet contains only three of the four compound terms in the above analogy: “Greek-Alphabet”, “Hebrew-Alphabet” and “Greek-Deity”. The lexical concept “Hebrew-Deity” is not listed as a WordNet entry, most likely because it is deemed to have little indexing value; while there are many Greek deities listed in WordNet that would structurally benefit from the clustering offered by the hypernym Greek-Deity, only one deity, Jehovah, is listed as having a Hebrew origin. Yet the concept seems logically well-formed, and a usage analysis (using the World Wide Web as a corpus) reveals that the term “Hebrew deity” has relatively widespread acceptance. The analogy suggests then that, on the basis of the similarity between Greek and Hebrew, WordNet should incorporate the lexical concept Hebrew-Deity. Were it to do so, its treatment of deities would become more systematic and balanced, with each proper deity (such as “Zeus”, “Mars” and “Jehovah”) instantiating a compound category that denoted its cultural basis.

Nonetheless, a simple proportional analogy like (6) may seem a weak basis on which to predict the existence of a new term. Consider the analogies of (7) and (8):

(7) Roman-Alphabet : Greek-Alphabet :: Roman-Empire : Greek-Empire

(8) Roman-Alphabet : Hebrew-Alphabet :: Roman-Empire : Hebrew-Empire

Here we seem to be predicating the possession of an empire on the existence of a unique alphabet, but an alphabet alone does not an empire make. The analogy of (7) holds true, since there is a historical entity called the “Greek Empire”, but the term “Hebrew Empire” can only be used metaphorically, perhaps to refer to the Jewish diaspora. However, this is not to say that analogy cannot be of use here, for what (7) and (8) fail to reveal is the variety of different analogies that support (7), and the comparative dearth of analogies that support (8). In addition to alphabets, the Greeks and the Romans both possessed their own mythologies, architectures, religions and deities. This semantic isomorphism suggests that if Rome possessed its own empire, it is at least meaningful to consider the possibility of a Greek empire also. The analogy in (7) is therefore strengthened by the lexico-conceptual fit between the concepts Roman and Greek, while the analogy in (8) is much weaker because of the lack of a coherent fit. This “fit” is not a measure of ontological closeness, but a measure of the overlap between the set of affordances possessed by both concepts. We can loosely estimate this set of affordances by observing the

lexical behavior of each term and how it relates to others. These observations will require us to define a set of basic term composition and decomposition operators along the following lines:

UM {X} : Usage as modifier: return a set of all compound terms such that the modifier of each is a member of the set {X}.

E.g., UM{Greek, Roman} = {Roman-deity, Greek-deity, ...}

UH {X} : Usage as head: return a set of all compounds such that the head of each is a member of the set {X}.

E.g., UH{Greek, Roman} = {Ancient-Greek, Times-Roman, ...}

M {X} : get modifiers: return the set of all modifiers of all compounds in {X}

E.g., M{Ancient-Greek, Times-Roman, ...} = {Ancient, Times, ...}

HX} : get heads: return the set of all heads of all compounds in {X}

E.g., H{Ancient-Greek, Times-Roman, ...} = {Greek, Roman, ...}

C ({X}, {Y}) : combination: return the set of all possible compound terms whose modifier is in {X} and whose head is in {Y}

E.g., C({Greek, Hebrew}, {Alphabet, Deity}) = {Greek-Alphabet, Greek-Deity, ...}

These operators allow us to dissect existing compound terms into their component parts (modifier and head), retrieve compound terms with a particular sub-component (modifier or head), and create novel combinations of these sub-components (modifiers crossed with heads). We can thus estimate the fitness of a novel compound X-Y in terms of the set of known compounds that support it, as follows (where L here denotes the set of all lexical items in the lexicon, i.e., all known terms):

$$\text{support-set}(X-Y) = C(M(UH\{Y\}), H(UM\{X\})) \setminus L$$

For example, consider the support set for the novel compound Hebrew-Deity from (6):

$$\begin{aligned} &\text{support-set}(\text{Hebrew-Deity}) \\ &= C(M(\{\text{Greek-Deity, Roman-Deity, Semitic-Deity, ...}\}, \\ &\quad H(\{\text{Hebrew-Alphabet, Hebrew-Calendar, ...}\})) \setminus L \\ &= C(\{\text{Greek, Roman, Semitic, Hindu, Celtic, Norse, ...}\}, \\ &\quad \{\text{Alphabet, Calendar, Lesson}\}) \setminus L \end{aligned}$$

= {Greek-Alphabet, Roman-Alphabet,
Roman-Calendar, Hindu-Calendar}

In effect, this set of four existing compounds represents the lexico-conceptual cross-product of the lexical concepts Hebrew and Deity. The larger the cross-product, the greater the potential interaction – and the greater the lexical fit – between both terms. One might think it strange that a lexical concept like Hindu-Calendar should support a term like Hebrew-Deity, but the intuition at work here is that deities and calendars appear to be differentiated in the same kind of way (e.g., culturally) and thus possess many of the same affordances.

Working backward from this formulation of fitness, we can formulate a generation mechanism for producing new compound terms from the old, one that implicitly incorporates the notion of lexical analogy. Consider that when a compound like Hebrew-Deity is generated from an analogy involving Hebrew-Alphabet, the head term is effectively modulated from Alphabet to Deity (see Veale et al. 2004; Hayes and Veale 2005). Head modulation thus offers an alternate perspective on the generation process. Imagine that $\text{newH}(X-Y)$ is a function that derives, via head modulation, a set of novel compound terms from an existing term $X-Y$. Using the operators above, newH can be formulated as follows:

$$\text{newH}(X-Y) = C(\{X\}, H(UM(M(UH\{Y\}))) \setminus \{Y\}) \setminus L$$

For example,

$$\begin{aligned} \text{newH}(\text{Muslim-Calendar}) &= C(\{\text{Muslim}\}, H(UM(M\{\text{Hebrew-Calendar, Hindu-Calendar ...}\}))) \setminus L \\ &= C(\{\text{Muslim}\}, H(UM\{\text{Hebrew, Roman, Hindu, ...}\})) \setminus L \\ &= C(\{\text{Muslim}\}, H(\{\text{Hebrew-Alphabet, Roman-Deity, ...}\})) \setminus L \\ &= C(\{\text{Muslim}\}, \{\text{Alphabet, Deity, Empire, ...}\}) \setminus L \\ &= \{\text{Muslim-Alphabet, Muslim-Deity, ...}\} \end{aligned}$$

The resulting set of speculative compounds must now be evaluated for lexico-conceptual fitness, using the support-set measure described earlier. At this point we expect those compounds with the greatest fit to be the best candidates for lexical innovation and subsequent admission to the lexicon. Before taking this final step, which could potentially corrupt the lexicon, we can apply a further fitness filter by demanding that each new term be present a given number of times in a given corpus (such as the WWW). Veale et al. (2004) report experimental findings which suggest that the probability of finding a newly generated term in a corpus such as the WWW increases with the size of the support set for that term.

The larger the support set, then, the safer it is to conclude that a lexical innovation is in fact meaningful.

If new compounds can be generated by modulating the head component of existing terms, it follows that generation can also proceed via a process of modifier modulation, whereby the modifier component of an existing term is modulated according to an implicit analogy. We can formulate modifier modulation as follows:

$$\text{newM}(X-Y) = C(M(UH(H(UM\{X\})))\backslash\{X\}, \{Y\}) \setminus L$$

For example,

$$\begin{aligned} \text{newM}(\text{Hebrew-Alphabet}) &= C(M(UH(H(\{\text{Hebrew-Lesson}, \text{Hebrew-Calendar}\})), \{\text{Alphabet}\}) \setminus L \\ &= C(M(UH(\{\text{Lesson}, \text{Calendar}\})), \{\text{Alphabet}\}) \setminus L \\ &= C(M(\{\text{German-Lesson}, \dots, \text{Muslim-Calendar}\dots\}), \{\text{Alphabet}\}) \setminus L \\ &= C(\{\text{German}, \text{French}, \text{Muslim}, \dots\}, \{\text{Alphabet}\}) \setminus L \\ &= \{\text{German-Alphabet}, \dots, \text{Muslim-Alphabet}, \dots\} \end{aligned}$$

So we speculatively create the compounds German-Alphabet because of an implicit analogy between Hebrew-Lesson and German-Lesson, and Muslim-Alphabet because of the implicit analogy between Hebrew-Calendar and Muslim-Calendar.

Alternately, we can conceive of a generation scheme in which both the modifier and head of an existing term can be modulated simultaneously, as formulated below:

$$\text{newMH}(X-Y) = C(M(UH\{Y\})\backslash\{X\}, H(UM\{X\})\backslash\{Y\}) \setminus L$$

This highly speculative formulation generates the cross-product of all modifiers that can apply to Y with all heads that can be modified by X. Again, many untenable combinations will be produced, but following Veale (2004), we can expect that those with a sizeable support set will be meaningful.

3.1 Experimental support

This belief is further supported by an experiment in which 100,000 novel compounds were chosen at random from the set of all compounds that can be created via the modulation of existing WordNet compounds. These new compounds, created using the formulations of newMH, newM and newH given above, are grouped into different categories according to the size of their support sets; for example, compounds with a support set of 5 other compounds are organised un-

der the category group-5, and so on. We can thus estimate the probability that a compound with a support set of size n will be validated on the WWW as that fraction of those elements of group- n that are so validated. In fact, there is a significant positive correlation (0.4) between n and the probability that an element of group- n will be validated via web-search. This correlation remains stable whether modifier modulation (newM), head modulation (newH) or simultaneous modulation (newMH) is used to generate the test data.

3.2 Phonetic analogies revisited

We are now in a better position to consider the phonetically-inspired morpheme-level analogies of Section 2. In each case we can view a multi-morphemic word as a compound term, composed of morphemes rather than words, such that one of these morphemes is modulated by an implicit analogy. Consider again the analogy of (2):

(2') astro-onomy : astro-onomer :: gastro-onomy : gastro-onomer

Now we can view “gastronomer” as a product of head modulation, where the head morpheme “-onomy” is transformed into the morpheme “-onomer” on the basis of an analogy with astronomy:astronomer. A similar process occurs in (3'), with the additional phonetic similarity between “astro-” and “gastro-” ensuring that the modulation, which is morphologically sound, also produces a euphonious result.

(3') astro-onomy : astro-onaut :: gastro-onomy : gastro-onaut

Note that the modulation perspective saves us from having to rationalise a relationship between astronomy and astronaut, allowing us instead to view them as words that share a common modifier “astro-”. This in turn allows us to exploit analogies like that of (4') where no such semantic relationship exists:

(4') astro-onomy : astro-dome :: gastro-onomy : gastro-dome

The terms “gastronaut” and “gastrodome” each have a singleton support set, corresponding to a single analogy of (3') and (4') respectively. In lieu of substantial support, however, these analogues are grounded by a phonetic similarity to their supports and this provides the requisite credibility for the new terms. That is, the similarity between “gastro-” and “astro-” is itself a support for the new terms. Note also that order is important, yielding a bootstrapping effect as new terms are incrementally accepted into the lexicon. For instance, if (4') is processed after (2') and (3'), the support for “gastrodome” can be determined as follows.

$$\begin{aligned}
 & \text{support-set}(\text{gastro-dome}) \\
 &= C(M(\{\text{astro-dome}\}), \\
 &\quad H(\{\text{gastro-onomy, gastro-onomer, ...}\})) \setminus L \\
 &= C(\{\text{astro-}\}, \{-\text{onomy, -onomer, -onaut}\}) \setminus L \\
 &= \{\text{astronomy, astronomer, astronaut}\}
 \end{aligned}$$

4. Ad hoc categories as analogical by-products

Thus far we have considered the deliberate and explicit creation of new terms and categories whose existence is predicated on an implicit analogy (i.e., where the analogy is implicit in the workings of newM, newH and newMH). We now consider the situation where a new lexical concept is created implicitly, as a by-product of the interpretation of an explicit analogy. For instance, consider the analogies in (9):

- (9) a. Zeus : Greek :: Jupiter : Roman
 b. Zeus : Greek :: ??? : Roman
 c. “Zeus is the Greek Jupiter”

The analogy of (9a) establishes an explicit mapping between Zeus and Jupiter and between Greek and Roman, suggesting that Zeus is the Greek equivalent of Jupiter. The variant in (9b) employs an elliptical form of analogy commonly found on scholastic aptitude tests, and requires us to provide the missing information; in effect, it equates to the question “What or Who is the Roman Zeus”? In contrast, the variant of (9c) assumes a compressed natural language form that can also be considered a metaphoric expression (e.g., see Hutton 1982).

The implicit relation common to (9a), (9b) and (9c) appears to be ??“deity of”: Zeus is a deity of the Greeks, while Jupiter is a deity of the Romans. However, consider the longer form of this analogy in (10):

- (10) Zeus is to Greek as
 a. Skanda is to Hindu
 b. Thor is to Norse
 c. Jupiter is to Roman
 d. Brigit is to Celtic
 e. Donar is to Teutonic

Each of the candidate pairings in (10) can be seen as instantiating the “deity of” relationship, so a more specialised relationship is clearly at work here. In fact, the correct relationship is “supreme deity of”, since this is the only conceptual rela-

tionship for the stem pairing that picks out just one of the five possible candidates. Now, WordNet contains the concept Deity, so one can imagine constructing the relationship “deity of” from this concept in a relatively straightforward fashion. But WordNet does not contain the concept Supreme-Deity, and for good reason: it is not a conventional collocation, and its meaning is simply a compositional function of existing terms. One of two situations must therefore hold: either the concept already exists but is not lexicalised; or else neither the concept nor its lexicalisation exists prior to the analogy. In either case, we can reasonably assume that the lexical term “supreme deity” is constructed especially to resolve the analogy.

Not all such analogies require us to construct new lexical concepts. Consider the analogy in (11), which can be seen as a close conceptual neighbour of (9a):

Ares : Greek :: Mars : Roman

Here it is the relationship “war god of” that connects Ares to Greek and Mars to Roman. In this case, however, WordNet does contain the lexical concept War-God, while its lexicalisation “war god” is such a conventional collocation that few would argue that it is constructed especially for the purpose of this analogy. However, this is not to say that the interpretation of (11) should be substantially different from that of (9) or (10). We can still presuppose that for each analogy, the same process is employed to construct a relational category between each concept in each pairing. In the case of (11), this relational category (War-God) will correspond to an existing lexical concept, while in (9) and (10) it will result in a lexical innovation (“supreme deity”) that may be added to the lexicon following an assessment of its support set or a corpus analysis.

The construction of these relational categories raises two key questions: first, where do the component parts such as “war”, “supreme” and “deity” come from; and second, why are these components, rather than others, selected? The lexicon or lexical ontology presumably plays a central role in resolving these questions, which further begs the question of what theory of the lexicon we should adopt. To remain as agnostic as possible, let us assume a rather simple, feature-theoretic view of the lexicon. Let F denote a function that maps a lexical concept onto a set of component features. Furthermore, let us assume that these features can be of one of two types. Taxonomic features, denoted with a $\uparrow$, are those that indicate the position of a concept in the lexical ontology. Associative features, denoted with a $@$, are those that predicate descriptive properties of the concept. For instance, consider Zeus again:

$F(\text{Zeus}) = \{\uparrow\text{deity}, @\text{Greek}, @\text{supreme}, @\text{mythology}, @\text{Olympus}\}$

Thus, Zeus is a deity that is Greek and supreme, associated with both mythology and Olympus. In contrast, we can define Jupiter as follows:

$$F(\text{Jupiter}) = \{\uparrow\text{deity}, @\text{Roman}, @\text{supreme}, @\text{mythology}, @\text{rain}\}$$

Jupiter is thus a deity that is Roman and supreme, associated with both mythology and rain (in the guise of Jupiter Pluvius). This feature-level decomposition suggests a means whereby new categorisations can be created for a given concept. Consider the following formulation of a function *alt*, which derives a set of alternate categorisations for a concept by constructing alternate compositions of elements in *F*:

$$\begin{aligned} \text{alt}(A) = & C(\{X \mid \text{UM}\{X\} \neq \{\} \\ & \wedge @X \in F(A) \wedge \uparrow X \notin F(A)\}, \\ & \{Y \mid \text{UH}\{Y\} \neq \{\} \wedge \uparrow Y \in F(A)\}) \end{aligned}$$

That is, the set of alternate categorisations of *A* comprises just those compound terms that can be created by combining the associative features of *A* that have in the past been used as compound modifier with the taxonomic features of *A* that have in the past been used as compound heads. The resulting compound terms are thus well-formed with respect to the lexicon and the language that it represents. Note that this formulation of *alt* prohibits the hypernymic terms of a concept (like $\uparrow\text{deity}$ for Jupiter) from serving as a modifier in any alternative categorisation of the concept, since this is a combination strategy rarely seen among English compounds.¹

Now, a simplistic view of analogy, based on the Aristotelian account (see Hutten 1982), might attempt to reconcile Zeus and Jupiter by seeking a common taxonomic feature (e.g., $\uparrow\text{deity}$) in both representations, but as demonstrated by (10), a genus term alone lacks discriminatory power. We need a common category that combines the Aristotelian notions of both genus and differentia. Given an analogical pairing *A*:*B*, we can construct this category using the function *adhoc*, formulated as follows:

$$\begin{aligned} \text{adhoc}(A:B) = & \{X-Y \mid X-Y \in (\text{alt}(A) \cap \text{alt}(B)) \\ & \wedge \neg(\exists P \uparrow P \in F(A)) \end{aligned}$$

1. Generally speaking, the modifier of a compound term denotes a property of the head (as in “wax paper”) or a concept from which a property is transferred to the head (as in “beehive hairdo”) or a concept to which the head relates via slot-filling (as in “harpoon gun”). In some hybrid compounds, both the modifier and the head denote a hypernym of the compound, as in “sofa bed”, but these are so rare as to be safely precluded from the current analysis. Were we to allow hybrid compounds, the formalism given here would surely over-generate; that is, precision would greatly suffer in exchange for modest gains in recall.

$$\begin{aligned} &\wedge \uparrow P \in F(B) \\ &\wedge \uparrow Y \in F(P))\} \end{aligned}$$

Expressed in English, *ad hoc* (A:B) generates a set of compound terms X-Y such that: i) X-Y is an alternative categorisation of both A and B; and ii) there is no other shared taxonomic feature of A and B (P, say) that is more specific than Y.

An analogy A:B::C:D is well-formed if precisely the same relationship holds between A and B and between C and D. For example, the analogy Zeus:Hindu::Jupiter:Roman is malformed because Zeus is not Hindu but Greek. Thus:

$$\begin{aligned} \text{well-formed}(A:B::C:D) &= (\exists M \quad M-B \in \text{alt}(A) \\ &\quad \wedge \quad M-D \in \text{alt}(C)) \\ \vee (\exists H \quad B-H &\in \text{alt}(A) \\ &\quad \wedge \quad D-H \in \text{alt}(C)) \\ \vee (\exists M_1 M_2 H_1 H_2 \quad M_1-H_1 &\in \text{alt}(A) \\ &\quad \wedge \quad M_1-H_2 \in \text{alt}(C) \\ &\quad \wedge \quad M_2-H_1 \in \text{alt}(B) \\ &\quad \wedge \quad M_2-H_2 \in \text{alt}(B)) \end{aligned}$$

The first disjunct covers the situations where B and D are super-ordinates of A and C (as in the analogy *ewe:sheep::hen:chicken* where coherence is given by the relations *female-sheep* and *female-chicken*). The second disjunct covers the situations where B and D are features of A and C (as in the analogy *Athena:Greek::Ganesh:Hindu*). The third disjunct, the most complex, covers those situations where B and D are in some sense antonyms of A and C (as in *wife:prostitute::umbrella:cab*). Now, well-formedness does not always imply solvability; for that, there must exist a relationship between A and C that is mirrored between B and D. Thus, given the analogy A:B::C:D, we additionally expect that it has a non-empty relational basis:

$$\text{basis}(A:B::C:D) = \text{ad hoc}(A:C) \neq \{\}$$

That is, the pairing A:C in a proportional analogy should share at least one relational category if A:B:C:D is to be considered a solvable analogy. As formulated above, *basis* may return a set containing a plurality of categories. In the case of analogies like (9a) and (11), it is sufficient that this be a non-empty set. But in the case of long-form analogies like (10), where a stem pairing must be matched with just one other in a group of candidate pairings, it may be possible that multiple candidate pairings share a non-empty relational basis with the stem pairing. In this case, one must choose the candidate with the strongest relational basis. Since each element returned by *basis* is a conceptual category, we can determine the discrimination strength of each category by considering it from an extensional

perspective. Given two categories in the relational basis of an analogical pairing, e.g., supreme-deity and Greek-deity, the strongest category is taken to be that which has the smallest extension (and which is thus the most discriminating). The extension of Greek-deity is larger than that of supreme-deity (108 members versus 6 members in WordNet), so we take supreme-deity to be the stronger category on which to ground an interpretation.

What of partial analogies like (9b), which form the basis of both examination questions (where a student must provide the missing information) and metaphoric allusions? In such cases, a suitable analogue must be retrieved to complete the analogy, using the available information as a retrieval cue. We can formulate a retrieval-oriented variant of basis as follows:

$$\text{basis}(A:B::???:D) = \{X-Y \mid \exists C @D \in F(C) \\ \wedge X-Y \in \text{adhoc}(A:C)\}$$

If the lexicon is sufficiently indexed, as one might expect in a structured lexical ontology, it should be relatively straightforward to identify C using D as an index.

5. Analogical retrieval in WordNet

The comprehensive scale of WordNet as a lexical database of English word meanings, with over 100,000 lexical concepts, allows us to put the intuitions and formulations of previous sections to the test. The specific task we propose in this section is that of analogical retrieval (see Veale 2003b; Veale 2004): given a lexical concept in one domain, such as “Zeus”, and a modifier that denotes another domain, such as “Roman”, we seek to retrieve those concepts in the modifier domain that are meaningful analogies for the original head concept. The retrieval task is thus a question-answering task, in which we attempt to find answers for queries such as “Who is the Norse Zeus?” and “Who is the Hindu Athena”. For balance, we shall conduct our test in two different domains of knowledge, namely deities and alphabets. The deities domain is quite well represented in WordNet, while structurally, the alphabetic domain is relatively impoverished. We shall demonstrate that the creation of ad hoc categories that are subsequently admitted to the lexicon can significantly improve the state of these impoverished domains.

We concentrate our efforts then on the noun section of WordNet, which contains over 70,000 taxonomically organised entries. In addition to this taxonomic information, WordNet associates a textual gloss with each entry, much like that offered by a regular dictionary. For example, WordNet associates the following information with the concepts Zeus, Jupiter, Alpha and Aleph:

Zeus: Taxonomy = {Greek-deity is-a deity is-a god ...}
 Gloss = "The supreme god of ancient mythology"
 Jupiter: Taxonomy = {Roman-deity is-a deity, is-a god...}
 Gloss = "(Roman myth) supreme god of Romans"
 Alpha: Taxonomy = {letter is-a character is-a written-symbol ...}
 Gloss = "the 1st letter of the Greek alphabet"
 Aleph: Taxonomy = {letter is-a character is-a written-symbol ...}
 Gloss = "the 1st letter of the Hebrew alphabet"

Unfortunately, WordNet does not offer an explicitly feature-theoretic description of each lexical concept, such as that provided by our function *F*. However, we can approximate the corresponding *F* for WordNet by assuming that the textual gloss of each concept is, in fact, a bag of associative features; we simply eject any non-content words (such as determiners, prepositions, and so on), and merge the resulting word set with the set of taxonomic parents that is explicitly provided by the WordNet. Thus, from WordNet we derive the following mappings for *F*:

F(Zeus): = {↑Greek-deity ↑deity @supreme @god @ancient
 @mythology}
F(Jupiter): = {↑Roman-deity ↑deity @Roman @supreme @god
 @Romans}
F(Alpha): = {↑letter ↑character @1st @letter @Greek
 @alphabet}
F(Aleph): = {↑letter ↑character @1st @letter @Hebrew
 @alphabet}

Applying the function *adhoc* to these representations, we obtain the following:

adhoc(Zeus:Jupiter) = {supreme-deity}
adhoc(Alpha:Aleph) = {1st-letter, alphabet-letter}

Note that the ad hoc concepts god-deity and letter-letter, though seemingly possible from the given values of *F*, are not created because of the definition of *alt* as formulated earlier (i.e., no taxonyms as modifiers). Note also that *adhoc* returns two different categories for the pairing of Alpha with Aleph. In this case, based on the extension of both categories, 1st-letter is deemed the stronger of the two. In fact, an extensional analysis reveals that the extension of 1st-letter (with just two members) is a proper subset of that of alphabet-letter (with 49 members), which suggests that 1st-letter is a specialisation of the category alphabet-letter.

Figure 1 illustrates the taxonomic structure of the letter domain in WordNet before any letter analogies (of the form Alpha:Greek::???::Hebrew) have been interpreted. Note the general paucity of organisational structure here: each letter

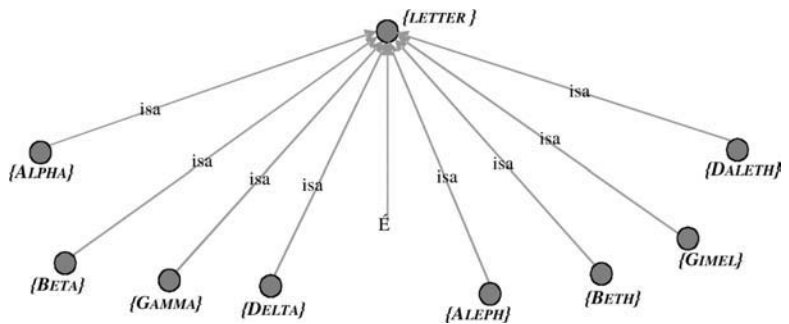


Figure 1. Structure of the Greek and Hebrew letters domain in WordNet

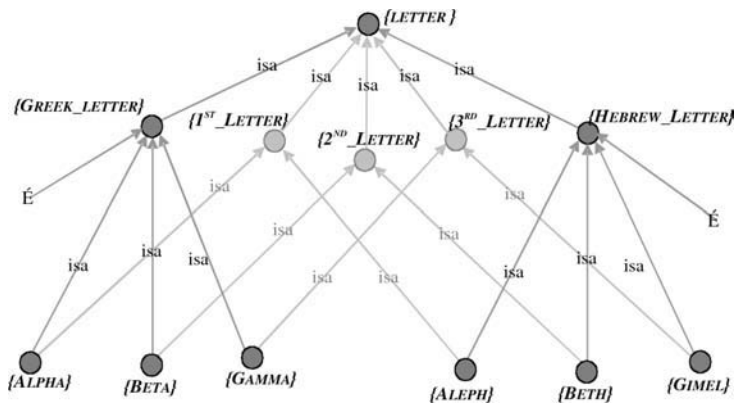


Figure 2. WordNet supplemented with new adhoc categories like Greek-letter, Hebrew-letter and 1st-letter

from each alphabet is forced to share the same super-ordinate category, letter, and no attempt is made to gather letters from different alphabets under separate super-ordinates.

This picture changes dramatically once each letter in the Greek alphabet is placed in analogical alignment with its corresponding letter in the Hebrew domain. Note that as the latter lacks vowels, a strict 1-to-1 alignment is not possible. Figure 2 illustrates the situation once the adhoc function has been allowed to create new lexical terms to cluster each pairing of letters under an analogically-specific category.

Table 1. Cross-domain variants are sought for each member of five deity pantheons

Ad hoc categories	Greek	Roman	Hindu	Norse	Celtic
Supreme-deity	Zeus	Jove	Varuna	Odin	N/A
wisdom-deity	Athena	Minerva	Ganesh	n/a	Brigit
beauty-deity, love-deity	Aphrodite	Venus	Kama	Freyja	Arianrhod
sea-deity*	Poseidon	Neptune	n/a	n/a	Lir
fertility-deity	Dionysus	Ops	n/a	Freyr	Brigit
Queen-deity	Hera	Juno	Aditi	Hela	Ana
war-deity*	Ares	Mars	Skanda	Tyr	Morrigan
hearth-deity	Hestia	Vesta	Agni	n/a	Brigit
Moon-deity	Artemis	Diana	Aditi	n/a	n/a
sun-deity*	Apollo	Apollo	Rahu	n/a	Lug

* WordNet contains the concepts Sea-God, War-God and Sun-God

6. Evaluation

We first consider the effectiveness of ad hoc category construction on the precision and recall of analogical retrieval in the WordNet deities domain. Table 1 presents the results of an experiment in which analogical variants are sought for the members of five different families of deity.

This experiment thus involves 20 different mapping tasks (i.e., Greek to Roman deities, Hindu to Norse deities, Celtic to Greek deities, etc.). The average precision of analogical retrieval across all tasks is 93%, while the average recall is 61%.

For the letter mapping experiment, an analogous Hebrew letter was retrieved for each Greek letter, and vice versa. The ad hoc categories created for each retrieval are of the form 1st-letter, 2nd-letter, and so on, and serve to pinpoint a precise analogue whenever one is available (that is, each ad hoc category has an extension containing precisely two members). The precision for the letter experiment is thus 100% (that is, no retrieval errors are made). Since the Greek alphabet has more letters than the Hebrew alphabet, recall is 100% for the Hebrew to Greek task, but only 96% for the Greek to Hebrew task (as the latter has one less letter than the former).

Explicit category creation in WordNet

Though we have described the process of ad hoc category creation as an implicit by-product of analogical reasoning, our formulations of alt and adhoc nonetheless

less allow us to exploit analogy as a deliberate mechanism of explicit category and term creation. For every lexical concept *A* in WordNet, we need simply consider those alternate categorisations (derivable via *alt*) that are also generated by at least one other concept:

$$\begin{aligned} \text{adhoc}(A:???) &= \{X-Y \mid X-Y \in \text{alt}(A) \\ &\quad \wedge \exists B A \neq B \wedge X-Y \in \text{alt}(B) \} \end{aligned}$$

In effect, we are generating alternate categorisations of a given concept that have the analogical potential to relate that concept to at least one other in the ontology. That is, we interest ourselves here only with those alternate categorisations that possess an extension of two or more members, and which might thus make non-trivial additions to the ontological lexicon to serve a genuine organisational purpose. For example, the alternate categorisation Greek-Wine constitutes a trivial addition to WordNet, since it serves to index a sole category member, Retsina. In contrast, the categorisation Italian-Wine serves to index at least three different members (sweet vermouth, soave and Chianti). By this measure, Italian-Wine serves a useful indexing and clustering role in the ontology and should be retained, while Greek-Wine serves no clustering role and should be discarded.² In the food domain alone, WordNet provides definitions for over 200 different terms whose gloss mentions a proper-named country like “Italy”, “Greece” or “Mexico”, so we should expect that the alternate categoriser (as formulated via *alt*) will pick out these national ties as features to be reified.

Applying the above formulation of *adhoc* to the 70,000+ noun concepts in WordNet, we obtain 8564 new and non-trivial compound categories. In total, these 8564 compounds differentiate 2737 different head concepts, suggesting that each head is differentiated in three different ways on average. Overall, the most differentiating modifier is “Mexico”, which serves to differentiate 34 different heads; for example, Mexico-Dish serves to group together Taco, Burrito and Refried-beans. The most differentiated head is “herb”, which is differentiated into 134 sub-categories such as Prickly-Herb, Perennial-Herb, European-Herb, etc.

To consider just a few other domains: sports are differentiated into team sports, net sports, court sports, racket sports and ball sports (surprisingly, but not meaninglessly, Bingo becomes categorised as a Ball-Game); constellations are

2. While Greek-Wine serves no useful clustering role, inasmuch as it serves to index just one concept, it may yet be seen as a useful addition to the ontology for reasons of symmetry. In an ontology that contains nodes like French-Wine, Italian-Wine and German-Wine, the addition of Greek-Wine, if only to index a single instance, would enhance the systematicity of the ontology (in that the Wine node would be consistently sub-organised by Nationality). The criteria considered in this section should thus be viewed as heuristics rather than hard constraints.

divided into northern and southern variations; food dishes are differentiated according to their nationalities and their ingredients, e.g., into cheese dishes, meat dishes, chicken dishes, rice dishes, and so on. As noted earlier, letters are differentiated both by culture, giving Greek letters and Hebrew letters, and by relative position, so that “Alpha” is both a 1st_letter and a Greek_letter, while “Aleph” becomes both a 1st_letter and a Hebrew_letter. Likewise, Deity is further differentiated into War_deity, Love_deity, Wine_deity, Sea_deity, Thunder_deity, Fertility_deity, and so on.

One can validly ask whether such terms are truly creative, for it seems that we comprehend linguistic creativity here in its broadest sense, that used by Chomsky (1957) to describe the potential of human language to generate (i.e., create) an unlimited number of valid word combinations. It would seem by our reckoning, then, that any novel combination of words that is syntactically and semantically valid should be considered creative. This criticism would certainly be apropos if the compounds under consideration were either entirely lexical or entirely conceptual. However, these compounds are both lexical and conceptual and are created relative to a lexical ontology in which they serve a useful organisational role. Compounds like “Strong-Drink” or “Love-deity” may seem mundane as linguistic artefacts in the context of general language usage, but from the context of a lexical ontology, they represent an insightful partitioning of a given conceptual space. Creativity requires clarity of perception, and the value of this insight can be seen most forcefully in the kinds of analogies that these new categories allow one to construct. For instance, the category Strong-Drink creates a cluster of diverse (but appropriate) bedfellows from espresso (strong coffee) to concentrated orange juice (strong juice) to whiskey (strong liquor). In turn, this cluster provides a firm lexico-conceptual basis for analogies of the kind whiskey:liquor::espresso:coffee.

7. Conclusions

With this paper I have attempted to provide a common formalisation – in terms of lexical composition and decomposition operators – for two different perspectives on the production of new lexical terms and categories. These perspectives are both analogically-motivated, and concern the explicit and implicit use of analogy in the implicit and explicit creation of new terms and categories. As such, it should be clear that we assign to analogy a central role in the mechanism of linguistic creativity.

Both perspectives create compound terms of the same form – simple modifier-head constructions – using lexical precedents to ensure that each term possesses both a linguistic and conceptual validity. However, while the outputs of

both processes may look similar, the processes themselves are quite different, and cover different parts of the lexico-conceptual space. The implicit analogy approach, for instance, is only capable of generating compounds that can be reached via modulation from some existing support base in the lexicon. In contrast, the explicit analogy approach works directly with the feature-theoretic representation of concepts in the lexicon, and can generate compounds that, while meaningful, may have an empty support set. Each approach is thus complementary to the other, and both taken together yield a creative reach that is beyond either alone.

However, it is clear that a fusion of both perspectives does not provide full coverage of the lexico-conceptual space, even when this space is limited to that of simple modifier-head constructions. Consider the terms “Gastropub” (a public house that serves restaurant-quality food), “metrosexual” (a heterosexual male with female grooming and fashion habits³) and “retrosexual” (a back-formation from “metrosexual” that describes the prototypical heterosexual male against which metrosexuality is defined as a reaction). These terms each combine a bound morpheme with a free morpheme, and while their structure is easy to analyse, it is extremely difficult to hypothesise an effective generation mechanism that does not simply combine every bound morpheme with every free morpheme in the lexicon. These compounds cannot be predicted either on the basis of existing compounds (via modulation and/or phonetic similarity⁴) or on the basis of conceptual features alone. Rather, since they are created for use in a particular communicative context, it is this context that provides the missing features that would make possible both the prediction of “Gastropub” and “Metrosexual” as valid words, as well as a broad ontological categorisation of their meanings (e.g., that gastropub is a kind of public house, or that metrosexual is a type of heterosexual male).

Since portmanteau words like “Gastropub” and “Metrosexual” comprise one of the most interesting varieties of modern lexical innovation (e.g., see Veale and O’Donoghue 2001), it would be a shame if this were all that one could conclude. Most likely, there exists a middle ground in which these terms might be, if not predictable from lexical structure, then constrained by lexical structure to the extent that the addition of automatic corpus analysis (using the WWW, say) might allow a computational system the ability to harvest such novel terms and categories as they arise in a cultural setting. As such, the analysis framework described

3. Intriguingly, the term “metrosexual” was first coined by the British journalist Mark Simpson in 1994 (see Simpson 1994), but lay dormant for the rest of the decade. The term underwent a resurgence in popular culture when used in a New York Times article in 2003.

4. It may be that “Gastropub” obtains some minor support from its phonetic neighbour “Gastropod” after it has been created using other means.

here should provide an adequate basis for interpreting novel portmanteau words if such words could be harvested automatically in lieu of being predicted automatically. One harvesting source that we are currently investigating is Wikipedia,⁵ an on-line open-source encyclopaedia that is constantly updated and modified by a veritable army of users. The popularity of Wikipedia makes it an ideal source from which to harvest new words as they gain prominence in the language, long before these words earn their place in conventional print dictionaries. For instance, Wikipedia offers a detailed entry for each of “Gastropub”, “Metrosexual” and “Retrossexual”, provides links between related terms, contains sufficient context to allow an automated system to construct an interpretation (e.g., the “Gastropub” entry mentions both public houses and gastronomy), and in some cases, even provides a pertinent analogy to explain the term (e.g., Wikipedia helpfully points out that the Gastropub, as commonly conceived, is the English equivalent of the French *brasserie*).

Exploring term creation in the context of resources such as WordNet and Wikipedia, which blur the traditional distinction between dictionary and encyclopaedia, constitutes an ongoing research programme that is predicated on the belief that term creation is a scaleable phenomenon through which one can explore creativity in general. That is, the issues in term creation run the gamut from phonological to conceptual, involving terms that range from the mundane to the humorous to the wildly creative. It is our hope that an understanding of the processes that underlie term creation may thus lead to a deeper understanding of creativity overall, one that can ultimately be exploited to build computational systems that exhibit genuine linguistic inventiveness.

References

- Attardo, S, Hempelmann, C. F. and Di Maio, S. 2002. “Script oppositions and logical mechanisms: Modeling incongruities and their resolutions.” *Humor: International Journal of Humor Research* 15–1: 3–46.
- Baron, J. 1977. “What we might know about orthographic rules.” In *Attention and performance VI*. S. Dornic (ed.), 557–572. Hillsdale, NJ: Erlbaum.
- Barsalou, L. 1983. “Ad-Hoc Categories.” *Memory and Cognition* 11(3): 211–227.
- Carpuat, M., Ngai, G. Fung and Church, K. W. 2002 “Creating a Bilingual Ontology: A Corpus-Based Approach for Aligning WordNet and HowNet.” In *Proceedings of GWC 2002, the 1st Global WordNet conference*, 284–292. Mysore: India.
- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
- de Bono, E. 1994. *Parallel Thinking*. Viking Press: London.

5. <http://www.wikipedia.org>

- Dong, Z. and Dong Q. 2006. *HowNet and the Computation of Meaning*. Singapore: World Scientific.
- Freud, S. 1905. *Der Witz und seine Beziehung zum Unbewußten*. Leipzig, Vienna: Dueticke. (Reprinted Frankfurt am Main: Fischer. English edition, 1976. *Jokes and their Relation to the Unconscious*. Harmondsworth: Penguin.)
- Hayes, J. and Veale, T. 2005. "Creative discovery in the lexical validation gap." *Journal of Computer speech and Language* 19(4): 513–523.
- Hutton, J. 1982. *Aristotle's Poetics*. New York, NY: Norton.
- Lenat, D. and Guha, R. V. 1991. *Building Large Knowledge-Based Systems*. Reading, Massachusetts: Addison Wesley.
- Pustejovsky, J. 1991. "The generative lexicon." *Computational Linguistics* 17(4): 409–441.
- Miller, G. 1995. "WordNet: A Lexical Database for English." *Communications of the ACM* 38(11): 39–41.
- Rumelhart, D. E. and Abrahamson, A. A. 1973. "A model for analogical reasoning." *Cognitive Psychology* 5: 1–28.
- Simpson, M. 1994. "Here come the mirror men." *The Independent* (London), November 15 edition, p. 22.
- Trask, R. L. 1996. *Historical Linguistics*. London: Edward Arnold.
- Veale, T. and O'Donoghue, D. 2001. "Computation and Blending." *Cognitive Linguistics* 11(3/4): 253–281.
- Veale, T. 2003a. "Dynamic Type Creation in Metaphor Interpretation and Analogical Reasoning." In *Proceedings of the International Conference on Conceptual Structure, Conceptual structures for Knowledge Creation and Communication*, LNAI 2746, 471–479. Berlin: Springer Verlag.
- Veale, T. 2003b. "The Analogical Thesaurus: An Emerging Application at the Juncture of Lexical Metaphor and Information Retrieval." In *Proceedings of IAAI'03, the 2003 Innovative Applications of Artificial Intelligence conference*, 137–142. Menlo Park, California: AAAI Press.
- Veale, T. 2004. "Creative Information Retrieval." In *Proceedings of CICLing'04. Lecture Notes in Computer Science*, LNCS 2495, 15–21. Berlin: Springer Verlag.
- Veale, T., Seco, N. and Hayes, J. 2004. "Creative Discovery in Lexical Ontologies." In *Proceedings of COLING'2004, the 20th International Conference on Computational Linguistics*, 1333–1338. Geneva, Switzerland. San Mateo, California: Morgan Kaufmann.
- Veale, T. 2005. "An Analogy-oriented Type Hierarchy for Linguistic Creativity." *Journal of Knowledge-Based Systems* 19(7): 471–479.
- Way, E. C. 1991. "Knowledge Representation and Metaphor." *Studies in Cognitive systems*. Amsterdam: Kluwer Academic Publishers.

Creative lexical categorisation in a narrative fiction

M^a Dolores Porto

Lexical creativity is more than just inventing new words; invented meanings for common words (neosemes) must also be considered. Both nonce formations and neosemes are characteristic of certain narrative genres, such as science fiction and fantasy, where the creation of new concepts demands new referring units. One of the strategies that readers commonly use to interpret these is lexical recategorisation, which is also the way we cope with neologisms in everyday speech. This paper examines how creative lexical categorisation works in a Fantasy novel, and how it helps the reader with the acquisition of new concepts and eventually with the construction of a different reality, that of a text world which may be dramatically different from the real one.

1. New words in fiction

Fiction is about invention, inventing a set of characters, their relations, feelings, personalities, speech and thoughts; inventing events in which those characters will be involved; and more than anything else, it is about inventing a text world where everything will take place. Naturally, all this involves new referents that need to be named, people and places, but also, depending on the kind of fiction, new concepts or ideas may be imagined and referred to. Consequently, fiction also involves the invention of new words.

In those fictional texts where the world depicted is mimetic to the real world, those we consider to be realistic, there is not much space for the creation of new concepts and their corresponding creative names, but the farther we go from the real world, the more likely it is that we will find invented terms and expressions for imaginative objects, concepts or actions. Thus, genres such as Science Fiction and Fantasy are usually packed with invented words, even to the point that this

creative lexicon becomes a mark of these genres (for further discussion on this point see Munat in this volume).

Peter Stockwell has considered this signalling function of neologisms in science fiction:

The use of the referential power of language to help create a visual reality in the mind's eye is, in science fiction, the totem use of new words to signal to the reader that something very clever, advanced, and technological is happening. [...] If you take the opening page of a few science fiction novels or especially short stories, you're likely to find new characters, new places, new machines, new social groups, new processes, and a host of new objects, all with names invented by the science fiction author. (Stockwell 2000: 117)

What is most striking about the invented words in fiction is that readers are generally capable of guessing their meanings and understanding them, even when not provided with a complete definition. Certainly, readers are not usually puzzled when they come across one of these nonce formations, but are able to infer a reasonable meaning on the basis of the context and add the new word to their mental lexicon. Then, if the word appears in the text again – and it generally does – the reader will be able to either confirm or correct the previous guess and thus arrive at a more precise meaning. And, what is more, readers do all this without really being aware of the whole process.

As a matter of fact, this is a very common process that we follow in our daily life. Aitchison (1994: 12) points out that people continually add new words to their mental lexicon, and, moreover, “they often create new words and new meanings for words while speech is in progress.” We must conclude then that the mechanisms readers activate to interpret new words while reading fiction are the same as those they use in daily life when they encounter new terms. New words are tentatively added to the mental lexicon, together with a guess as to the meaning and possible uses. If we never encounter the word again, it will probably be forgotten. Instead, if the word reappears in several other contexts and circumstances, it will become more and more salient and the meaning will be enriched.

This is an automatic process and requires an efficient storage mechanism. Apparently, the way in which speakers of a language organise their internal lexicon is not simply a long list of words with their corresponding meanings as in a dictionary. On the contrary, each word in our mind is linked to many others, and these to still others, forming a huge network of information, the *human word-web* (Aitchison 1994: 97). In this web, once a word is activated, many others may be called into play as necessary. A stored list of words would probably prove to be a rather slow and inefficient way of managing all the information needed during speech in progress, but the web model explains how a word can provide a

variety of related information beyond the purely linguistic meaning attached to it. As Dirven and Verspoor (1998:25) put it, in a single word there is “‘the whole world’ or at least all the experiences we have of our world that have somehow been categorised linguistically.”

We will see that this is also true for the invented words in a fictional narrative. When the reader acquires a new word, instead of simply storing it as a separate entry in the mental lexicon, it is inserted in one of the already existing categories and so, in spite of being a recent acquisition, a whole network of concepts is activated because the reader will strive to link it to other familiar words.

As long as created neologisms look pronounceable [...], readers often try to interpret them in terms of words they already know. (Stockwell 2000: 124)

2. Lexical categorisation

The mental lexicon, as well as all our encyclopaedic knowledge of the world, is actually organised in terms of prototypical categories. As Geeraerts (1997:113) points out “prototype categories have structural stability and flexible adaptability to cope with changes and informational density.”

This means that the system allows us, on the one hand, to store and organise a large amount of information associated with each word and, on the other, it also allows modifications when new information is added, without causing a collapse of the entire system.

This is especially interesting in fiction, since the new information provided by the narration, of whatever nature (concepts, words, meanings...), is only acquired on a temporary basis. Prototypical categories enable readers to interpret what is new by relating it to their previous knowledge, before incorporating it into the existing categories. Finally, when the reading is complete, the reader discards the recently acquired information and returns to the primitive organisation, leaving the whole conceptual system the same as before reading.¹

1. Probably the whole cognitive model of the world will never be the same as before reading, since the “discarding” of the new concepts and terms acquired during the reading will never be complete. I would claim that there is always some sort of trace left, a footprint of the temporary recategorisation that can be reused if needed, so that the whole system will never be exactly the same. When Elena Semino considers the temporary or permanent effects of “cognitive reordering” induced by the metaphors in a text-world construction, she points out that “even the most transient of effects enables us to contemplate, at least momentarily, a new way of conceiving the world” (Semino 1997:213). Also, the new vocabulary in a fiction can give rise to some “cognitive reordering” and give a new perspective on some common concepts that might be taken

Furthermore, the possibility of integrating new words and meanings in existing categories, even if only temporarily, contributes to the make-believe of every fiction, i.e. it supplies the credibility for the story by inducing readers to blend what they know about the real world with the fictional one.

In order to understand how the categorisation of new words and meanings is performed by the reader of a fiction, it will be useful to have a wider comprehension of the nature and characteristics of prototypical categories.

Prototypical categories have been extensively studied in Cognitive Linguistics (Labov 1973; Rosch 1975 and 1978; Lakoff 1987; Langacker 1987; Tsohatzidis 1990; Taylor 1995; Geeraerts 1997). Unlike the traditional Aristotelian categories – conceived as closed compartments where members are *in* or *out* depending on how they fulfil certain conditions – prototypical categories are considered to have *fuzzy boundaries*, so it is possible to distinguish different degrees of category membership (Rosch 1975). All the members of the category are organised around a *prototype* or best example, but some members are more similar or closer than others to the prototype.² In other words, they are more central members of the category, while others are more peripheral, closer to the boundaries, and could even be shifted to a different category. For this reason, membership in a prototypical category is a matter of degree (Langacker 1987; Taylor 1995). This also means that categories can overlap, so that one item can be included in more than one category, or even be in different positions in relation to its proximity to the respective prototypes.

Analogy is the most obvious means of linking a member to the prototype, but metaphor and metonymy are also common means to extend the categories.³ Also, it must be remembered that some members are not directly connected to the prototype but to other members of the category. Following an example by Wittgenstein on the category of GAMES, Rosch and Mervis (1975) introduced the

beyond the narration itself. Obviously, a thorough discussion is beyond the scope of this paper, so let us simply agree that the new information is *discarded* at the end of the reading.

2. The prototype can be a best example or a schema (Langacker 1987), i.e. an abstract representation that contains the common features of the members of the category, but, as Taylor (1995) points out,

...categorisation by prototype occurs developmentally prior to categorisation by schema [...]. The increasing abstractness required of schematic representations suggest that schemas may only be accessible to more sophisticated, reflective language users (1995:67)

a primitive categorisation, such as that made by children, for example.

3. See Lakoff (1987), Langacker (1991), Taylor (1995), Talmy (2000).

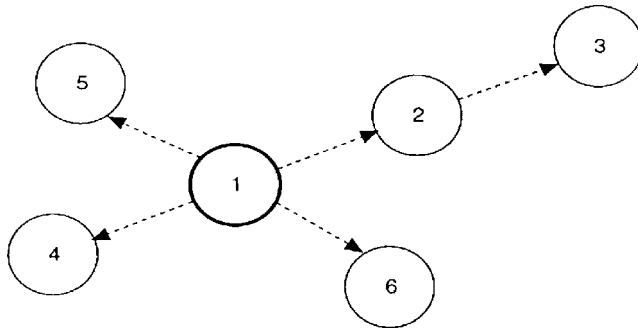


Figure 1. Radial categories as seen by Cuenca and Hilferty (1999)

concept of *family resemblances* to explain how all the members of a category are related to other members and only a few of these are linked directly to the prototype itself (see Taylor 1995 and Ungerer and Schmid 1996). As a result, prototypical categories can be visualised as radial categories (Lakoff 1987). Cuenca and Hilferty (1999) provide a figure of what a radial category might look like and how it can be expanded (Figure 1).

In the figure we can see how number 3 is not directly linked to the prototype (number 1), so it holds a peripheral position, but is still part of the category via its link to number 2.

The consequences of conceiving the mental lexicon as a web of prototypical categories rather than a long list of words and definitions are extremely interesting for the study of creative lexis in fiction. This network organisation can explain how the invented words found in a narrative can evoke others in the reader's mind and thus allow him/her to draw on a wide range of senses and associated meanings from the moment the new word is encountered, and later can be redefined, moving around inside the category. We will see that this system also explains how semantic change is possible within the scope of a novel without necessarily altering the meaning of the word with which it was originally stored.

When studying lexical categories, two perspectives are possible. We can either study the category of senses that make up the meaning of a word, or the category of words that can be chosen to refer to a concept. The first is a semasiological approach, moving from words to meanings, and explains polysemy and the relationship between the various senses of a word, whereas the second is onomasiological and goes from meanings to words and describes synonymy, variation and the relationship between the words that make up a lexical field (Geeraerts 1997; Dirven and Verspoor 1998: 25ff.; Štekauer 1998). We will see that creative categorisation is possible in both approaches to the lexicon. Creativity in semasiological categories produces new meanings for common words, *neosemes* (see Stockwell

2000: 119), whereas onomasiological creativity provides the right interpretation for neologisms or nonce formations.

Both neosemes and neologisms typically occur in fiction and allow the reader to construct new concepts. However, the analysis of lexical categorisation in fiction can take us even further. It is possible to reorganise the internal structure of a conventional, well-known category, as well as to combine terms from several other categories in order to produce a new one, and the final result will be the creation of a new concept, or at least, a new perception of something familiar that will make it appear different.

3. The Narrative: A fantasy novel

In order to analyse the way in which creative categorisation works, I will present some examples taken from a Fantasy novel called *Forging the Darksword* (Weis and Hickman 1988).

As mentioned earlier, genres like science-fiction and fantasy lend themselves to the study of lexical creativity in narrative because they display worlds full of new or invented concepts that require names. Further, science fiction often presents invented objects or processes that must sound scientific or pseudo-scientific. Fantasy, on the contrary, creates words that sound original or unique and the concepts they express are distant from the real world as we know it. Magic and fantastic creatures like dragons, elves and orcs are the basic elements of Fantasy novels, and credibility is a question of trust.⁴ Mostly, the words invented for the occasion are nonce formations that will never reappear in another text and only occasionally can they be interpreted by analysing their component constituents. So the only clue to understanding these words is to be found in the context, that is, the text itself and the genre to which it belongs. The reader's primary strategy for providing these novel words with meaning is that of categorisation.

Forging the Darksword is a Fantasy novel about an imaginary world, *Thimhallan*, where everybody is born with magical powers. It is, then, a thoroughly magical place, in which social position depends on both the amount and the kind of magic that an individual possesses, e.g. wizards are the most powerful and so in charge of the government, magicians and alchemists are craftsmen, warlocks are warriors and law enforcers, etc. After birth, a religious ceremony, a sort of baptism, is held to determine the kind of magic the baby has been born with and therefore his or her future function in this extremely rigid society. However, one

4. "that willing suspension of disbelief for the moment, which constitutes poetic faith" (Coleridge 1956: 168)

baby fails all the tests because he possesses no magic powers. In a world where magic is an essential quality, life without magic is equivalent to death, so this child will have to hide his lack of magic powers from the others, otherwise he would be considered a threat to their world.

4. New meanings for words: Semasiological categorisation

The meaning of any linguistic expression is organised in very complex prototypical categories, composed of all possible senses of the word connected among themselves by analogy, and by metaphorical or metonymical projections (see Section 2 above). In this way, the meaning of any word comprises not only a purely referential meaning, but also a great deal of information about the possible uses of the word, the contexts in which it can be used, the speaker's encyclopaedic knowledge about the expression in relation to the world he lives in, along with links to other related categories.

This conception of meaning as a vast, complex category goes against the traditional idea of one "literal" meaning and several other derivative senses. As explained above, every category is organised around a prototype, so in any semantic category there will always be a central, prototypical sense, which is the most salient, frequent and conventionalised.⁵ But the choice of a sense to interpret a word in speech will be determined by the context and the negotiation between the speakers (Langacker 1999: 237), which means that its salience may vary according to the different contexts in which the expression occurs.

An example of a semantic category, actually a small fragment of what the whole category would be like, is given by Langacker (1991) for the word *ring* (Figure 2).

In this representation, the box outlined in bold shows the schematic prototype, "circular object", and the most salient specification or best example, "circular piece of jewellery".

The most important conclusion resulting from this dynamic conception of meaning is that there is not only one literal, true meaning that speakers draw on

5. It must be noted, however, that prototypes are not always objective, fixed and clear cut, so it is not possible to reduce the meaning to the prototype. In Langacker's words:

A speaker's knowledge of a lexical item cannot in general be reduced to a single structure, such as the prototype or the highest-level schema. For one thing, not every lexical category has a single, clearly determined prototype... The conventional meaning of a lexical item must be equated with the entire network, not with any single node.
(Langacker 1991: 3)

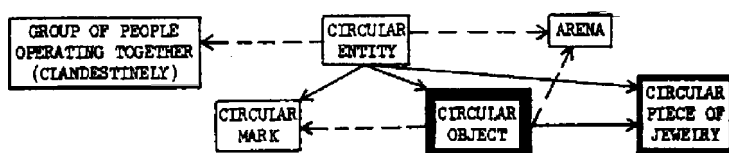


Figure 2. A partial representation of the senses of *ring* (Langacker 1991)

before considering other possibilities. The whole category constitutes the meaning, so when any one sense is activated, it is all partially active, as if in standby, in support of the meaning.⁶ Thus, if any corrections or adjustments are necessary, it is relatively easy to complete the meaning, or even to shift from one sense to another following the clues provided in the text. This is a common process in everyday speech, but particularly significant for the interpretation of new meanings in fiction because, as we will see in the examples from the novel, new meanings are always related to conventional ones; otherwise readers would be unable to interpret them.

5. The semantic change of *dead*

Forging the Darksword presents a new sense for the word *dead*. Consider the following excerpts taken from the novel:

- (1) In fact, the only one present who was not in agreement that the baby was Dead appeared to be the baby himself, who was screaming frenziedly (p. 9).
- (2) He thought of the hard life the young man had led, the constant struggle for survival, the need to hide the fact that he was truly Dead (p. 371).

The prototypical sense of *dead* in everyday speech, “without life”, is the first one that comes to mind if we encounter the word in isolation. It is also the first entry in dictionary definitions and is directly accessible from our experience of the world. This sense is related to others such as *dead fingers*, *a dead place*, *dead fire*....⁷

The Oxford English Dictionary (1978) presents thirty-one senses for the word *dead*, grouped in five main sections, while other dictionaries (*The Longman Dic-*

6. Williams (1992) reports some experiments that evidence how the contextually irrelevant meanings of polysemous adjectives also become partially active during the process of interpretation.

7. In order to determine the salience of a particular sense, Dirven and Verspoor (1998:31) suggest the following procedure:

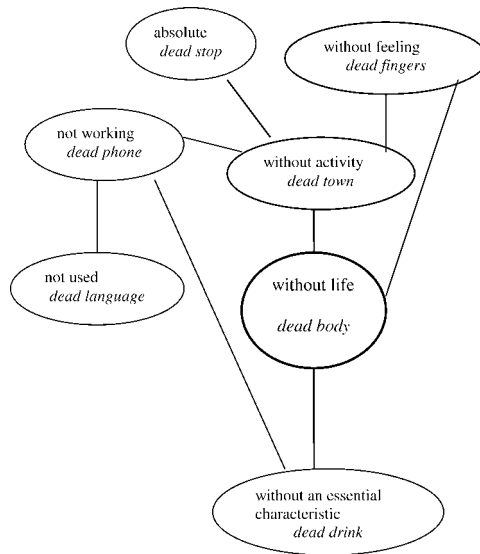


Figure 3. A partial representation of the meaning of *dead* as a category of senses

tionary of Contemporary English 1987, *The New Shorter Oxford English Dictionary on Historical Principles* 1993, or *The New Oxford Dictionary of English* 1998) reduce these to two or three main semantic groups. With no claim to being exhaustive, Figure 3 shows a graphic representation of some of the most common senses that constitute the meaning of *dead*.

The prototypical sense of *dead* is activated the first time we encounter the word in the first chapter, and that interpretation is reinforced by the context: A religious ceremony is being held in the cathedral and everybody is wearing mourning clothes. However, there are some signals that suggest something else is going on: the word *Dead* is always written with a capital letter and there is a strange emphasis in stating that everybody “agreed” with the fact that “the child was *Dead*”. Even so, the context is so unambiguous that the reader tends to ignore those signs. However, two pages further on, the reader is forced to readjust his/her interpretation:

In order to establish the salience of a sense, we can look at what particular sense comes to mind first, we can make a statistical count as to which use is the most frequent, or we can look at which sense is more basic in its capacity to clarify the other senses.

On the same topic, Williams (1992) admits for his experiments that the salience (*meaning dominance*) of one sense or another is determined by the order in which they appear in dictionaries, and Tuggy (1999) defends the validity of the linguist’s intuition to decide on this point because it would be an endless job to experimentally check the prototypicality of the various senses for every word.

- (3) In fact, the only one present who was not in agreement that the baby was Dead appeared to be the baby himself, who was screaming frenziedly (p. 9).

This is what Brandsford and Johnson (1973) call a *semantic anomaly* or the impossibility of matching the implied meaning of an utterance with our knowledge of the world. There is nothing in the conventional meaning of *dead* as represented in Figure 3 that can be applied in this case, so the reader will be forced to seek a new sense for the word through *recontextualisation*.⁸

However, in the process, the reader will not just ignore or abandon the conventional senses of *dead*, which are still activated, and substitute them for the new one. Instead, the new sense suggested by the reading will enter the category constituted by all the previous senses and so become part of the meaning of the word. In order to be included in the previous category, the new sense must be related to one or more of the members of that category as shown in Figure 4:

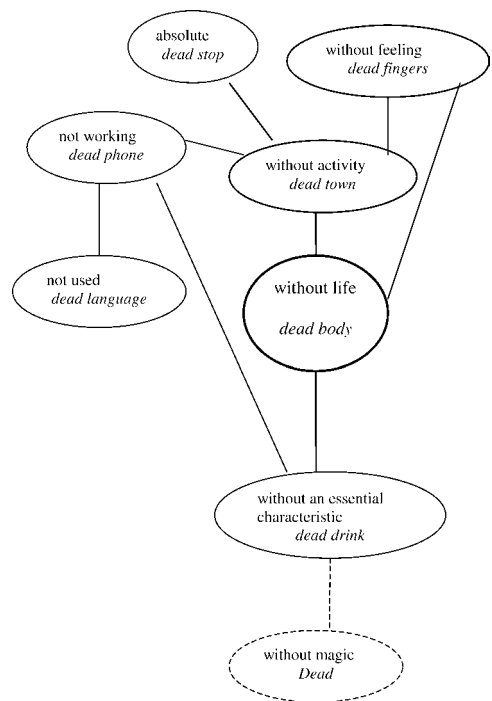


Figure 4. Acquisition of a new sense for the word *dead*

8. “Recontextualisation is where the peculiarities of the text world affect the semantic field of a word” (Stockwell 2000: 121).

Magic is vital in the world of *Thimhallan*: it is everywhere and all babies are born with some magical powers. Therefore, the idea that someone could lack magic and still be alive seems impossible, because for the inhabitants of this world “Life is Magic” (p. 23). The new sense can be fitted into the category by linking it to the already-existing sense “without an essential characteristic” as in *dead fire* or *dead colour*. All the other senses in the category remain and they are actually activated elsewhere in the novel:

- (4) “He’s dead,” Saryon answered without emotion. “Died about ten years ago, when I was fifteen. A wasting illness” (p. 31).
- (5) The sounds of autumn’s dry, dead leaves – blown by a chill wind – skittering and rustling through the empty streets (p. 268).
- (6) Suddenly Simkin lurched forward, his dead weight dragging Saryon to the floor (p. 210).

Then, as the story continues, the new sense is reinforced and moves towards the centre of the category. It becomes the most frequent sense in the novel: there are 78 occurrences of *Dead* with a capital letter, whereas *dead*, which comprises all the other senses, occurs 58 times. Besides the frequency, the salience of the sense “without magic” is also higher than that of other senses, i.e. it is the first sense that comes to mind when the word appears in the text because this is a story about a world of magic, whose main character is a *Dead* young man. Furthermore, that salience is supported by adverbials like *truly*, which often accompany the word.

- (7) He thought of the hard life the young man had led, the constant struggle for survival, the need to hide the fact that he was truly *Dead*. (p. 371)

Therefore, “without magic” becomes the central member of the category, the prototype (within the context of this novel), the first sense that comes to mind when the word appears in the text (Figure 5).

As a consequence, not only a new member has been added to the category, but a recategorisation has taken place. Certainly, this distinction is extremely important for the interpretation of the narrative, because categories allow us to organise our knowledge of the world. Thus a recategorisation can provide a different perspective, even a different reality; in this case it has helped to create an imaginary world in which it is worse to live without magic than to be without life, because even the dead in this world possess some magic.

- (8) I am *Dead*, truly *Dead*. No magic stirs within me, less than what is in a corpse, if we believe the legends of the ancient Necromancers, who were able to communicate with the souls of the dead.

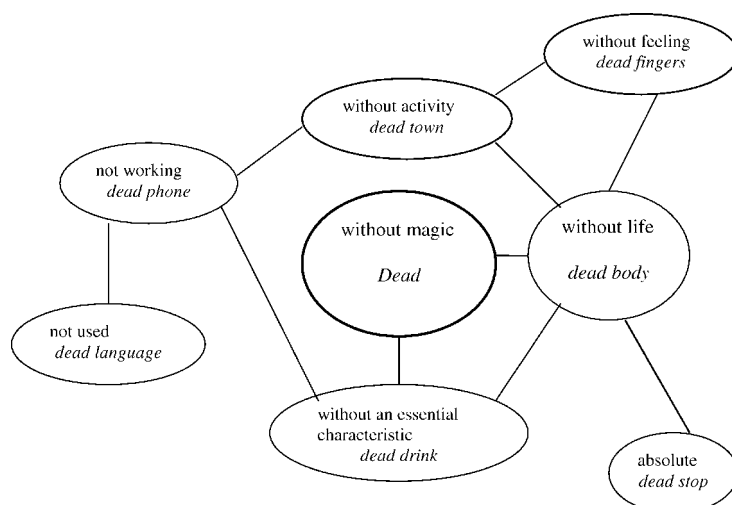


Figure 5. Recategorisation of the meaning of dead with the new sense as the central member of the category

6. New lexical fields: Onomasiological categorisation

Lexical categories in an onomasiological approach are similar to what have traditionally been known as lexical fields, but the study of lexical fields as categories allows a better understanding of phenomena like synonymy, antonymy and variation.

Lexical field analysis tends to describe systems of 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. (Geeraerts, et al. 1994: 3)

The conceptual relations that occur between words in a lexical field are similar to those between the senses of a word (semasiology) (Dirven and Veerpoor 1998: 37). Again, we will find a prototype, which is the first word that comes to mind to refer to a given concept, as well as several peripheral, less common words, and also fuzzy boundaries in the category, which means that the borderline members might be included in other close categories.

In terms of creative categories from an onomasiological perspective, I will consider three possibilities:

1. The construction of new categories with words invented for the purpose, although they may also include a few common words from other categories.

2. The recategorisation of members that usually belong to other categories but that have come together to create a new one.
3. The reorganisation of common categories in which the members are all the same but the prototype changes, as does the salience of most of the words. As a result, some peripheral members move towards the centre of the category, whereas other very salient terms become marginal, even borderline terms.

6.1 Construction of a category of invented words: Specialisation in magic

As stated earlier, the world depicted in *Forging the Darksword* is inhabited only by magicians, and their position in the society is strictly determined by the kind and amount of magic they are born with. Thus, in such an extensive lexical field, it is necessary to designate the different kinds of magicians, depending on their power and social position. From the very beginning of the story, the importance of this field is signalled by a significant accumulation of what might be considered synonyms in common speech:

- (9) All of the *wizards*, *magi*, and *archmagi* who floated in a shimmering circle above the marble floor [...]. All of the black-robed *warlocks* [...] hovered at their assigned posts [...]. All of the *thaumaturgists* – catalysts – who stood humbly upon the blue floor (p. 7).

The occurrence of these words in separate sentences and referring to people in different positions in the scene marks a clear distinction in their meanings. It is the first clue for the reader that some adjustments must be made in the familiar category of *MAGICIANS*.

Throughout the story there are more than thirty terms used to refer to different types of magicians. Some of them are existing words, often with a new sense that has been added to their habitual meaning, and others are nonce words. Table 1 is a collection of those terms, and it also summarises the functions of each kind of magician in the story.

Note that the expressions in the third column below designate the function, but they are also often used in the text to name a type of magician, as there is a one-to-one correspondence between types and functions. Thus we can refer to the same person as “the warlock,” “the *duuk-tsarith*” or “the enforcer,” which demonstrates the fuzziness in the limits of lexical categories, since these expressions (*war masters*, *stone shapers*, *healers*...) are not members of the category of *MAGICIANS*, strictly speaking, but can be used as such.

Table 1. Members of the lexical field of magician in *Forging the Darksword*

KINDS OF MAGI	SUBTYPES	FUNCTION	
Diviners		Time Magi	(they see the future)
Necromancers	Theurgists	Spirit Shapers	(they talk to the dead)
Archmagi	Kan-Hanar		(In charge of commerce)
	Sif-Hanar	Weather Magi	(they make weather)
Witches & warlocks	Duuk-tsarith	Enforcers	
	Dkarn-duuk	War Masters	(warriors)
Wizards/wizardesses	Albanara		(In charge of Government)
Conjurers	Quin-alban	Craftsmen	
Magicians	Pron-alban	Wood Shapers	
		Stone Shapers	
Alchemists	Mon-alban	Craftsmen, metal workers	
	Field magi	Farmers	
	housemagi	Servants	
Druids	Fihanish	Field druids	(they take care of plants)
	Mananish	Healers	(they cure minor illnesses)
	Theldara (pl. Theldari)	Shamans	(they cure major illnesses)
Illusionists		Artists	
Thaumaturgists	Sharak-Li	Catalysts	(they work with healers)
	Thon-Li	Catalysts and Corridor Masters	
	House catalysts		(They work with house/field magi)
	Field catalysts		
Sorcerers		Technologists	

In the real world, there is no need to specify different types or subtypes of magicians. There is no distinction between different types of warlock, for instance, especially since no such thing exists in the world we know. Thus almost all the words in the second column are invented, providing new categories of terms that have been created to denote specialisations in magic. Even though invented, most of these new words are formed systematically, according to the rules of productive English word formation, thus respecting already familiar morphological patterns. Consequently, the reader will find them easier to interpret.

In the first place, these invented words are all hyponyms and, consequently, most of them are compounds or derivations. Subordinate, specific terms are always more descriptive than basic level terms in order to signal the specification that makes them different (Dirven and Veerpoor 1998:40 and ff.; Ungerer and Schmid 1996: 86ff.). For this reason, compounding and derivation are so frequent at this level, and this facilitates the task of interpretation for the reader. On the other hand, loanwords are also very frequent in hyponymy. The use of foreign

words as hyponyms fails to retain the descriptive function and with it the immediate identification of the referent, but, in contrast, it stresses the existence of something so new that it requires a strange-sounding name.⁹

Both kinds of hyponymy, descriptive and pseudo loanwords, are found in the category of SUBTYPES OF MAGICIANS. There are compounds like *housemagi*, *field magi*, *house catalysts* and *field catalysts*; there are pseudo derivatives such as *pron-alban*, *mon-alban*, *fihanish*, *mananish*, etc. and many of these, even if invented, sound like possible loanwords, with unusual affixes and plural inflections, as in the case of *theldara* / *theldari*.

The reader can identify some regular patterns of affixation: all thaumaturgists take the suffix *-Li*, craftsmen in different fields take the suffix *-alban*,¹⁰ *archmagi* take the suffix *-hanar*, minor druids take *-nish* and warlocks have *duuk* as an affix, or perhaps as a stem, since the position varies, but this is impossible to determine (and probably irrelevant). The point is that the mechanisms become familiar to the reader and so it is possible to categorise all these nonce words just as any other neologisms we encounter in the real world.

As for the interpretation of these words, it seems obvious that the analysis of the constituents is not enough, because both stems and affixes are invented and have no relation with any existing words. Therefore, it will be mostly the context that provides meaning. In fact, these novel formations are very often followed by a phrase in apposition that explains the meaning. If the term occurs frequently enough throughout the text so that the reader can acquire the word and its meaning, the apposition is later removed, as in the case of salient terms like *duuk-tsarith*. Otherwise, the apposition remains:

- (10) ...the *Pron-alban*, wizards skilled in the arts of shaping stone and earth, took hold of the river... (p. 148)
- (11) ... who work with the *Pron-alban*, the Stone Shapers, or the *Mon-Alban*, the Alchemists... (p. 316)

More transparent formations, such as *housemagi*, *field magi*, etc. are probably interpreted by analogy with common words like *housemaid*, along with the analysis of their constituent parts and, of course, with help from the context:

9. Of course, there are hyponyms that are neither compositional nor loanwords, but far fewer than we might think. Ungerer and Schmid (1996:96) point out that many seemingly unanalysable subordinates are actually historically derived from loanwords or metaphorical projections, and provide examples such as *daisy* ("day's eye") or *dandelion* (from French "lion's tooth").

10. There is even a mini-process of grammaticalisation, as the suffix *-alban* would appear to be derived from the word *albanara* that designates the most powerful magicians in that particular group and, therefore, in charge of the government.

- (12) ... memories of childhood stories told to him by the housemagus when she put him to bed. (p. 193)

Analogy is, then, a major aspect in the interpretation of nonce words, probably more important than morpheme analysis, which will only be carried out if analogy fails to provide a meaning for the term (Aitchison 1994: 131). When confronted with a new word, readers will make use of all possible resources and strive to find a meaning for the word, so analogy is not only established among morphemes such as in *housemagi*, but phonemes also have a role in this process of categorisation in order to arrive at an acceptable interpretation.¹¹ Notice that the affix for the subtypes of warlocks is *-duuk*, which reproduces a significant phoneme of the word, /k/, whereas the suffix for thaumaturgists, all of them catalysts, is *-Li*, recalling the syllable /li/ from the hypernym. Even the subtypes of druids, a word that evokes old Nordic cultures like Finnish, Swedish and Danish, take the same suffix *-ish*.¹²

Readers may not be fully conscious of these phonological relations, but still the links are likely to be created automatically.

The creation of a whole system, on the basis of recognisable patterns of word formation, instead of just some unlikely or unconnected novel expressions, enables the reader to interpret these inventions. It also confers credibility to the invented concepts, because they are based on a productive system of word creation. Thus the reader can imagine that there exist other subtypes of thaumaturgists or druids in that fantastic world that do not appear in the narrative, but could theoretically emerge in subsequent volumes, and if so, they would appear with the suffixes *-Li* or *-nish*, respectively.

6.2 Recategorisation of common words to form a new category: Human motion verbs

In the world we know, people walk. *Walk* is the prototypical term to refer to human movement from place to place. From an onomasiological perspective, the category WALK includes a vast number of other verbs: people can walk quickly (*stride*, *run*, *march...*), slowly (*wander*, *stroll...*), quietly (*tiptoe*, *sneak*, *creep*), heavily (*stamp*, *plod...*), unsteadily (*stagger*, *totter...*), etc. However, in the fantastic world of the

11. For further discussion on the role of analogy in word creation, see Veale in this volume.

12. Vallès (2003) considers that paradigmatic relations among words contribute to the organisation of the categories in the mental lexicon, and can also be productive. Those paradigmatic relations are mostly morphological in her analysis; however, it is also possible to establish phonological relations, even if they would be far less productive (Vallès 2003: 145).

novel, magicians usually move in the air and only occasionally touch the ground. Consequently, there is a whole new category of terms to refer to “motion in the air”, equivalent to the category WALK that refers to “motion on foot”. Certainly, we cannot expect that this new category, developed in a single novel, will be as vast as WALK, but it will have a significant number of members. Table 2 shows the members of the category and the number of occurrences in the text:

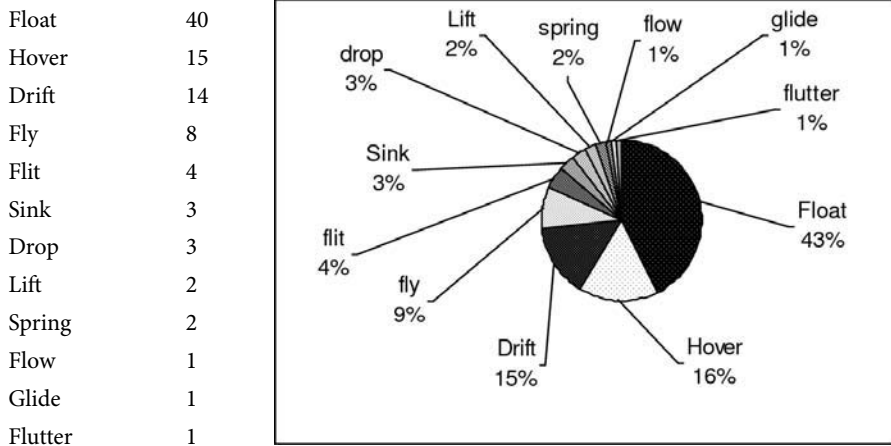


Figure 6. Frequency of the members of the invented category FLOAT in *Forging the Darksword*

Contrary to what we might think, the most frequent term is not *fly*, but *float*. The cognitive models developed by each word are quite different from each other. Our common idea of the concept FLY often includes the image of a pair of wings, as in birds, insects or planes, or at least stretching arms as if they were wings. However, the inhabitants of this fantastic world are human beings, and they have no wings. They just move in the air without any particular effort, just as naturally as we walk. Furthermore, the model of FLY is also commonly associated with the idea of “speed”, as evidenced in some common metaphorical extensions of the meaning of *fly*: *time flies* or *he went flying up the steps*. For this reason, *fly* is only used in the novel for other magical creatures gifted with wings, or else it means a very fast motion:

- (13) Four faeries, two holding his feet and two more flying over his head...
(p. 210)
- (14) Corridors branched off in all directions, but Simkin flew past them without pausing. (p. 213)

On the other hand, *float* will prove to have a more general meaning as occurs with *walk* in our common lexicon. Also, its meaning, connected with floating on

the water, conveys images of “weightlessness” and “effortlessness”, which are very convenient for the mental picture of magicians floating in virtue of their magical powers, whereas *fly* sounds ‘purposeful’ and more ‘demanding’. In addition, *float* can also be accompanied by prepositions to indicate direction, as in the case of *walk*, and in this way further extend its meaning:

- (15) Members of the Royal Household floated up to the Emperor to offer their condolences and their sympathy... (p. 16)
- (16) He watched the people float past, their thin robes rippling around their bodies as they went about their daily business. (p. 58)
- (17) Mosiah started to float forward with some idea of intervening but his father stopped him with a look. (p. 115)

Thus, because of its range of meaning and because of the cognitive model provided, *float* will become the default term, i.e. the prototype of the category, which can be used in place of other less salient and more specific terms. In general, the more specific the meaning of a word, the less frequent and salient it becomes.

As for the other members of the category, they all come from the conventional categories of MOTION IN THE AIR and MOTION ON WATER, linked to the prototype by both senses in the meaning of *float*.

Therefore, *fly* indicates “speed”, as explained above, and *flow* communicates the image of “regularity”. If there is a vertical direction in the motion, then *spring*, *lift*, *drop* and *sink* are used, along with “up” and “down” respectively, whereas *hover* suggests absence of motion, thereby occupying roughly the same position as *stand* in the category WALK:

- (18) who maintained their attitude of cool aloofness and strict attention to duty as they hovered at their assigned posts. (p. 7)

Likewise, *glide* and *drift* convey the idea of “lack of a definite direction”, as in *stroll*.

- (19) Look with fear upon those black-robed figures that glide silently about the Royal Palaces.

And whenever the action in the text entails some sort of agitation and confusion, *flit* or *flutter* are chosen, due to their association with the rapid movement of the wings of insects and semantic extension to indicate nervousness and excitement:

- (20) The people flitted from house to house, talking in frightened, subdued voices. (p. 361)

A possible representation of the new category is shown in Figure 6.

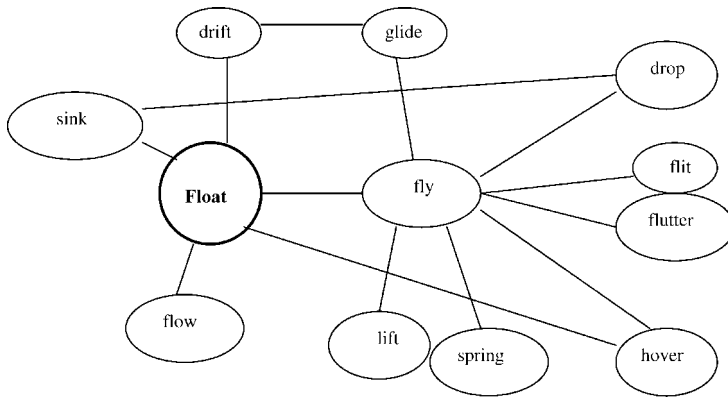


Figure 7. A schematic representation of the semantic category of HUMAN MOTION IN THE AIR in *Forging the Darksword*

The creation of the new category does not involve the elimination of the conventional category WALK in the same way that the new sense acquired for *dead* did not totally eradicate the previously existing senses of the word. Nevertheless, the only terms remaining are those that emphasise the opposition between FLOAT, a light, graceful motion, and WALK, a purposeful, strenuous one. As a result, the terms in this category most frequently found in the text, apart from *walk*, are *stumble* and *trudge*:

- (21) Seeing the overseer hesitate, the catalyst frowned and trudged over to stand beneath where the observer floated in the air. (p. 77)

Consequently, the new category for MOTION IN THE AIR has been generated from a change in the prototype. When *float*, rather than *fly* takes on the central position in the category, the whole concept is seen from a different perspective. Moving in the air is not necessarily faster than walking, but it represents a 'natural' and effortless movement that clearly contrasts with the idea of walking, trudging and stumbling on the ground. Once again, creative categorisation goes beyond the mere invention of a word, by presenting a whole new perspective on an already existing concept.

6.3 Reorganisation of a conventional category: MAGICIANS and MAGI

There is yet another possibility as regards creative categorisation: the rearrangement of an already existing category, that is, a prototype shift and an overall reorganisation of the rest of the members in terms of frequency and proximity to the new prototype.

Table 2. Frequency of the members in the semantic category of *MAGUS*

Magus / magi	131	(42 / 89)
Warlock / witch	122	(118 / 4)
Wizard / wizardess	52	(46 / 6)
Sorcerer	52	
Druid	33	
Diviner	8	
Conjurer	4	
Alchemist	3	
Illusionist	3	
Thaumaturgist	3	
Necromancer	2	
Archmagus / archmagi	2	(1 / 1)
Magician	1	

Table 1, in Section 5.1 above, included all the terms appearing in the novel serving to designate the different kinds of magicians, and we have already seen how the terms in the second column, the subordinate level, form a new category of nonce words.

As for the words in the first column, all of them, except *archmagus*, which has a transparent meaning anyway, are words which exist in our lexicon, even if rare. We might then conclude that the lexical field of *MAGICIANS* is an already existing category. However, we will see that in the novel the word *magus* moves to the central position, which, in turn, causes an overall rearrangement of the rest of the members in terms of frequency and proximity to the new prototype.

The number of occurrences of the members of the category is shown in Table 2.

It is evident that the order of the frequencies is exactly the opposite of what occurs in normal usage. For instance, the *Oxford Thesaurus* (1997) does not have an entry for *magus*, *wizard* or *warlock*, whereas those cited as the most frequent are *magician*, *conjurer* and *illusionist*; *wizard*, *sorcerer*, *magus*, *necromancer*, *enchanter* and *warlock*, follow in this order. Also, the online lexical database *WordNet*, compiled by The Cognitive Science Laboratory at Princeton University, provides the following frequency order: *sorcerer*, *magician*, *wizard*, *necromancer*, *thaumaturge*, *thaumaturgist*, which again places *magician* and *sorcerer* among the most common words for the concept and leaves *warlock* and *magus* out of the classification. We may thus conclude that the prototype most commonly oc-

curring in everyday use is *magician*, a word with a wide range of meanings, from fantastic wizards to show-business illusionists. Nevertheless, for the readers of fantasy novels, the prototype would certainly be *wizard*, which is a far more common term in the genre.¹³

Yet, as shown in Table 3, the most frequent term in this novel is *magus* (pl. *magi*), a very unusual word in English, as evidenced by the Latin plural form. The closest association, perhaps the only one, that the reader can establish is with *the Three Magi*, the Wise Men from the East that, according to the tradition, visited Jesus in Bethlehem. This association evokes images of “wisdom”, “eastern exoticism” or even “religion”, probably stronger links than those with “magic”, in spite of the common root of the word with *magic*, *magical* or *magician*.

Not only frequency, but also salience determines a prototype, and *magus* is also the most salient member of the category. In fact, it is the default word for the semantic field; anyone who has some magical powers, regardless of the amount or kind, is a *magus*.¹⁴

On the contrary, *magician*, the most common word in normal usage, is shifted to a very peripheral position, both in the number of occurrences in the text (only one!) and in the role that magicians play in the story, that of craftsmen, closer to alchemists and conjurors than to the prototype. Similarly, *illusionist*, the only kind of magician we acknowledge in the real world, is also displaced to the periphery of the category, and regarded as a minor member with a low frequency of occurrence.

The other most frequent terms are *wizard* and *warlock*, the most powerful kinds of magi and therefore in charge of government and law enforcement. It is not surprising that they are the closest to the prototype. As has been pointed out, wizards are prototypical in Fantasy novels.

Nevertheless, it is not only the lexical field that is restructured, so that peripheral terms become more central and vice versa. The meanings of the words also undergo a semasiological recategorisation. Many of these words are considered to be near synonyms in ordinary speech, as we do not usually distinguish between, for example, *conjurors* and *sorcerers*. However, as I have already pointed out, there is an obvious intention of marking differences among them from the beginning of the story, as evidenced in fragment (9).

The ordinary associations of these words as they are used in actual speech are exploited and new senses are added. In this way, druids, the wise men in Celtic tribes, become the doctors in the story; illusionists, commonly associated with the

13. Regarding the best means of determining the prototype of a category, see Note 5.

14. This is the reason why it does not appear in Table 1.

idea of simulation for purposes of entertainment, become the artists; alchemists, traditionally capable of transforming lead into gold, are metal workers, whereas wizards, usually connected with wisdom and power (consider *Merlin* or *Gandalf*) are the rulers of *Thimhallan*. Once again, the meanings are not entirely altered, but extended and then reorganised.

Also, as concerns the somewhat fuzzy limits of the category, there are two kinds of magi that could hardly be considered as such: *sorcerers* and *thaumaturgists*: *Sorcerers* are *technologists*, i.e. they have some magical powers but they have chosen not to use them, preferring technology to magic. As for the *thaumaturgists*, their function in the society depicted in the novel is to act as catalysts: no magic is possible without their presence, but they cannot perform any magic themselves. Both these terms will be placed, then, at the periphery of the category.

The choice of such a rare word as *magus* for the prototype of the category produces on the reader an estrangement that is in accord with the intention of signalling the gap between the real world and the fantastic one in the novel. The magicians of *Thimhallan* are distinct from anything in our previous experience or reading.

Consequently, in the same way that a shift in the prototypical sense of a semiological category results in a new word, because “the semantic change performed in some common words make them work as new words” (Stockwell 2000:119), likewise the reorganisation of the category *MAGICIANS* produces a new category of *MAGI*.

7. Conclusions

Our mental lexicon is organised in prototypical categories. This system has proved to be both flexible and stable enough to deal with the huge amount of information that is associated with each word. It also allows speakers to make constant adjustments whenever new information is provided, as well as create and interpret new words and meanings in everyday speech. If categorisation is the way we deal with neologisms in common speech, it must also be the way to cope with nonce formations in fiction.

Beyond what we commonly understand as creative lexical items, namely nonce formations and invented words, we have seen that the ascription of new meanings for common words (neosemes) or the reorganisation of conventional lexical categories are powerful tools for lexical creativity, as they both result in common expressions referring to a different concept. So, for instance, the concept of *MAGUS* acquired through the reading of *Forging the Darksword* is a very specific, detailed one – a magus is born with magical powers, can float in the air,

etc. – far from the model of MAGUS that the reader is likely to have held before the reading.

We have seen several examples of creative categorisation, both semasiological and onomasiological, in a fantasy novel. The resulting categories have proved to be a powerful instrument for the interpretation of the whole text, as they provide the reader not only with new concepts, but also with different perceptions on some common concepts, allowing him/her to construct a text world which is distinct from the real world. This is, after all, the main purpose of a Fantasy narrative.

References

- Aitchison, J. 1994/1987. *Words in the Mind. An Introduction to the Mental Lexicon*. Oxford UK and Cambridge USA: Basil Blackwell.
- Brandsford, J. D. and Johnson, M. K. 1973. "Considerations of Some Problems of Comprehension." In *Visual Information Processing*, W. G. Chase (ed.), 383–438. New York: Academic Press.
- Coleridge, S. T. 1956. *Biographia Literaria*. 1817. George Watson (ed.). London: Dent.
- Cuenca, M. J. and Hilferty, J. 1999. *Introducción a la Lingüística Cognitiva*. Barcelona: Ariel.
- Dirven, R. and Vespoor, M. 1998. *Cognitive exploration of Language and Linguistics*. Amsterdam: John Benjamins.
- Geeraerts, D. 1997. *Diachronic Prototype Semantics*. Oxford: Clarendon Press.
- Labov, W. 1973. "The Boundaries of Words and their Meaning." In *New Ways of Analysing Variation in English*, C. J. Bailey and R. W. Shuy (eds), 340–373. Washington D.C: Georgetown University Press.
- Lakoff, G. 1987. *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*. Chicago, London: University of Chicago Press.
- Langacker, R. W. 1987. *Foundations of Cognitive Grammar. Vol. 1*. Stanford: Stanford University Press.
- Langacker, R. W. 1991. *Concept, Image and Symbol. The Cognitive Basis of Grammar*. Berlin and New York: Mouton de Gruyter.
- Langacker, R. W. 1999. "The Contextual Basis of Cognitive Semantics." In *Language and Conceptualization*, J. Nuyts and E. Pederson (eds), 229–252. Cambridge: Cambridge University Press.
- Rosch, E. 1975. "Cognitive Representations of Semantic Categories." *Journal of Experimental Psychology*, General 104: 193–233.
- Rosch, E. 1978. "Principles of Categorisation." In *Cognition and Categorisation*, E. Rosch and B. Lloyd (eds), 27–48. Hillsdale, N.J. and New York: Lawrence Erlbaum Associates.
- Rosch, E. and Mervis, C. 1975. "Family Resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 15: 346–378.
- Semino, E. 1997. *Language and World Creation in Poems and Other Texts*. London and New York: Longman.
- Štekauer, P. 1998. *An Onomasiological Theory of Word-Formation in English*. Amsterdam/Philadelphia: John Benjamins.

- Stockwell, P. 2000. *The Poetics of Science Fiction*. London: Longman.
- Talmy, L. 2000. *Toward a Cognitive Semantics*. Cambridge, Massachusetts, London: The MIT Press.
- Taylor, J. R. 1995. *Linguistic Categorisation. Prototypes in Linguistic Theory*. Oxford: Clarendon Press. [1st edition 1989].
- The Longman Dictionary of Contemporary English*. 1987. London: Longman.
- The New Oxford Dictionary of English*. 1998. Oxford. Clarendon Press.
- The New Shorter Oxford English Dictionary on Historical Principles*. 1993. Oxford: Clarendon Press.
- The Oxford Dictionary and Thesaurus*. 1997. Oxford: Oxford University Press.
- The Oxford English Dictionary on Historical Principles*. 1978. Oxford: Clarendon Press.
- Tsohatzidis, S. L. (ed.). 1990. *Meanings and Prototypes. Studies in Linguistic Categorisation*. London and New York: Routledge.
- Tuggy, D. 1999. "Linguistic Evidence for Polysemy in the Mind: A Response to William Croft and Dominiek Sandra." *Cognitive Linguistics* 10 (4): 343–368.
- Ungerer, F. and Schmid, H.-J. 1996. *An Introduction to Cognitive Linguistics*. London, New York: Longman.
- Vallès, T. 2003. "Lexical Creativity and the Organization of the Lexicon." *Annual Review of Cognitive Linguistics* 1: 137–160.
- Weis, M. and Hickman, T. 1988. *Forging the Darksword*. New York, London, Sidney: Bantam Books.
- Williams, J. N. 1992. "Processing Polysemous Words in Context: Evidence for Interrelated Meanings." *Journal of Psycholinguistic Research* 21: 193–218.
- WordNet www.cogsci.princeton.edu (last accessed: 27th June 2005).

Sociopolitical effects on creativity

Occasional and systematic shifts in word-formation and idiom use in Latvian as a result of translation

Andrejs Veisbergs

This chapter will examine the way in which creativity in word-formation and idiom use has been changing as a result of translations from English, effecting stable changes in Latvian word forming processes.

Translation, being first and foremost a form of cross-cultural communication, affects target culture, target thinking, target language, as well as target language norms and conventions. Lexical innovation is one way that a language adapts to new concepts. Though it is ruled by fixed norms and conventions, these, however, may change as they are affected by the language situation and social constraints. The paper will discuss and exemplify the types of creative language and the shifts in creative wordplay in Latvian that have resulted from outside influences.

1. Creative use, nonce use, wordplay

Creativity in changing word forms is embedded in the language – words and phrases in a language can be varied, changed, manipulated and played with, within the norms of the specific language system. Creativity also ensures the possibility of language change and often it is the first step towards such change. Word forms and fixed phrases can be varied for stylistic reasons, and in some cases these changes are not merely a one-off nonce phenomenon but affect more general changes in the language system (*langue* as opposed to *parole*). These creative changes may vary according to the norms and conventions of use of a given language.

The difference between norms and conventions (Chesterman 1997:55ff.) is of some importance as regards the issue under examination. Creative use of language in individual cases may in fact break both the conventions and norms of use of the established language unit. On the other hand, the means/techniques of

creative use are also part of the language system and as such can be considered fully conventional and normative. Thus we presume that substitution of lexical or morphological components in language units is possible in any language, perhaps a language universal. But the ways and means of such transformations may have a different distributional pattern in various languages and are subject to the norms of the given language. For example, adding a diminutive suffix to almost any noun or to nominal compounds in idioms would fully correspond to the language norms of Latvian while the same contents would generally be expressed by analytical means (adding an adjective) in English. Also the use of diminutives in general is far more frequent in Latvian than in English.

Speech conventions of English would, for example, presume a relatively more frequent use of idiomatic transformations in certain text types (e.g. newspaper and magazine headlines) while their use is much rarer in Latvian. In this sense conventions would come very close to what can be called expectancy norms, the breaking of which would create a certain unease. Yet expectancy norms and conventions are not engraved in marble; they are flexible and depend not only on tradition and popular perceptions but also on cultural evolution, on the effect of contact with other languages and the sociolinguistic situation.

It is not easy to cover all possible types of creative use of vocabulary and idioms or find a single term to encompass all, especially as it is necessary to distinguish between individual nonce use, sometimes called speech transformations, and those changes that become adopted by the language community and enter the general lexicon. **Transformations, modifications, adaptations** are terms that are perhaps too general for our purposes. **Occasional change/use** is usually applied only to phraseology and, as the word suggests, implies individual, contextual changes of an idiom's meaning or structure and meaning for stylistic purpose (Veisbergs 1997). The same is true of **instantial** use, which is also a contextual phenomenon (see Naciscione 2001). Further, **nonce** use as a term tends to exclude those items (normally not idioms) that (may) become neologisms and enter the language as stable elements.

Another term often used to describe linguistic creativity is **wordplay**. Wordplay is commonly described as being "based somehow on a confrontation or clash of two meanings" (Alexieva 1997: 138), ambiguity being a central feature. Most transformed elements in the initial stage constitute a kind of wordplay. Though used widely, wordplay is also a rather vague term and there is little agreement on its characteristics and limits. **Pun**, another frequently employed term (sometimes used interchangeably with wordplay), is actually only one subtype of wordplay, normally involving play with meaning(s).

What is characteristic of all types of creative use is the effect (stylistic, expressive, semantic) of defeated expectancy (Riffaterre 1960). It functions as an "at-

tention grabber” (Redfern 1986) – attracting the attention of the reader/listener to a specific point in the text. Creative language is also used to bind **a specific**, innovative meaning to established language material (words, idioms). For the receiver there is often the added pleasure of solving the puzzle – decoding the novel creation – as in the case of nonce use: the usual/initial word form is normally not lost, but simply contrasted in the receiver’s mind with the innovation.

There exists a multitude of classifications, types and subtypes of wordplay that researchers have identified and classified (see Hausmann 1974; Petrotta 1991: 20; Fill 1992: 552; Alexander 1997). These classifications, however, usually tend to reflect the specific aspect that is being researched (wordplay), rather than systematise such a broad and variable phenomenon. As Redfern has pointed out (1986: 82) “classifications...produced more confusion than enlightenment.”

This paper will discuss only structural changes in words and phraseology, while a relatively similar phenomenon – that of semantic variation or semantic manipulation – will not be considered, since polysemy, homonymy and homophony are normally used for contextual stylistic purposes and do not affect the language system as such. Accordingly, the basic types of creative transformations which we will examine are:

- Creative compounds, compound phrases, acronyms, blends, e.g.:
Guintelligence, berluscronies, yuppie – huppie, muppie, curruppy, rumpy, buppy, lombards, paytriotism, baggravation, to bagonise, Alice-in-Wonderland idea
- Distortions, graphic changes, paronymy, e.g.:
*Do your teeth a **flavour** (positive effects of toothpaste);*
*U2 (you too); **to basque in glory** (about having a holiday in Basque land);*
***Suite dreams** (Daily Mirror. 10.6.2000)(about apartments);*
***ECological** (Scanorama 7-8.1995) (about Umberto Eco)*
***Cereal killers.** (Wallpaper 5.2005.) (about cereals)*
- breaking of structural stability in idioms (substitution of components, phonological changes (see Kuiper’s paper in this book), insertion, addition, allusion, zeugma, ellipsis, contamination), e.g.:
Where there is a will there is a fraud.
So priceless a bird in hand is worth two in the bush.
He killed two birds with one diplomatic stone.
They are still looking for the birds in the bush.
She lost her money and her heart.
There is no smoke without...
He is up to the neck ... in the soup.

- jocular definitions, spoonerisms, limericks, etc., e.g.:
archeologist = a man whose career lies in ruins
twin sister = sin twister

It should be noted that there is a general expansion and growth of the use of these devices/techniques. However, only some of these creative word forms (in their concrete form) enter the general language. On the whole, creative use of phraseology remains an occasional, nonce phenomenon as the stylistic effect is the result of the contrast between the language form and meaning of the idiom and its contextually changed form.

2. Latvian from isolation to “awakening”

It was the Baltic German-speaking elite that in fact systematised and codified Latvian and created so-called early written Latvian as early as the 17th century. These philologists were supplanted by native Latvian literati and linguists in the second half of the 19th century, who actually formed the nation (as in several other East European countries) and established Modern Latvian literary language as a combination of early written Latvian with the modern spoken language.

Since then Latvian has had a very stable (but somewhat rigid) system of word-formation and idiom use which, while theoretically admitting occasional and systematic change, rarely practiced it. One could say that Latvian was a very rule-governed and conservative language with its linguists advocating elements of purism, a fully understandable position determined by the naturally defensive stance of a small language community which is always threatened by the tendentially dominant regional contact languages (in this case, German and Russian).

Also, in the 20th century interwar period of independence, one rarely encountered creative word-formation and idiom use. Traditional Latvian word-formation used (and continues to use) compounding, derivation and borrowing as basic patterns, with abbreviation and semantic shifts being less frequent, and relatively few cases of historical conversion. Clipping was admitted only in colloquial language while blending was generally unknown. Also some of the above “approved” types of word formation were subject to restrictions – in compounding, for example, etymological hybrids tended to be avoided, new terms were (and still are) generally transparent compounds, etc. Idioms were generally rigidly frozen units. Transformations, use of novel word-formation patterns and nonce words for stylistic purpose was unusual.

The fifty years of Latvian history under the USSR brought about further ossification of the linguistic potential and alignment with Soviet Russian (which

became almost the sole contact and intermediary language). To a large extent this could be seen as a result of Soviet straight-jacket language policies, for which suspicion and censorship of any ambiguity or play with meaning and structure was rife. In fact, one can speak of three languages in this period:

One was the official, bureaucratic, severe soviet style, a purified and to some extent Russianised type of Latvian used in all media and in the small chinks of popular science that Latvian was allowed to preserve.

The other was the Latvian of culture (literature, theatre, etc.) which generally remained intact and preserved a relatively high quality. However one could also see certain ossification in the “good” Latvian – ideological and linguistic censorship and internal self-censorship made the literary language timidly careful in obeying the norms. Wordplay was a rare phenomenon, new words and loans were carefully avoided. Any new loans from Western languages were, as a rule, the same as those adopted in Russian (for example, translations of Western literature into Latvian only after they had been translated into Russian and thus authorised as ‘safe’).

The third language was colloquial Latvian – live, spontaneous, ever-changing, slightly adulterated by Russian. This language could be heard, but rarely seen in printed form or in the media and was somewhat frowned upon by the linguists and intellectuals.

The spheres of language use where neologisms and transformations are normally frequent were either nonexistent or limited to orthodox varieties (advertising, sophisticated jokes, journalese, media titles, etc.). The careful eye of state censorship suspected any innovations and even unapproved metaphors, branding them as dangerous and “reactionary”, as can be seen in a GDR textbook on style (Faulseit 1975: 113):

Heute, wo die Lehren von Marx und Lenin der Menschheit den Weg in die Zukunft weisen, ist eine dem Expressionismus aehnliche Behandlung des Wortmaterials als reaktionaeer einzuschaetzen. (ibid. 174)

Today when the teachings of Marx and Lenin show mankind the road to the future, such Expressionism-like dealing with word material is to be evaluated as reactionary. (translated by A. V.)

Creativity naturally tended to be viewed in the dichotomy of creative versus established norms. The established norms, as Hermans points out (2000: 13), imply a course of action generally accepted as proper or correct or appropriate. And linguistic creativity, being unpredictable and novel, thus not previously approved,

was naturally the first victim of censorship and consequently of self-censorship of writers, journalists and translators. The censor's position was often that of self justification and self preservation – when something could not be 'sanctioned' by safe and approved sources, it had to go.

A sprinkling of innovative elements existed, mainly in fiction and in literary translation as well as, more occasionally, in orally transmitted jokes (so called anecdotes) about the soviet reality:

- (1) – *What is **expressionism**? – An **express** train full of **Zionists***
(s-z change in English makes it untranslatable (A. V.)).
- (2) – *What is that new material they use in building Intourist hotels?*
***Microconcrete**?*
– *Yes, it consists of 20% of **concrete** and 80% of **microphones**.*

Nevertheless, the general tendency was to mute linguistic creativity, especially if/as it did not fully meet traditional Latvian norms and conventions.

Speaking of the role of translators in shaping linguistic conventions, one has to recognise the limits to the translator's independence (especially under a totalitarian system where all publishing is controlled, monitored and supervised). Translation cannot escape the well-known tendencies of most political structures to control at least in part the relationships between the territory and other cultural units (Lambert 1995: 100), but there were some exceptions (proving the general rule). Thus A. A. Milne's "Winnie the Pooh" had an excellent translation with novel and elegant transfer of linguistic creativity, because it was considered by the authorities a classic and an innocent book for children.

On the other hand, the translation of Vonnegut's "Cat's Cradle" (1963) into Latvian in 1973 was so extensively purged and amended politically, ideologically and linguistically that it was retranslated by the same translator in 2002 and published with a statement on the cover that the new version did not contain any censorial restrictions.

Creative use of words and idioms in translations was often reduced or simplified, when not substituted by descriptive explanations. To illustrate this technique, a few examples from the Latvian translation of John Fowles' "Ebony Tower":

- (3) ***Picasso's** name had been actually pronounced as "**pick-arsehole**"*
Translation: Pikaso vārdu vēcuks pat izrunāja tā, ka angļu valodā tas ieguva nepieklājīgu nozīmi (Faulzs 1982: 28).
Gloss: Picasso's name the old man pronounced so, that in English it gained an impolite meaning.
- (4) *High treason, that's all. **Mess of scientific pottage. Sold the whole bloody shoot down the river.***

Translation: Valsts nodevība – redz, kas tas ir. Zinātniska putra. Nolaiduši dibena visu to kuģīti (Faulzs 1982:43).

Gloss: High treason – that's what it is. Scientific porridge. They have let the ship sink.

This is not to criticise the particular translation which reflects the dominant conventions of the time. Yet such an approach in a way created, or at least served to sustain, these conventions.

Political change in the early 90s (the so-called Awakening) when the country regained independence and was opened up to outside/Western influences, led not only to a cultural reorientation, acculturation and massive influence of English, as well as massive borrowing to fill linguistic lacunae, but also influenced deeper linguistic structures and phenomena as well as conventions of language use. This process can be viewed as consisting of two parallel stages that were:

1. brought about and initiated through translation
2. developing independently in native Latvian usage.

Among the various resulting linguistic shifts in Latvian, we will be analysing:

- general democratisation of language use with an influx of the colloquial element in standard language (clippings, lower strata), freer use of “unestablished” or novel linguistic material;
- more extensive use of idiom transformations;
- the growth of creativity in word formation: nonce use, graphic manipulations, etc.;
- appearance of formerly nonexistent compound phrases in Latvian.

3. Translation, globalisation and language

We can also look upon the issue of lexical creativity in terms of discussions of minor and major cultures and languages in translation theory, where an asymmetrical cultural exchange takes place. As suggested by Toury (1995), dominant cultures tend to impose their own linguistic and cultural conventions in the translation of minority language cultures. But we can also observe the reverse process – that of minor cultures absorbing (willingly or unwillingly) the dictum of the more pervasive or prestigious cultures, modifying their conventions according to the changing situation, hegemonistic pressures of the prestigious languages, etc. As linguistic processes are less subjected to conscious rationalising than many others, they reflect the essence of the cultural processes more clearly (Hymes 1983:24). Thus English, apart from its general weight, in a way symbolis-

ing the free world, carrying with it the additional novel attraction of the hitherto forbidden fruit, had an immense impact on the “information-starved” masses of the former soviet space. Also, minor cultures tend to be more familiar with the major ones (Zauberga 2005:68) than vice versa, hence, more prone to linguistic borrowing, imitating and adapting. In addition, globalisation seems to support foreignising tendencies in translation.

In the linguistic area this is usually evidenced in such processes as interference and borrowing, as well as in the takeover of certain contact language patterns. All these linguistic influences have been present in the case of Latvian-Russian contact, and can also be seen in more recent Latvian-English contact, as reflected in the massive borrowing that has taken place. Yet, as regards shifts in conventions and impact on language structures, the issue is far from clear. Should the activation of the potentialities of a language under the influence of another language or culture be viewed as an interference – that is, as a detrimental phenomenon – or as linguistic and cultural imperialism (Venuti 1995)? Or, rather, should it be seen as a general standardisation of language with a consequent loss of cultural uniqueness (Munat 2004:115)? Perhaps the answer lies in the question – does such vitalisation/activation mean an expansion of the stylistic potential of the language, or does it represent a progressive impoverishment of traditional means (e.g. word formation patterns, standard approaches to translation methods)? In other words, do we see an externally conditioned change in traditional patterns, or enhancement of linguistic potential inherent in the language? Depending on the answer, the process could be viewed as a negative or a positive one. In the case of Latvian, it seems that the latter may be closer to the truth.

The shift in policies of translation signified a huge reorientation toward the western world and toward English in particular. For example, statistics show the following patterns of narrative translations into Latvian, which have not changed much since the mid nineties. The general growth and the distribution is interesting:

1985	209 books translated; 140 Russian : 9 English (proportion 15: 1)
1994	519 books translated; 40 Russian : 243 English (proportion 1 : 6)
1996	493 books translated; 29 Russian : 267 English (proportion 1 : 9)
1998	670 books translated; 53 Russian : 385 English (proportion 1 : 7)
2000	653 books translated; 68 Russian : 392 English (proportion 1 : 6)

2004 679 books translated; 68 Russian : 359 English
 (proportion 1 : 6)

Translations of fiction from other languages today are below the level of Russian. English translations on the average constitute two thirds of the total. It is interesting to note that during the soviet times translations from Russian also constituted two thirds of all fiction translations.

Yet narrative translations constitute only the tip of the iceberg in this linguistic shift, as the proportion of fiction among translated texts consumed by the readers/listeners today is extremely small. In fact, we live in a translated world where about 70% of all texts read and heard by a typical Latvian are translations (Veisbergs 2005). Taking into account the dominant Latvian tradition of loyalty and fidelity to the source text, such a proportion cannot but have a huge impact on Latvian.

As English began to replace Russian as the main contact and intermediary language, it brought not only direct linguistic influence in the form of loans, but also certain readjustments in linguistic conventions of Latvian. The agent of this change seems to have been English, which is textually a much more pliable and creative language, and where nonce-use has a prominent function as part and parcel of the conventions of language use. Thus translations from English were the primary vehicle of this change.

3.1 Democratisation

This cultural transition to a more open linguistic policy brought about one all-encompassing change in Latvian – a more colloquial style in the media and freer use of substandard lexis in printed media (formerly taboo language). It would be impossible to say whether this is the result of the new contact situation and English speech conventions or merely the result of what could be called the democratisation of society and, along with it, a shift in conventions of language use, or both factors combined. Many people in Latvia still disagree with this amalgamation of styles, the disintegration of hitherto clearly demarcated, separate styles. There was and there still exists a blend of styles that can often be considered inappropriate to a given context. Yet, like the creativity issue, this can also be seen as an actualisation of something which is always present in languages; colloquial and rude registers naturally existed under the Soviets, but would not be seen or heard in official media. This shift in usage consequently affected some word formation patterns, e.g. clippings, which were formerly a sign of oral colloquial language, were now spreading into more neutral standard texts. Clippings in Latvian (being an inflectional language) are usually accompanied by diminutive suffixes (-itis,

-inš) which add to the colloquial character of these formations that abound in the media: **ķinītis** (kinoteātris/kino (cinema)), **tīnis/tīnītis** (tīneidžers (teenager)), **ēdūzis** (ēdnīca (canteen)), **letiņš** (latvietis (Latvian)), **busiņš** (autobuss (bus), **dīpītis** (deportētā persona (deported person)), **trams** (tramvajs (tram)):

(5) *Svaigs ir tikai tā krāsojums, kas rotā **tramu**, šo pilsētas vilcienu.*

***tramam** pat izdomāja krustmāti... (Diena, 25.11.1994.)*

Gloss: It is only the painting that decorates the tram, this railway of the towns. a godmother was even invented for the tram.

3.2 The early stages of creativity through advertising

Advertising is an area where language creativity has blossomed. Advertising is a metaphor of our time, manipulating and changing values through language and, at the same time, shaping the language. Outdoor advertising was perhaps the most visible sphere (not everybody reads fictional works or even newspapers) where people came face to face with the new developments in language. As Loc-mele (2001: 47) points out “one can draw parallels between language change and advertising change in Latvia, as well as find peculiarities of change in the Latvian culture by looking into language of advertising”. From the rare Soviet advertisements in a rather Spartan style (usually consisting of a picture and the product name), the consumer was suddenly faced with striking and aggressive advertising campaigns that employed both direct transfer methods and translated commercials that were linguistically untargeted to the Latvian context (often intended for the large neighbouring Russian market). Despite these teething problems, the mistakes, the word-for-word translations of many ads did attract the attention of the public, though not the type of attention the advertisers had hoped for. People frequently analysed and commented on the absurdities of the creative texts they were now subjected to.

Thus in 1997, shortly before Easter, British Airways filled outdoor signboards with a poster “*Izjūti augšāmcelšanās prieku*” (*feel the joy of take-off*), which inadvertently (?) used wordplay with **augšā celšanās** (*take off*) and **augšāmcelšanās** (*biblical resurrection*). When the Latvian Church authorities protested, the posters were changed overnight to the more direct non-biblical form “*Izjūti augšā celšanās prieku*”. (*feel the joy of taking-off*). However, the initial ad and the commotion created attracted even more attention to the issue (perhaps that was what the advertiser wanted) and earned the ad the Golden Hammer for best ad adaptation at the end of the year.

Yet after the first clumsy attempts, advertising became more localised and creative. Advertising often employed various allusions to familiar words and phrases,

e.g. the audio company “Pioneer” used the slogan of “*Gatavāks nekā jebkad*” (More ready than ever) which alluded to the Red pioneer movement slogan of Soviet times “*Vienmēr gatavs*” (Always ready). This, however, often went hand in hand with a kind of popular etymology trend that reinterpreted new company names. The Latvian branch of the construction company “Paroc” waged an aggressive campaign with the title “*Paroc Latvija*” that was generally interpreted as “*Paroc Latviju*” (*Dig up Latvia*). A **Duracel** battery ad was interpreted as “*Dura cel*”, *dura nes septiņreis tālāk*” (*the fool raises, the fool carries seven times further*).

“**Whiskas**” cat food, there being no phonetic similarity between “whiskers” and the Latvian equivalent, compensated by finding a phonetic similarity in the description: “*Whiskas – viss, kas nepieciešams*” (*all that is needed*).

Unusual, visually striking compounds appeared, e.g. **Viscaurbraukšana** (FF, 1. 1997: 10) (all-through-going) in an ad for snowboards.

Today in Latvia advertising uses a variety of linguistic devices to attract attention, e.g. a recent advertisement for one of the chain stores:

- (6) *Ķer tās cenas šonedēļ*. (the word **tās** in green characters, the rest black)

Gloss 1: *Ķertās cenas šonedēļ* = Crazy prices this week.

Gloss 2: *Ķer tās cenas šonedēļ* = Catch the prices this week.

Thus advertising was a major channel through which Latvian usage conventions were modified and universalised. This seriously extended both the frequency of creative use and its scope, consolidating these types of creativity in the Latvian medium.

The third major influence was that of magazines and the daily press that now obtained most of their information from English (through translations, adaptations and latent/hidden translations, i.e., rewrites from foreign sources with few adaptations, presented under the signature of a Latvian writer). Either intentionally or due to carelessness, these messages in the press contained numerous instances of creative language transferred into Latvian. This is an interesting phenomenon as, strictly speaking, hidden translations need not be based on close rendition of source-text linguistic material, so that alternative means of expressing the same ideas could be found. Yet it seems that the authors/translators wanted to transfer the nonexistent elements into the target language. The film and TV industry followed suit, as for example, in **Great sexpectations** – translated using a semicalque: *Lielās sekspektācijas*. The crucial element is a blend of universally used **sekss** and a somewhat rare loan **ekspektācijas**, which presumes that most of the Latvians will already be familiar with the latter term. The usual translation of Dickens’ novel in Latvian is “*Lielās cerības*”).

- (7) *Darbaholiķu sievas ir visdrošākais laimests. Nodokļu ievācējs – “darbaholiķu” un “karjerholiķu” sievu speciālists* (Atmoda Atpūtai 30.7.1995),
 Gloss: *Wives of workaholics are the easiest to get. Tax collector – an expert of workaholics and careerholics’ wives.*
- (8) ... apvija tikai **yes man** – **jā-ļautiņu** aura (D. Īvāns. *Gadījuma karakalps*. 1995: 117).
 Gloss: ...was surrounded only by yes-men – Yes people aura.

3.3 Creativity in idiom transformations.

Along with borrowing of words and idioms, Latvian tended to expand occasional/instantial use of idioms, both in translated and native language texts. The means of occasional transformations in Latvian are similar to those in English and other languages. Occasional transformations of idioms are intentional, subjectively and stylistically motivated transformations (Gläser 2001: 142).

The modifiability of idioms (see Čermak, 2001) and the systemic character of these modifications are among the fundamental traits of phraseology. The effect of such transformations results from the contrast between the transformed speech variant and the normal reading of the idiom in its unchanged form, which defeats the reader's expectation (see discussion in Kuiper, this volume). The expressive effect of breaking the language norms was pointed out by Charles Bally as early as 1921 and it constitutes a way of refreshing and lightening the often trite and hackneyed dead metaphors that pervade the language. The means and types of occasional transformations are part of the language and are re-activated each time they are used in a concrete speech situation. No new occasional transformation types have been noted in Latvian; the change lies rather in an enormous increase in the frequency of their use. While in the past the frequency of occasional change could be estimated as very low (but corresponding to the conventions of use at the time), now it can be considered relatively high. As in English (Moon 1998: 290; Hermeren 1999) and other languages (Coffey 2001), this phenomenon is rife, mainly in advertising, journalese and fiction. There are now Latvian texts that are saturated with occasional change, yet according to small-scale interviews concluded by the author, they are not viewed as inappropriate or striking by the average Latvian reader.

To exemplify this phenomenon of linguistic transfer, let us look at just one example. In the early nineties a generally unknown phrase, *skelets skapī* (*a skeleton in the cupboard*), within a few years became one of the most frequently-used idioms in Latvian texts. The earliest occurrences of the expression generally ap-

pear in texts dealing with foreign countries; the expression is also accompanied by an explanation, yet even at this early stage, it frequently undergoes creative transformations, e.g.

- (9) *Nekas neliecina, ka arī šīs kuplās ģimenes **skapī** slēpjas kāds “skeletons”, proti, noslēpums...*
 Gloss: *Nothing tells one that also in the cupboard of this big family is hidden a “skeleton”, namely, a secret*
- (10) *Tā ka – ne ģimene bez sava **skeleta skapī**, kaut arī šis **skeletons** šoreiz nav nedz asins, nedz kāds cits grēks, vien mērķu un ideālu pieticība.*
 Gloss: *Thus no family without its skeleton in the cupboard, though this skeleton is no blood or other crime, only humbleness of goals and ideals.*
- (11) *Dāņi nebūt nav vienīgie, kas azartiski metušies **skeleta kauliņu apsūkāšanā** (Santa 1989,9)*
 Gloss: *Danes are not the only ones who go for sucking the little bones of the skeleton with abandon.*

The early use occasionally presents Latvian variants of the idiom (**corpse in the cupboard**):

- (12) *Hičkokiskā mērcē pasniegta ģimenes drāma ar “**mironi skapī**”. (Izklaide. 16.11.2000)*
 Gloss: *A family drama with a corpse in the cupboard offered in a Hitchcock-like dressing.*

However, in the 21st century the phrase has been fully integrated into the Latvian phrasal lexicon, and now appears in original Latvian texts, although it almost always appears in a modified form:

- (13) *KNAB iespējamās vadītājas skapī **uzpeld skeletons** (www.delna.lv 15.9.2003)*
 Gloss: *A skeleton appears/floats in the cupboard of the KNAB potential boss.*
- (14) *Kalnietei arī **skeletons skapī**. Ūdrei varbūt nav skeletons, taču ir pilns skapis ar nullēm. (www.klab.lv 4.8.2004)*
 Gloss: *Kalniete has a skeleton in the cupboard. Udre does not have a skeleton, but her cupboard is full of zeros.*

In Latvian the plural form seems to be more frequent:

- (15) *Jaunā laika **skapim** pieteikušies 2 **skeletoni** (Diena 7.10. 2003)*
 Gloss: *2 skeletons have applied for the New Era cupboard.*

- (16) **Skeleti skapī** (*Diena*, 24.3.2004)
Gloss: *Skeletons in the cupboard*
- (17) **Skeleti jaukuma skapī** (*Diena* 27.1.2005)
Gloss: *Skeletons in the cupboard of charm.*
- (18) **Baltijas skeleti skapī** (*Latvijas Avīze*, 18.8.2005)
Gloss: *Baltic skeletons in the cupboard.*

The idiom is also frequently accompanied by the verb ‘to rattle’. Thus a look at the Latvian internet pages on any single day may present up to eight “rattling skeletons in the cupboard” on unrelated, purely Latvian topics, suggesting that the idiom has been fully absorbed in the language, e.g.:

- (19) **Kas grabinās skapī?** (*Zemgales ziņas*, 28.7.2005)
Gloss: *What is rattling in the cupboard?*
- (20) **Vai tās vadības skapī negrab kāds skelets?** (www.politika.lv)
Gloss: *Does not a skeleton rattle in the cupboard of the management?*

Creative use of this idiom has recently increased even more markedly. Thus in an article in a women’s magazine (*Ieva* 10.8.2005), written by a native speaker who interviewed three native Latvians, this idiom appears 20 times, practically always with instantial variations. The idiom is introduced by the interviewer, but it is picked up by all participants and appears in the title as well:

- (21) **Skelets skapī** (*A skeleton in the cupboard*) followed by a short introduction:
- (22) **Skelets skapī. Tā to mēdz saukt. Bet ko tad, ja kāds jauši vai nejauši sāk virināt skapja durvis? Un ko tad, ja... tu skeletam pēkšņi uzduries pati?**
Gloss: *A skeleton in the cupboard. That is what it is called. But what if someone by chance or intentionally starts opening the cupboard doors? And what, if... we come upon the skeleton ourselves?*

First sentence:

- (23) *Protams, man arī ir daži.*
Gloss: *Of course, I also have several. (skeletons A.V.)*

Further on in the text the image of the skeleton continues to recur, taking various twists and turns, performing the function of textual cohesion. In fact, it becomes a backbone structure of the text. And though the stories are quite different, it is the sustainability (Naciscione 2001: 64) of this idiom in all its creatively manipulated forms that serves the discourse needs and actually functions as a form of cohesion in the article:

- (24) *Ja kāds saka – viņam **skapī nav nekāda skeleta**, tad zini: viņam **tie ir pagrabā**. Un **daudz**.*
 Gloss: If someone says he does not have a skeleton in the cupboard, you know: they are in the basement. And many.
- (25) *Tas ir **mans skapis, mans skelets**, mana dzīve. To, ko es daru ar saviem **skeletiem**, pastāstīšu vēlāk.*
 Gloss: This is my cupboard, my skeleton, my life. What I do with my skeletons, I will tell you later
- (26) *...izstāstīsi noslēpumu – **skelets no skapja izgaišs?***
 Gloss: ..if you tell the secret – **the skeleton will disappear from the cupboard?**
*..tiekam galā ar saviem **skeletiem un skeletiņiem**.*
 Gloss: ..we cope with our **skeletons and little skeletons**.
*Šoreiz runāsim par **vīriešu kārtas skeletiem**.*
 Gloss: This time let us talk about the **male skeletons**.
*Tas ir mūsu **kopīgais skelets**.*
 Gloss: This is our **joint skeleton**.
*Man pašai to **skeletu ir tik daudz**, ka reisēm baidos – “**skapis**” būs jāpērk lielāks.*
 Gloss: I, myself, have **so many skeletons**, that I sometimes worry – I will have to buy a bigger “**cupboard**”.

And the last sentence:

- (27) *Līdz tu pati notīci, ka **nekāda skeleta tai jūsu skapī nav**.*
 Gloss: Until you start believing that **there is no skeleton in your cupboard**.

Concluding this section, it can be observed that creative use of idioms in Latvian has reached a frequency that would never have been expected ten or fifteen years ago. This can be viewed as a serious shift in idiom usage conventions.

3.4 Creativity in word formation

The most significant outside influence seems to be in the sphere of word formation where creativity has been extended into hitherto unthinkable areas. At the beginning of this trend, most creative formations appeared in translated texts, but in the past five to ten years they have also been appearing in native Latvian texts, suggesting a greater penetration of neologisms and their patterns. Originally, practically all new formations had English counterparts, mostly compounds and blends. Others were direct loans:

holideģenerēšanās (holidegeneration), bolivuda (Bollywood), vidiots (vidiot), seksperts (sexpert), kleptokrātija (cleptocracy), lukanomika (Lukanomics), kokakolonizācija (Cocacolonisation),

some semiloans (a mixed compound consisting of a loan and a translated lexeme):

eksvīrs (ex-husband), darbaholiķis (workaholic), karjerholiķis (careerholic), sierburgers (cheeseburger),

and some translation loans:

jā-ļautiņi (yes-men),

(28) *Cik ilgi izvilks **Lukanomika**. (Klubs, 9, 1998) (Lukashenko + economics)*

Gloss: How long can Lukanomics survive.

(29) ***Kleptokrātija** kā demokrātijas uzlabotā forma (Diena 30.08. 1998);*

Gloss: Cleptocracy as an enhanced form of democracy.

(30) *– Pie kura sevis pieskaiti? – Hm... pie **vidiotiem** (Rīgas Balss, 2.11.1994.)*

Gloss: To whom do you belong? – Hm... to vidiots.

(31) *Paldies par Nē – **kokakolonizācijai** (Diena, 29.11.2000).*

Gloss: Thank you for No to Cocacolonisation.

Yet, gradually, in a parallel process, we have witnessed a growing number of original Latvian compounds and blends (a fair share of these has foreign components that were long ago borrowed in Latvian; this seems to point towards a general feeling that the model is imported and its use safer or more usual with non-native lexemes). Some examples:

- Compounds

Varastrīce ((political) powerquake) on the analogy of zemestrīce (earthquake),

eiroremonts (eurorepairs)

standard word for “quality, modern repairs” since the early nineties

Natoremonts [NATO remonts] (NATO repairs) on the analogy of eiroremonts.

Natoremonts, however, was an ironic term referring to those officials who used NATO integration money to purchase and renovate their apartments,

(32) *Čečenija izraisa **varastrīci** Krievijas armijas virsotnē. (Labrīt 23.12.1994).*

Gloss: Chechnya created a powerquake at the top of the Russian army.

(33) *Eihmaņa sarīkotajam **natoremontam** būs sekas. (Diena 13. 10. 1998);*

Gloss The natorepairs organised by Eihmanis will have consequences.

- Partial blends

Zaķutēka [*Zaķusalas diskotēka*] (*Zaķu island disco*),
humpala [*humanitārā palīdzība*] (*humanitarian assistance*)
zemsavieši [*Zemnieku savienība*] (*people of Farmer's Union*)
baltbats [*Baltijas bataljons*] (*Baltic battalion*)
Latbats [*Latvijas bataljons*] (*Latvian battallion*)
murgologs [*murgi logs*] (*nightmarelogue*)
Latvegasa [*Latvijas Lasvegasa*] (*Latvia's Las Vegas*),
Migrorajons [*migrantu rajons*] (*migrant region*) on the analogy of
Mikrorajons (*microregion – peripheral dwelling region*)
Šmerļavuda [*Šmerlis Holivuda*] (*Šmerlis – Riga outskirts where the*
Latvian cinema industry centre was located)

(34) **Latvegasa** 1994 (*Rīgas Balss* 26.10.1994)

(35) **Šmerļavudas** stāstiņi (*Rīgas Balss* 19.10.1994)

Gloss: Šmerļawood stories.

(36) Kurš **murgologs** gan nosauktu Prustu... (*Diena* 16.5.2001)

Gloss: What nightmarelogue would call Proust...

(37) ...viņam liela pieredze **humpalu** gādāšanā. (*Rīgas Balss* 20.10.1995.)

Gloss: ... he has a great experience in organising humanitarian assistance.

(38) **Eihmaņgeita**. (*Diena* 15.10.1998);

Gloss: Eihmangate

- Full blends

ceļarāma [*Latvijas ceļa panorāma*] (*Party „Latvia's way” Panorama*),
sliktenis [*slikts liktenis*] (*bad fate*),
ļeņineklis [*Ļeņins piemineklis*] (*Lenin monument*),
cūkmens [*cūka betmens*] (*pig Batman*) a media image of a nature polluter,
Putinočets [*Putins Pinočets*] (*Putin Pinochet*),
Ģimnastrāde [*ģimnāzija estrāde*] (*grammar school stage*),
Gastronauts [*gastronomija astronauts*] (*gastronomy astronaut*):

(39) TV **Ceļarāmā**. vienīgais pareizais ceļš ir Latvijas ceļš. (*Rīgas Balss*, 23.03.1994).

Gloss: In TV Wayrama the only correct way is Latvia's Way.

(40) Vai Krievijā valdīs **Putinočets**? (*Lauku Avīze* 23.3.2000).

Gloss: Will Russia be ruled by Putinochet ?

(41) Mans liktenis ... **Sliktenis**... (*Una* 2005. febr.)

Gloss: My fate...Bad fate.

- (42) *Orālā kloāka Latvija svin **pedofilgeitu** ar īsti histērisku vērienu. (Diena 30.9.1999).*

Gloss: The oral gutter Latvia celebrates its paedophilegate in a really hysterical way.

Wordplay based on word formation has now become a common occurrence, penetrating discourses in TV and cinema, where the necessity of rapid decoding presumes that the public is accustomed to the playful use of language. Thus the TV serial “Sex in the city” is characterised by the pervasive use of suggestive idiomatic language and wordplay, most of which is imaginatively integrated into Latvian in the translation.

Graphic wordplay has also become commonplace, including paronymic substitution of letters or sounds for purposes of wordplay as in: **valdības sastārdišana** (*valdības sastādīšana* – composition of the government; *saārdīšana* – destroying); **migrorajons** substitution of a letter in **mikrorajons** (microregion, dwelling region), which results in a blend, **migrants region**;

- (43) *Rīgas **migrorajoni** bagātīgi piebira ar čaukstošām papīra lapiņām (Īvāns. Gadījuma karakalps. 1997.210. lpp);*

Gloss: Riga migroregions were littered with rustling leaflets.

To describe NASA failures with the space shuttle programme the newspapers have used

- (44) ***NASAnāca.** (TV Diena. 26.7.2005)*

The orthographic change and substitution of a letter reads as “NASA came”, but alludes to **nesanāca** (It did not succeed). A few weeks later another wordplay appeared:

- (45) ***NASAlūza** (TV Diena 6.8. 2005),*

which would translate as *NASA broke down*. Similarly:

- (46) *NESCAFE un AIZNESCAFE (TV Diena. 26.7.2005) NESCAFE would read in Latvian as *bring the coffee*, AIZNESCAFE as *take away the coffee*.*

Most of the wordplay involving occasional or nonce use of vocabulary centers on two types of linguistic units: phrases and compounds. These are either familiar units which can be considered as clichés, often hackneyed and trite, but as a result of the creative manipulations they gain new life, as in the following:

- (47) *Skaista latvju **tauku** meita (TV Diena, 14.5.2005),*

Gloss: A beautiful Latvian fat-maiden.

where **tautu meita** (*folk-maiden*) has been changed into **tauku meita** (*fat maiden*).

- (48) *Pēdu manu lielu pēdu!* /49 izmēra īpašnieks (TV Diena 20.augusts).

In a well-known rhyme from a Latvian folksong **bēdu manu lielu bēdu** (trouble, what big trouble I have) **bēda** is substituted by **pēda** (foot). And the bracketed part of the sentence says *owner of size 49*, which suggests he finds it difficult to get shoes of his size.

The second category of creative wordplay includes topical, fashionable linguistic items, usually with an ironic and sarcastic tone. Thus an American TV serial about the Hilton sisters on a farm attracted much ridicule in Latvia and was subject to various sarcastic criticism, e.g.

- (49) *30 dienas spermā.* (TV Diena 29.4. 2005)

Gloss: *30 days in the sperm.*

Substitution of **fermā** (*farm*) by **spermā** (*sperm*).

- (50) *Smaga riebuma pakāpe – 4 promiles.* (TV Diena 20.08.2005).

Gloss: *A heavy degree of disgust – 4 promilles.*

Substitution of **reibuma** (*intoxication*) by **riebuma** (*disgust*)

- (51) *Pašas adidas zeķītes.* (TV Diena 29.4. 2005)

Gloss: *Self made socks.*

Substitution of **aditas** (*self made*), by **ADIDAS**.

In a TV counter-ad **Jaunais laiks** (*New Era* – party name, in itself an allusion to a famous poem by the great Latvian poet Rainis) became **Ļaunais laiks** (*Evil era*) by changing a letter.

There are some contexts in which wordplay is particularly frequent. Usually these are topical issues of political and cultural life that attract wide attention. These are then played upon by the media, linking the politicians' statements with what many people think about them, thus joining in one linguistic unit a contrast or a paradox between the words and perceived deeds. For example, when the government announced that only iodised salt would be sold in Latvia, sparking off a shopping bout of "ordinary salt", the newspapers responded with various kinds of wordplay, playing on the homonym of Latvian **jods** (*iodine* and the *Devil*)

- (52) *Pie joda. Dodiet sāli* (TV Diena 7.7. 2005)

Gloss: *To hell/devil. Give us the salt.*

Much wordplay with word formation devices occurs on the internet and in chat rooms where its full potential, as well as the resourcefulness of "amateurs", can be seen. Thus the portal www.ass.lv plays with the word 'ass': in all the subtitles

one has a red morpheme **ass** (meaning *sharp* in Latvian), frequently adding an extra 's' where it does not exist in some other words in the Latvian language, e.g. **viskasste**, (box for everything), also allusion to **miskaste** (rubbish bin), **assumi** (sharpness), **sarunass** (conversations).

Some previous nonce words have gained wide currency in the language and can be considered as having entered the Standard Latvian lexicon. A case in point is **nacbols** (Nacionāl boļševiks – national Bolshevik):

(53) *Nacbola lode. Diena 7.5.2002.*

Gloss: Natbol's bullet

Finally, systemic novelty has even entered into the traditionally conservative stronghold of Latvian – that of terminology. If **kaplete** (*kapsula+tablete, capsule tablet*) is viewed by some as an imported blend then, for example, the more genuinely Latvian **mēstule** (*mēslu+vēstule, junk e-letter*) serves as proof that the expansion of new word formation patterns has been extended to all styles and registers of the language.

3.5 Compound phrases

Phrasal compounds of the type *on-the-spot creations, will-she-or-won't-she-get-the-guy comedy* were in fact nonexistent in Latvian before the nineties and this model of hyphenated compound phrases has certainly been imported. Hyphenation is rarely used in Latvian, and can generally be viewed as a novel phenomenon even between novel prefixes and the root lexeme, as in **e-pasts** (*e-mail*), **e-pārvalde** (*E-government*), **i-banka**, (*internet bank*) all of which appeared contemporaneously with phrasal compounds. But these are stable loans and not within the scope of this study.

"Temporary" or "improvised" nonce phrasal compounds, which are quite popular in English were, however, not usually translated as compound phrases, but were initially used almost exclusively in native language texts. In this way, the innovation bypassed the usual first "translation" stage, and was borrowed as a technique to be used with native Latvian material without being absorbed through numerous translations.

There were probably several reasons for borrowing this model: novelty and imitation in the first place, but also, as in English, this word formation pattern is a form of expression that allows meanings to be condensed in a brief expression. These are usually, but not exclusively, adjectives, which provide the writer with an almost unlimited stock of novel and also graphically obtrusive expressive modifiers. Different from English, where linguists have urged restraint in the use of

hyphenated compounds, no objections to this form have so far been registered in Latvia. These expressions tend to appear mainly in newspapers and magazines.

- (54) Rūdītais **lai – ar – kādi – vēji – pūš** politiķis Nikolajs... (Diena, 7.3.1994.)
 Gloss: The weathered let-any-winds-blow politician Nikolajs
- (55) ..tā sniedz cilvēkiem priekšstatu par **realitāti – aiz – realitātes**. (Grāmatu Apskats, 1994, 13.)
 Gloss: It provides people with an idea about reality-behind-reality.
- (56) “**Pirmais-kas-lieto**” valstīs tiesības stājās spēkā.... (Intelektuālā īpašuma aizsardzība. 1995)
 Gloss: In the first-to-use countries the law takes effect...
- (57) Ar demokrātisko **divtūkstoš-un-cik-tur-balsošanu** izšķirtos... (Diena Izklaidei 12.3. 1999)
 Gloss: With democratic two-thousand-and-what-not-voting it would come to
- (58) ..kādas un kam ziepes sagatavojis **ne-politiskais-pusžurnālists-it-kā-producers** ar iesauku Arābs-Labvakars (Diena 30.9.1999)
 Gloss: ..what trouble has been brewed by non-political- halfjournalist-as-it-were- producer with a nickname Arab- Goodevening.

Hyphenation is occasionally used also for stylistic purposes, where it divides existing compounds for purposes of wordplay, e.g.

- (59) **DZIR-NAV-NIEKI** (Ventas Balss 9.7.2005)

The word **dzirnavnieki** (millers) in this article about millers is used in the title in a hyphenated form, creating a certain intrigue as to why this has been done. Decoding, however, is made possible in the last sentence, where the word is split into three:

- (60) Jo tas **dzir(d) nav nieki**.
 Gloss: For that one hears is no joke.

After having creating suspense, the meaning of the hyphenated title is finally explained to the reader. Wordplay as a result of creative manipulation of the word has again fulfilled the function of textual cohesion.

Finally, capitalisation of letters or the use of a different colour for part of the word to mark wordplay has become fashionable. A poster drawing attention to the exodus of Latvians to Ireland (Irija) as guest-workers, makes use of the linguistic potential of the word Siberia (Sibīrija) with its connotations of Soviet deportations.

(61) *Latvju vīri SibĪRĪJĀ.*

Gloss: *Latvian men in SibERIA.*

4. Conclusion

Thus we have seen how outside influences have activated latent structural and semantic features of the Latvian lexicon by first having an effect on translations (both acknowledged and unacknowledged), then moving into advertising discourse (both translated and original), ultimately affecting original texts (in the media, the electronic medium, chat lines), thereby liberalising formerly conservative and rigid use of lexical and idiomatic material.

Linguistic creativity has expanded the stylistic boundaries of Latvian, making texts much more attractive and stylistically effective. The growth of the scope of creativity has been attested both in the expanded use of certain stylistic means (idiom transformations, use of clippings, nonce compounding) that previously were little exploited, and in the appearance of new stylistic devices such as native blends or compound phrases. The latter linguistic devices, though imported, have occupied a niche in the Latvian linguistic system and are now used in various speech domains without representing a threat to the identity of the Latvian language or its use.

References

- Alexander, R. J. 1997. *Aspects of verbal Humour in English*. Tuebingen: Gunter Narr Verlag.
- Alexieva, B. 1997. "There Must be Some System in this Madness." In *Traductio, Essays on Punning and Translation*. D. Delabastita (ed.), 137–154. St Jerome Publishing.
- Bally, C. 1921. *Traite de stylistique française*. Heidelberg: Winter's Universitaetsbuchhandlung.
- Chesterman, A. 1997. *Memes of Translation*. Amsterdam: John Benjamins.
- Coffey, S. 2001. "Disturbing the Form-Meaning Nucleus of Multiword Units: Data and Issues." *Studi Italiani di Linguistica Teorica e Applicata*. 2: 215–228.
- Čermak, F. 2001. "Substance of Idioms: Perennial Problems. Lack of data or Theory." *International Journal of Lexicography*. Vol. 14. Nr. 1: 1–20.
- Faulseit, D., Kuehn G. 1975. *Stilistische Mittel und Moeglichkeiten der deutschen Sprache*. VEB Bibliographisches Institut. Leipzig.
- Faulzs, D. 1982. *Melnkoka tornis*. Riga: Liesma.
- Fill, A. 1992. "“Visions and Revisions”: A type of Pun Based on Idiomatised Complex Words." In *Language and Civilisation: A Concerted Profusion of Essays and Studies in Honour of Otto Hietsch*. C. Blank et al. (eds). Vol. 2, 551–563. Frankfurt-on-Main, etc.: Peter Lang.

- Gläser, R. 2001. "The Stylistic Potential of Phraseological Units in the Light of Genre Analysis." In *Phraseology. Theory, Analysis, and Applications*. A. P. Cowie (ed.), 125–144. Oxford, OUP.
- Hausmann, F. J. 1974. *Studien zu einer Linguistik des Wortspiels*. Tübingen: Max Niemeyer.
- Hermans, T. 2000. "Norms of Translation." In *Oxford Guide to Literature in English Translation*. P. France (ed.), 10–15. Oxford.
- Hermeren, L. 1999. *English for Sale: A Study of the Language of Advertising*. Lund: Lund University Press.
- Hymes, D. H. 1983. "Notes toward a History of Linguistic Anthropology." In *Essays in the History of Linguistic Anthropology*. D. H. Hymes (ed.), 1–57. Amsterdam, Philadelphia: John Benjamins.
- Lambert, J. 1995. "Literatures, translation and (De) colonisation." In *Translation and modernisation. Proceedings of the 13th Congress of the international Comparative Literature association*. 98–118. Tokyo.
- Locmele, G. 2001. "Means of Manipulation in Latvian Advertising – Products of Change." In *Contrastive and Applied Linguistics*. A. Veisbergs (ed.), 47–54. Riga: University of Latvia.
- Moon, R. 1998. *Fixed Expressions and Idioms in English*. Oxford: Clarendon Press.
- Munat, J. 2004. "A Case Study in Cross-cultural Translation: TINTIN in English and Italian." In *Contrastive and Applied Linguistics* 12. A. Veisbergs (ed.), 101–120. Riga. SVTN.
- Naciscione, A. 2001. *Phraseological Units in Discourse*. Riga: LAC.
- Petrotta, A. J. 1991. *Lexis Ludens. Wordplay and the Book of Micah*. New York: Peter Lang.
- Redfern, W. 1986. *Puns*. Oxford: Basil Blackwell.
- Riffaterre, M. 1960. "Stylistic Context". *Word* 15: 154–174.
- Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins.
- Veisbergs, A. 1997. "The Contextual Use of idioms, Wordplay and Translation." In *Traductio. Essays on Punning and Translation*. D. Delabastita (ed.), 155–176. Manchester: St Jerome Publishing.
- Veisbergs, A. 2005. "Tulkojumspiediēns – guvums vai apdraudējums?" *Valsts Valodas Komisija. Raksti*. 1.sējums: 187–201. Riga: VVK.
- Venuti, L. 1995. *The Translator's Invisibility*. Routledge.
- Zauberga, I. 2005. "A Knock on the Door: On the Role of Translated Literature in Cultural Image Making." *ACROSS Languages and Cultures*. Vol. 6, issue 1: 67–77.

Critical creativity

A study of 'politically correct' terms in style guides for different types of discourse

Roswitha Fischer

This article aims to contribute to a more general understanding of the inter-relations between lexical creativity, language policy and the discursive domain. It focusses on 'critical creativity', which puts existing orders into question and offers new alternatives. A case in point are expressions which in the past have been subsumed under the term 'political correctness', some of which have recently been codified by style guides of bias-free usage. When comparing guidelines for academic writing, official documents and newspaper language, it is apparent that the respective discourse type strongly determines the number and kind of recommendations given. An inherent need for new, specialised vocabulary facilitates the implementation of the suggested alternative expressions.

1. Introduction

This article does not so much deal with imagination and its linguistic potential, but rather with its inhibitions and limitations. By examining the impact of censorship on lexis, it aims to contribute to a more general understanding of the inter-relations between lexical creativity, language policy and the discursive domain. It focusses on 'critical creativity', which arises from a critical assessment of existing orders and proposes alternatives. A case in point are expressions which in the past have been subsumed under the term 'political correctness', some of which have recently been codified by style guides of bias-free usage. When comparing guidelines for academic writing, official documents and newspaper language, it becomes evident that most of the recommendations are based on existing lexical structures and make use of the possibilities of polysemy and meaning shift. Furthermore, the respective discourse type strongly determines the number and kind

of recommendations given. Apparently, an inherent need for new, specialised vocabulary facilitates the implementation of the suggested alternative expressions.

2. Creativity as social practice

There are several approaches to creativity in existence today, which can be subsumed under psychological and sociocultural approaches (Carter 2004: 17–52). The point of view adopted here is that creativity is both an individual and a social practice. It is a result of the interaction between the individual speaker/writer, the discourse variety and the forms and styles which are acceptable in the respective culture or society. The individual creative act has to be acknowledged by others and institutionalised in a certain environment, otherwise it will not be available to a wider range of users. When creativity serves the purpose of criticising the existing conditions and offering an alternative point of view, it can be called *critical* (Carter 2004: 47). In contrast to other kinds of creativity, critical creativity is very restricted, because it is not based on what could be done but rather on what should not be done. It does not so much emerge from playfulness but from thorough analysis and wilful thinking. We would expect the degree of creativity, which is preceded by a negative sign, as it were, to be quite low. Nevertheless, it will be interesting to see how much, and in what way, creativity is still possible under these difficult circumstances.

‘Politically correct’ terms can be considered a case of critical creativity. On the grounds that certain minorities are discriminated against, and that language has an effect on attitudes, non-biased expressions are introduced in order to replace other established but discriminatory terms and phrases. Many representatives and proponents of social movements and of political correctness have made suggestions for non-discriminatory expressions, but only a limited number of these has achieved any historical value, standing the test of time and becoming institutionalised in the wider speech community or parts of it. In this chapter, I will deal with the type of critical creativity called political correctness that has increasingly effected linguistic change in society, as opposed to personal creativity, whose sphere of activity is more restricted. In particular, I shall be looking at those terms that have found their way into certain varieties of specialised discourse. One might think that creativity and specialised discourse represents a contradiction in terms since specialised lexis has been characterised *inter alia* as mono-referential, conservative and lacking emotion (Hoffmann 1984). Furthermore, formal written technical language is not supposed to be creative, but should be fixed, objective, valid and reliable. Nevertheless, some politically correct terms have been integrated into the specialised discourses of public written texts – apparently reflecting a felt need.

Indeed, some of these have been codified through institutional guides and reference works of bias-free usage. These guidelines mostly aim at the implementation of bias-free language in public written discourse. The main domains of language concerned are academic publications, the language of national and international political institutions, and the print media. It is my purpose in this article to examine whether, and in what way, these three types of discourse are affected by critical creativity.

In addition to the psychological and sociocultural approaches to creativity, there exist the concepts of *linguistic creativity* and *linguistic productivity*. These two specialised terms have to be understood against the background of generative morphology, which distinguishes between rule-governed and non rule-governed word-formation processes. While productivity refers to rule-governed word-formation, the speaker's/writer's creativity is unpredictable and not governed by rules. Being rule-based, the productive word-formation processes are part of morphology. In contrast, creativity is perceived as a non-regular, intentional activity, and thus not determined by the grammar.

However, productivity and creativity cannot be kept entirely separate. Morphological rules may also be unproductive (cf. the English suffix *-al*), and not all unproductive formations are irregular such as the German *-er* plural (Haspelmath, 2002: 41). Haspelmath has further pointed out that morphological rules play a creative role in the invention of new words not listed in the lexicon. Moreover, whether a word becomes accepted as normal or not depends on a variety of circumstances, most of which have nothing to do with productivity. Since the focus of this article is on the sociopolitical conditions affecting the implementation of new linguistic conventions and not on productive word-formation rules, it seems best to disregard the distinction between linguistic productivity and creativity altogether and turn directly to the analysis and interpretation of existing lexical-morphological patterns.

3. Politically correct terms as a case of critical creativity

Political correctness is today a widely used 'buzzword' in the Western world. The concept has its roots in the American social movements of the '60s and '70s, which not only criticised legal discrimination but also opposed stigmatisation through language (Fischer 2002). These movements eventually triggered extensive societal changes, including programs for linguistic sensitisation and more respectful attribution of names. In the '80s and '90s a second wave of concern for equal treatment of minorities began to take its course among leftist groups on US campuses, spreading through the universities and the national and international mass media.

At that time the phrase 'political correctness' became widely used as an umbrella term for a number of heterogeneous topics, extending from general attempts to support social minorities to revisions of the literary canon at universities and affirmative action programs. Since then the expression, with its corresponding adjectival construction 'politically correct', has become discredited. It now tends to be taken as a disparaging term for allegedly radical liberals who uphold the rights of feminists, multiculturalists, Afrocentrists, and other disadvantaged minorities. According to Sabine Wierlemann (2002: 39f., 99ff.), the concept of political correctness has been exploited by conservative politicians and the media (see also Fischer 2005). Caricature, mockery and trivialisation have also contributed their share to the shaping of public opinion.

Because politically correct substitutes are an attempt to deliberately influence a language, they can also be considered as instances of language planning. According to Daoust (1997: 438ff.), the study of language planning was at first mainly concerned with internal aspects of language and was eventually refined to encompass not only linguistic and sociolinguistic but also economic and political aspects of the integration of language into society. In addition, socio-psychological aspects of language behaviour were also taken into account, and attitudes came to be regarded as an important driving force for linguistic maintenance and change.

Language planners picture an ideal sociolinguistic situation that should be generated in the future. Motivated by a certain ideology, they consider it necessary to change the attitudes and norms of speaking or writing in a given direction. Envisioning an ideal world free of inequality and injustice, the proponents of political correctness promote language purged of discriminatory or disparaging expressions. They are in fact purists, because they intend to cultivate and codify the language in accordance with an ideal model of 'pure' language by the elimination of elements that do not fit into this model.

Despite ongoing controversies over the attitudes and aims of the proponents of political correctness, some remarkable linguistic changes have nevertheless come about. A number of new language norms proposed by the advocates of political correctness have gained recognition, partly due to the increasing influence of American social movements on the government machinery, and partly due to the authorities of American universities and their legal advisors. Nowadays a considerable number of international, governmental, judicial and other public institutions, as well as the media agencies, provide guidelines for bias-free writing. The successful implementation of some terms and meanings gives rise to the supposition that societal feelings, judgments, and attitudes have also undergone certain changes.

4. Style guides

4.1 General reference books of bias-free usage

There are a number of extensive guidebooks of bias-free usage which contain suggestions for non-discriminatory language. Well-known examples are those by Rosalie Maggio (1991), Anne Pauwels (1991) and Henry Beard & Christopher Cerf (1992). These comprehensive manuals aim at a wide readership, including writers, editors, public speakers and virtually anybody who wants to avoid offensive, imprecise or misleading language. These guides include thousands of alternatives for words and phrases that are considered to be biased against people on the basis of sex, race, age, sexual orientation, disability, ethnicity, class, or religion. In addition, they often give definitions of key concepts (such as *homophobia*), quotations, explanations and etymologies. Most start with an introduction explaining the relationship between language and society and stating the need for linguistic changes. Some are divided into chapters dealing with gender, race, disabilities, etc., first giving general explanations and then commenting on specific items. Others prefer to list the terms as entries in alphabetical order with further commentaries included whenever necessary. Thus the user of the guide can either familiarise him/herself with the basic concepts of bias-free usage, or just look up a word or phrase when needed.

While manuals such as Maggio and Pauwels are written in a serious tone and ask for the replacement of biased words with terms that treat all people respectfully, Beard & Cerf's handbook also intends to entertain and amuse the reader and even makes fun of the movement for political correctness, which can be considered a reaction to the concern for equal treatment in the United States in the 80s and 90s. In this book, the idea of non-biased usage is taken to the extreme, and many nonce-formations are included in a jocular vein. Examples are *gingerbread person* instead of the common *gingerbread man*, *significant other* instead of *girlfriend*, and *botanical companions* instead of *flowers*. Nevertheless, established politically correct items such as *firefighter* or *chair* are also included, and therefore Beard and Cerf's handbook is suitable for the writer who is concerned about linguistic discrimination. Sometimes it is the chosen style guide recommended by institutions (perhaps for lack of familiarity with other alternatives).

To sum up, these comprehensive guidelines serve as general reference works for the uncertain user who might want to check on current usage. The recommendations given are non-binding, and it can be assumed that their influence on the style of the average writer is rather negligible. In contrast, style guides published by certain official institutions for their own employees can be expected to have a much higher impact on the writing styles within a given professional field.

These serve not so much as guidelines but as prescriptions to be followed. For this reason we would expect to find far fewer recommendations in style guides for specific discursive domains. Furthermore, we would assume that the number and the kinds of alternatives given for different fields of discourse vary according to purpose, content, and context. Our task now is to relate the findings in the respective style books to the needs and the intentions of the type of discourse involved and to establish whether it is possible to arrive at a more comprehensive understanding of the interrelations between lexical creativity, language policy and the discursive domain.

4.2 Guidelines for public written discourse

Though the use of politically correct language was originally intended to pervade all areas of linguistic usage, it has mainly affected certain subdomains of written language. In particular, there are three specialised discourses into which PC terms have been more or less successfully introduced: academic writing, official documents and media language. For these discourse types, a number of guidelines have been published, which in itself is quite remarkable and indicates their important role in society. According to Gotti (2003:24), specialised discourse can be defined as a language typical of a specialised community stretching across the academic, the professional, the technical and the occupational areas of knowledge and practice. Discourse types can be distinguished according to a variety of interrelated text-internal and text-external factors. Apart from the disciplinary domain, the level of specialisation and the relationship between the participants in the communication need to be considered. The style guides examined are all examples of instructional texts. The level of specialisation of these guidelines for both writers of academic papers and of official documents can be called intermediate, in that specialists address non-specialists to explain notions pertaining to their discipline (the highest level of specialisation being that of experts addressing other specialists, cf. Gotti 2003:26). In contrast, newspaper style guides provide information of a technical nature for both journalists and non-specialists, and are thus the least specialist-oriented – strictly speaking, they do not involve a truly ‘specialist’ use of language.

The style guides examined mostly consist of general recommendations for the composition of texts, including spelling, disputed usage, and jargon. Some headers or paragraphs also deal with politically correct language. These match the many usage guides that are on the market for the common language user, for instance *The Oxford Guide to English Usage* or *A Dictionary of Modern English Usage*. It would be interesting to compare the various style guides for different

fields of discourse with the guides for general usage and relate the findings to differences in language, content and situational factors. Here, however, I shall restrict my focus to style guides for the three discourse practices mentioned above and examine the recommendations given for bias-free writing. Before beginning the comparison, however, an introduction to the selected style guides is necessary.

Ten style guides were chosen, most of which are available on the internet. The selection was made more or less randomly. Since few institutions provide their users with stylistic advice, I selected only the style guides that could actually be found via a search engine and that also suited my purpose; i.e. they had to include at least some considerations about non-biased language. When available, the recommendations of international or widely acknowledged institutions were preferred as a guarantee of wide distribution and recognition.

For academic writing, the internationally used *MLA Handbook for Writers of Research Papers* was chosen; in addition, the online recommendations of three universities with different academic orientations were considered, viz. the *Kent Publications Style Guide* from the University of Kent (Communications and Development), the *Editorial Style Guide* from Sheffield Hallam University (Marketing and Development) and the *College Style Guide* from the Royal College of Surgeons of England in London.

With regard to government domains, style guides of the European Commission, UNESCO, the United Nations Economic Commission for Europe,¹ and guidelines of the Government of Queensland (Australia) were considered.

Finally, newspaper language is represented by *The Times Style and Usage Guide* and *The Guardian Style Guide*, which are the most frequently consulted style guides for journalistic writing. The recommendations of the various style guides were then compared with respect to category (general, gender, disability, diseases, mental health, ethnicity, sexuality, age, inter alia), the expressions to be used or avoided and discourse type, including purpose and readership.

4.2.1 Quantitative differences

As expected, the numbers of recommendations found in the style guides of discourse fields are fewer than those in the general style books. Assuming that a general style book of non-discriminatory language has approximately 3000 entries on average, the items listed in the style guides examined for specific discourse fields are actually very few in number. With an average of 60 terms, they amount to only 2% of the total. This indicates that the implementation of politically correct

1. The *Daily Style Guide* was developed by the Economic Commission for Europe to achieve professional standards of information presentation in *The Daily*, which is Canada's official Statistics release bulletin, published under the authority of the Canadian Minister of Industry.

terms in restricted domains is not rigorously applied, and it seems to be strongly influenced by the purpose, content and attitudes involved.

All the style guides examined present advice on items to be avoided and usually suggest alternatives. Often general recommendations are given, for instance advice to avoid giving offence to women or to avoid defining people by their medical condition. For the three text types, the following frequencies were obtained for (a) general remarks about discriminatory usage, (b) the avoidance of certain items and (c) proposed alternatives:

Table 1. Style guide recommendations for academic writing

	MLA	Kent	Sheffield Hallam	College of Surgeons	Sum
General advice	2	3	–	3	8
Expressions to be avoided	4	37	9	5	55
Expressions to be used	4	24	10	6	44
Sum	10	64	19	14	107

Table 2. Style guide recommendations for official documents

	EC	UNESCO	UNECE	Queensland	Sum
General advice	3	2	4	7	16
Expressions to be avoided	3	62	10	14	89
Expressions to be used	6	195	24	16	241
Sum	12	259	38	37	346

Table 3. Style guide recommendations for news language

	Times	Guardian	Sum
General advice	2	4	6
Expressions to be avoided	28	43	71
Expressions to be used	25	86	111
Sum	55	133	188

The tables show that the frequencies in the different categories vary significantly with respect to the style guide examined. While, for instance, the MLA, Sheffield Hallam University and the Royal College of Surgeons of England do not include more than 20 items, the *Kent Publications Style Guide* enumerates three times as many expressions. The enormous number of items of gender-neutral language given in the UNESCO guidelines is also quite remarkable. It seems to be the case

that the number of recommendations given is a very individual matter, regardless of discourse domain or other social factors, and depends on the initiative and efforts taken by the respective authorities and the compilers of the guidelines. Nevertheless, some interesting tendencies can be observed. The academic guidelines tend to be quite vague in their suggestions, focussing on the avoidance of terms rather than on the presentation of alternatives. Generally it is recommended to avoid outdated terms that either stereotype or stigmatise, the reason given being that society has undergone certain changes that should be reflected by language. The *MLA Handbook* expresses itself most cautiously, saying that:

careful writers of research papers avoid language that implies unsubstantiated or irrelevant generalizations about such personal qualities as age, economic class, ethnicity, sexual orientation, political or religious beliefs, race, or sex.

(*MLA Handbook* 60)

thus evading its own responsibility.

In contrast to the avoidance strategy of academia, governmental and international organisations are more likely to suggest several substitutes for one item. Apart from the Queensland Government, these organisations mainly focus on gender-neutral language and on occupational titles in particular. Their recommendations originate in the wish to endorse equal treatment and, given that language also shapes thinking, to anchor it in language. Clear instructions are given, such as "Use your judgment in choosing noun forms to emphasise or deemphasise gender" (EC) or "Include all people in general references by substituting asexual words and phrases for 'man' words" (UNECE). Despite the imperatives used, the organisations stress that responsibility lies with the user, which reflects the main principles of human rights and fundamental freedoms pursued by these organisations. The first quotation given above is an example of this attitude, and the following commentary from the UNESCO Guidelines also demonstrates the spirit of these organisations very clearly:

The aim of this booklet is not to abolish certain words or to alter historically established texts; nor is it suggested that these guidelines be followed to the letter. For the sake of equality, however, writers are asked in every case to pause and consider the alternatives.

(*Guidelines* ³1999:6)

As for the two style books of newspaper language, the frequencies are divided. While the suggestions to avoid and to substitute expressions appear to be balanced in *The Times*, *The Guardian* instead favours giving, on average, two alternatives for one item. Nevertheless, the recommendations given in both style guides are presented in a straightforward manner, and they cover the whole range of categories of discriminatory usage. It is noticeable that newspapers seek not only

to reflect changes in society (similarly to the style guides of academia) but also to incorporate certain values, such as the “newspaper’s values” (*The Guardian*) and “traditional virtues of ‘style’” (*The Times*). *The Times Style and Usage Guide* even raises the question of correctness:

This upgraded and expanded guide aims to be a ready reference source for any writer or reader, professional or casual, who cares enough about the language to use it correctly. (Austin 2003: i)

The values conveyed are in line with the concerns of the readership to whom ‘correct’ usage in newspapers is important, as demonstrated by the many letters to the editor about allegedly misused words or incorrect grammatical constructions. This attitude evidences the editors’ concern about their readers’ opinions. Since the readers are customers, they are inclined to satisfy their demands. *The Times* even encourages feedback “from readers who are convinced that we have failed to cull the archaic or that we have tried too hard to be ‘stylish’” (*The Times*: i). What also becomes clear is the fact that the style guide is directed not only to trained journalists but also to the non-specialist. The purpose is to reach a wide readership, and thus the guidelines do not aim at a truly ‘specialist’ use of language. This wide target audience may also explain the fact that the newspaper guidelines are the only ones among the style guides examined that cover the whole range of politically correct items, dealing with gender, disability, diseases, mental health, ethnicity, sexuality and age. They thus serve as a more general guide for usage, and can be compared to the common usage guides, dictionaries and grammar books on the market – which is a clear sign that newspaper agencies have a role in the codification of language.

4.2.2 *Qualitative differences*

Though all style guides suggest using non-discriminatory language and avoiding giving offence, the items listed vary considerably. Most often the focus lies on gender-neutral usage, though two institutions list hardly any gender terms while including expressions dealing with disability, ethnicity and age. In three cases, general recommendations are given at the expense of actual examples, and in one case general advice is totally absent.

The cross-section of items is largest in the four style guides of academic writing. No single items occur in more than one of the guides. All of them deal with gender-related terms, but only two include other categories. Concerning gender, it is generally advised to use generic terms that include women, for instance *chair*, *police officer*, or *humanity*. However, the Royal College of Surgeons of England differs from the others in proposing the use of *chairman* and *chairwoman* and advising against *chair* and *chairperson*. It also recommends avoiding *female* and

male and using *women* and *men* instead, which probably means making use of *woman doctor* rather than *female doctor* for example. When one considers the other categories, both the University of Kent and Sheffield Hallam University support the strategy of replacing nouns with the corresponding adjective and noun, for instance *deaf people* instead of *the deaf*, or *an Asian woman* instead of *an Asian*. Nevertheless, the choice of items appears to be quite eclectic and suggested alternative expressions are somewhat arbitrary, for instance *mixed race* for *half caste* (Sheffield) or *spouse* for *partner* (Kent).

Of the four guidelines produced by political institutions, three consider only gender terms, which can partly be explained by the aims of the respective organisation. As an educational, scientific and cultural organisation, UNESCO explicitly seeks to avoid discrimination in all fields, notably in the area of education. Since girls and women still suffer severe discrimination because of their gender, especially in developing countries and the Islamic world, UNESCO "seeks to transform behaviour and attitudes that legitimise and perpetuate the moral and social exclusion of women". On the basis of the *Universal Declaration of Human Rights* (1948) and the *Convention on the Elimination of all Forms of Discrimination against Women* (1979), UNESCO deems it necessary to change language behaviour, in that it shapes our thinking and creates certain unwanted images. The declared concerns of UNESCO make the comparatively large number of recommended items (259) readily understandable.

The Daily Style Guide from the United Nations Economic Commission for Europe shows some overlap with UNESCO's recommendations, but there are two important differences to be noted. Firstly, with only 38 items, the list of terms in the UNECE is much shorter, which shows that their concern about gender equality is less pressing than UNESCO's and this can also be associated with the economic orientation of the UNECE. Secondly, several alternative suggestions for the expressions *man-hours*, *executive* and *manager* are provided by the UNECE, which again can be explained by the economic focus of the organisation. The UNECE's list of items is rather disappointing, since it gives only general recommendations. The entries of the Queensland Government seem to have been chosen very eclectically; the main concern is apparently to "focus on the person rather on the distinguishing feature", which explains the somewhat idiosyncratic advice to use *academic* instead of *bespectacled academic*, *carpenter* instead of *muscular carpenter* and *typist* instead of *paraplegic typist*.

Both newspaper guides stand out due to their diverse and thorough lists of entries. In comparison to *The Times' Style and Usage Guide*, *The Guardian's Style Guide* is more liberal, more elaborate and more entertaining. *The Times* allows *blacks* as a noun as well as *actress*, while *The Guardian* only accepts *black* as an adjective and suggests *actor* as opposed to *woman actor*. *The Times* prefers *chair-*

man and *chairwoman* over *chair* and *chairperson*, but the *Guardian* favours *chair* and advises against *chairman* and *chairwoman*. While *The Times* suggests *woman* in preference to *girl*, the *Guardian* is more precise in stating that *girl* is only acceptable for persons under 18, and thus suggests a semantic narrowing of the term. Whereas *The Times* lists only *Inuit* (plural), the *Guardian* also mentions *Inuk* (an individual). It is noteworthy that the *Guardian* includes some cartoons and also humorously advises against the use of the term *political correctness* “on the grounds that it is, in Polly Toynbee’s words, ‘an empty rightwing smear designed only to elevate its user’”.

As a provisional result, it can be stated that the relatively low number of politically correct recommendations given in the style guides indicates a rather low degree of institutionalisation of these terms and phrases. Institutionalisation and codification seem to be detrimental to creativity in all its various manifestations. In the following I will explore whether and in what way striving for ‘correctness’ can be seen as detrimental to lexical innovation.

5. Creative word-formation patterns

In general, strategies proposed to replace allegedly biased expressions with more neutral terms involve the morphological, syntactic and semantic levels of language. Some suggested alternatives in English are, for instance, the use of the pronouns *he* and *she* – or *they* and *we* – to replace the generic *he* for both men and women, or the phrase *person of mixed ancestry* instead of *half-breed*; other recommendations are the use of impersonal expressions as in “*the search for knowledge has led to...*” instead of *man’s search for knowledge has led him to...*, or the substitution of *firefighter* for *fireman* or the expression *woman* instead of *lady*. In the following discussion I will focus primarily on the morphology and semantics of the suggested lexical items.

It is difficult to make generalisations about the creative patterns of the proposed non-discriminatory expressions in the style guides examined, since the recommendations vary considerably and concern different domains of use, often affecting only individual items. Nevertheless, some emerging patterns can be observed. The analysis of PC expressions in a selection of style guides for public discourse suggests that three main patterns prevail, those of *unification*, *specification*, and *euphemism*.

5.1 Unification

Unification is largely restricted to non-sexist items and consists in the suppression of gender specifications. It is employed to establish linguistic equality by discarding asymmetrical constructions. Unification is achieved either by truncation (i.e. leaving out the verbal constituent that marks gender), using an already existing gender-neutral synonym or creating a new expression.

a) *Truncation*. This strategy involves the reduction of a composite word (or phrase) to its base (or head), and it is applied to certain occupational and other human agent nouns that mark the female sex either by means of suffixes such as *-ess*, by compounds containing *woman*, or by phrases with *female*. For instance, *poetess* should be substituted by *poet*, and *woman doctor* or *female doctor* to *doctor*. The marked female forms are viewed as discriminatory, since they imply that female is not the 'norm' in the respective contexts, and they also have a trivialising effect (Pauwels 1998: 110) in that they contribute to a view of linguistic dependency of the female on the male. The uses of *doctor*, *poet* and the like for both male and female proponents can be considered a special case of semantic extension.

The word *stewardess* represents an exception to this pattern, as it was not reduced to *steward* as a generic term. Instead, both the gender-marked terms were subsumed under *flight attendant* or *cabin crew*.

In rare cases, truncation is also used in contexts other than gender stereotypes. For instance, the style guide of the Queensland Government, Australia, recommends *doctor* in place of *Asian doctor* and *given name* or *surname* for *Christian name*.

b) *Synonymy*. Synonyms or synonymous phrases are suitable as alternative referring expressions. Nevertheless, synonyms may differ with respect to collocations, stylistic level (including register), connotations, and context. Their use may also be restricted with respect to region, topic or medium. Sometimes the recommended alternative denotes only one of the meanings of a polysemous word. For instance, while *mankind* and *humanity* are common nouns for the human species, *humanity* also refers more specifically to *humankind*, as opposed to animals, gods or machines, and to the quality of being kind. The revival of words no longer in use also falls into this category.

In the style guides examined, the strategy of synonymy is mainly applied to expressions with the morpheme *man* (and, as a consequence, to a few newer, parallel formations with *woman*). These formations portray *male* as the norm since women are considered to be subsumed under 'man' and are thus underrated. In order to

avoid this discrimination, *policeman* is replaced by *police officer*, *mankind* by *humanity*, *humankind* or *human race*,² and *businessman* by *manager* or *executive*.

c) *Neologisms*. Apart from discarding gender-specific nouns by either truncation or synonymy, a new word may be coined. Firstly, the two morphemes *man* and *woman* can be substituted by the suffix-like constituents *person* or *people* (*businessperson*, *business people* instead of *businessman*). Secondly, *man* and *woman* can be left out altogether, resulting in a new expression, as in *chair* (for *chairman* or also *chairwoman*) or *anchor* (American English for *anchorman*). Thirdly, they can be replaced by other nouns, such as *attendant* (as in *flight attendant*) or *fighter* (as in *fire fighter*). Finally, the whole expression may be discarded and a new, neutral expression introduced. The only example found in the style guides is the honorific title *Ms*, which replaces *Mrs* and *Miss* in order to avoid reference to the marital status of women, thereby eliminating the asymmetry between *Mrs* and *Miss* on the one hand and *Mr* on the other.

Two exceptions to the general pattern of *unification* seem to be the pairs *chairman* / *chairwoman* and *businessman* / *businesswoman*, on which the style guides express differing opinions. The more conservative guidelines prefer gender-specific expressions over *chair* or *chairperson* and *businessperson*. The majority of the guides, however, adhere to the usual pattern of *unification*.

In concluding this section, it appears that the main intention of the language planners is the *avoidance* of sexist expressions, replacing them with already existing forms or phrases. Apart from the establishment of *(-)person* and *(-)people* as new suffixes and a handful of newly institutionalised words such as *firefighter*, the recommendations do not seem very innovative – quite the contrary: it appears that few changes are considered necessary, and these should be as inconspicuous as possible. Thus lexical creativity can only be found in a few individual cases.

The *unification* strategy is also supported by the morphological structures of the English language. English lacks grammatical gender and a common gender-suffixing practice, so that the unmarked gender-neutral form suggests itself (the suffix *-ess* is not applied regularly). In contrast, for instance, the German language, being highly inflected, distinguishes grammatical gender and makes regular use of the suffix *-in* denoting women. As a consequence, gender-specificity is the preferred strategy in German (cf. Fischer 2004).

2. Since *man* is not a morpheme in words like *human* and *humanity*, it does not need to be replaced.

5.2 Specification

This strategy aims at a more detailed description of a person's characteristics regarding ethnicity, age, gender, and other traits. It often goes hand in hand with the avoidance of pejorative terms, such as *Eskimo* and *Amerindian* or *Red Indian*. Some style guides suggest that the writer should indicate the name of the ethnic group or community to which a person belongs. The use of a clear and precise formulation is generally recommended, and specific examples or alternative expressions are only occasionally provided. For instance, the term *girl* should not be used for women over 18, the *elderly* should be at least 65 years old, which are cases of semantic narrowing. The abbreviation OAP (for *old age pensioner*) should not be used since it mingles age and economic status. The expressions in this group are comparatively few, and there are no new coinages recommended.

Another type of specification can be called *attribution*. This strategy of non-discriminatory usage mainly concerns disability, diseases and mental health. It recommends the use of certain words only as adjectives and not as nouns, because it is considered offensive to define people by their physical characteristics or condition of health. Among the adjectives that should not be used as nouns are: *disabled*, *dyslexic*, *diabetic*, *deaf* and *epileptic*. A possible option is to use the paraphrases *a person with XY* or *a person has XY*, as in the expressions *a person with dyslexia* or *a person has dyslexia*. The corresponding adjectives may of course be used as in: *She is dyslexic*. Dispreferred terms from other areas of politically correct usage are *male / female*, *black*, *heterosexual / bisexual / homosexual / gay*. Exceptions are the adjectives / nouns *lesbian* and *elderly*. Again, the strategy is one of avoidance, and the reader is encouraged to make use of alternative established morphological and phrasal structures. No new coinages are listed or even recommended.

5.3 Euphemism

A number of words are considered pejorative, since they stereotype or stigmatise certain groups of people on the basis of their distinguishing features. These words are re-evaluated and replaced by an expression that is assumed not to be derogatory and is therefore more socially acceptable, i.e. by a euphemism. Listed expressions and their suggested alternatives are *Eskimo(s)* > *Inuit / Inuk* (see above), *half caste* > *mixed race*, *non-whites* > *Blacks* or *African-Americans*, *lady* > *woman*, *old people* > *elderly people* and *invalid* or *handicapped* > *disabled*.³ It is recommended that some words be avoided, but no alternative expressions or phrases are pro-

3. The symbol > is here used for 'to be replaced by'.

vided – see for instance *crippled*, *spastic*, *loony*, *psycho*, *schizo*, and *immigrant* in *The Guardian Style Guide*.

In synthesis, it may be observed that amendments found in style guides for public written language aim first and foremost to achieve equal treatment of people by avoiding offensive or stigmatising expressions and by employing alternative neutral words or phrases that are already established in the language. In a few cases, and when a suitable alternative is apparently not available, a new expression may need to be introduced. A new formation can serve as a model for others and may trigger analogous coinings, as in the case of the composites with *people* and *person*, as well as ethnic compounds, for instance *Native American* and *African-Caribbean*, based on the formation *African-American*.⁴ At first sight, it may seem surprising that so few neologisms are created by comparison with the numerous synonymous expressions, but it must be remembered that personal creativity is only acceptable within a certain social and cultural environment. Since the alternative synonyms are already available (and used), it is more likely that they will be adopted in preference to totally new coinages.

6. Conclusions

The recommendations of the style guides will now be related to the three aspects of language involved: prescriptivism or language planning, discourse type and creative word-formation patterning.

Concerning the implementation of non-discriminatory language, lexical creativity is clearly lacking. The guidelines for writers of academia and of national or international organisations – both of which can be placed at an intermediate level of specialisation – are very careful about voicing any opinion. The only exception is UNESCO's *Guidelines on Gender Neutral Language*. The stylebooks for the writing of research papers are the most tolerant and cautious and mainly leave the decision up to the writer. This strategy is quite remarkable if one considers that it was the universities in the late 80s and early 90s that promoted political correctness in language. Their *laissez-faire* approach today can only be explained by the reactions following the introduction of speech codes, including law suits on the basis of the freedom of speech as defined in the American constitution. The advice given by official institutions such as the EC or the UNECE seems to be the most conservative. Finally, the newspaper manuals, which cannot be considered specialised discourse in a strict sense, appear disguised as general usage guides of

4. This well known expression no longer seems to be the predominant expression (at least in British English), instead *Black* has had a revival.

the English language and thus stand in the prescriptive tradition of the 18th and 19th centuries. This is especially true for *The Times*, and to some extent also for the *Guardian*.

Regarding discourse type and critical creativity, it can be said that non-discriminatory lexical items mainly find their way into specialised discourse when they are thematically motivated, and thus become part of the respective jargon. For instance, new occupational titles have become necessary following the creation of new jobs, and since women have now entered many of the professions, the titles formerly reserved for men – and containing the morpheme *man* – had to be supplemented or generalised (thus the generalisation strategy became the favoured procedure for mainly linguistic reasons, as explained above). The specialised discourse of official texts thus had to be extended to include new developments in the economic and political arena and consequently new words as well. The main examples are formations with *-people* and *-person* for obvious reasons. It seems likely that the activism of feminists promoted or perhaps even triggered these changes, but without the need for new lexemes within specialised discourse they would have been of little use. Correspondingly, non-biased language use with respect to ethnicity, health or age does not play a part, since it is generally not even mentioned. We can therefore record the fact that language planning in specialised discourse can only be put into effect if it meets the needs for new technical vocabulary, otherwise it will be largely ignored.

Nevertheless, exceptions are possible if the suggestions are backed up by a strong approval of such change on the part of the authorities in charge and with a high motivation to take linguistic action. A case in point is the UNESCO style guide, but again only gender terms are affected. There is also considerable space for individual preferences, as can be seen in the guidelines of the University of Kent, Sheffield Hallam University and the Queensland Government. But we can assume that their rather idiosyncratic suggestions will not have far-reaching or lasting effects.

Newspaper discourse cannot be considered as truly specialised (see above). The wide audience aimed at in the style books of *The Times* and the *Guardian* is not seen to be in need of specialised jargon but, rather, of general advice. These books tell the writer or the interested reader in a simplistic or even amusing way what is wrong or correct, and what is preferred or dispreferred. Of the guidelines examined, only these two covered all areas of bias-free usage and did so quite extensively. Since the media have contributed a great deal to the introduction and the spread of politically correct terms, they are now nourishing the child they have brought into the world. Apart from a number of thematically motivated items (i.e. mainly job titles), the preoccupation with non-biased language usage seems to be created largely by domestic factors (e.g. house styles, regionalisms,

etc.) and maintains the moralising prescriptive tradition that was earlier thought to have been discarded.

As far as creative patterning is concerned, it does not seem that specific style guides favor particular types of word-formation patterns. Therefore, the type of word-formation involved does not play an important role in the recommendation of politically correct terms. Nevertheless, it is evident that some alternatives stand the test of time more than others. All guidelines agree that *man* as a generic term is out and should be substituted by a neutral term, and an already existing expression is to be preferred over other possible alternatives. As regards naming and occupational titles, the gender-neutral term is preferred to the specific male or female item – though a certain rivalry does exist between these two possibilities, as the recommendations of *The Times* and the Royal College of Surgeons reveal.

The gender-neutralisation strategy is also supported by the morphological structures of the English language. There is a certain reluctance to use adjectives either as modifiers or as nouns, which is also in line with the priority of unification over specificity.

Apart from a few neologisms and some cases of semantic extension and semantic narrowing, the suggestions given do not seem to be very creative, since they are mostly recastings of existing expressions. It can thus be concluded that language purism and the policy of censorship have a negative effect on lexical, or even critical, creativity. Language seems to offer many choices, so that certain expressions can be avoided without any great loss. The existing alternatives then absorb the meanings and replace the censored expressions. Even new coinages are based on already existing ones, such as *Ms.* (in form similar to *Mrs.* and *Miss*, and perhaps a blend) and composites ending in *-people* or *-person* (analogous to formations ending in *-man* or *-woman*).

The strategy of falling back on proven means can also be considered a skillful method of achieving language implementation: familiar terms may not be rejected as easily as ‘true’ neologisms, especially with respect to the controversial issue of political correctness. In the history of nations, language prescriptivism has always been a laborious undertaking, especially if initiated by grassroots language planners. Even when language changes are imposed by governments or official institutions, the negative attitude of the community might bring a halt to their usage. For an implementation of non-discriminatory language change it thus seems reasonable to rely on structures and meanings that are already established in the language. This approach also corresponds to the general opinion in sociolinguistics that language change usually occurs when two alternatives have existed for some time side by side and one of them eventually predominates.

Finally, it must be noted that language purism often has the effect of turning into a kind of creativity that ridicules censorship. Some manuals like the already

cited handbook by Beard & Cerf take the idea of non-biased usage to extreme limits, thus deriding the whole movement of political correctness. But their playful ad hoc formations are neither a case of critical creativity nor of institutionalisation and will thus be left to future research.

References

- Austin, T. 2003. *The Times Style and Usage Guide*. London: HarperCollins.
- Beard, H. & Cerf, C. 1992 (updated 1994). *The Official Politically Correct Dictionary and Handbook*. New York: Villard Books / Random House.
- Carter, R. 2004. *Language and Creativity. The Art of Common Talk*. London: Routledge.
- College Style Guide. 2004. The Royal College of Surgeons of England, London. Retrieved June 1, 2005, from http://www.rcseng.ac.uk/services/publications/annals/style_guide.html.
- Daooust, D. 1997. "Language planning and language reform." In *The Handbook of Sociolinguistics*, F. Coulmas (ed.), 436–452. Oxford: Blackwell.
- The Daily Style Guide*. United Nations Economic Commission for Europe (UNECE). Developed by Communications Division. 1996. Retrieved June 1, 2005, from http://www.unece.org/stats/documents/media/style_manuals/1.e.html.
- Editorial Style Guide*. Sheffield Hallam University, Marketing and Development. Retrieved June 1, 2005, from <http://www.shu.ac.uk/services/marketing/styleguide>.
- English Style Guide. A Handbook for Authors and Translators in the European Commission*, 2005 [1982]. European Commission Directorate-General for Translation. Retrieved June 1, 2005, from http://europa.eu.int/comm/translation/writing/style_guides/english/style_guide_en.pdf.
- Fischer, R. 2002. "Die Etikette der Gleichheit. *Political Correctness* im englischen Sprachgebrauch." *Blick in die Wissenschaft. Forschungsmagazin der Universität Regensburg*. No. 14: 56–62.
- Fischer, R. 2004. "Coach-Frau, Frau Coach oder Coacherin? Wie Sprachstruktur geschlechtergerechten Sprachgebrauch beeinflusst." In *Adam, Eva und die Sprache. Beiträge zur Geschlechterforschung*, K. M. Eichhoff-Cyrus (ed.), 176–190. Mannheim: Dudenverlag.
- Fischer, R. 2005. Rev. of *Political Correctness in den USA und in Deutschland*, by S. Wierlemann 2002, Berlin: Erich Schmidt Verlag. *Amerikastudien / American Studies* 49.4: 611–612.
- Gibaldi, J. 2003. *MLA Handbook for Writers of Research Papers*. New York: The Modern Language Association.
- Gotti, M. 2003. *Specialized Discourse. Linguistic Features and Changing Conventions*. Bern: Peter Lang.
- The Guardian Style Guide*. 2004. Guardian Newspapers Ltd. Retrieved June 1, 2005, from <http://www.guardian.co.uk/styleguide>.
- Guidelines on Gender Neutral Language*. 1999 [1987] UNESCO. Retrieved June 1, 2005, from <http://unesdoc.unesco.org/images/0011/001149/114950mo.pdf>.
- Haspelmath, M. 2002. *Understanding Morphology*. London: Arnold.
- Hoffmann, L. 1984. "Seven roads to LSP." *Special Language – Fachsprache* 6 (1–2): 28–38.

- Kent Publications Style Guide*. 2004. University of Kent, Communications and Development Office. Retrieved June 1, 2005, from <http://www.kent.ac.uk/cdo/publications/corporate/style-guide/>.
- Maggio, R. 1991. *The Bias-free Word Finder. A Dictionary of Non-discriminatory Language*. Boston: Beacon Press.
- Pauwels, A. 1991. *Non-discriminatory Language*. Canberra: Australian Government Publishing Service.
- Pauwels, A. 1998. *Women Changing Language*. London: Longman.
- Style Guide*. Government of Queensland: EQ Publishing. Department of Education and the Arts. Retrieved June 1, 2005, from <http://education.qld.gov.au/publication/production/print/style.html>.
- Wierlemann, S. 2002. *Political Correctness in den USA und in Deutschland*. Berlin: Erich Schmidt.

Name index

A

Abrahamson, A. 191
Adams, D. 22, 23, 27
Adams, V. 116, 121, 124, 140,
147, 152, 167
Aitchison, J. 214, 228
Alexander, R. 95, 101, 102, 241
Alexieva, B. 240
Algeo, J. 116, 146, 151, 152, 153
American Speech 123, 131
Amis, K. 20
Aronoff, M. 16, 63, 96, 147
Asimov, I. 169, 175, 183
Attardo, S. 104, 191
Austin, J. 7
Austin, T. 272

B

Baayen, H. 42, 43, 50, 61, 63,
64, 65
Bally, C. 250
Baron, J. 191
Barsalou, L. 189
Bauer, L. 3, 16, 17, 18, 27, 35,
124, 152, 153, 165, 166, 168,
176, 180
Bauman, R. 110
Beard, H. 267, 281
Beaugrande, R. de 4
Belmore, N. 138
Benczes, R. xiv
Benveniste, E. 28
Bergen, B. 55
Biber, D. 4
Boas, H. 168
Botha, R. 165
Brandsford, J. 222
Brekke, H. 28

Brown, G. 4
Burger, H. 93
Bybee, J. 167

C

Carroll, L. 118
Carpuat, M. 190
Carter, R. xiv, 264
Cerf, C. 267, 281
Čermak, F. 250
Chesterman, A. 239
Chiaro, D. 95, 99
Chomsky, N. 94, 96, 209
Clark, E. 3
Clark, H. 3
Coffey, S. 250
Collot, M. 138
Cowie, A. 102
Cruse, D. 138, 154
Crystal, D. 9, 23, 33, 34, 35, 138,
139, 141, 143, 144, 147, 150,
154, 155
Cuenca, M. 217

D

Dahl, R. 163, 175–177, 178, 184
Daoust, D. 266
de Bono, E. 190
De Brabanter, P. 28, 31, 33
Dederding, H. 17
Dick, P. 163, 171–173, 175, 178
Di Sciullo, A. 16, 20, 165
Dirven, R. 215, 217, 220–221,
224, 226
Dong, Q. 190
Dong, Z. 190
Downing, P. 3, 19
Dressler, W. 4, 32, 167, 180

E

Esser, J. 4
Everaert, M. 94

F

Faulseit, D. 243
Fill, A. 241
Fillmore, C. 96
Fischer, R. 7, 167, 169, 170,
263–282
Forster, K. 128
Fowles, J. 244–245
Fraser, B. 95
Freud, S. 191
Fromont, R. 39, 44

G

Geerearts, D. 215, 216, 217, 224
Ghomeshi, J. 25, 26
Giegerich, H. 34
Gläser, R. 93, 96, 250
Gordon, E. 40, 44
Gotti, M. 268
Gramley, S. 149
Gregory, M. 44
Gries, S. 119, 127
Guha, R. 190
Guttenplan, D. 123
Guy, G. 41

H

Halliday, M. 4
Hamilton, P. 175, 183
Haspelmath, M. 265
Hay, J. 7, 39–57
Hayes, J. 190, 197
Herbert, F. 175, 183
Heringer, H. J. 21

- Hermeren, L. 250
Herring, S. 138, 144
Hickman, T. 218
Hilferty, J. 217
Hoffmann, L. 264
Hohenhaus, P. 5, 6, 7, 15–38, 143, 166, 167, 168
Holmes, J. 145
Hornby, N. 21, 24, 26, 27, 30, 33
Hutton, J. 200, 202
Hymes, D. 245
- J
Jackendoff, R. 25, 31
Johnson, K. 222
Jurafsky, D. 44, 47
- K
Kastovsky, D. 3, 17, 19, 22, 164, 170
Kay, P. 96
Kehoe, A. 62
Kelly, M. 116, 119
Keune, K. 52
Koopman, H. 96
Kubozono, H. 116, 119
Kuiper 7, 8, 93–112
- L
Labov, W. 41, 216
Lakoff, G. 216, 217
Lamb, S. xiv
Lambert, J. 244
Langacker, R. 165, 216, 219, 220
Lavoie, L. 47
Lehrer, A. 6, 8, 115–134, 179
Lenat, D. 190
Lennon, J. 94
Lipka, L. xv, 3–12, 17, 19, 22, 23, 36, 164, 166, 167, 168, 170, 179
Locmele, G. 248
Lodge, D. 32
López Rúa, P. 6, 8, 119, 137–159
- M
Mackin, R. 97
Maggio, R. 267
Marchand, H. 3, 5, 9, 23, 124, 170
- Melčuk, I. 95, 97, 98
Mervis 217
Miller, G. 190
Milne, A. A. 177, 178, 180, 183, 244
Moon, R. 93, 250
Morley, B. 62
Morris, D. 123
Motsch, W. 17
Munat, J. 9, 21, 28, 163–185, 246
- N
Naciscione, A. 93, 95, 240, 252
Nicolas, T. 97
Nunberg, G. 97
- O
O'Connor, M. 96
O'Donoghue, D. 210
Okrand, M. 23
Ortner, L. 22
Orwell, G. 124
- P
Pacey, M. 63
Paolillo, J. 93
Pauwels, A. 267, 275
Pawley, A. 100
Petrotta, A. 241
Plag, I. 52, 116, 119, 125, 129, 152, 166, 180
Pope, R. xiv
Porto Requejo, D. 9, 113–126
Pullum, G. 167
Pustejovsky, J. 189
- Q
Quirk, R. 140, 149
- R
Radford, A. 16
Raskin, V. 22
Redfern, W. 241
Renouf, A. 9, 61–89
Riffaterre, M. 240
Roberts, A. 172, 173
Rosch, E. 216, 217
Rosemann, P. 29
Rumelhart, D. 191
- S
Saffran, J. 43
Safire, W. 123
Saukkonen, P. 20
Scalise, S. 32
Schmid, H.-J. 20, 22, 217, 226, 227
Schreuder, R. 42
Schultink, H. 165
Searle, J. 4
Seco, N. 190, 197
Semino, E. 167, 215
Seppänen 19
Serra, R. xiii
Shi, R. 47–48
Simpson, M. 210
Sinclair, J. 68
Smith, M. 144
Soehn, J.-P. 96
Spiegel, E. 4
Steele, H. 123, 131
Stein, G. 124
Štekauer, P. 17, 30, 169, 170, 175, 180, 217
Stockwell, P. 167, 169, 170–171, 173, 176, 177, 214, 215, 217, 222, 234
Strang, B. 4
Syder, R. 100
- T
Taylor, J. 216, 217
Tournier, J. 3
Toury, G. 245
Toynbee, P. 274
Trask, R. 191, 192, 193
Tuggy, D. 16, 221
- U
Ungerer, F. 217, 226, 227
- V
Vallès, T. 228
Van Lancker, D. 93
Veale, T. 6, 10, 189–212
Veisbergs, A. 6, 10, 239–261
Venuti, L. 246
Veres, C. 127
Verstraten, L. 96

Vespoor, M. 215, 217, 220–221,
224, 226
Vonnegut, K. 244

W

Warren, B. 124
Way, E. 189
Weinreich, U. 98
Weis, M. 218

Welte, W. 15, 29
Werry, C. 144, 156
Wierlemann, S. 266
Wightman, C. 46
Wilcox, C. 8, 93, 102–103,
109, 110
Williams, E. 16, 20, 96, 165
Williams, J. 220, 221
Wladowa, E. 17

Wood, A. 144
Wray, A. 93

Y

Yule, G. 4

Z

Zauberga, I. 246
Zwicky, A. 167

Subject index

- A
- abbreviations x, 138, 139, 140, 141, 143, 149, 150
 - About a Boy* 21, 24, 26, 27, 30
 - accessibility condition 95
 - acronym / acronymy xiv, 3, 5, 6, 8, 64, 69, 80, 150–151, 152, 173, 179, 241
 - definition of 141
 - ad hoc (*see also* categories)
 - concepts 192, 205
 - formations (*see also* nonce forms) 18, 132, 169, 281
 - patterns 180
 - adaptations 3, 240, 249
 - advertising 8, 23, 179, 243, 248–250, 260
 - affixation (*see also* prefixes, suffixes) 6, 41, 43, 45, 64, 147–149, 152, 157
 - patterns of 227
 - affordances 195, 197
 - allusion 5, 6, 7, 9, 10, 71, 74, 85, 88, 115, 204, 241, 248
 - alphabetism 141 (definition), 150, 151
 - ambiguity 240, 243
 - of boundaries 64
 - lexical 27
 - multiple 110
 - referential 180
 - structural 101, 109, 111
 - analogy
 - analogical retrieval 204–209
 - by-products of 200–204
 - in compound creation 192, 254
 - in derivation 193
 - direct 27–28
 - explicit 192, 200, 208, 209, 210
 - false 194
 - implicit 194, 198, 199, 200, 209, 210
 - in inflection 192
 - (role in) interpretation 228
 - lexical 10, 189, 192, 197
 - phonetic 194
 - proportional 195, 205
 - semantic 176
 - (role in) term creation 194–200
 - (role in) WF 27, 167, 189, 192–194, 211, 228
 - anaphoric reference 19, 20
 - Aristotelian
 - account 202
 - category 216
 - attention seeking devices (ASDs) xv, 5, 6, 7, 8, 10, 23–24, 166, 178
- B
- base form 95
 - The BFG 163, 175–178, 179
 - blends (*see also* clippings, splinters) 5, 26, 115–132
 - compound blends 253
 - copulative 152
 - domains of occurrence 128–131
 - orthography of 120
 - partial 255
 - phonology of 117, 119–120
 - processing of 126–128
 - structure of 117–119
- C
- borrowing 155, 173, 242, 245, 246
 - bound
 - base 124
 - forms 30
 - morpheme 121, 210
 - word 96, 100
 - calques 109 (*see also* paraphrase)
 - capitalisation 259
 - cartoon art xiv, 6, 8, 56, 93–111
 - cataphoric reference 20
 - categories
 - ad hoc 189, 190, 191, 192, 200–204, 205, 207
 - Aristotelian 216
 - compound 208
 - conceptual 192, 203
 - conventional 218, 230, 231
 - creative 224
 - lexical 9, 10, 217, 225, 234
 - onomasiological 175, 224
 - ontological (*see also* ontologies) 190
 - prototypical 9, 215, 216, 217, 219, 234
 - radial 217
 - semantic 172, 173, 219
 - semasiological 217, 234
 - superordinate 203, 206
 - syntactic 101
 - categorisation (*see also* recategorisation)
 - analogical 192
 - creative 9, 217, 218, 231, 235

- lexical 9, 213–235
- of new words 216, 227
- onomasiological 224–234
- ontological 189, 210
- primitive 216, 218
- semasiological 219, 235
- of texts 4
- CELEX lexical database 50, 53
- censorship (*see also* Soviet Russia) 243–244, 263, 280
- channels
 - of communication 137
 - electronic 139
- children's literature (*see* genres)
- clipping (*see also* blends, splinters) 116–175, 128, 131, 153–154
 - conventional 141
 - embedded 118
 - fore clipping 152, 153
 - hind clipping 152, 153
 - mid clipping 151, 152
 - as stylistic feature 175
- clues xiii, 20, 33, 126, 218, 220, 224
- code (*see also* subcode)
 - electronic 138
 - restricted 139
 - shared 35
- code-switching 121
- cognitive 22, 55, 170, 191, 215, 229, 230
- cohesion 4, 6, 252, 259
- coining / coinages, (*see also* neologisms, nonce forms) ix, 7, 16, 18, 39, 40, 42, 48, 53, 54, 61, 83, 87, 117, 124, 128, 146, 151, 165, 167, 168, 169, 175, 277
- collocate / collocation 9, 63, 74, 78, 79, 82, 83, 97, 100, 180, 190, 201, 277
- collocational
 - frameworks 68
 - partners 82
 - patterns 78
 - profile 63, 74, 83
- combining forms (*see also* neo-classical) 64, 117, 124, 140, 154, 173–174
 - semantics of 125–126
- communication (*see also* cross-cultural, electronic communication)
 - CMC (computer-mediated) xiv, 33, 144
 - in-group 143
 - SMS (short message service) (*see* text messaging)
- competence (*see also* langue) 163, 164, 170
- complex words (*see* words)
- compositional / compositionality (*see also* decompositional) xi, 31, 40, 41, 42, 55, 97, 98, 168, 174, 194, 201
- compounds 63, 64, 68, 82–85, 115, 116, 117, 124, 127, 142, 148, 154, 156, 173, 174, 175, 177, 178, 179, 196, 197, 198, 202, 208, 209, 210, 226, 227, 241, 249, 253, 256, 259, 275, 278
 - abbreviated 125, 152
 - clipped 141, 152, 153, 154, 155
 - coordinate or copulative 119
 - deictic 21
 - dummy 19, 20
 - episodic 21
 - fused 174
 - hybrid 202
 - identical constituent (ICCs) 7, 25, 26, 27
 - letter 153
 - nominal xi, 154, 189, 240
 - phrasal 31, 33, 258
 - syntactic or exocentric 119
- computer (*see also* electronic communication)
 - based ontology xv
 - jargon 145, 156
 - speak 8, 124
- concept creation xiv, xv, 22, 34, 195, 209
- conditions 28
 - accessibility 95
 - conventional 98, 108
 - recoverability 96, 101, 104
 - restricted 98
 - sociopolitical 265
 - truth 97
 - usage 98
- consonant clusters 51, 175
- constituent(s)
 - adjunct 97
 - initialised 154
 - integration 152
 - lexicalised 96, 99, 100, 101, 102
 - optional 96, 97, 100
 - root 170
- constraints
 - phonological 177
 - phonotactic 139
- contact language 242, 246
- context dependency 30, 35, 169
- context/contextual
 - clues (*see also* clues) xiii, 20
 - environment 25
 - information/knowledge 26, 28, 108
- contrastive reduplication (CR) – (*see* identical constituent compounds)
- conventions 239, 244, 245, 246, 247
- conversion 4, 32, 64, 137, 140, 174, 242
 - delocutive 28, 29, 30
- corpus/corpora 7, 8, 9, 25, 39, 44, 60, 61, 62, 67, 87, 102, 103, 117, 119, 137, 140, 163, 192, 195, 197
- co-text xiii, xv, 4, 28, 29, 30, 168
- creativity
 - critical 7, 263, 264, 279, 280
 - definitions of x, 70, 165

- diachronic 62, 63
 individual 164, 264, 279
 rule-changing 164, 165
 rules of 74
 as social practice 264–265
 societal 164, 264
 v.s. productivity (*see* productivity)
 credibility (*see also* verisimilitude) 216, 218, 228
 cross-cultural
 communication 239, 241
 Cyc ontology 190–191
- D**
 decoding 142, 143, 241, 256
 decomposable/decompositional
 (*see also* compositional) 7,
 39, 40, 41, 42, 43, 44, 47, 48,
 52, 54, 55, 170
 default term 230
 deformation
 artistic or creative 8, 93,
 95, 96, 98–102, 108, 109,
 110, 111
 phonological 99, 103–104,
 178
 semantic 101–102, 107, 178
 structural 99–101, 102,
 104, 107
 deictic/deixis 3, 4, 15, 19, 20,
 21, 26
 democratisation 242, 247
 derivation (*see also*
 affixation) 4, 16, 30, 62, 63,
 64, 66, 76, 77, 82, 137, 147, 166,
 172, 175, 177, 178, 193, 226, 231
 deviance 18, 23, 30, 168
 diachronic xv, 7, 8, 17, 61, 62,
 63, 86, 88, 120, 168
 dictionary (*see also*
 knowledge) 191, 211
 diminutives 240, 248
 disambiguation xi, 9, 26, 17,
 175, 180
 discourse
 specialised 264, 268, 278,
 279
- types 263, 268, 269, 278, 279
 disjuncts 203
 distortion 156, 176, 178
 domains of language use 5, 10,
 240, 248, 268, 269, 270, 217
 dual route model 42
 dynamic lexicology 3, 5
- E**
Eastenders 29
 electronic communication x,
 5, 6, 7, 8, 27, 137, 138, 139, 140,
 143, 144, 146–156, 157
 encyclopaedia (*see* knowledge)
 environment (*see* context,
 socialcontext)
 eponym 155
 estrangement 165, 173, 179, 234
 euphemism 7, 151, 274, 275–276
 euphony 194
 evolution (*see also* language
 change) 7, 168
 expectancy norms 240
 extensional analysis 203–204,
 205
 extra-grammatical 167, 180
- F**
 family resemblance 217
 fantasy fiction (*see* genres)
Fawlty Towers 28, 34
 foregrounding 17, 23, 24, 179
 foreignising in translation 246
Forging the Darksword 218,
 220, 225, 234
 formulae/formulaic 31, 93, 98,
 138, 139, 148
 frame
 collocational 68
 contextual 25
 situational 20
 frozenness 209
 function words 47
 functions of WF xv, 5, 7, 8, 9,
 15, 17, 18, 19, 21–33, 155, 163,
- 164, 166, 167, 169, 171, 174–175,
 179, 225, 240, 247, 252, 259
 functional sentence perspective
 (FSP) 6
 fuzzy
 boundaries/categories 152,
 216, 224, 225, 234
 concepts 16
- G**
 gender terms (*see also*
 stereotype) 272, 273, 275,
 276, 279, 280
 genres xiv, 4, 8, 10, 15, 33, 34,
 128, 132, 144, 181, 233
 children's literature 9, 163,
 175–178, 179
 fantasy fiction xv, 7, 9, 146,
 213, 218–219, 232, 235
 science fiction 5, 7, 9, 22, 23,
 27, 35, 146, 163, 164, 166,
 170–175, 177, 178, 179, 180,
 213, 214, 218
 geographic variation 139
 globalisation 10, 245, 246
- H**
 hapax 166
 heuristics 208
 homograph 99, 105
 homonym 99, 102, 104, 124,
 257
 homophone 8, 99, 104, 130,
 139, 141, 142, 143, 149
 letter homophone 143
 number homophone 149,
 150, 151, 157
How to be Good 33
 HowNet 190
 humour (*see also* ludic
 language) xiv, 4, 5, 7, 8, 22,
 24, 65, 70, 83, 84, 85, 101, 102,
 103, 109, 122, 146, 151, 157,
 210, 274
 hybrids 202, 242
 hyperbole 87, 171
 hypernym 195, 202, 228

hyponym 99, 125, 165, 226, 227
 hypostatisation 7, 15, 22–23,
 35, 167

I

ideology 266
 idiolect 176, 178, 179
 idioms 32, 101, 240, 242, 244,
 245, 250, 260
 idiom use 239, 242, 253, 239
 infixing 118, 178
 inflections / inflectional 66,
 68, 69, 70, 82, 148, 178, 191,
 192, 193, 227, 247
 in-group bonds 8, 21, 138, 143,
 156, 157
 inhibitors 66–69, 87
 initialisation 8, 139, 141, 143,
 149, 150, 151, 156
 instancial
 use 240, 250
 variation 252
 institutionalisation 5, 6, 7, 8, 17,
 35, 143, 166, 169, 171, 179, 264,
 274, 276, 281
 intentions (*see* speaker/writer
 intentions)
 Internet (*see also* WWW) 143,
 150, 154
 Internet Relay Chat
 (IRC) 140
 language 144
 users 139, 144, 146
 intersemioticity 4
 intertextuality 4, 5, 6, 10

J

jargon 8, 32, 124, 138, 139, 145,
 148, 155, 156, 268, 279
 computer 8, 145
 hacker 144, 147, 156
 jokes 10, 79, 191, 243, 244

K

knowledge (*see also* dictionary)
 domains of 204
 encyclopaedic 6, 191, 211,
 215, 219
 episodic 21

speaker 219
 world 222, 223

L

language (*see also* Latvian,
 Soviet Russian)
 bias free 263, 264, 265, 266,
 267, 269, 279, 281
 change xi, 239, 248, 280
 community 240, 242
 contact 242, 246
 creative 239, 241, 247, 248,
 249
 natural 191, 200
 planning (*see also* prescrip-
 tivism) 266, 278, 279
 play 8, 35, 93, 179, 256
 (*see also* ludic language,
 word play)
 policy 5, 263, 268
 resources 61, 95
 system (*see also* langue) 16,
 165, 170, 239, 240, 241
 taboo 247
 target 239, 249
 universal 240
 usage 209, 279
 langue (*see also*
 competence) 164, 180
 lateral thinking (*see also*
 conventional meaning) 190
 Latvian language xv, 10, 239,
 242, 243, 245, 247, 249,
 250, 254, 258, 260
 clippings 247
 compounds 245, 254
 idiom use xi, 250, 253
 norms, conventions 240,
 244, 247, 249
 social transformation 10
 substandard 247
 word formation 242
 wordplay 239, 256
 lexical (*see also* ontologies)
 access 42
 decision task 127
 exchange 100, 105–106, 109
 fields 217, 224–225, 226, 227,
 232, 233

fit 197

substitution 99, 100, 101,
 102, 104–105, 240

lexicalisation 5, 6, 10, 52, 192,
 201

lexicon

mental 8, 9, 10, 94, 95, 100,
 102, 108, 110, 166, 190, 214,
 215, 217, 228, 234

phrasal 95, 251

lexis

creative 217
 diachronic perspective 62
 pseudo scientific 177
 specialised 264

life-cycle (of words) 9, 62, 72,
 75, 81, 87

literalisation (of idioms/
 PLIs) 102, 103, 107

loan processes 3

loan words 5, 10, 64, 108, 146,
 155, 172, 173, 175, 226, 227, 243,
 247, 249, 253, 254, 258

LOB Corpus 4

London Lund corpus 4

Lucky Jim 20

ludic language (*see also* humour,
 wordplay) 8, 9, 10, 137, 147,
 163, 179

M

malapropism 104, 194

masked primes 127–128

maximal onset principle 45, 48

meaning

compositional 97, 98, 194
 connotative 77
 conventional 219, 222
 dominance (*see also* sali-
 ence) 221
 extensional 147, 149
 literal 101, 219
 referential 219
 transparent 232

media xiv, xv, 5, 74, 75, 79, 80,
 88, 115, 128, 243, 247, 260,
 265, 266, 279

British media x, 7, 9, 61, 75

- memory 179
 retrieval from 40
 short term 168
- metafunctions xiv, 7, 15, 24–25,
 26, 27, 29, 32, 33
- metalinguage 4, 6, 25
- metamorphology 167
- metaphor and metonymy 3, 5,
 9, 36, 62, 102, 132, 171, 172, 179,
 191, 215, 216, 219, 243
- metaphorical extension 9, 70,
 74, 157, 229
- modifiers 80, 82, 101, 196, 202,
 206, 208, 210, 211, 280
- modulation 198, 199, 210
 head 197, 199
 modifier 198, 199
- morpheme
 bound 121–123, 125, 210
 boundary 43, 44, 51, 64
 free 40, 132, 210
- morphological
 gradient structure 44
 patterns 226
 processes 165, 166, 175
 structure 7, 39, 276, 280
 theory xiii, 3, 11
- motivation 9, 22, 61, 146, 170
 functional 163, 178–179
 ludic 151, 157
 phonetic/phonological 163,
 172, 175, 176, 178
 pragmatic 146, 157, 167
 semantic 177
 sociolinguistic 147, 157
 stylistic 163, 170, 178
- MUDs/MUDDer (multi-user
 dimension) 33–34, 144,
 147, 150
- N**
- naming xiv, 6, 7, 9, 15, 17, 18, 19,
 21, 22, 28, 34, 35, 36, 169
- naming units (NUs) 6, 21, 169,
 170, 174, 175
- neighbours 126, 127, 201
- neo-classical combining forms
 (CFs) (*see also* combining
 forms) 64, 124, 152
- neologisms (*see also* coinages,
 nonce forms, novel
 formations) 3, 7, 8, 9, 63, 66,
 75, 79, 86, 115, 121, 123, 124,
 128, 131, 166, 170, 179, 194,
 213, 214, 215, 218, 227, 234,
 240, 243, 253
 vs. nonce formations 17, 18,
 164, 165, 168–169, 171
- neoseme xv, 9, 171, 213, 217,
 218, 234
- Netspeak 144, 147
- newspapers (*see also*
 media) 10, 62, 77, 116, 248,
 256, 257, 259, 271, 272
- New Zealand English (*see also*
 ONZE) xi, 7, 39, 40, 44,
 52, 55
- non-biased (*see* language)
- nonce forms (*see also* coinages,
 novel formations) xiii, 7, 17,
 55, 155, 163, 164, 165, 168, 170,
 176, 179, 225, 227, 228, 232,
 242, 258
- No Phrase Constraint (NPC)
 31
- norms 239, 240, 242, 243, 250,
 266
- notational terms 4, 15–16, 18,
 164
- novel formations (*see also*
 coinages, nonce words,
 neologisms) xiii, xiv, 7, 9,
 10, 63, 115, 116, 117, 121, 126,
 128, 129, 132, 163, 164, 166, 167,
 169, 170, 171, 173, 174, 175, 176,
 177, 178, 179, 180, 181, 198, 209
- O**
- Observational Linguistics
 (OL) 9, 10
- onomasiological approach (*see*
also categories) 6, 9, 175,
 217, 224, 228
- onomatopoeia 142, 175, 176,
 178
- ontologies (*see also* categories)
 xv, 7, 189, 190, 191, 192, 201,
 204, 208, 209
- onymic units (*see also* proper
 names) 176
- ONZE (Origins of New Zealand
 English Corpus) 39, 44
- P**
- paradigmatic 228
- paramorphology 167
- paraphrase (*see also*
 calques) 33, 84, 108, 109,
 277
- parody 5, 10, 22, 23, 27, 84, 156
- parole (*see also*
 performance) 164, 239
- paronym/paronymic 104, 241,
 256
- parsing 42, 43, 101, 102, 107, 117
- participant observation
 (PO) 8, 10
- performance (*see also*
 parole) 15, 16, 110, 163, 164,
 165, 170, 180
- perlocutionary 179
- phonetic
 analogy 194
 detail xi, 44, 54, 55
 length of un- 31–55
 reduction 47, 48
 respellings 142, 149, 150
 salience 41
 similarity 85, 150, 156, 194,
 199, 210, 249
- phonology of blends 117,
 119–120
- phonotactics xv, 39, 43, 44, 48,
 50, 51, 199
- phrasal
 compounds 26, 31, 32, 33,
 258
 lexical items (PLIs) 8, 31,
 32, 93–112
 lexicon 95, 102, 251

- phraseology (*see also* phrasal
 lexical items) 240, 241,
 242, 250
 playful formations (*see also* ludic
 language, word play) 9, 166,
 167, 176, 180, 281
 political correctness xi, 7, 151,
 263–281
 polysemy 217, 241, 263
 polyword NPs 173, 174
 portmanteau words (*see also*
 umbrella term) 210, 211
 Prague School 6
 prefixes / prefixation 7, 9, 22,
 30, 35, 65, 74, 124, 141, 144,
 147, 148, 154, 174, 258
 derivational 65, 174
 length of 40, 43, 52, 53
 negative 148
 novel 258
 patterns of 9
 un- 39–55
 premorphology 167
 prescriptivism 143, 278, 279,
 280
 product vs. process 169
 process (*see* word formation)
 processing 8, 43, 126–128, 179
 productivity (rule-governed) 3,
 16, 18, 55, 63, 140, 165, 167,
 226
 vs. creativity xiv, 3, 8, 9, 15,
 16, 17, 24, 26, 61, 62, 63
 proper names (*see also* onymic
 units) 6, 26, 64, 65, 123, 173
 prototypes 165, 216, 219, 223,
 230, 232
 pseudo
 derivation 227
 loanwords 227
 puns, punning 70, 71, 72, 85,
 99, 102, 104, 107, 108, 109,
 167, 179
 purism 242, 266, 280
- R**
 recategorisation
 grammatical 172
- semantic xv, 172
 semasiological 233
 syntactic 17, 35, 36
 of words 228
 recontextualisation 173, 222
 recoverability (*see also*
 condition) 96, 100, 101, 104
 reduction (*see also* vowel
 reduction) 3, 275
 clause 172
 letter 139, 141, 149, 150, 157
 phrase 148
 prefix 45
 reduplication 25, 33, 155
 register 4, 6, 10, 15, 20, 26, 145,
 153, 173, 247, 258, 275
 reinterpretation 151, 249
 relative frequency 43, 48, 54
 reorientation 10, 245, 246
 resegmentation 120, 122, 125
 respelling (*see* phonetic
 respelling)
*Restaurant at the End of the
 Universe* 27
 restricted collocation 97, 100
 restrictions (*see also* slot)
 conventional 100
 selectional 101
- S**
 salience (*see also* phonetic
 salience)
 of category membership
 225, 227, 230, 233
 as meaning dominance 74,
 219, 220, 221, 223, 225
 relative to base word 43
 schema (*see also* prototypes)
 as abstract representation
 216
 image schema 189
 prototype 219
 scholastic aptitude tests 200
 science fiction (*see* genres)
 selection set 97, 100
 semantic (*see also* analogy,
 categories, deformation,
 motivation, semasiology)
 anomaly 174, 176, 222
 change 217, 220, 234, 242
 extension 124, 230, 275,
 280
 fields 105, 127, 222, 233
 head 119, 152
 narrowing 274, 277, 280
 recategorisation xv, 215
 structure 152, 190
 substitution 172
 semasiological perspective
 (*see also* categories,
 recategorisation) 9, 217
 sense spectra 139, 157, 220–223
 shortening 47, 137, 149–154, 173
 degrees of 153
 devices/methods 6, 139,
 140, 141, 142, 144, 156
 rules 142, 143
The Simulacra 163, 171–173, 175
 slang 139, 144, 145, 146
 slot(s) 72, 96, 100
 filling 202
 open slot 72, 100
 restrictions 96, 100, 101, 106
Small World 32
 SMS (short message service) (*see*
 text messaging)
 social
 constraints ix, 239
 context or environment xiii,
 xiv, xv, 52, 55, 87, 278
 factors 271
 groups 166, 214
 issues 64, 66
 movements 265, 266
 role of PLIs 94
 stigmatisation 7
 trends 169
 sociocultural xiv, 7, 264, 265
 sociolinguistic 88, 137, 147, 157,
 240, 266, 280
 sociopolitical xv, 265
 sound
 association 176, 177
 imitation 172, 178
 overlap (*see* constituent
 integration)
 play 178
 similarity 193

- substitution 256
 symbolism 23, 175, 177, 180
 source words 116, 117, 118, 119, 120, 121, 125, 126, 127, 128, 132
 frequency of 127
 Soviet Russia(n)
 bureaucratic style 248
 censorship 243
 contact language 243
 language policy 243
 translation from 211
 speaker/writer intentions xiii, 7, 10, 94, 99, 102
 specialised (*see also* discourse, jargon)
 fields 143, 146
 functions 8, 36
 lexis/vocabulary 15, 263, 264, 265
 users 138
 specification 7, 140, 144, 226, 274, 275, 277
 speech acts 4, 25, 29, 30
 speech rate 45, 46, 48
 splinters (*see also* blending, clipping) 115, 116 (def.), 117, 118, 119, 120–125, 127, 132, 151, 153, 154
 as bound morphemes 121–125
 final 119
 initial 119, 123
 novel 121
 Spoonerism 99, 103, 176, 242
 stereotype (n) (*see also* gender terms) 31, 32, 275
 (v) (*see also* stigmatise) 7, 31, 32, 33
 stereotypical 32, 33, 32
 stigmatise 271, 277
 structural (*see also* ambiguity, deformation)
 change 241
 compositionality 174
 features or properties 169, 260
 parallelism 156
 representation 94, 107

 style(s) 15, 243, 247, 264, 279
 colloquial 10, 247
 markers 9, 18, 163, 175
 register vs. style 145, 258
 style guides 6, 143, 243, 263, 267–281
 writer/authorial style 70, 163, 175
 stylistic(s) xv, 9, 164
 devices 72
 effects 23, 70, 171, 242
 function 169, 173, 179, 239, 40, 241, 259
 motivation 163, 170, 178, 180, 181
 subcodes 8, 138, 143
 substitution (*see also* lexical, semantic, sound)
 of letters 126
 of sounds 99, 256
 of words 99
 suffixes /suffixation (*see also* affixation) 144, 149, 153, 172, 227, 228, 265, 276
 derivational (*see also* derivation) 174, 177, 178
 diminutive 240, 247
 gender 275, 276
 invented 174, 178
 length of 41
 pseudo 174
 superordinate (*see* categories)
 support set 196–199, 201, 210
 syllable
 length 46, 48, 49, 50, 51
 weight 46, 47, 49
 synchronic xiii, 165
 synonymy 7, 63, 217, 224, 225, 275, 276, 278
 syntactic reorganisation (*see* (re)categorisation)

 T
 taxonomies (*see also* categories) 4, 95, 129, 171, 191
 text linguistics 4, 164
 text messaging (SMS) x, 8, 138, 139, 140, 141, 142–144, 145, 150, 156

 text type (TT) 4, 5, 23, 240, 270
 text world xiv, 20, 32, 167, 171, 179, 213, 235
 textual functions 6, 17, 164, 179
 textuality 4, 16, 20, 21, 22
 Time 24, 29
 timed experiments 127
 time-series graph 81
 time-series plot 71
 tokens (*see also* types) 44, 45, 47, 53, 54, 62, 164, 171
 transfer
 linguistic 250
 semantic 3
 transformations
 creative 10, 241, 251
 idiom 240, 245, 250–253, 260
 occasional 250
 social 10
 speech 240
 transformational-generative grammar 6, 16, 31
 translation 10, 243, 244, 245, 246, 247, 249, 256, 258, 260
 approaches to 246
 hidden 249
 literary 244
 policies xv, 5, 246
 theory of 245
 translation loans 254
 truncation 7, 151, 275, 276
 types (*see also* categories tokens) 3, 9, 10, 140, 151, 163, 164, 175, 178, 181, 241, 242, 249

 U
 umbrella term (*see also* portmanteau) 166, 266
 unification strategy 7, 274, 276, 280
 universal franchise (of words) 87
 universalisation strategies 137, 142, 249
 usage conditions 98

 V
 variant 65, 68, 69
 analogical 207

- creative 72, 76, 87
- derivational 62
- dialectal 80
- inflectional 68, 69
- orthographic 69
- playful 9
- spelling 81
- verisimilitude (*see also*
credibility) 163, 173, 179
- vogue words 65, 71, 86
- vowel reduction 47, 48

- W**
- WebCorp 62
- web pages 8, 73
- websites 75, 123
- Wikipedia 211
- word manufacture 18, 165, 178
- word(s) (*see also* coinages,
loans, neologisms, nonce
forms, novel formations)
- actual vs. potential 165,
170, 180
- complex 6, 10, 36, 43, 55,
163, 165, 166, 168, 170, 172,
174, 176, 177, 179, 180
- invented 213, 214, 215, 217,
225, 234
- simplex 6, 42, 165, 167, 170,
172, 174, 176, 177, 178, 179,
180
- storage 41–44, 214
- Word-Based Hypothesis – *see*
No Phrase Constraint
- word formation (WF) (*see also*
functions of) 3, 5, 6, 9,
16, 17, 20, 22, 30, 31, 32, 34,
26
- analogy-based 24, 27
- creative 16, 26, 164
- nonce 18
- patterns of 21, 178, 228
- phrasal 32
- processes xiii–xiv, 5, 7, 8,
10, 31, 138, 157, 164, 174,
179, 239
- rules 164, 166, 167, 168, 169,
179, 180
- WordNet xi, 7, 10, 190, 192, 193,
195, 198, 201, 204, 205, 207,
208, 211, 232
- wordplay (*see also* ludic
language) 5, 9, 10, 23, 24,
35, 61, 70, 82, 103, 151, 156, 163,
239, 240, 241, 248, 256, 257,
258, 259
- world creation (*see text worlds*)
- world wide web (WWW) (*see*
also Internet) 9, 137, 177

- Z**
- zero derivation (*see* conversion)
- zeugma 241

In the series *Studies In Functional And Structural Linguistics* the following volumes have been published thus far or are scheduled for publication:

- 59 **DREER, Igor:** Expressing the Same by the Different. The subjunctive vs the indicative in French. 2007. xxx, 272 pp.
- 58 **MUNAT, Judith (ed.):** Lexical Creativity, Texts and Contexts. 2007. xvi, 294 pp.
- 57 **DAVIS, Joseph, Radmila J. GORUP and Nancy STERN (eds.):** Advances in Functional Linguistics. Columbia School beyond its origins. 2006. x, 344 pp.
- 56 **JING-SCHMIDT, Zhuo:** Dramatized Discourse. The Mandarin Chinese *ba*-construction. 2005. xxii, 337 pp.
- 55 **BEEDHAM, Christopher:** Language and Meaning. The structural creation of reality. 2005. xiv, 225 pp.
- 54 **ŠTEKAUER, Pavol:** Meaning Predictability in Word Formation. Novel, context-free naming units. 2005. xxii, 289 pp.
- 53 **EDDINGTON, David:** Spanish Phonology and Morphology. Experimental and quantitative perspectives. 2004. xvi, 198 pp.
- 52 **GORLACH, Marina:** Phrasal Constructions and Resultativeness in English. A sign-oriented analysis. 2004. x, 151 pp.
- 51 **CONTINI-MORAVA, Ellen, Robert S. KIRSNER and Betsy RODRÍGUEZ-BACHILLER (eds.):** Cognitive and Communicative Approaches to Linguistic Analysis. 2004. viii, 389 pp.
- 50 **VACHEK, Josef:** Dictionary of the Prague School of Linguistics. Translated from the French, German and Czech sources. In collaboration with Josef Dubský. Translated by Aleš Klégr, Pavlína Šaldová, Markéta Malá, Jan Čermák and Libuše Dušková. Edited by Libuše Dušková. 2003. x, 216 pp.
- 49 **HLADKÝ, Josef (ed.):** Language and Function. To the memory of Jan Firbas. 2003. x, 339 pp.
- 48 **REID, Wallis, Ricardo OTHEGUY and Nancy STERN (eds.):** Signal, Meaning, and Message. Perspectives on sign-based linguistics. 2002. xxii, 413 pp.
- 47 **MARTÍN-VIDE, Carlos (ed.):** Issues in Mathematical Linguistics. Workshop on Mathematical Linguistics, State College, PA, April 1998. 1999. xii, 214 pp.
- 46 **ŠTEKAUER, Pavol:** An Onomasiological Theory of English Word-Formation. 1998. x, 192 pp.
- 45 **MARTÍN-VIDE, Carlos (ed.):** Mathematical and Computational Analysis of Natural Language. Selected papers from the 2nd International Conference on Mathematical Linguistics (ICML '96), Tarragona, 1996. 1998. xviii, 391 pp.
- 44 **JESSEN, Michael:** Phonetics and Phonology of Tense and Lax Obstruents in German. 1999. xx, 394 pp.
- 43 **ANDREWS, Edna and Yishai TOBIN (eds.):** Toward a Calculus of Meaning. Studies in markedness, distinctive features and deixis. 1996. xxviii, 432 pp.

