

Vito Pirrelli, Ingo Plag and Wolfgang U. Dressler

Word knowledge in a cross-disciplinary world

Abstract: This editorial project stemmed from a 4-year period of intense interdisciplinary research networking funded by the **European Science Foundation** within the framework of the **NetWords** project (09-RNP-089). The project mission was to bring together experts of various research fields (from brain sciences and computing to cognition and linguistics) and of different theoretical inclinations, to advance the current awareness of theoretical, typological, psycholinguistic, computational and neurophysiological evidence on the structure and processing of words, with a view to promoting novel methods of research and assessment for grammar architecture and language usage.

The unprecedented cross-disciplinary fertilization prompted by a wide range of scientific and educational initiatives (three international workshops, two summer schools, one main conference and over a hundred grants supporting short visits and multilateral exchanges) persuaded us to pursue this effort beyond the project lifespan, spawning the idea of an interdisciplinary handbook, where a wide range of central topics on word knowledge and usage are dealt with by teams of authors with common interests and different backgrounds. Unsurprisingly (with the benefit of the hindsight), the project turned out to be more challenging and time-consuming than initially planned. Cross-boundary talking and mutual understanding are neither short-term, nor immediately rewarding efforts, but part of a long-sighted, strategic vision, where stamina, motivation and planning ahead play a prominent role. We believe that this book, published as an **open access volume**, significantly sharpens the current understanding of issues of word knowledge and usage, and has a real potential for promoting novel research paradigms, and bringing up a new generation of language scholars.

Keywords: interdisciplinarity, word knowledge, word usage, language units, statistical and computer modeling, levels of understanding, between-level mapping, linking hypotheses, scale effects

Vito Pirrelli, Italian National Research Council (CNR), Institute for Computational Linguistics, Pisa, Italy

Ingo Plag, Heinrich-Heine-Universität Düsseldorf, Department of English, Düsseldorf, Germany

Wolfgang U. Dressler, University of Vienna, Department of Linguistics, Vienna, Austria

1 Context

Scientists are nowadays faced with a few important discontinuities with the past: (a) an exponentially growing rate of technological innovation, (b) the ever-increasing availability of multimodal data, (c) an increasing disciplinary specialization, involving the danger of being blind to interdisciplinarity, and (d) a pressing demand for problem-oriented interdisciplinarity. 19th century medical practitioners based a diagnosis upon visiting their patients. For a 21st century medical doctor, patient encounters are complemented by a number of sophisticated diagnostic techniques, ranging from radiography, PET and MEG to ECG, EEG and ultrasound. This is what contemporary medicine is about: creating new objects of scientific inquiry by multiplying and integrating different information sources.

21st century language scientists are no exception. They can benefit from an equally large array of technological tools tapping linguistic information at unprecedented levels of range and detail. They know that words, phrases and utterances are not just mental representations or convenient descriptive devices grounded in introspection and informants' intuition. They are multidimensional objects, emerging from interrelated patterns of experience, social interaction and psychological and neurobiological mechanisms. Investigation of these objects calls for integration of manifold information sources at a conceptual and functional level.

In this book, we strive to understand more of words in language by squarely addressing a number of questions underlying the relationship between speakers' knowledge of words, evidence of the way speakers use words in daily communicative exchanges and psychological and neurofunctional correlates of word usage. How are words processed in working memory? Are they stored in long-term memory as a whole or rather composed 'on-line' in working memory from stored sub-lexical constituents? What role is played in this process by knowledge-based factors, such as formal regularity and semantic transparency, and usage-based factors, such as perceptual salience, familiarity and frequency? Does word-level knowledge require parallel development of form and meaning representations, or do they develop independently and at a different pace? How do word meanings function and combine in daily communicative contexts, and evolve through learning? What types of lexical knowledge affect on-line processing? Do the dramatic typological differences in word structure across world languages impact on processing and acquisition? And how will a thorough investigation of such differences change lexical models worked out on the basis of a single language? Finally, what neurobiological patterns of connectivity sustain word processing and storage in the brain? And how can they break down as a result of neurological damage or disorders?

Any serious effort to address these questions needs to ultimately be based upon recognition that words define a multifactorial domain of scientific inquiry, whose thorough investigation requires synergic integration of a wide range of disciplines. Of late, a few independent lines of scientific inquiry appear to lend support to an integrative approach to the study of the mental lexicon:

- In line with a view of word knowledge as an **interface domain**, the architecture of the mental lexicon is better understood as resulting from the dynamic integration of multiple levels of information (Jackendoff 2002); correlation of these levels, albeit indirect and possibly non-linear, enforces constraints and mutual dependencies that are not justified on single-level grounds (Elman 2004, 2009). This view is not incompatible with a principle of representational modularity, segregating linguistic information according to levels of representation (Jackendoff 2000, 2007). Nonetheless, it conceives of lexical knowledge as emerging from the unique, distributed network of stored associations among fragments of disparate representations, including constructions, idioms, proverbs and social routine clichés (e.g. Arnon et al. 2017; Arnon and Snider 2010; Bannard and Matthews 2008; Grimm et al. 2017; Tremblay and Baayen 2010; Siyanova-Chanturia et al. 2017; Vespignani et al. 2009).
- Word processing requires a two-way **interactive perspective**, whereby the speaker can anticipate what the hearer needs to be provided with in order to obtain the intended perlocutionary effects, and, in turn, the hearer can predict what may be offered in the ongoing spoken or written communicative interaction (Huettig 2015; Pickering and Garrod 2013; Riest et al. 2015); communicative factors include Theory-of-Mind states (Milligan et al. 2007; Wellman 2002) and perspective taking (Brown-Schmidt 2009), contextual and co-textual embedding and transparency of words (Marelli et al. 2017; Mikolov et al. 2013; Mitchell and Lapata 2010), especially of neologisms and occasionalisms (Mattiello 2017; Plag 2018), choice between synonyms, lexical and morphological differences between child-directed and adult-directed speech (Kilani-Schoch et al. 2009; Saxton 2008, 2009; Taylor et al. 2009), paraphrases, and simultaneous top-down and bottom-up processing strategies (Ferro et al. 2010; Kuperberg and Jaeger 2016; Pickering and Garrod 2007, 2013; Smith and Levy 2013);
- Accordingly, word processing is modelled as the task of optimal resolution of multiple, parallel and possibly conflicting constraints on complex lexical structures, where top-down expectations, based on past experiences and entrenched memory traces, combine, in on-line processing, with the bottom-up requirements of input stimuli (Berger et al. 1996; Kukona et al. 2011; Seidenberg and MacDonald 1999; Spivey and Tanenhaus 1998; Tabor and Tanenhaus 1999);

- This is in keeping with a **Maximization of Opportunity Principle** for word processing: different processing strategies are applied simultaneously, and preference for one strategy over another is opportunistically given on the basis of task-based requirements, or compensatory mechanisms offsetting contingent failures caused by language impairments or production/perception errors and other types of noise (Libben 2005, 2016);
- All these perspectives are compatible with the hypothesis of an **indirect correspondence** between low-level principles of word processing/organization and their brain localization (Clahsen 2006; Hasson et al. 2018; Pirrelli 2007; Price 2017). On this view, complex language functions are not localized to specific brain regions, but are rather the emergent property of the interaction of parallel distributed networks of densely interconnected regions (D’Esposito 2007; Price 2010, 2012). In this context, the functional anatomy of language cannot be deduced from a high-level conceptualization of the way language is understood to work in the brain, but it requires a deep understanding of the functional interaction of concomitant low-level processing principles and associative mechanisms (Hasson et al. 2018, Pirrelli et al. this volume).
- Over the last 20 years, the anatomy of language has been investigated with neuroimaging techniques (e.g. PET and fMRI) and brain areas associated with language processing have been identified consistently (Ben Shalom and Poeppel 2008; Hickok and Poeppel 2004; Price 2010, 2012, 2017). Future studies will undoubtedly be able to improve the spatial and temporal precision with which functional regions can be located (see Davis 2015, for the neuroanatomy of lexical access). Nonetheless, assuming that our current understanding of the general picture is correct, the main task for future research will be to specify the details of the inter-region organization and computational operations.

2 Content

In this volume, experts of various disciplines look at common topics from complementary standpoints, to discuss and understand what can be learned from integrating different approaches into converging perspectives. Most chapters are jointly authored by at least two experts from different fields, not only to bring together evidence from different domains but, more importantly, to make these domains talk to each other, with a view to gaining a deeper understanding of the issues focused on in the chapter.

The book is structured into three parts. Part 1: Technologies, Tools and Data (covering chapters 2 through to 5) is chiefly devoted to the methodological

pre-requisites to interdisciplinary research on languages: technologies, tools and data. Its focus ranges over the contribution, goals and limits of computer simulations, statistical techniques for multidimensional data analysis and modeling, neuroscientific experimental paradigms and tools, and shared data and data infrastructures. Part 2: Topical Issues (including chapters 6 through to 11) deals with topical issues in word inquiry, including the morphology-phonology interface, inflection, derivation, compounding, lexical semantics and morpho-pragmatics. Finally, Part 3: Words in Usage (chapters 12 through to 17) contains an overview of classical theoretical approaches to the dualism between word storage and processing, together with more focused contributions on word usage issues, zooming in on multilingual lexica, word reading, word acquisition, errors in morphological processing and developmental disorders in word competence. In what follows, we provide a concise introduction to the main topics harped on in each chapter, with a view to highlighting converging trends, actual and potential interactions, as well as prospects for cross-fertilization.

2.1 Outline

Chapter 2, on psycho-computational and algorithmic models of the mental lexicon, delineates a clear connection between word frequency distributions and information theoretical measures for word families, statistical correlations over behavioral evidence (e.g. wordlikeness ratings and reaction times), principles of discriminative learning, and integrative algorithmic models of word storage and processing. However tightly interrelated, this heterogeneous bundle of evidence has traditionally been in the purview of distinct domains of scientific inquiry such as corpus linguistics, psycholinguistics, machine language learning, computational linguistics and serial cognition. By going through Marr's (1982) levels of understanding of complex systems, Vito Pirrelli, Marcello Ferro, Claudia Marzi, Franco Alberto Cardillo, Harald Baayen and Petar Milin show that approaching all these issues from a learning perspective sheds light on their potential for integration, while defining a fruitful line of research in the years to come.

Chapter 3, by Jacolien van Rij, Nemanja Vaci, Lee H. Wurm and Laurie Beth Feldman, is a guided tour to some of the most successful statistical techniques for psycholinguistic data modelling to date, from ANOVA to Generalized Additive Models. It addresses, step by step, a wide range of methodological issues that are only occasionally discussed in the technical literature at this level of depth. In spite of its apparent technicality, the chapter will thus be beneficial to non-expert as well as more advanced users of statistical packages for analysis of language data. We believe that these techniques are bound to become

part and parcel of the methodological tool-kit of any language scientist, as witnessed by the growing awareness of the importance of quantitative data, even within theoretical frameworks that proved, in the past, more reluctant to accept usage-based data as part of their empirical evidence.

In Chapter 4, Paola Marangolo and Costanza Papagno provide a clear, comprehensive introduction to the best-known protocols and techniques for investigating the neurophysiological reality of words in the brain, using aphasia as a case study. Whereas in earlier times language brain substrates could only be studied indirectly, through correlation of cerebral lesions with dysfunctional behavior, today functional neuroimaging allows direct *in vivo* visualization of cerebral activity. This opens up unprecedented, exciting opportunities in investigating the neurobiology of language, to offer rich evidence that distinct cerebral areas process different word classes. Nonetheless, a couple of caveats are in order here. First, in using neuroimaging methods, one must be aware of their inherent limitations. Methods that are based on the study of perfusion and metabolism (such as PET and fMRI) detect neural activity only indirectly, based on local blood flow. On the contrary, recordings of event-related potentials with electroencephalography can detect neural activity directly, with optimal temporal resolution, but poor spatial precision. A better understanding of the brain dynamics involved in word processing is thus likely to require a combination of techniques with different temporal and spatial resolutions. Secondly, establishing a causal relationship between a language task and the activation of a specific brain region should be assessed with care, since several uncontrolled variables can produce a misinterpretation of results. For example, localization of a verb-specific (as opposed to noun-specific) brain region can in fact be due to effects of morpho-syntactic processing, such as subject-verb agreement checking, rather than to a pure, categorical effect. In fact, language-driven interpretations of the involvement of specific cortical areas in an experimental task could (and, according to some scholars, should) be replaced by more parsimonious explanatory accounts, postulating basic or domain-general computations (Hasson et al. 2018; Price 2017). As the number of linguistic and extra-linguistic variables can be extremely large, Marangolo and Papagno suggest that a closer interaction of neurobiological models with both low-level computer models and high-level cognitive linking hypotheses can provide fruitful, top-down constraints on the interpretation space.

The important issue of producing and sharing high-quality multimodal evidence of elicited as well as unelicited language production/recognition, is addressed in Chapter 5, where Emmanuel Keuleers and Marco Marelli discuss at some length the complex and delicate nature of what they appropriately call “the language data ecosystem”. They focus on the often-neglected fact that data are never produced in a vacuum, but are always the by-product of a

complex interaction between scientific goals, methodological stances and analytical tools. Awareness of this deep interdependency is key to pushing progress in our field. Only by getting a clearer view of the shortcomings of analyses exclusively based on data that are elicited in tightly controlled experimental conditions, scholars can hope to address fundamental questions concerning the neurobiology of language usage in more ecological settings.

Chapter 6, by Sabine Arndt-Lappe and Mirjam Ernestus, deals with the relation between morpho-phonological alternations and lexical storage and processing. There is a long tradition of structurally and theoretically oriented studies of morpho-phonology that have explained phonological alternations in complex words in the form of rules (or similar mechanisms). More recently, however, a growing body of evidence has accrued that morpho-phonology may be closely linked to how speakers and listeners process complex words. The authors discuss several morpho-phonological alternations and demonstrate what we can learn from these alternations about the storage of complex forms. Existing theoretical and computational models are evaluated in the light of psycholinguistic evidence. Ultimately, it seems that alternations can only be explained if we assume lexical storage of at least some alternants.

In dealing with inflection as a central component of morphological competence, the authors of Chapter 7 set themselves the ambitious goal of focusing on the role of formal contrast in marking functional differences in the syntactic distribution of inflected words. Claudia Marzi, James Blevins, Geert Booij and Vito Pirrelli discuss the way storage of frequent forms can interact with generalization strategies that compensate for lack of input evidence in the low-frequency range. Both morphological and constructional information are assumed to be stored in long-term memory, in keeping with a view of lexical representations as highly context-sensitive. This is in line with recent psycholinguistic evidence reported, among others, in Chapter 6 of this volume, showing how much information is actually accessible in the mental lexicon, both in terms of the phonetic details stored for each word, and in terms of how morphologically-complex words are actually stored as (possibly) independent lexical units.

In Chapter 8, Ingo Plag and Laura Wither Balling cast a very wide net on the extremely rich and variegated evidence on derivatives and derivational processes coming from as diverse research areas in language sciences as phonetics, theoretical linguistics, psycholinguistics, neurolinguistics and computational linguistics. Such a bird's eye view allows for careful assessment of widely held assumptions, as well as more contentious issues, while charting those yet unexplored territories in morphological derivation that may offer fruitful prospects of converging progress in the years to come. In particular, the authors observe that theoretical linguistics has typically over-emphasized representational issues at the expense of

processing issues, with psycholinguistics and neurolinguistics being more, if not exclusively, concerned with the latter (i.e. behavioral evidence of the human word processor). Such a discipline-oriented bias made theoretical linguistics relatively blind to the relevance of formal contrast for word recognition irrespective of the formal means by which it is enforced (i.e. whether morphemically or not). On the other hand, more brain-oriented language disciplines turned out to be relatively blind to issues of word production, with comparatively sparser attention being paid to how sublexical constituents are combined to produce whole word meaning in derivatives.

Morphological compounds bear witness to the advantages of taking a multidisciplinary perspective on a common pool of data. In Chapter 9, Gary Libben, Christina Gagné and Wolfgang U. Dressler keep their focus on both representational and processing issues. From this two-fold perspective, compounds appear to be linguistic objects of a quintessentially dual nature. On the one hand, their meaning is intimately associated with their lexical wholes. Such constructional effects are “both greater than the sum of their parts and greater than the division of their wholes”, requiring some form of “weak compositionality” (Baroni, Guevara and Pirrelli 2007) mostly dictated by paradigmatic relations holding between overlapping members of the same compound family (as opposed to combinatorial principles of syntactic composition). At the same time, the processing of compounds calls for activation of their constituents as distinct units, with more transparent compounds, i.e. those compounds whose form and meaning are more directly amenable to the form and meaning of their constituent parts, being the easiest to process.

Issues of lexical semantics are the specific focus of Chapter 10, illustrating, in a somewhat exemplary way, the benefit of comparing different perspectives on the same subject area, and weighing up their respective strengths and weaknesses. Paolo Acquaviva, Alessandro Lenci, Carita Paradis and Ida Raffaelli provide a comprehensive overview of very diverse models of lexical meaning. Coverage includes the traditional, structuralist view of word meanings as forming part of a systemic network of value contrasts/oppositions; the symbolic rule-based approach of generativist tradition; the investigation of concept formation as rooted in cognitive primitives like space and geometry; more recent distributional approaches, where meanings are points in a multidimensional space defined by the distribution of words in context. All these models appear to articulate different, and in some cases, irreconcilable answers to fundamental questions about the nature of lexical meaning. It would be rather naïve to claim, however, that they offer just complementary and inevitably incomplete rival perspectives on the vast, elusive realm of lexical semantics. In the end, all these aspects need be reconciled and accounted for within a unitary, analytical

framework, able to integrate the results of different approaches, including data of typologically different languages, experientially-based evidence, results of computer simulations using word distributions in context, and results of psycholinguistic and neurolinguistic experimental paradigms.

Chapter 11 focuses on the relation between morphology and pragmatics. In contrast to models that take morphology and pragmatics to be always secondary in being based on the morphosemantics of the respective categories, Lavinia Merlini Barbaresi and Wolfgang U. Dressler argue that at least some of these categories (e.g. evaluatives, such as diminutives and augmentatives) have a basic pragmatic meaning, a claim which is incompatible with the assumption of external modularity between grammar and pragmatics. Although emotion may be heavily involved in the pragmatic meanings of morphopragmatic categories, pragmatic meanings cannot be reduced to the presence of emotion. The chapter goes beyond evaluatives and pragmatic devices of reduplication (both grammatical and extragrammatical) which are most extensively discussed in the literature, to also include honorifics, excessives and pragmatic uses of plurals, as well as many other categories of word formation and inflection. The roles of lexical pragmatics, sociopragmatics, corpus linguistic approaches and new developments in pragmatics are also discussed in some detail.

The dualism between storage and computation in morphology is focused on in Chapter 12, where Martina Penke and Antonio Fábregas scrutinize competing theoretical frameworks of lexical competence, to assess theoretical predictions in the light of some of the major word processing effects that have been identified in psycholinguistic research over the last decades. In particular, they discuss two of the most established behavioral findings to date: (i) the relative insensitivity of regularly inflected forms to token frequency effects in word processing, and (ii) the stronger perception of morphological structure in regulars as opposed to irregulars. Somewhat surprisingly, these findings appear to cut across two of the main theoretical dimensions governing the contemporary debate on morphology: namely, the opposition between lexicalism and neo-constructionism, and the item-and-arrangement vs. item-and-process dualism. According to the authors, both A-morphous Morphology and Minimalist Morphology prove to be compatible with evidence that humans process regulars and irregulars differently. Nonetheless, they appear to take opposite sides on the theoretically crucial question of what morphological units are stored in the mental lexicon and what units are produced by rules. This suggests that the relationship between principles of grammar organization (e.g. lexicon vs. rules) and processing correlates (storage vs. computation) is not as straightforward as the “direct correspondence” hypothesis (Clahsen 2006) has claimed in the past. Differential processing effects may in fact be the complex outcome of the

non-linear interaction of uniform learning and processing principles. Since modelling such interaction may well exceed the limits of both theoretical conceptualizations and box-and-arrow models of cognition, settling these theoretical issues will call for advanced sources of experimental evidence (e.g. computational and neuropsychological models of language behavior) and more sophisticated experimental paradigms (e.g. discriminating between morpho-phonological and morpho-syntactic effects in word processing).

In Chapter 13, Madeleine Voga, Francesco Gardani and H el ene Giraud investigate multilingualism from a two-fold perspective: the psycholinguistic modeling of the bilingual (and multilingual) lexicon, and the role of language contact in language change. In both domains, the co-existence of lexical items belonging to different languages and, possibly, to different morphological systems, raises a number of non-trivial questions on structural and processing counts. What sort of interaction governs the two sets? Does similarity of forms play a prominent role in this dynamic relationship? Or is rather similarity of meaning involved here? Or just a combination of the two, as with classical accounts of morphological relatedness? Is such a relationship symmetrical or asymmetrical, and what prevents items from one language from interfering with items belonging to the other language in daily communicative practice? The authors go on with establishing an interesting parallelism between the L1–L2 contrast in bilingualism (as well as the factors governing the L1–L2 interaction), and the synergic opposition between a recipient language and a source language in the literature on language contact and change. Interestingly, the two oppositions share a number of properties: (i) the gradient asymmetry of their relationship, accountable in terms of both frequency effects and the entrenchment of connections between the lexical and the conceptual level of speakers' word knowledge, (ii) the prominent role of word families in spreading cross-linguistic activation, and (iii) the sensitivity of systemic co-activation to pragmatic factors. In fact, all these interactive effects appear to be influenced by the specific pragmatic force of speakers' utterances, and their perlocutionary effects. The authors conclude that, in spite of persisting differences in methodology, terminology and goals, the material continuity of multilingual evidence in both domains lends support to a unifying view, and encourages a converging perspective in their scientific investigation.

Chapter 14 focuses on the connection between reading skills and morphological competence from a psycholinguistic, neuropsychological and computational perspective. Marco Marelli, Daniela Traficante and Cristina Burani start with an overview of evidence supporting the classical morpheme-as-unit view of lexical representations in the mental lexicon, together with the developmental literature supporting the idea that morphological awareness is an age-related,

emergent aspect of word processing. Effects of both semantic and frequency-sensitive modulation of morpheme access, as well as evidence of the global organization of lexical and sublexical constituents in large word families, and the context-sensitivity and task-dependency of behavioral findings based on established experimental protocols, jointly suggest that morphological effects may not require dedicated processing modules and storage units. Rather, these effects can be accounted for by general-purpose mechanisms for time-serial processing, coupled with the ability to track down and generalize statistically-strong form-meaning patterns. Reading skills can take advantage of these general abilities. At the same time, the age-related development of these abilities can largely benefit from increasing literacy levels. On a more general, methodological note, the authors point out that it is increasingly difficult to explore such a complex interaction of multiple, concurring factors through traditional experimental protocols. Computational simulations can nowadays dynamically model the interaction of several factors in the context of a specific task, thereby allowing one to weigh up and inspect their individual influence as well as their joint, interactive effects, at unprecedented levels of accuracy. It is only to be expected that large-scale computational simulations will play an important role in the investigation of morphological effects in reading in the years to come.

In Chapter 15, Emmanuel Keuleers, Dorit Ravid and Wolfgang U. Dressler deal with morphology and lexicon acquisition in children's first three years of life, by zooming in on a few focal points from an interdisciplinary perspective. The fundamental advantage of taking a broader perspective on issues of morphology acquisition is that integration of different viewpoints can shed light on the inherent limitations of domain-specific findings. Theoretical linguistic frameworks have long offered conceptual scaffolding for describing children's linguistic behavior in a structured, systemic way; and they will likely continue to do so in the near future. However, the further assumption that theoretical concepts and classification criteria developed for descriptive purposes are mapped linearly onto developmental evidence is misconceived. The scientific ability to identify minimal linguistic units and fundamental principles for their combination should not be confused with the hypothesis that language behavior can be understood by only observing the behavior of elementary units. Scale and complexity effects are ubiquitous in complex systems. It would be highly surprising if they were not observed in the acquisition of the most complex communication system in nature. Besides, what theoretical linguists call categories can be mentally structured along a similarity gradient. Likewise, some nearly instantaneous generalization processes in language development, apparently due to rule application, can in fact be the outcome of a continuous process of memory self-organization. Finally, it is difficult to over-estimate the contribution of the

information-theoretic notions of entropy and communication code to understanding how children learn words in context, and the proper role that frequency, stimulus discrimination and concept development play in the process.

In Chapter 16, Thomas Berg carries out a fine-grained analysis of morphological errors in speech, by assessing the causal factors involved, and their theoretical implications. Errors may be triggered by morphological competition of the mistaken target with (i) words that appear in the context being uttered (*in praesentia*), or (ii) paradigmatically-related companions of the target (*in absentia*). The author emphasizes the important role played by lexico-semantic factors in weighing up the strength of paradigm relations and, ultimately, the degree of accessibility of morphological structure and the competition between paradigmatically-related words. From this perspective, derivation and inflection are conceptualized as two opposing points in a cline going from the more lexical to the more grammatical end of the language spectrum. The availability of derivational paradigms vs. inflectional paradigms is crucially modulated by lexical semantics. Since members of the same derivational family share less lexico-semantic content than members of the family of inflected forms of the same lemma (or inflectional paradigm), the former belong to “weaker”, less accessible “families” than the latter do. A similar line of argument also allows one to draw a principled distinction between phonologically conditioned allomorphs (as with English –s plural marker) and morphologically (and lexically) conditioned allomorphs (as with *foot* and *feet*). Phonological allomorphs require involvement of two processes only: ordering and contextual accommodation of segmental material. Morphological allomorphs, on the other hand, call for an extra process of lexically-conditioned selection, involving a further processing cost, and making morphological allomorphs more prone to errors.

Developmental disorders offer a spacious window onto the neurobiological reality of word knowledge and its complex interaction with general cognition. In Chapter 17, Mila Vulchanova, David Saldaña and Giosué Baggio persuasively show that language disorders can hardly be associated with highly specific grammatical deficits. What may appear as a deceptively selective difficulty in language usage, such as the production of inflected regular forms by children with Language Impairment, are in fact subject to language-specific variation, depending on subtle factors such as the complexity of an inflectional system, the size, formal variety and frequency distribution of its paradigms, or the perceptual salience of morphological markers. Likewise, semantic problems in lexical development may be associated with general receptive deficits, as well as non-verbal IQ, maternal education level and language learning deficits, such as effects of increased lexical competition in the mental lexicon of language impaired children. The general emerging picture seems to suggest that the

patterns of dysfunctional language behavior observed in children with language disorders reflect the complexity, subtlety and robustness of the language system, rather than a broad dissociation between language and cognition.

3 Lessons to be learned

The thoughts and evidence offered in this book elucidate a number of non-trivial methodological and theoretical points in word knowledge. By way of a summary, we recap here a few take-home points.

In approaching interdisciplinary issues in language inquiry, there is a common two-way misconception, which scholars should be aware of and warned against. On the one hand, we contend that it is simply wrong to see theoretical language models as inevitably partial and incomplete, waiting for external evidence from mind and brain sciences to validate them. On the other hand, it would be just as wrong to see psychological and neurological methods of language inquiry as invariably in need of linguistic concepts and classification criteria that were developed by theoretical models for different purposes. Both views strike us as the misconstrued, or at best preconceived, by-product of a persisting lack of disciplinary crosstalk.

In too many cases, psycho- and neurocognitive evidence seems to cut across the theoretical frontline between rival linguistic accounts. In the last few pages, we pointed out that the decade-long theoretical confrontation between Item-and-Process and Item-and-Arrangement morphologies, probably the most influential dimension of classificatory variation among competing models of word competence, proves to be orthogonal to the wealth of psycholinguistic and neurolinguistic evidence on human processing behavior accrued over the last decades. No matter whether affixes are conceived of as ontological units existing independently of lexical items and stored alongside with them, or, rather, as on-line processing effects of word production/recognition, both views can be reconciled with evidence of human selective sensitivity to lexical structure and frequency effects. This by no means implies that the theoretical distinction going back to Hockett (1954) is irrelevant. In fact, Hockett's concerns appear to be supported by the mounting awareness that, contrary to classical generative assumptions, knowledge of "what" (stored representations) and knowledge of "how" (processing principles) can hardly be decoupled in the debate on what constitutes human word knowledge (Pirrelli 2018). In the post-Bloomfieldian flourishing of word models, Hockett's prescient concerns have been largely misinterpreted as supporting some architectures for word processing at the expenses of some others.

Likewise, the use of predefined morphological categories and established theoretical frameworks to understand the physiology and development of language skills in children seems to suggest that the role of mind and brain sciences is to simply validate existing linguistic categories and theories. This is another misconception. With a few exceptions, theoretical approaches were established outside the field of child language acquisition, and, in most cases, predate it. Data and categories from linguistics are certainly key to carving out areas of language development and defining scientific objectives. Nonetheless, linguistic terms and categories should be used with extreme care in the context of language acquisition, and should be validated against the specific goals and independent methodological requirements of this research domain. The simplifying assumption that linguistic categories are, as such, a reflection of the child's mind (rather than a working hypothesis) is as dangerous as blatantly ignoring these categories.

Marr's (1982) hierarchy of epistemological levels of understanding of complex systems offers a valid methodological scaffolding for discussing interdisciplinary issues in language sciences on a principled footing. The hierarchy is useful to distinguish between knowledge of what we do when we use language (Marr's "computational" level), knowledge of how we do it (his "algorithmic" level), and knowledge of how this ability is neurobiologically "embodied" in the brain (his "implementational" level). The distinction was intended to emphasize that each such level can, in principle, be investigated independently, through its own concepts and level-specific objects of inquiry. Nonetheless, a full understanding of a complex system ultimately requires integration of multiple perspectives, with each level being assessed on its own merits, for its intended goals and limitations, but with acquisitions from one level constraining acquisitions of all other levels.

To illustrate, due to the dominant focus of theoretical linguistics on the basic units of language and the laws for their combination, linguists have laid more, if not exclusive emphasis on representation issues, with processing issues being comparatively neglected. The approach is in sharp contrast with the psycholinguistic and neurolinguistic prevalent concerns with behavioral and physiological evidence of the human processor. Pace Clahsen (2006), however, it is highly unlikely that the two perspectives can be related mutually through some form of direct correspondence. Evidence of different time-scale effects in the behavior of complex dynamic systems should warn us against the search for straightforward one-to-one relationships between either linguistic representations (Marr's computational level) and their behavioral correlates (Marr's algorithmic level), or observable processing effects (Marr's algorithmic level) and their neuroanatomical localizations (Marr's implementational level). The properties

of the whole linguistic system may well be constrained and shaped by the properties of its parts. But its causal dynamics are inherently multileveled (Corning 2004). This means that the performance of each part, and its functional role, can only be understood in terms of its interaction with other parts and the whole system. Inter-level mapping rarely implies the simple extrapolation to level Y of properties holding at level X. More often, it is a matter of discovering entirely new laws and concepts, and requires a creative shift from quantitative to qualitative differentiation (Anderson 1972).

All this may sound somewhat discouraging. After all, direct inter-level relationships would be by far simpler to understand and investigate than multileveled, non-linear relationships. But there is room for some hope when it comes to language. Recent advances in the technological and analytical weaponry of language sciences promise to provide the level of material continuity between empirical data and functional modeling that constitutes an essential precondition to concrete methodological unification of neighboring language domains. Looking at the boundary between cognitive psychology and neuroscience, for example, the advent and development of neuroimaging technology permitted *in vivo* investigation of the functional interconnection between brain data and psychological evidence, thus establishing a direct explanatory link and a causal continuity between observations and hypotheses in the two domains. Likewise, cognitively-motivated computational models of language processing, however admittedly simpler than the still poorly understood human mechanisms they are intended to simulate, may assist scientists in decomposing a complex process into a handful of interacting sub-processes, and may enable carrying out experiments under more favorable and controlled conditions than those holding for experiments with human subjects. Decade-long developments in Recurrent Neural Networks learning complex language units have proved instrumental in addressing a few open issues about the psychological nature of classical linguistic categories and basic units. In particular, the connectionist idea that storage and processing are two interlocked steps of a unique learning dynamic appears to provide an elegant solution to the linguistic conundrum of the appropriate ontological status of sublexical constituents (as either storage or processing units). This unifying view lends support to Poggio's (2010) claim that (language) learning is key to the appropriate methodological unification of Marr's epistemological levels. Accordingly, units in language crucially depend on the way they are acquired, organized and used by humans. Any form of ontological realism is, in this connection, rather dubious.

On a more analytical front, linear and non-linear regression models for the quantitative scrutiny of multifactorial language data, have considerably freed language data collection from the strict methodological constraints of prior

hypothesis testing, dispensing with the need for a tightly controlled and balanced protocol of data elicitation based on a clear experimental design. Such freedom in data elicitation, combined with the huge support of information and communication technologies to digital storage and cooperative efforts for data creation, has spawned the innovative development of “megastudies” (Keuleers and Marelli 2020, this volume) specifically designed to maximize utility, availability and reusability of behavioral data.

In our view, all these developments will have an increasingly large impact on data modeling for linguistic and psycholinguistic research. In a similar vein, distributional semantic models (Acquaviva et al. 2020, this volume; Jones et al. 2015; Landauer and Dumais 1997; Mikolov et al. 2013) have proved to be able to quantitatively assess the role of linguistic context in shaping word meanings, and in guiding speakers’ expectations about the typical events nouns participate in, and the typical arguments verbs subcategorize for. Accordingly, a distributional, graded interpretation of word meaning similarity is bound to have a considerable impact on psycholinguistic accounts of morpho-semantic opacity/transparency effects in word processing (Dressler 2005; Kilani-Schoch and Dressler 2005; Mayerthaler 1981), moving away from Frege’s (1891) logical principle of compositionality, according to which an expression is either fully transparent or opaque. In Marelli, Gagné and Spalding’s (2017) CAOSS model, for example, relational effects in compound interpretation are modeled as the by-product of nuanced operations across patterns of word distributions.

Similarly, the information theoretic notion of entropy has been used to model the discriminative power of words in context, thereby offering a quantitative measure of the elusive notion of salience against a background of contextual events (Keuleers et al. 2020, this volume). In addition, the *Low Entropy Conjecture* (Ackerman and Malouf 2013) is based on the role of implicative paradigmatic relations in allowing speakers to infer an unseen inflected form from its paradigm companions. The fact that, cross-linguistically, inflectional paradigms tend to exhibit low expected conditional entropy, i.e. low uncertainty in intra-paradigmatic inference, can thus be interpreted as meeting some basic learnability requirements. Once more, insights from information theory and from computational modeling of usage-based theories have made it possible to see competing views and diverging perspectives subjected to critical assessment on experimental grounds. We welcome this as an important precondition to rapid progress in the field.

To our knowledge, no other existing single publication covers, in such a highly complementary and interdisciplinary way, as many different approaches to word knowledge and usage as the present volume does. We are deeply grateful to all contributing authors for sharing with us the view that interdisciplinary

crosstalk is indeed possible, and for taking much of their time and effort to prove its merits. We believe that this book will be beneficial for diverse types of readers and we hope its open access publication will make its impact and influence wide and durable. Young researchers, who already see a clear advantage in the synergic integration of traditionally segregated competences, will find, here, useful material and pointers for developing a truly interdisciplinary curriculum. Single-domain specialists, interested in knowing more about how their expertise can contribute to understanding issues of common interest when approached by other disciplines, will look for methodological guidelines and open issues to be investigated through interdisciplinary cooperation. Finally, both specialist and non-specialist readers will be offered easily accessible, state-of-the-art information, covering interconnected areas of lexical expertise that are rarely discussed and comparatively assessed within a single book.

Acknowledgments: We gratefully acknowledge the European Science Foundation Research Networking Programmes, and the former ESF Standing Committee for the Humanities for their great foresight and support, and for making it possible to publish the present editorial effort as an open access volume with generous funding of the NetWordS project in the years 2011–2015. We express our hope that similar programmes will continue to be launched and funded in Europe in the years to come.

References

- Ackerman, Farrell & Robert Malouf. 2013. Morphological organization: the low conditional entropy conjecture. *Language* 89. 429–464.
- Arnon, Inbal, Stewart M. McCauley & Morten H. Christiansen. 2017. Digging up the building blocks of language: Age-of-acquisition effect for multiword phrases. *Journal of Memory and Language* 92. 265–280.
- Arnon, Inbal & Neal Snider. 2010. More than words: frequency effects for multiword phrases. *Journal of Memory and Language* 62. 67–82.
- Anderson, Philip W. 1972. More Is Different. *Science* 4047 (177). 393–396.
- Bannard, Colin & Danielle E. Matthews. 2008. Stored word sequences in language learning: the effect of familiarity on children's repetition of four-word combinations. *Psychological Science* 19 (3). 241–248.
- Baroni, Marco, Emiliano Guevara and Vito Pirrelli. 2007. NN compounds in Italian: modelling category induction and analogical extension. *Lingue e Linguaggio*, 6 (2). 263–290.
- Ben Shalom, Dorit & David Poeppel. 2008. Functional anatomic models of language: assembling the pieces. *The Neuroscientist* 14 (1). 119–127.
- Berger, Adam L., Vincent J. Della Pietra & Stephen A. Della Pietra. 1996. A maximum entropy approach to natural language processing. *Computational linguistics* 22 (1). 39–71.

- Brown-Schmidt, Sarah. 2009. The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin and Review* 16 (5). 893–900.
- Clahsen, Harald. 2006. Linguistic perspectives on morphological processing. In Dieter Wunderlich (ed.), *Advances in the Theory of the Lexicon*, 355–388. Mouton De Gruyter.
- Corning, Peter A. 2004. *Holistic Darwinism*. University of Chicago Press.
- Davis, Matthew H. 2015. The Neurobiology of Lexical Access. In Gregory Hickok & Steven L. Small (eds.), *Neurobiology of Language*, 541–555. Associated Press.
- D’Esposito, Mark. 2007. From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 362 (1481). 761–772.
- Dressler, Wolfgang U. 2005. Word-formation in Natural Morphology. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of Word Formation*, 267–284. Springer.
- Elman, Jeffrey L. 2004. An alternative view of the mental lexicon. *Trends in Cognitive Sciences* 8 (7). 301–306.
- Elman, Jeffrey L. 2009. On the Meaning of Words and Dinosaur Bones: Lexical Knowledge Without a Lexicon. *Cognitive Science* 33. 1–36.
- Ferro, Marcello, Dimitri Ognibene, Giovanni Pezzulo & Vito Pirrelli. 2010. Reading as active sensing: a computational model of gaze planning during word recognition. *Frontiers in Neurobotics* 4 (6).
- Frege Gottlob. 1891. *Funktion und Begriff*. Pohle.
- Grimm, Robert, Giovanni Cassani, Steven Gillis & Walter Daelemans. 2017. Facilitatory Effects of Multi-Word Units in Lexical Processing and Word Learning: A Computational Investigation. *Frontiers in Psychology* 8 (555).
- Hasson, Uri, Giovanni Egidi, Marco Marelli & Roel M. Willems. 2018. Grounding the neurobiology of language in first principles: The necessity of non-language-centric explanations for language comprehension. *Cognition* 180. 135–157.
- Hickok, Gregory & David Poeppel. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92. 67–99.
- Hockett, Charles F. 1954. Two models of grammatical description. *Word*, 10. 210–231.
- Huetting, Falk. 2015. Four central questions about prediction in language processing. *Brain Research* 1626. 118–135.
- Jackendoff, Ray. 2000. Fodorian Modularity and Representational Modularity. In Yosef Grodzinsky, Lewis P. Shapiro & David Swinney (eds.), *Language and the Brain. Representation and Processing*, 3–30. San Diego: Academic Press.
- Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press.
- Jackendoff, Ray. 2007. *Language, Consciousness, Culture: Essays on Mental Structures*. The MIT Press.
- Jones, Michael N., Jon A. Willits & Simon Dennis. 2015. Models of Semantic Memory. In Jerome R. Busemeyer, Zheng Wang, James T. Townsend & Ami Eidels (eds.), *Oxford Handbook of Mathematical and Computational Psychology*, 232–254. (Oxford Library of Psychology). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199957996.013.11>
- Kilani-Schoch, Marianne, Ingrida Balčiunienė, Katharina Korecky Kröll, Sabine Laaha & Wolfgang U. Dressler. 2009. On the role of pragmatics in child-directed speech for the acquisition of verb morphology. *Journal of Pragmatics* 41. 219–239.
- Kilani-Schoch, Marianne & Wolfgang U. Dressler. 2005. *Morphologie naturelle et flexion du verbe français*. Narr.

- Kukona, Anuenue, Shin-Yi Fang, Karen A. Aicher, Helen Chen & James S. Magnuson. 2011. The time course of anticipatory constraint integration. *Cognition* 119 (1). 23–42.
- Kuperberg, Gina R. & T. Florian Jaeger. 2016. What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* 31 (1). 32–59.
- Landauer, Thomas K. & Susan T. Dumais. 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review* 104 (2). 211–240.
- Libben Gary. 2005. Everything is psycholinguistics: Material and methodological considerations in the study of compound processing. *Canadian Journal of Linguistics* 50. 267–283.
- Libben, Gary. 2016. The quantum metaphor and the organization of words in the mind. *Journal of Cultural Cognitive Science* 1. 49–55.
- Marelli, Marco, Christina L. Gagné & Thomas L. Spalding. 2017. Compounding as Abstract Operation in Semantic Space: Investigating relational effects through a large-scale, data-driven computational model. *Cognition* 166. 207–224.
- Marr, David (1982). *Vision*. San Francisco: W.H. Freeman.
- Mattiello, Elisa. 2017. *Analogy in Word Formation. A study of English neologisms and occasionalisms*. De Gruyter.
- Mayerthaler, Willi. 1981. *Morphologische Natuerlichkeit*. Athenaeum.
- Mitchell, Jeff & Mirella Lapata. 2010. Composition in distributional models of semantics. *Cognitive science* 34 (8). 1388–1429.
- Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S. Corrado, & Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In Christopher J. C. Burges, Leon Bottou, Zoubin Ghahramani & Kilian Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013*, 3111–3119.
- Milligan, Karen, Janet Wilde Astington & Lisa Ain Dack. 2007. Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Development* 78 (2). 622–646.
- Pickering, Martin J. & Simon Garrod. 2007. Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences* 11 (3). 105–110.
- Pickering, Martin J. & Simon Garrod. 2013. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences* 36. 329–392.
- Pirrelli, Vito. 2007. Psycho-computational issues in morphology learning and processing. An overture. *Lingue e Linguaggio* 6 (2). 131–138.
- Pirrelli, Vito. 2018. Morphological Theory and computational linguistics. In Jenny Audring & Francesca Masini (eds.), *The Oxford Handbook of Morphological Theory*, 573–593. Oxford University Press.
- Plag, Ingo. 2018. *Word-formation in English*. Cambridge University Press.
- Price, Cathy J. 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences* 1191 (1). 62–88.
- Price, Cathy J. 2012. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62 (2). 816–847.
- Price, Cathy J. 2017. The evolution of cognitive models: From neuropsychology to neuroimaging and back. *Cortex* 107. 37–49.
- Poggio, Tomaso. 2010. Afterword. Marr's Vision and Computational Neuroscience. In David Marr, *Vision*, 362–367. MIT Press.

- Riest, Carina, Annett B. Jorschick & Jan P. de Ruiter. 2015. Anticipation in turn-taking: mechanisms and information sources. *Frontiers in Psychology*, 6: art. 89.
- Saxton, Matthew. 2008. What's in a name? Coming to terms with the child's linguistic environment. *Journal of Child Language* 35 (3). 677–686.
- Saxton, Matthew. 2009. The inevitability of Child Directed Speech. In Susan Foster-Cohen (ed.), *Language acquisition*, 62–86. Palgrave Macmillan.
- Seidenberg, Mark S., Maryellen C. MacDonald. 1999. A Probabilistic Constraints Approach to Language Acquisition and Processing. *Cognitive Science* 23 (4). 569–588.
- Sivanova-Chanturia, Anna, Kathy Conklin, Sendy Caffarra & Edith Kaan. 2017. Representation and processing of multi-word expressions in the brain. *Brain and Language* 175. 111–122.
- Smith, Nathaniel J. & Roger Levy. 2013. The effect of word predictability on reading time is logarithmic. *Cognition* 128 (3). 302–319.
- Spivey, Michael & Michael K. Tanenhaus. 1998. Syntactic ambiguity resolution in discourse: Modelling the effects of referential context and lexical frequency. *Journal of Experimental Psychology Learning Memory and Cognition* 24 (6). 1521–1543.
- Tabor, Whitney & Michael K. Tanenhaus. 1999. Dynamical Models of Sentence Processing. *Cognitive Science* 23 (4). 491–515.
- Taylor, Nicole, Wilberta Donovan, Sally Miles & Lewis Leavitt. 2009. Maternal control strategies, maternal language usage and children's language usage at two years. *Journal of Child Language* 36 (2). 381–404.
- Tremblay, Antoine & R. Harald Baayen. 2010. Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In David Wood (ed.), *Perspectives on formulaic language: Acquisition and communication*, 151–173. Bloomsbury Publishing.
- Vespignani, Francesco, Paolo Canal, Nicola Molinaro, Sergio Fonda & Cristina Cacciari. 2009. Predictive Mechanisms in Idiom Comprehension. *Journal of Cognitive Neuroscience* 22 (8). 1682–1700.
- Wellman, Henry M. 2002. Understanding the psychological world: Developing a theory of mind. In Usha Goswami (ed.), *Handbook of childhood cognitive development*, 167–187. Oxford, England: Blackwell.