

The European Science Foundation (ESF) is an independent, non-governmental organisation, the members of which are 80 national funding agencies, research performing agencies, academies and learned societies from 30 countries.

The strength of ESF lies in the influential membership and in its ability to bring together the different domains of European science in order to meet the challenges of the future.

Since its establishment in 1974, ESF, which has its headquarters in Strasbourg with offices in Brussels and Ostend, has assembled a host of organisations that span all disciplines of science, to create a common platform for cross-border cooperation in Europe.

ESF is dedicated to promote collaboration in scientific research, funding of research and science policy across Europe. Through its activities and instruments ESF has made major contributions to science in a global context. The ESF covers the following scientific domains:

- Humanities
- Life, Earth and Environmental Sciences
- Medical Sciences
- Physical and Engineering Sciences
- Social Sciences
- Marine Sciences
- Materials Science
- Nuclear Physics
- Polar Sciences
- Radio Astronomy Frequencies
- Space Sciences

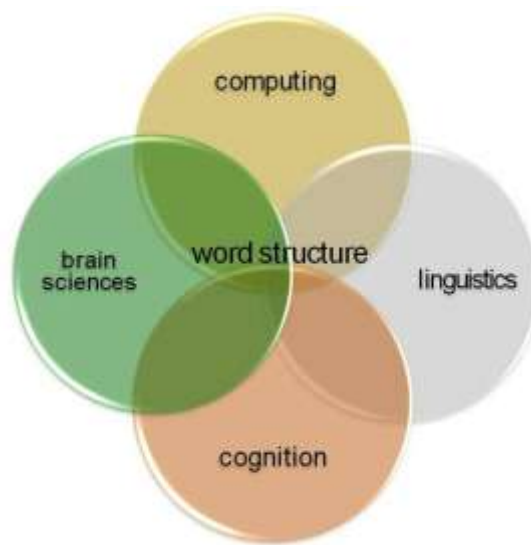
Introduction

Morphologically complex words are common to all European languages. They represent a fundamental part of what we mean by human language knowledge and the basic building blocks of language productivity. Nonetheless, words remain a challenging realm of scientific inquiry, at the interface between lexicon and grammar, requiring integration of a number of orthogonal disciplines and approaches, ranging from psycho- and neuro-linguistics, to theoretical, variationist and historical linguistics, to memory processes and computational models of (sub)symbolic processing.

Scientists all over Europe are currently pursuing important lines of work on word structure, mostly supported by nationally-funded projects or bi-lateral cooperation programmes. There nonetheless seems to be a growing need for a larger-scale integrated European effort, focusing on common medium-term objectives, to promote interdisciplinary cross-fertilisation and synergy, and to optimise research investments in terms of more convergent and complementary efforts. The European research scenario is particularly conducive to these goals, due to the robustly empirical character of its methodological stance and the unique range of relevant scientific domains where European scientists appear to have, at present, a huge potential for major breakthroughs.

By bringing together experts of various research fields (from brain sciences and computing to cognition and linguistics) and of different theoretical inclinations, **NetWordS** intends to advance our current awareness of the 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 physiology. This will be achieved through knowledge networking and dissemination and scientific meetings organised over a four-year period.

Moreover, the Research Networking Programme aims to have a highly interdisciplinary profile, to promote training and development of young scientists through short visits and exchange grants, and to encourage the integration of new partners.



The running period of the ESF NetWordS Research Networking Programme is four years, from May 2011 to April 2015.

Scientific context

Words are the basic building blocks of language productivity, establishing the most immediate connections between language and our conceptualisation of the outside world. Besides, they represent complex interface units, which are not only parts of larger constructions such as phrases or sentences, but are themselves, in all European languages, made up out of simpler meaningful sub-lexical constituents, such as roots and affixes.

Such a dual status of morphologically complex words, at the interface between lexicon and grammar, raises a number of fundamental questions, many of which are still unanswered. 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? Do both knowledge-based factors, such as formal regularity and semantic transparency, and usage-driven factors, such as word length and frequency, play any role in this? Does word-level knowledge require parallel development of form and meaning representations, or do the latter develop independently at a different pace? How do word meanings function and combine in communicative contexts and evolve through learning? Does lexical knowledge affect on-line processing? Do the dramatic differences in word structures across the languages of Europe impact on processing models worked out on the basis of a single language? What neurobiological patterns of connectivity sustain word processing and storage in the brain?

Almost all levels of language knowledge and processing (from phonology to syntax and semantics) are known to be affected by knowledge of word structure at varying degrees. A better understanding of the human strategies involved in learning and processing word structure thus lies at the heart of our comprehension of the basic mechanisms serving both language and cognition and is key to addressing the three fundamental challenges for the study of the

physiology of grammar that are described in more detail below.

Lexicon and Grammar

According to dual-route approaches to word structure, recognition of a morphologically complex input word involves two interlocked steps: i) preliminary full-form access to the lexicon, ii) optional morpheme-based access of sub-word constituents of the input word, resulting from application of combinatorial rules taking care of on-line word segmentation. Algorithmically, step ii) is taken if and only if step i) fails to find any matching access entry in the lexicon. Such a view, recently challenged by several scholars, rests on the hypothesis of a direct correspondence between principles of grammar organisation (lexicon vs. rules), processing correlates (storage vs. computation) and localisation of the cortical areas functionally involved in word processing. Although such a direct correspondence is probably the most straightforward model of the grammar-processing relation, it may only be the artefact of outdated views of memory as rote storage. In fact, other theoretical models have put forward a more nuanced indirect correspondence hypothesis. For instance, in the Word-and-Paradigm tradition, fully inflected forms are associatively related through possibly recursive paradigmatic structures, defining entailment relations between forms. Any serious appraisal of such an indirect correspondence requires extensive empirical testing on a wide array of morphologically rich languages of the sort spoken in Europe, and is likely to exceed the limits of both human intuition and box-and-arrow models of cognition. We believe that increasing availability of multi-lingual data sets and computer models of language learning and processing will have much to say in this respect in the near future.

Another fundamental open issue is how theoretical models relate to neurobiologically-grounded models and theories of word structure. Recent evidence of automatic sub-lexical segmentation of monomorphemic words triggered by pseudo inflectional endings lends support to a less deterministic and modular view of the interaction between

entrenchment of lexical units with productivity, internal structure and ease of interpretation.

These aspects agree with well-established psycholinguistic evidence that language comprehension is highly incremental, with readers and listeners continuously updating the meaning of utterances as they parse them. Much recent research suggests that language comprehension can be highly predictive, as long as the linguistic and non-linguistic context supports these predictions. Prediction can also be used to compensate for problems with noisy or ambiguous input and may explain the human advantage in parsing morphologically irregular forms (where morphosyntactic and morpholexical features are marked through extended exponence) over morphologically regular forms (where a morphological exponent systematically follows a full stem).

A parsimonious explanation of anticipatory mechanisms of language comprehension is that prediction uses some components for language production. There is indirect empirical evidence pointing in this direction: listeners activate the appropriate articulatory cortical areas for tongue and lips while listening to speech and brain areas that are

associated with production during aspects of comprehension from phonology to narrative structure. This is in keeping with evidence of activation of mirror neurons in monkeys by perceptual predictions and perceived actions, but may also be understood as involving context-sensitive language ‘emulators’. In turn, anticipatory mechanisms of language comprehension may be closely related to mechanisms for Short Term Memory content rehearsal such as Baddeley’s phonological loop.

All of this points to a converging trend between computational and cognitive lines of scientific inquiry, supporting the view that grammar and lexical competence are acquired through minimal steps, shaped up by performance-driven factors such as memory limitations, frequency-based sensitivity, and modality-specific constraints, ultimately blurring the dichotomy between language knowledge and usage.

Words and Meanings

By exchanging words in ecological settings, we share, assess, modify, extend and structure our ‘semantic memory’. Yet, the nature and content of such memory, the

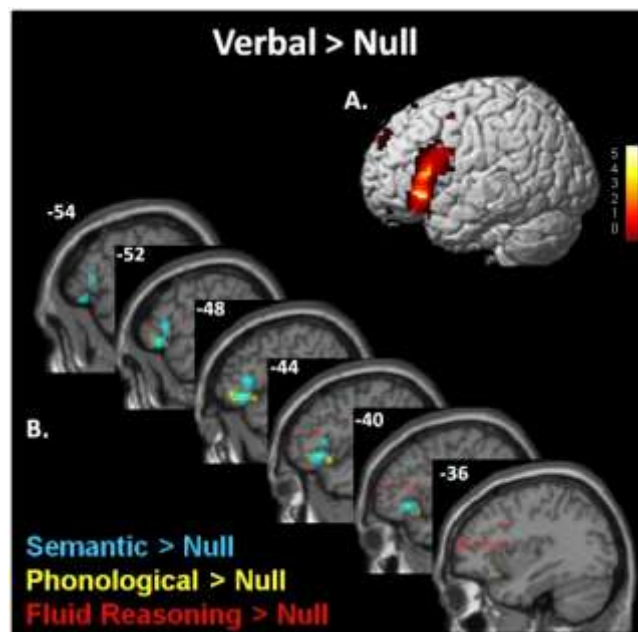


Figure 2 – Left lateral PFC involvement in verbal fluid reasoning. (A) Brain rendering showing activations for Verbal Fluid Reasoning > Null contrast for all participants. (B) Left lateral PFC regions active in the semantic reasoning task (cyan), in the phonological reasoning task (yellow), and common fluid reasoning regions across verbal and visual domains (red). x coordinates are provided for sagittal views. (reproduced © courtesy of Basque Center on Cognition, Brain and Language, Donostia, San Sebastián, Basque Country, Spain).

principles of its associative organisation and internal structure, the developmental role of the dynamic interaction between linguistic form, meaning and sensing are among the most controversial issues in the current linguistic and neuro-cognitive debate.

Suggestions in the literature range from relatively abstract representations, including hierarchical semantic networks and lexical conceptual structures, to more concrete perceptual- or motor-based representations. Each of these approaches faces difficulties. Abstract representations elude the issue of symbol interpretation by severing meaning from our system of experiences of the external world. On the other hand, linguistic units can combine and behave distributionally in ways that are not strictly predictable from their semantic properties. Inferences, sense extensions, metaphors and processes of concept composition and coercion show that grounded sensory motor knowledge does not suffice to account for our ability to extract meaning from language. Intermediate hypotheses need to be entertained and empirically assessed, casting meaning as abstract, schematic representations, based on linguistically articulated, structured knowledge and word co-occurrences in large text samples, which are nonetheless

embodied in human perceptual and motor systems. Researchers working in a neurocomputational framework have recently addressed issues of semantic knowledge arising from patterns of combinatorial information using more brain-like neural network simulations.

Interpretation of Noun-Noun compounds such as 'bread knife' requires integration of the meaning representations associated with the two constituent nouns and independently accessed from the lexicon. However, it has recently been shown that access to conceptual representations is considerably more dynamic and context-sensitive, so that the whole construction appears to prompt a process of selective activation of contextually-relevant semantic properties. From a computational standpoint, constraint-satisfaction approaches made the interesting suggestion that the interpretation of a complex construction makes use of pre-compiled, schematised information, memorised in the mental lexicon and applied probabilistically.

These aspects bring in the issue of interactive negotiation of referential and intentional word meanings in the process of learning word usages in daily communicative exchanges. Lexical pragmatics investigates the processes by which linguistically-specified

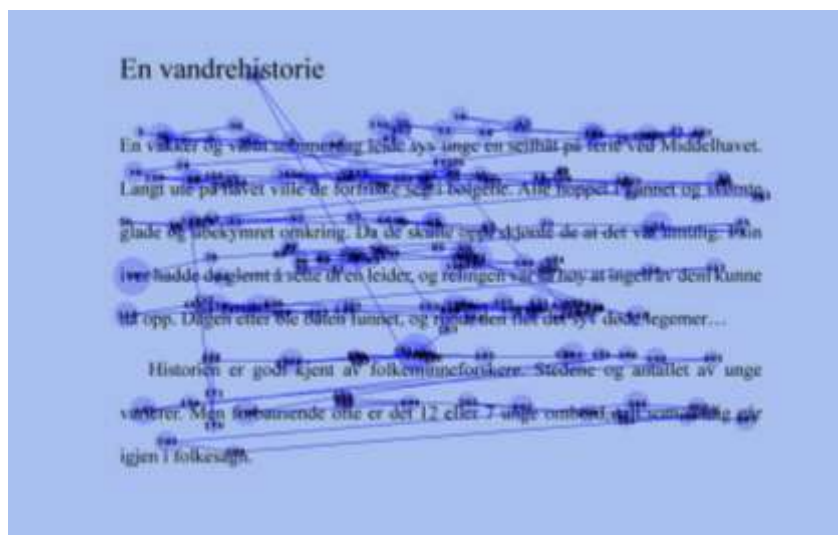


Figure 3 – Text processing in dyslexia. An example of reading patterns from a study comparing dyslexic and non-dyslexic participants. Dots represent fixations and the dots' sizes are related to fixation length. Numbering of and lines between dots indicate the order of the fixation sequence. (reproduced © courtesy of Language Acquisition and Language Processing Lab, Norwegian University of Science and Technology, Trondheim, Norway)

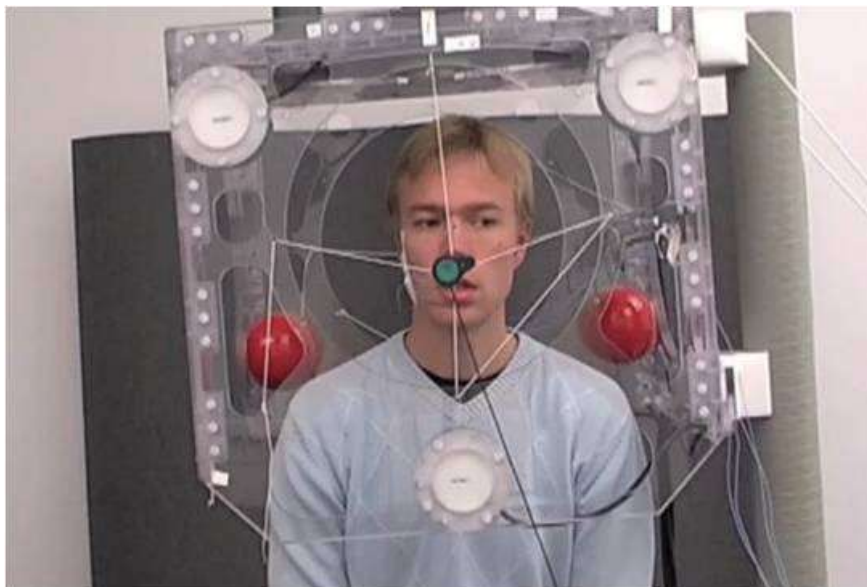


Figure 4 – Articulograph AG500 allowing sensors placed on the tongue, lips and reference points to be tracked in 3D using an electromagnetic field with high spatial and temporal resolution (reproduced © courtesy of Humanities Lab, Lund University, Sweden)

(i.e. literal) word meanings are modified in use on the basis of factors related to pragmatic competence, such as knowledge of the specific communicative context, knowledge about the co-conversant(s), knowledge about the specific ongoing task and general knowledge of the world. Mediation of all these factors is key to understanding the ontogenesis of word meaning and its creative usage in daily conversation, as illustrated by so-called 'one-off compounds' such as Downing's 'apple juice seat' example.

- * exploring the implications of domain-specific approaches for other fields
- * testing claims by broadening the empirical basis for their support
- * examining whether extensions of theoretical claims can be developed
- * promoting interdisciplinary cross-fertilisation and synergy
- * focusing on common medium-term objectives
- * optimising research investments in terms of convergent and complementary efforts.

Aims and Objectives

Europe has firm and deep roots in as diverse disciplinary fields as theoretical models of language architecture, brain sciences, cognitive modelling, language development, short-term and long-term memory processes, psycho-computational models of lexical processing and storage, predictive models of language behaviour, machine learning, diachronic, diamesic and diastratic evidence of language varieties. The present healthy condition of European research makes time ripe for a larger-scale cross-disciplinary European effort into word structure aiming at:

There is growing awareness that interdisciplinary cooperation in this area will have much wider chances of success than traditional specialist work in highly focused knowledge domains and that it will consolidate European excellence in the field. By bringing together European (and non-European) experts in complementary knowledge areas, **NetWordS** rises to this challenge by setting common research priorities, developing joint training programmes and establishing virtual cross-disciplinary laboratories and research infrastructures. Collaboration will unfold through the following steps:

- * discuss and develop consensual word representations in context
- * establish common experimental protocols and suggest novel ones
- * take stock of and integrate data based on the large array of European languages
- * transfer best practice in use of new computational and statistical techniques for lexicon modelling
- * share experimental evidence, software and equipment
- * facilitate, through community building, the development of optimum cross-disciplinary and cross-linguistic research strategies
- * prompt and extend collaboration between partners and link European activities with the wider community world-wide.

Programme activities and instruments

Scientific Partnerships

NetWordS promotes development of interdisciplinary transnational scientific partnerships through **short-visit grants**, that are assigned yearly on the basis of **open calls** for short-term project proposals. Scholars taking part in interdisciplinary activities funded through NetWordS grants convene periodically to discuss and disseminate results. Mature results are also expected to be disseminated and published through newsletters, existing major international conferences, journals and edited books. Short-visit grants are also geared towards planning focused collaborative work, with a view to catalysing credible large-scale proposals within more application-oriented European projects and initiatives.

Training

NetWordS promotes training of young scientists through two instruments: **exchange grants** and **summer schools**. Exchange grants are meant to cover long stays (up to 6 months) of promising junior researchers at leading hosting labs. Grants are assigned yearly on the basis of **open calls**.

Two summer schools, planned to take place in **2012** and **2014**, are aimed at bringing up a new generation of PhD students and young researchers with a truly interdisciplinary

background, and shaping new ways of thinking that are conducive to cross-disciplinary breakthroughs.

Scientific Meetings

NetWordS organises yearly **workshops** on inter-disciplinary issues in word structure. The usual time frame is between late November and early December. A major final conference is planned to take place in 2015.

Discussion and Networking

To maximise synergy, NetWordS sets itself the goal of defining a priority list of shared research topics of general interest (e.g., word reading, interpretation of complex nominals, developmental acquisition of inflection, etc.) that will be addressed and discussed from the perspective of the three above-listed challenges (Lexicon and Grammar, Word Knowledge and Word Use, Words and Meanings) through NetWordS **internet forums**. Forums are intended to stimulate discussion on domain-specific approaches and explore ways of integrating and extending current approaches also through data sharing.

Global dimension

NetWordS promotes a global world-wide collaborative dimension through established links with other cognate research initiatives such as the *Mental Lexicon Research Group* in Canada.

Joining NetWordS

NetWordS encourages **integration** of new European partners through an open cooption scheme. For the latest information about the programme and if you wish to know more about how to join NetWordS, consult the NetWordS websites or use the email contact:

www.esf.org/networks

www.networks-esf.eu

info-networks@ilc.cnr.it

Funding Organisations

ESF Research Networking Programmes are principally funded by the Foundation's Member Organisations on an *à la carte* basis. NetWordS is supported by:

- **Fonds zur Förderung der Wissenschaftlichen Forschung Österreich (FWF)**
Austrian Science Fund, Austria
- **Fonds voor Wetenschappelijk Onderzoek-Vlaanderen (FWO)**
Research Foundation Flanders, Belgium
- **Nacionalna zaklada za znanost (NZZ)**
National Foundation for Science, Croatia
- **Det Frie Forskningsråd – Kultur og Kommunikation (FKK)**
The Danish Council for Independent Research – Humanities, Denmark
- **Suomen Akatemia**
Academy of Finland / Research Council for Culture and Society, Finland
- **Centre National de la Recherche Scientifique (CNRS)**
National Centre for Scientific Research – France
- **Deutsche Forschungsgemeinschaft (DFG)**
German Research Foundation, Germany
- **Országos Tudományos Kutatási Alapprogramok (OTKA)**
Hungarian Research Fund, Hungary
- **Irish Research Council for the Humanities and Social Sciences (IRCHSS)**
Ireland
- **Consiglio Nazionale delle Ricerche (CNR)**
National Research Council, Italy
- **Norge Forskningsråd (NCR)**
Research Council of Norway, Norway
- **Slovenská Akadémia Vied (SAV)**
Slovak Academy of Sciences, Slovak Republic
- **Javna agencija za raziskovalno dejavnost Republike Slovenije (SRA)**
Slovenian Research Agency, Slovenia
- **Departamento de Educacion, Universidades e Investigacion**
País Vasco, Spain

- **Forskningsrådet för Arbetsliv och Socialvetenskap (FAS)**
Swedish Research Council, Sweden

NetWordS Steering Committee

Vito Pirrelli (*Chair*)

Institute for Computational Linguistics
Italian National Research Council
Pisa, Italy
Email: vito.pirrelli@ilc.cnr.it

Wolfgang Ulrich Dressler

Austrian Academy of Sciences
Vienna, Austria
Email: wolfgang.dressler@univie.ac.at

Walter Daelemans

University of Antwerp
Antwerp, Belgium
Email: walter.daelemans@ua.ac.be

Ida Raffaelli

University of Zagreb
Zagreb, Croatia
Email: ida.raffaelli@ffzg.hr

Hanne Ruus

Department of Scandinavian Studies – University of Copenhagen, Denmark
Email: haru@hum.ku.dk

Krista Lagus

Aalto University, School of Science
Aalto, Finland
Email: krista.lagus@tkk.fi

Hélène Giraud

Laboratoire CLLE – Maison de la Recherche
Université de Toulouse le Mirail
Toulouse, France
Email: helene.giraud@univ-tlse2.fr

Ingo Plag

University of Siegen
Siegen, Germany
Email: plag@anglistik.uni-siegen.de

Gábor Prószéky

Pázmány Péter Catholic University
Budapest, Hungary
Email: proszeky@itk.ppke.hu

Paolo Acquaviva

School of Languages and Literatures
University College Dublin
Belfield, Dublin, Ireland
Email: paolo.acquaviva@ucd.ie

Mila Dimitrova Vulchanova

Norwegian University of Science and Technology
Trondheim, Norway
Email: mila.vulchanova@ntnu.no

Radovan Garabík

Slovak Academy of Sciences
Bratislava, Slovak Republic
Email: radog@juls.savba.sk

Tomaz Erjavec

Department of Knowledge Technologies
Jozef Stefan Institute
Ljubljana, Slovenia
Email: tomaz.erjavec@ijs.si

Manuel Carreiras

Basque Centre on Cognition, Brain and Language
St. Sebastian, Spain
Email: m.carreiras@bcbl.eu

Carita Paradis

Centre for Languages and Literature
Lund University
Lund, Sweden
Email: Carita.Paradis@englund.lu.se

Programme Coordinator

Claudia Marzi

Institute for Computational Linguistics
Italian National Research Council
Pisa, Italy
Email: claudia.marzi@ilc.cnr.it
coordination-networks@ilc.cnr.it

ESF Liaison

Barry Dixon

Science Officer

Marie-Laure Schneider

Administration

European Science Foundation
1 quai Lezay-Marnésia
BP 90015
67080 Strasbourg cedex
France
Tel: +33 (0)3 88 76 71 39
Fax: +33 (0)3 88 37 05 32
Email: sch-rnp@esf.org

For the latest information on this Research
Networking Programme consult the
NetWordS website
www.networks-esf.eu

or visit the following ESF link
www.esf.org/networks

or write to info-networks@ilc.cnr.it
