Interdisciplinarity in language sciences:
an integrative perspective

vito pirrelli

Communication Physiology Lab
Institute for Computational Linguistics
National Research Council – Pisa, Italy

ComPhys
physiology of communication

Net WordS

the context

- the current international scientific and technological scenario presents us with two important discontinuities with the past:
  - *exponential growth of technological innovation*
  - *growing interdisciplinary integration*

- need to adjust objectives and methodologies to a changing knowledge context
  - *evolving methodology*
  - *inter-personal and interdisciplinary collaboration*
  - *problem-centred focus*

- "the world has complex problems, universities have departments ..."
  - *global warming*
  - *sustainable economic growth*
  - *... language as a means to acquire, create and transfer knowledge*

language in a global context

- language is the most fundamental human means to process, acquire, create and transfer knowledge
- multi-level, multi-scale phenomenon
  - *physical level: what is the material substance of language?*
    - sounds
    - articulatory mechanics
    - physical channels and their properties
  - *neurophysiological: how is this implemented in the brain-ware?*
    - which brain regions are involved?
    - how do they interact?
  - *psychological level: how do humans use language?*
    - perception
    - production
  - *formal level: how is this information structured*
    - levels of embedding
    - combinatorial and distributional properties
  - *computational level: how is language processed?*
    - storage and processing
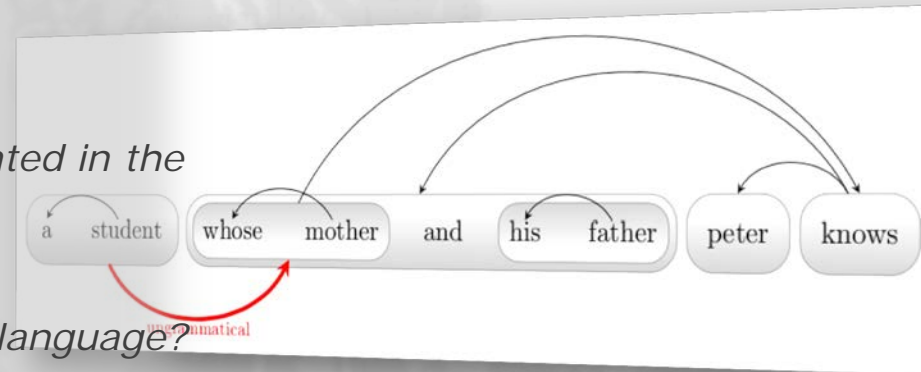 - *cultural level: how does language affect communication and culture?*

physical interactions -> speech processing

- *physical level: what is the material substance of language?*
  - **sounds**
  - **articulatory mechanics**
  - **physical channels and their properties**

- *neurophysiological: how is this implemented in the brain-ware?*
  - which brain regions are involved?
  - how do they interact?

- *psychological level: how do humans use language?*
  - perception
  - production

- *formal level: how is this information structured*
  - levels of embedding
  - combinatorial and distributional properties

- *computational level: how is language processed?*
  - **storage and processing**

- *cultural level: how does language affect communication and culture?*

speech processing

- – *automated speech recognition*
  - • voice user interfaces
  - • speech to text technologies for narrative texts
  - • voice recognition or speaker identification

- – *speech synthesis*
  - • assistive technology for people with visual impairment
  - • dyslexia
  - • speech impairment
  - • second language acquisition

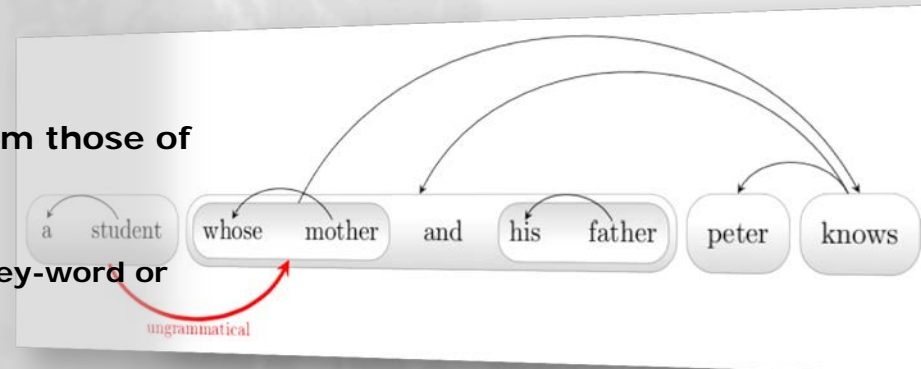# logico-computational interactions -> natural language processing

– *physical level: what is the material substance of language?*
  • sounds
  • articulatory mechanics
  • physical channels and their properties

– *neurophysiological: how is this implemented in the brain-ware?*
  • which brain regions are involved?
  • how do they interact?

– *psychological level: how do humans use language?*
  • perception
  • production

– ***formal level: how is this information structured***
  • **levels of embedding**
  • **combinatorial and distributional properties**

– ***computational level: how is language processed?***
  • **storage and processing**

– *cultural level: how does language affect communication and culture?*

logico-computational interactions -> natural language processing

- *machine translation*
- *automatic summarization*
  - produce a readable summary of a text
- *word sense disambiguation*
  - tell occurrences of '*bank*' of a river from those of '*bank*' as a financial institution
- *information retrieval*
  - find a set of documents relevant to a give key-word or topic (e.g. "global warming")
- *information extraction*
  - extract from documents relevant pieces of information (e.g. what is "information extraction" about?)
- *optical character recognition*
  - given an image of a printed text, determine the corresponding text
- *text readability*
  - how easy to read is a text?
- *… many more*

anthropo-sociological interactions -> language communication

- *physical level: what is the material substance of language?*
  - sounds
  - articulatory mechanics
  - physical channels and their properties

- *neurophysiological: how is this implemented in the brain-ware?*
  - which brain regions are involved?
  - how do they interact?

- **psychological level: how do humans use language?**
  - **perception**
  - **production**

- *formal level: how is this information structured*
  - levels of embedding
  - combinatorial and distributional properties

- *computational level: how is language processed?*
  - storage and processing

- **cultural level: how does language affect communication and culture?**

# anthropo-sociological interactions -> language communication

- *language as social interaction*
  - speaker's anticipation of hearer's needs
  - prediction-based processing
  - child-directed vs. adult-directed speech

- *language evolution and language change*

- *language and learning:*
  - how do we learn abstract concepts?
  - how do we learn novel words?

- *does language shape up our thoughts?*

- *language and culture*

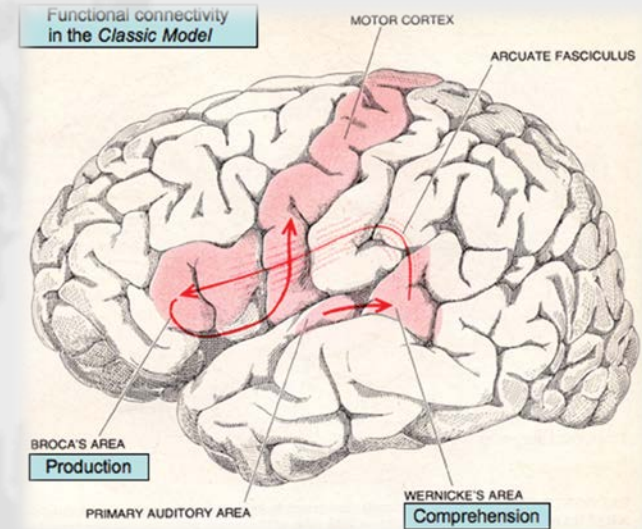# integrative interactions -> emergence from the brain

- *physical level: what is the material substance of language?*
  - sounds
  - articulatory mechanics
  - physical channels and their properties

- **neurophysiological: how is this implemented in the brain-ware?**
  - **which brain regions are involved?**
  - **how do they interact?**

- **psychological level: how do humans use language?**
  - **perception**
  - **production**

- *formal level: how is this information structured*
  - levels of embedding
  - combinatorial and distributional properties

- **computational level: how is language processed?**
  - **storage and processing**

- *cultural level: how does language affect communication and culture?*
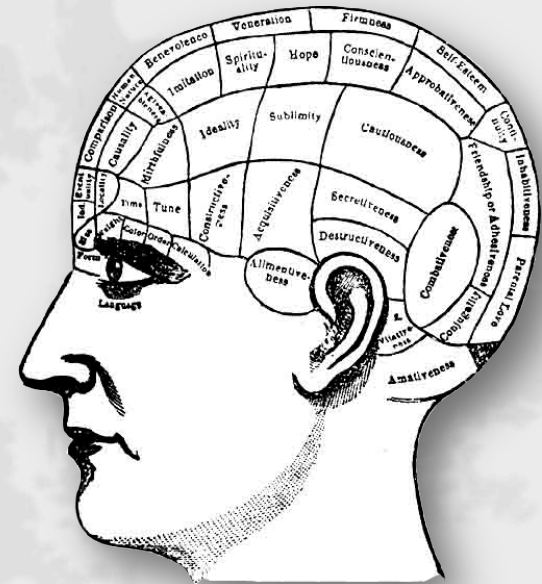
integrative interactions -> emergence from the brain

– the anatomy of language has been investigated with **functional neuroimaging techniques** (PET, fMRI) for more than 20 years now, and brain areas associated with language processing have been identified consistently, leading to fairly uncontroversial conclusions

– the greatest challenge ahead of us is to understand how different brain regions **interact** with one another in language comprehension and production

– assuming that our current understanding of the general picture is true, the main task for future research will be to specify the details of the within-stream organization and computational operations and their functional relationship with specific aspects of language behaviour

Functional connectivity in the *Classic Model*

MOTOR CORTEX

ARCUATE FASCICULUS

BROCA'S AREA
Production
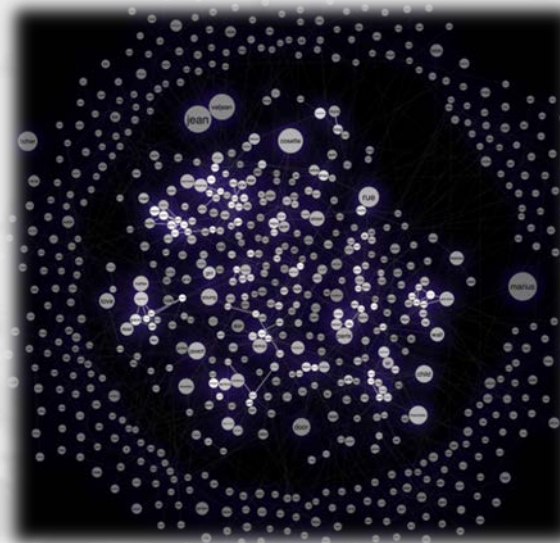
PRIMARY AUDITORY AREA

WERNICKE'S AREA
Comprehension

word self-organisation

- are words simply listed?
- unstructured?
- … or rather structured?

macht
gemacht
gefragt
kataba
yaktubu
book
handbook
deridere
ridiamo
scartabelliamo

lexical emergentism

- in accordance with "morphological emergentism", **all word forms are memorised** in the lexicon

- the extent to which a word is processed as morphologically **complex** is based on the degree of **probabilistic support** that the structure of the word receives from its family of **morphological neighbours**

- **morphological structure** is a gradient property emerging through acquisition from principles of **word-based** self-organisation

3 principles
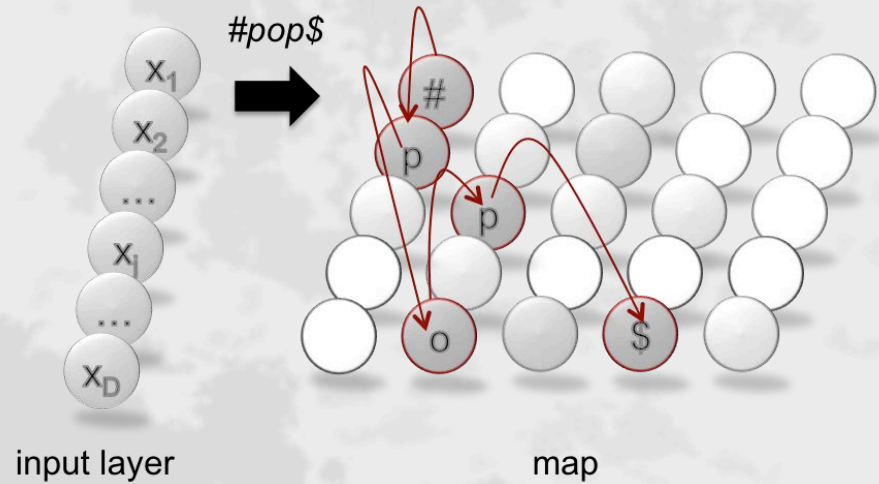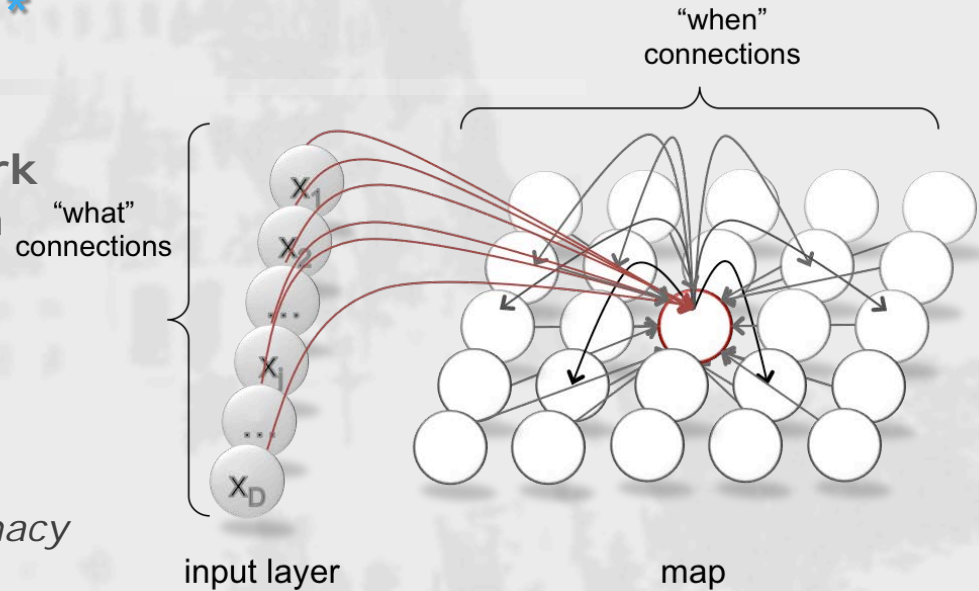# of word self-organisation

- **frequency**
  - *any lexical model must assume that accessing a lexical item*
    - **modifies** the item's representation
    - increases the **probability** that the item will be accessed in the future

- **symbol identity**
  - ***similar symbols** activate overlapping memory traces*

- **time**
  - *symbol representations are time-bound and context sensitive*

Hebbian brain maps *

- a **neuro-computational framework** for paradigm-based word acquisition

    – *words encoded on input layer as **temporal sequences** of symbols*

    – *map nodes activate **concurrently** at each time tick and **compete** for primacy*

    – *over training, nodes **specialise** for specific input symbols in specific contexts*

"when" connections

"what" connections

x_1
x_2
...
x_i
...
x_D

input layer

map

x_1
x_2
...
x_i
...
x_D

#pop$

#
p
p
o
$

input layer

map

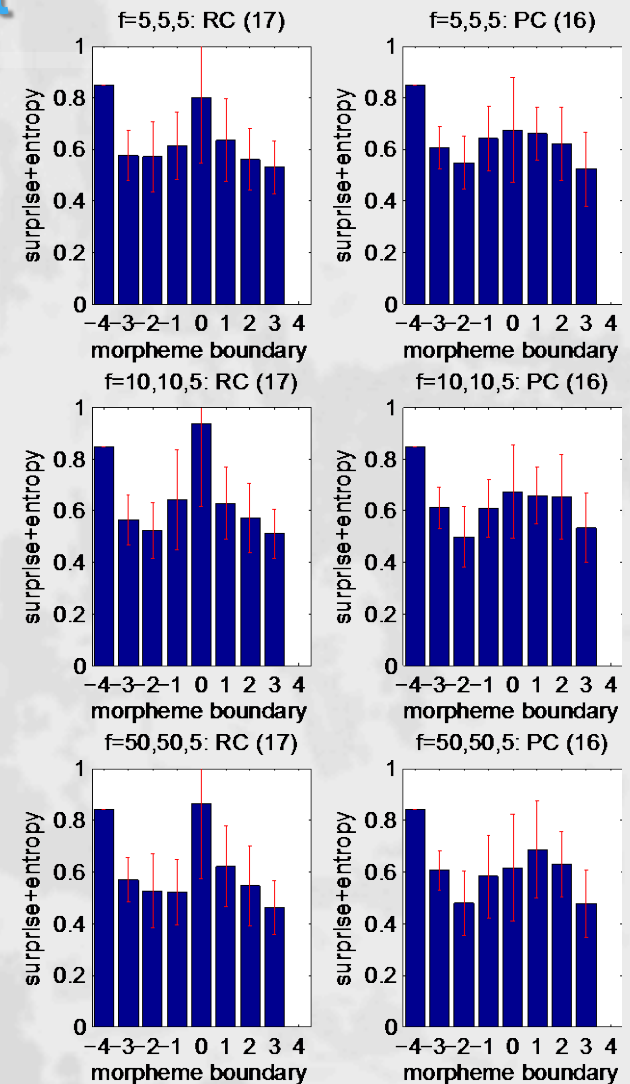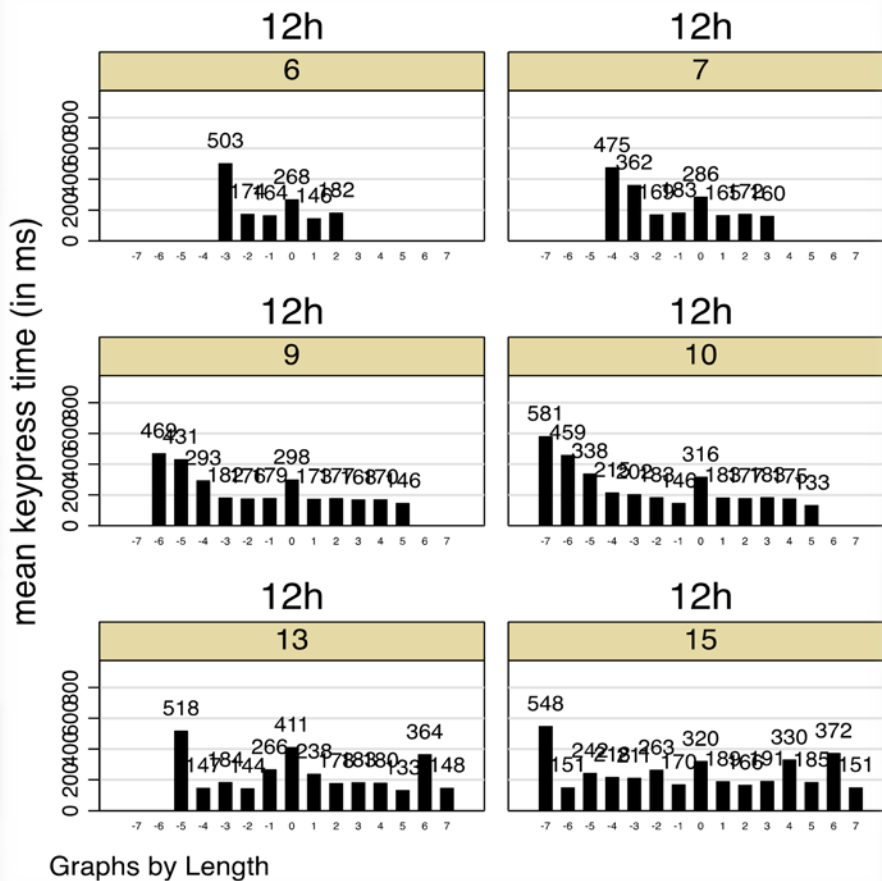# lexical memory as node activity
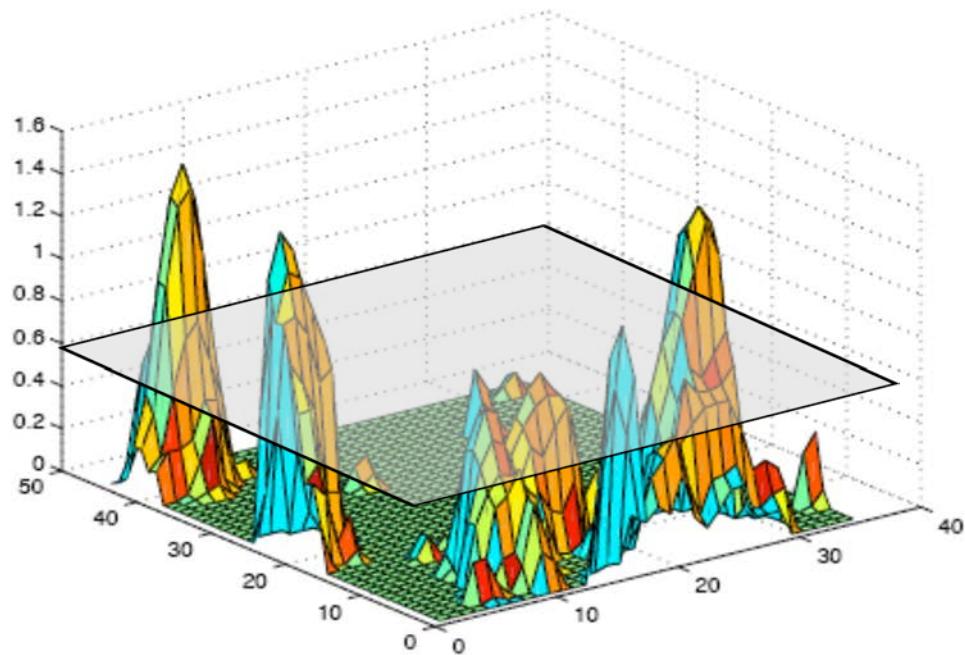
#,G,EH,T,IH,NG

typing "car wash" vs. "carpet" *

multidisciplinarity in 3 ways

- **multidisciplinarity** draws on knowledge from different disciplines but stays within the boundaries of those fields

  – *sequential and goal-oriented*
    - multidisciplinarity is used in medical practice, where collaboration of biologists, chemists, pharmacists, opticians, psychologists etc. is required; however, a multidisciplinary application-oriented research does not replace or shift the research boundaries of each discipline

  – *technology-driven*
    - likewise, when a particular science uses, in an essentially instrumental way, sophisticated technologies, then that science–technology is inherently multidisciplinary: making use of computational tools and modern techniques of digital reproduction does not make, as such, principles of contemporary philology any different from traditional philological practice

  – *mathematical formalization*
    - cases of successful transfer of structures and mathematical theorems to other sciences, such as the application of statistical methods to psychological and epidemiological contexts, do not generate as such disciplinary integration

inter-disciplinarity

- **interdisciplinarity** analyses, synthesizes and harmonizes links between disciplines into a coordinated and coherent whole

- it has to do with shared **methodology** and effective causal **continuity** between data and between operations defined on data

- new **functional magnetic resonance technologies** are very important to understand the connection between neurophysiology and psychological evidence; by studying metabolic activity of brain areas "in vivo", they allow us to understand what areas are active when the subject is performing a particular task, thus paving the way to unification between the two domains
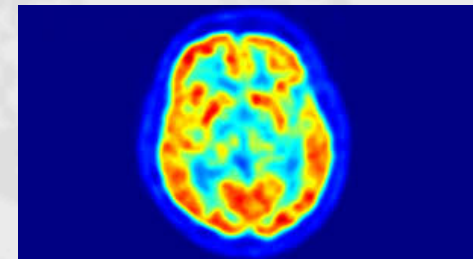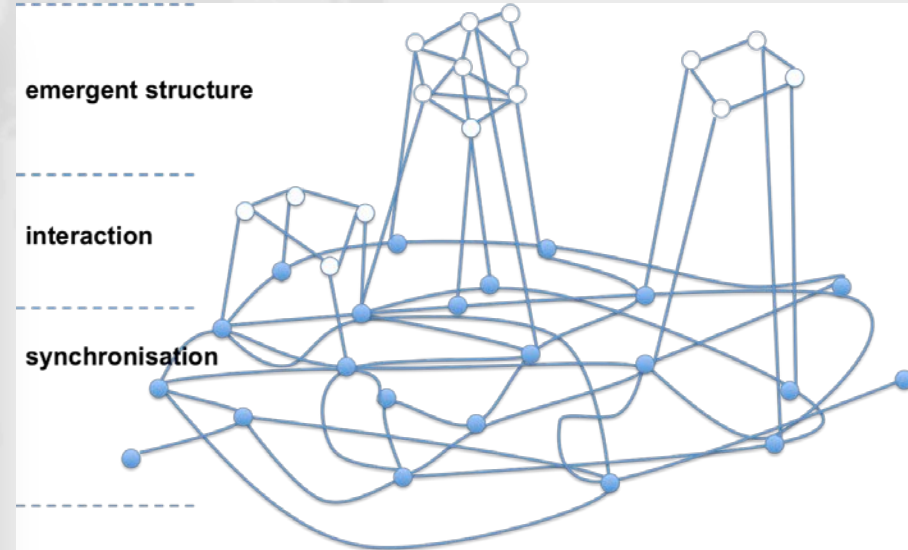
trans-disciplinarity

- **trans-disciplinarity** integrates the natural, social and health sciences in a pragmatic goal-oriented context, and in doing so transcends each of their traditional boundaries
    - *researchers from different disciplines work jointly to develop and use a shared conceptual framework that synthesizes and extends discipline-specific theories, concepts and methods to create new approaches to address a common problem*
- integration of disciplines and non-academic participants
- development of integrated knowledge and theory among science and society

towards unification in language sciences?

- **scientific methods** are characteristic of each science because they depend on specific operations, relationships and ontological units within each field

- certain **methods** (e.g. computer modelling) lately achieved considerable popularity and showed a great potential in shedding light on some outstanding, unresolved issues in one particular domain

- **incorporation** of a novel methodology into a consolidated field does not per se require a fundamental reconceptualization of the boundaries of that domain
  - *computational linguistics is not per se different from traditional descriptive or theoretical linguistics unless computer modelling is used to address issues lying on the edge of two language domains, i.e. on the verge of emergence*

integrative emergence and domain boundaries

- the concept of **emergence** does not require that what emerges exists previously sub-merged such as dolphin in the sea

- an emergent behaviour/property can appear when a number of simple entities operate in an environment, forming more complex behaviours as a collective

- emergent behaviours can occur because of intricate causal relations across different scales and feedback, known as interconnectivity

- complex behaviour or properties are not a property of any single such entity, nor can they easily be predicted or deduced from behaviour in the lower-level entities, and might in fact be irreducible to such behaviour

emergent structure

interaction

synchronisation

a transdisciplinary perspective

- joining forces in the language domain will not only lead to considerable progress in our theoretical understanding of the physiology of communication, but will also be conducive to more effective ways to **help** real people engaged in their daily communicative exchanges.

- integrating **data** analysis and modelling with sound **theoretical insights** and bio-engineering and computer **technologies** will help us build assistive tools in as diverse areas as:

    – *multilingualism, cultural integration, and the importance of input and socio-economic factors;*

    – *language as key to knowledge access and knowledge acquisition;*

    – *language teaching and language learning at school;*

    – *diagnostic and therapeutic protocols for language disorders;*

    – *language for effective accessible communication and creative thinking;*

    – *language and cognitive well-being (diagnostic protocols for cognitive disorders and disturbances, e.g. in elderly people, autistic children, dyslexics etc.)*