

# Interfaces of sociolinguistics: Cognition and big data

Jean-Pierre Chevrot

- 1 - Lidilem Lab, Université Grenoble Alpes, Grenoble, France
- 2 – Dynamic Network Team, Rhône-Alps Institute of Complex Systems (IXXI), Ecole Normale Supérieure, Lyon, France

Contact : [jean-pierre.chevrot@univ-grenoble-alpes.fr](mailto:jean-pierre.chevrot@univ-grenoble-alpes.fr)



Phonetik und Phonologie im  
deutschsprachigen Raum (P&P)  
2018, Universität Wien, 6-7  
September



# Summary

**1/ Sociolinguistics and variationist sociolinguistics**

**2/ Sociolinguistics and cognitive science**

**Example : the field of sociolinguistic acquisition**

**3/ Sociolinguistics and Data science**

**Example : a study of French sociolinguistic variables on Twitter**

**4/ Risks and benefits of interdisciplinary research**

# 1 Sociolinguistics



# Sociolinguistics

## General aim of sociolinguistics

The understanding of the interaction between language and society

### Three research traditions in sociolinguistics

Variationist sociolinguistics	Linguistic Anthropology	Sociology of language
<p>Understanding languages as <b>variable evolving systems</b>, due to their internal dynamics, the cognitive constraints, contact with each other, and their links with social organization, which is itself evolving, composite and multi-layered (Laks, 2013).</p>	<p>Understanding language as a <b>“cultural resource”</b> and speaking as a “cultural practice”</p>	<p>Study of both language and society at the <b>macro level</b> (nation states or social classes)</p>
<p><b>Quantitative methods</b> describing the usage of sociolinguistic variables in different social groups, regions, settings</p>	<p><b>Qualitative methods</b> (ethnographic participant observation, audiovisual recording, interviews, etc.) to observe intrinsically social and cultural speech events.</p>	<p><b>Census and broad surveys</b> describing the state of language diversity and inequality (e.g. number of indigenous languages in South America)</p>

# Sociolinguistics

## General aim of sociolinguistics

The understanding of the interaction between language and society

### Three research traditions in sociolinguistics

#### Variationist sociolinguistics

Understanding languages as **variable evolving systems**, due to their internal dynamics, the cognitive constraints, contact with each other, and their links with social organization, which is itself evolving, composite and multi-layered (Laks, 2013).

**Quantitative methods** describing the usage of sociolinguistic variables in different social groups, regions, settings

#### Linguistic Anthropology

Understanding language as a **“cultural resource”** and speaking as a **“cultural practice”**

**Qualitative methods** (ethnographic participant observation, audiovisual recording, interviews, etc.) to observe intrinsically social and cultural speech events.

#### Sociology of language

Study of both language and society at the **macro level** (nation states or social classes)

**Census and broad surveys** describing the state of language diversity and inequality (e.g. number of indigenous languages in South America)

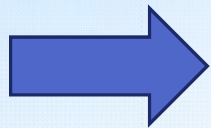
# The framework of variationist sociolinguistics

A two-minute summary of five decades of research...

## **Linguistic heterogeneity and within-language variation : a non controversial assertion**

Depending on their situation in the social or geographical space, speakers of the same language speak different dialects of this language.

## **The pioneering work of William Labov conferred a scientific status to the fundamental heterogeneity of languages.**



Notions of variation, variable, variant

Creation of variationist (socio)linguistics

## **Variables : points within the linguistic system where the speaker can say the same thing in different ways.**

The variants are “identical in reference or truth value, but opposed in their social and/or stylistic significance”

# The framework of variationist sociolinguistics

Examples of sociolinguistic variables (studied in different languages at different linguistic levels)

	Language	Linguistic Level	Example of study
<b>Variable pronunciation of -ing</b> [in] vs. [iŋ] (playing / playin')	English	Phonology	Fischer, 1958; Labov, 2006
<b>Variable omission of the first morpheme of the negation which surrounds the verb</b> <i>je mange pas</i> vs. <i>je ne mange pas</i>	French	Morphology	Armstrong, 2002
<b>Variable interrogation in French</b> <i>Sont-elles là ?</i> Verb-Pronoun <i>Est-ce qu'elles sont là ?</i> Particle-Pronoun-Verb <i>Elles sont là ?</i> Pronoun-verb + rising intonation	French	Syntax	Coveney, 1997

# The framework of variationist sociolinguistics

## 1/ The variants often carry social information about the speaker and the context (socio-indexical information)

Standard variants associated to ...	Non-standard variants associated to ...
social prestige, high education level, professional ambition... They may be valorized.	social skills, solidarity or loyalty towards the group, masculinity... They may be stigmatized.

## 2/ The frequency of use of the variants depends upon several factor

The sociodemographic characteristics of the speaker	Elements of the interactional context
more standard variants in women, older speakers, individual with a higher level of education and those who are peripheral in the local social network, etc.	more standard variants in formal situations, speaking about work-related topics or to a prestigious addressee, to increase the social distance and decrease connivence, etc.

For a more flexible view on socio-indexicality and usage of variants : Campbell-Kibler, 2008; Eckert, 2008 (*inter alia*)



# 2

## Sociolinguistics and cognitive science Example : the field of sociolinguistic acquisition

The meeting of  
sociolinguistique  
and cognitive  
sciences in the  
70s and 80s...



# Sociolinguists and social scientists address mental entities

Since the beginning, social science uses concepts which refer to mental entities

Emile Durkheim : Collective representation

Karl Marx: Ideology

Max Weber: Subjective meaning

Kaufman & Clément, 2011

## The same is true of sociolinguistics

Social meaning, indexical meaning

Communicative intention

Social judgement or evaluation

Selection of sociolinguistic variants

Mental states

Mental processes

# The same mental state can be constructed as two scientific objects : cognitive vs. social

## Confidence as a cognitive fact (Cognitive economics)

- Punctual event : change in the mental state when an individual decides to engage in cooperation
- Correlation with punctual behavior, cerebral activation and oxytocin production

## Confidence as a social fact (History and sociology)

- Progressive establishment of a mutual commitment
- One individual/group accepts making herself more vulnerable
- The other is committed not to exploit the vulnerability
- Social norms create mutual expectations

# Long-standing convergences between cognitive and social science



Social psychology, McDougall, 1908

Social cognition, Fisker & Taylor, 1984

Cognitive sociology, Cicourel, 1974

*Social neuroscience, Cacioppo & Berntson, 1992*

*Cognitive sociolinguistics, Kristiansen & Dirven, 2008*

*Sociolinguistic cognition, Campbell-Kibler, 2010*

Kaufman & Clément, 2011; Lahire & Rosental, 2008; Molénat, 2012;  
Morgan & Schwallbe, 1990

# Five areas of contact between variationist sociolinguistics and cognitive science

Chevrot, J.P., Drager, K. & Foulkes, P. (in press). Sociolinguistic variation & cognitive science, special issue of *Topics in cognitive science*

## Cognitive sociolinguistics

Kristiansen & Dirven, 200

- Interaction between language and cognition
- Grounded on corpus-based studies
- Draws on the three working hypotheses of Cognitive linguistics: language is not modular, grammar is a conceptualization of the world, linguistic structure emerges from language usage.
- **Languages and conceptualization of the world vary across cultural and social diversity**

# Five areas of contact between variationist sociolinguistics and cognitive science

Chevrot, J.P., Drager, K. & Foulkes, P. (in press). Sociolinguistic variation & cognitive science, special issue of *Topics in cognitive science*

## Sociolinguistic cognition

Campbell-Kibler, 2010;  
Loudermilk, 2013

- Explores the cognitive and cerebral mechanisms underpinning the ability to memorize sociolinguistic variation, to implement it during speech production, and to process it during speech perception
- Based on experimental methods from psycholinguistics and social cognition (elicitation tasks, reaction time and eye-tracking experiments, neuroimaging, social priming)

# Five areas of contact between variationist sociolinguistics and cognitive science

Chevrot, J.P., Drager, K. & Foulkes, P. (in press). Sociolinguistic variation & cognitive science, special issue of *Topics in cognitive science*

## Computational modelling

Hruschka et al., 2009; Stanford & Kenny, 2013

- Dynamic modelling of populations of 'agents' interacting with one another and sharing simulated sociolinguistic
- Major value : to test, in a concentrated time frame, the long-term effect of linguistic, social and cognitive constraints on language variation and change that are hard to control in experiments or in the usual conditions of language use
- e.g. size and structure of populations, long-term effect of attentional bias toward leaders, etc.

# Five areas of contact between variationist sociolinguistics and cognitive science

Chevrot, J.P., Drager, K. & Foulkes, P. (in press). Sociolinguistic variation & cognitive science, special issue of *Topics in cognitive science*

## Comparative study of variation in animal communication

Henry, Barbu, Lemasson & Hausberger, 2015

- In the vocalization of certain species (e.g. birds, marine mammals), varieties referred to as “dialects” by Darwin (1859).
- Strong analogies with sociolinguistic variation in humans (e.g. regional and social “dialects” functioning as social passwords’ that indicate belonging to a group)
- Variation could ensure adaptive benefits via group cohesion and social recognition in different species



# Five areas of contact between variationist sociolinguistics and cognitive science

Chevrot, J.P., Drager, K. & Foulkes, P. (in press). Sociolinguistic variation & cognitive science, special issue of *Topics in cognitive science*

## Language acquisition

Anderssen, Bentzen & Westergaard, 2011; Chevrot & Foulkes, 2013; De Vogelaer, Chevrot, Katerbow & Nardy, 2017; Lacoste & Green, 2016

First and second language acquisition of sociolinguistic variation and in sociolinguistic variation

Focus on this area that illustrates the meeting of cognitive and social science

# Four communities/fields of research

Chevrot, J.-P. & Ghimenton, A. (to appear). Bilingualism and bidialectalism, in Lourdes Ortega & Annick de Houwer (eds), *The Cambridge Handbook of Bilingualism*

Research on...	Speakers	Contexts
<b>1. Child acquisition of dialectal varieties of the first language</b>	First language learners in contact with several varieties of one language	Communities where one language is dominant
<b>2. Child multilingual acquisition</b>	First language learners confronted with varieties of what is considered as different languages (overlaid on dialectal variation in each language)	Multilingual communities
<b>3. Acquisition of L2 sociolinguistic variation during second language acquisition</b>	Second language learners confronted with L2 sociolinguistic variation	Study abroad, Migration FL classroom
<b>4. Lifelong second dialect acquisition</b>	First or second language learners confronted with a second or a third dialect of the same language	Geographical (move to a new area) or social mobility Political change (fall of Berlin wall).

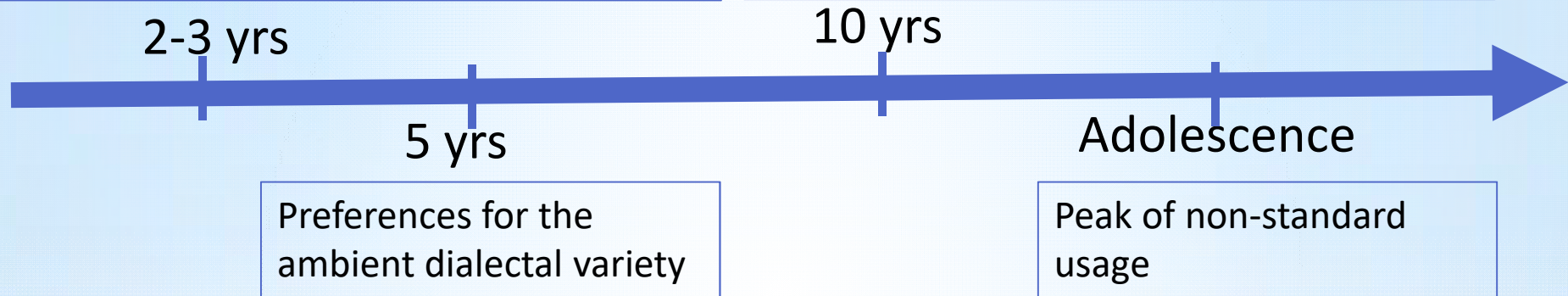
# Key findings from these fields

Chevrot, J.-P. & Ghimenton, A. (to appear). Bilingualism and bidialectalism, in Lourdes Ortega & Annick de Houwer (eds), The Cambridge Handbook of Bilingualism

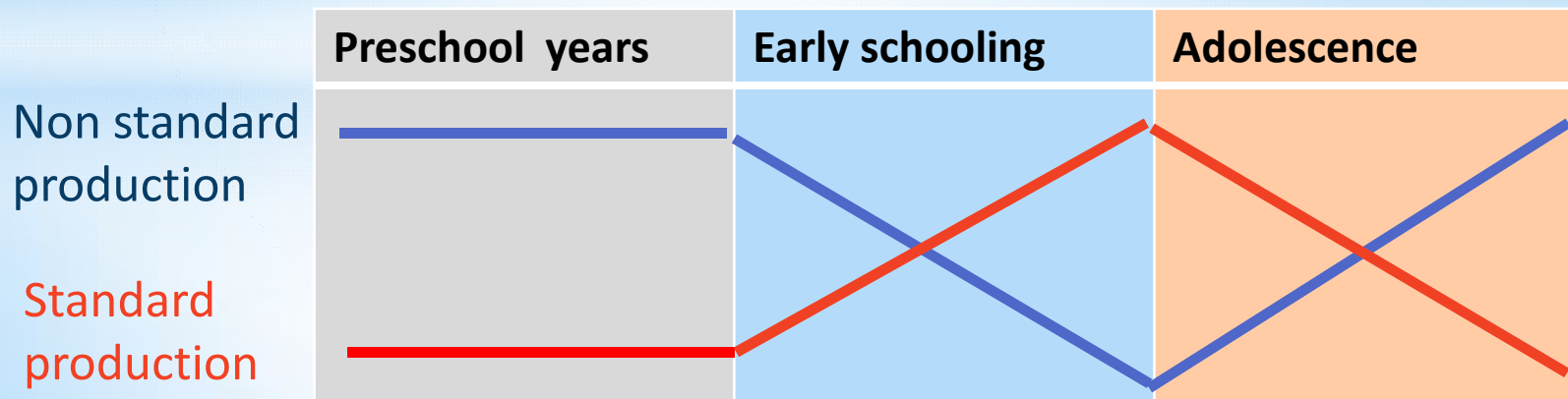
## Areas 1 & 2 - The child is confronted with within or between language variation

- Production of the dialectal variants of the social or regional ambient variety
- Stylistic ability during family interaction

Preferences align with adult categorization based on the social value of language (prestige variety vs. other varieties)



## Double crossing pattern in production



# Key findings from these fields

Chevrot, J.-P. & Ghimenton, A. (to appear). Bilingualism and bidialectalism, in Lourdes Ortega & Annick de Houwer (eds), *The Cambridge Handbook of Bilingualism*

## Area 3 - Learners acquiring sociolinguistic variation in L2

1/ L2 learners show negative attitudes to the non-standard varieties of L2

... But exposure to a greater dialectal variation increases learners' perception and production !

2/ L2 learners underuse non-standard variants when learning occurs in the classroom.

The favoring effect of context on acquisition of non-standard varieties  
*Naturalistic context > study abroad > immersion curriculum > FL classroom*

3/ Influence of social factors on L2 sociolinguistic acquisition

<b>Social status</b>	Higher status learners use more frequently the standard variants of the L2
<b>Social network</b>	L2 learners with L1 speakers in their network are less standard (Gautier, 2016)
<b>Stylistic ability</b>	- Learners generally do not manifest the ability to modify the frequency of standard / nonstandard variants according to the context. - When they do it, they undershoot the native nonstandard target

# Key findings from these fields

Siegel (2010) summarized in Chevrot & Ghimenton (to appear)

## Area 4 - Learners acquiring a second dialect (D2)

1/ D2 acquisition is difficult compared to L2 acquisition

- The similarity between D1 and D2 linguistic systems favors transfers
- Intercomprehension is guaranteed. D2 acquisition is motivated by the desire to be viewed as a local.
- Identity issues may inhibit D2 acquisition (loyalty toward the D1 community or negative reaction from the D2 community)

2/ Most important factors influencing D2 acquisition

<b>Age</b>	Optimal age of acquisition: the mid-teens or younger for lexicon and morphology and 7 years or younger for phonology
<b>Social interaction</b>	Integration in the social network of the D2 community favors D2 acquisition
<b>Identity</b>	How the learner identifies to D2 community speakers influences D2 acquisition.

# Focus on the acquisition of phonological variation : a review

Nardy, Chevrot, & Barbu, 2013

We analysed 30 studies on phonological variation published from 1964 to 2011.

Developmental turn during the 90's

- Before : Children patterns of variation are seen in light of established adult sociolinguistic regularities
- After : increasing concern with developmental issues (earlier age groups, theoretical hypotheses to explain the developmental patterns)

# Focus on the acquisition of phonological variation : a review

Nardy, Chevrot, & Barbu, 2013

## 1/ Appearance of adult-like patterns and effects of age and input

Topics	Number of studies and languages	Children age	Main tendencies
<b>Social background</b>	7 studies English, French, Spanish	[3-10]	All the studies converge: from the age of 3, the higher the social status of the parents, the more standard variants the child produces. See also, Barbu et al. (poster, this session)
<b>Context (style)</b>	10 studies; children aged English, French, Spanish	[3-12]	All the studies but one converge: the children use more frequently the standard variants in the formal context (from the age of 3-4 within family exchanges).
<b>Gender</b>	11 studies English, French	[2-10]	All possible tendencies found : - 7 studies (ages 2-9): No gender effect - 2 studies (ages 6-10): More standard variants in girls - 2 studies (ages 3 -7): More standard variants in boys
<b>Input</b>	2 studies English	[3-4]	- Mothers of girls address more standard variants to their 2-4 year-old daughters than mothers of boys (Foulkes et al., 2005). - At the age of 3, correlations between the use of non standard variants by mothers and by their children (Smith et al.2007).
<b>Evaluation of variants</b>	6 studies English, French	[2-12]	- 4 studies : Evaluation based on context or status takes place between 9 and 12 years of age. - 1 study using a very simple task (pointing the puppet who speaks correctly ): 5-6 year olds from upper-class families distinguish between standard and non standard variants (Nardy, 2008)

# Focus on the acquisition of phonological variation : a review

Nardy, Chevrot, & Barbu, 2013

## 2/ Hypotheses and new research questions

**The bases for social stratification and stylistic flexibility of sociolinguistic usage are laid early within the family interactions**

→ How and when do these bases meet the norms and values shared by the community ?

**The gender difference seems establish later**

→ How does this difference combine with the process of gender socialization ?

**A lot of work should be done in two directions**

- What kind of cognitive device does underpin sociolinguistic acquisition? Variable rule vs. usage-based schema; Role of awareness; Mechanisms linking social and linguistic information ?
- What is the influence parents, peers, and teachers



# Sociolinguistics and Data science

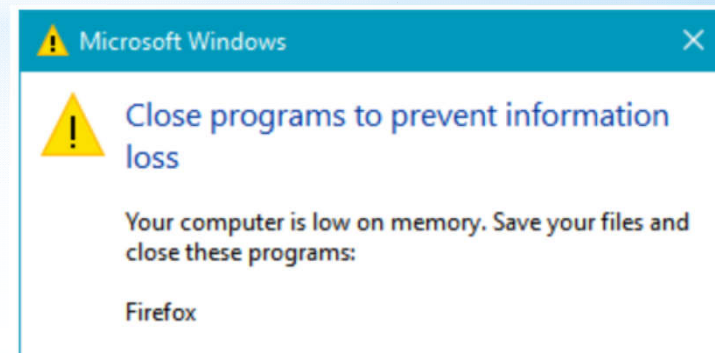
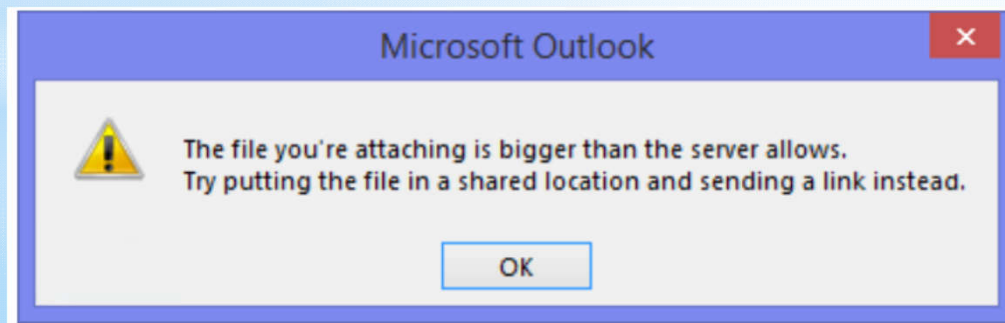
## Example : a study of French sociolinguistic variables on Twitter



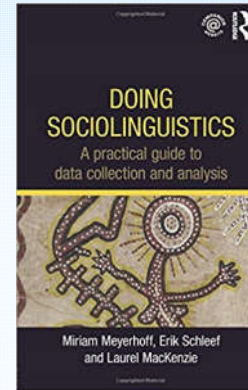
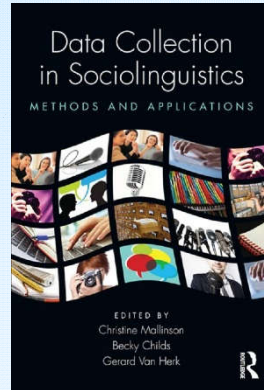
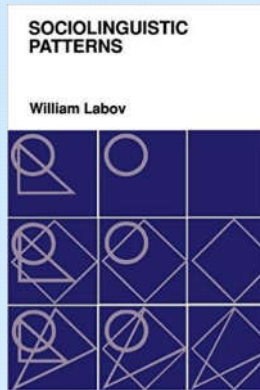
*You cannot email this data to a colleague. You can't even download it on your computer. This is data on an unprecedented impossibly mind boggling massive scale. - Kenneth Benoit (2015)*

From the website of Josef Fruehwald, Univ. of Edinburgh : **Big Data and Sociolinguistics**

[https://jofrhwd.github.io/papers/plc39\\_2015/#/](https://jofrhwd.github.io/papers/plc39_2015/#/)



# From its beginning, sociolinguistics strongly emphasized the need for data



**Quantity** - Data allowing quantitative analysis of language variation in relation with social features

**Quality** - Data that are « good enough » for representing the actual language usage of the speech community

**Methods to overcome the observer's paradox**

*“ to find out how people talk when they are not being (...) observed; yet we can only obtain this data by systematic observation”*

# Sociolinguistics rapidly joined the emerging computational social science

Lazer et al. (2009)  
[www.davidlazer.com](http://www.davidlazer.com)

## Our new ability to collect, analyze and model massive datasets on the behavior of individuals within collective entities

Use of sensors (proximity sensors, wearable audio recorders, etc.)

Automatic, permanent and unsupervised collection of digital data from...

- The digital communication : blogosphere, social media, e-mail exchange, etc.
- The recording of human activity : bank transaction, location of mobile calls, booking of cars, bicycles or rooms, peer-to-peer services, etc.

A way to move beyond the observer's paradox ?



Crucial political, ethical and privacy issues

# Computational sociolinguistics

Nguyen et al. (2016). Computational Sociolinguistics: A Survey, *Computational Linguistics* 42(3), 537-593.  
[http://www.mitpressjournals.org/doi/full/10.1162/COLI\\_a\\_00258](http://www.mitpressjournals.org/doi/full/10.1162/COLI_a_00258)

## A subfield of Computational linguistics

Study of language variation and change in large databases from digital interactions (blogs, social media)

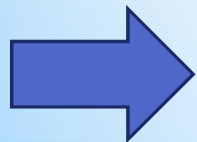
- Language data: generally written language of the digital communication. For example: studies on the written counterpart of spoken sociolinguistic variables (e.g. *in* spelling for the (ing) variable)
- Social data : network links and social profile of the users

# Computational sociolinguistics contributes to the same trends

Several studies reduplicate well-known results observed in face-to-face spoken interaction

Résultats	Références
Effect of age, gender, region, ethnicity, etc.	Bryden et al., 2013; Eisenstein, 2015; Gonçalves and Sánchez, 2014; Magué et al., 2015
Dynamics of language innovation and loanwords	Altmann et al., 2011; Garley and Hockenmaier, 2012; Eisenstein et al., 2014
Convergence amongst connected people	Danescu-Niculescu-Mizil et al., 2011; Tamburrini et al., 2015

- Can digital data reveal new sociolinguistic patterns ?
- What is the risk of bias and how to overcome it ?
- What is the benefit for sociolinguistic theory ?

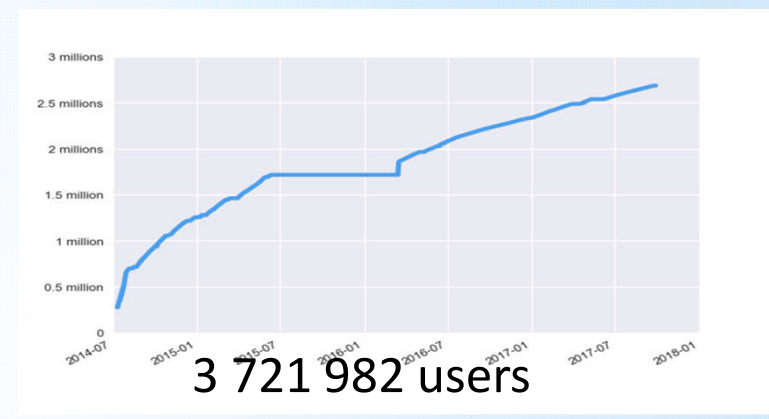
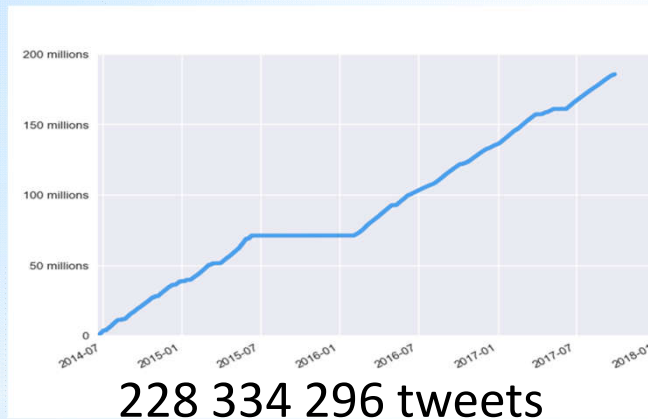
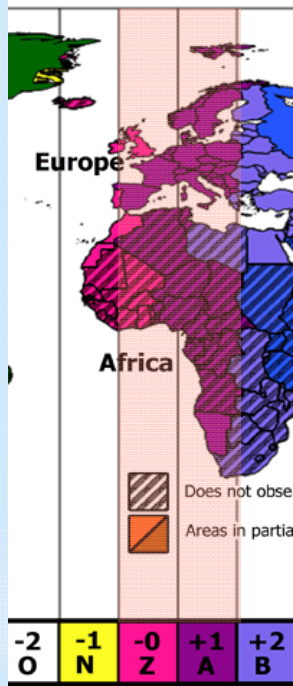


**First results of the Sosweet project**

**Four-year collection of approx. 15%-25% of the tweets produced in 3 years ...**

- Written in French
- Users located in the GTM and GTM+1 areas → French-speaking countries
- + Recording of the network of follower/followee links between users

### Current state of the database



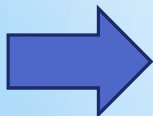
**An interdisciplinary project combining four teams**

Sociolinguistics + NLP + Data science + Corpus linguistics

### First results on French sociolinguistic variables



Abitbol, J., Karsai, M., Magué, J.P., Chevrot, J.P. and Fleury, E., 2018, Socioeconomic dependencies of linguistic patterns in Twitter: a multivariate analysis. In Proceedings of ACM WWW conference, Lyon, France, April 2018 (WWW'18),

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>



# A well-known sociolinguistic variable : Negative *ne*

<b>Standard negation</b>	<b>NE + verb + Part. 2</b> : <i>pas</i> ‘not’, <i>jamais</i> ‘never’, <i>rien</i> ‘nothing’, etc.)	Il <b>ne</b> boit <b>jamais</b> de vin ‘He never drinks wine’
<b>Non standard negation</b>	<b>Deletion of ne</b>	Il $\emptyset$ boit <b>jamais</b> de vin

Standard	Non Standard
 <p>Emmanuel Macron ✓ @EmmanuelMacron</p> <p>‘I do not have the same fights as you but I'm as free as you’</p> <p>Je n'ai pas les mêmes combats que vous mais je suis aussi libre que vous. #LEmissionPolitique</p>	 <p>Mister X ✓</p> <p>‘I seem to have lost track: is Cannes vs. Netflix the new Taxi vs. Uber?’</p> <p>J'ai pas suivi le truc, Cannes vs Netflix c'est le nouveau Taxi vs Uber ?</p>

## Several studies on spoken French and 3 studies on digital communication

Factors that influence NE realisation	References
More NE realisation in <b>older speakers</b>	Ashby, 2001; Blattner & Williams, 2011; Hansen & Malderez, 2004
More NE realisation in the more <b>formal settings</b> and <b>emphatic contexts</b>	Armstrong, 2002; Williams, 2009; van Compernelle, 2009
More NE realisation <b>if the subject of the verb is a NP</b> (vs. a clitic) and <b>with certain Part. 2</b>	Ashby, 1981; Armstrong & Smith, 2002; Coveney, 1996; Hansen & Malderez, 2004, van Compernelle, 2008
<b>Effect of Socioeconomic status : no convergence</b>	Blattner & Williams, 2001; Williams, 2009; Hansen & Malderez

# An additional linguistic feature : the -s / -x plural ending in French spelling

Plural marking of written French in the nominal phrase : mute letters -s or -x at the end of the adjective and the noun.

With –s ending	Without –s ending
 <p>Le Labo des savoirs @Labodessavoirs</p> <p>‘Heated debate among researchers! Scientific battles at #JSUN2017 animated by @maximelabat’</p> <p>Les <u>chercheurs</u> sont chaaaauds ! Les battles scientifiques de la #JSUN2017 animées par @maximelabat</p>	 <p>Miss Y</p> <p>‘At high schoolers’ demonstration, it’s the cop [missing plural ending] who intervene when it is the supervisor [missing plural ending] for us and that makes the difference’</p> <p>Au monôme des lycées en ville c’est <u>les policier</u> qui interviennent alors que nous c’est <u>les surveillant</u> c’est la ou tu vois la différence</p>

Omitting mute plural -s is the most frequent error in French writing

Lucci & Millet, 1995

Correlation between standard usage of the plural ending and social status, in preteens, teens and adults

Brissaud, 1999; Lucci & Millet, 1995; Totereau et al. 2013



# How to infer socioeconomic status (SES) of Twitter users?

Using INSEE\* data for transforming GPS localization into SES proxy ...

**INSEE : French National Institute of Statistics and Economics Studies**

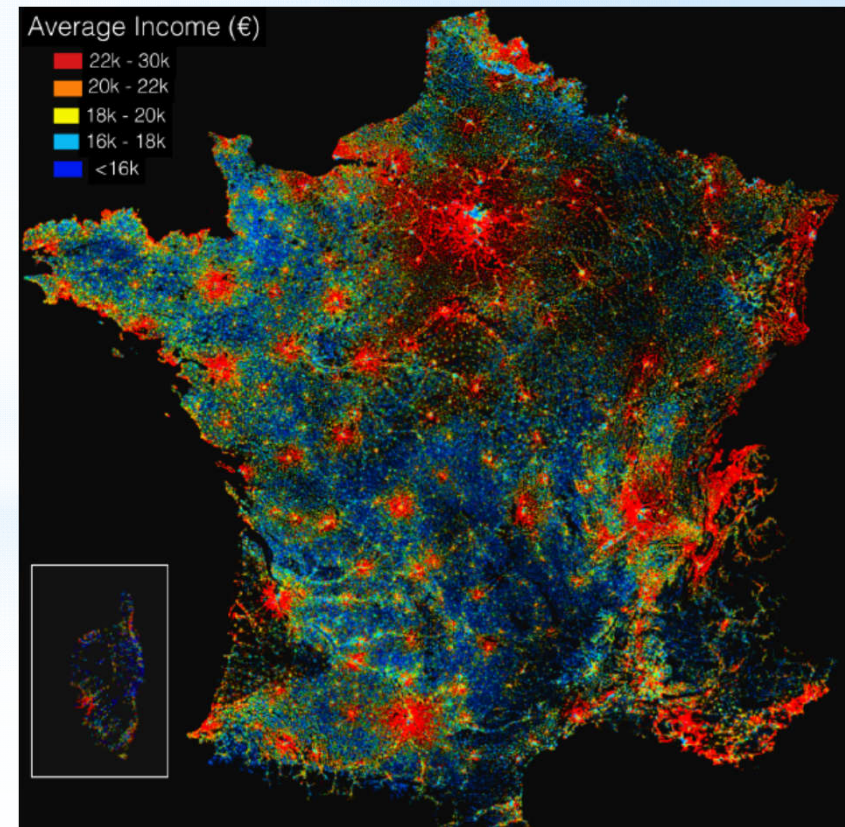
Set of sociodemographic aggregated indicators for each 200m x 200m square areas across France

## In each square

- Mean yearly income (winsorized mean, without extreme values)
- Rate of homeowners
- Population density

## Working hypothesis

Living in an area with higher income mean, with bigger rate of owners and with smaller density (single-family house) is associated with higher SES

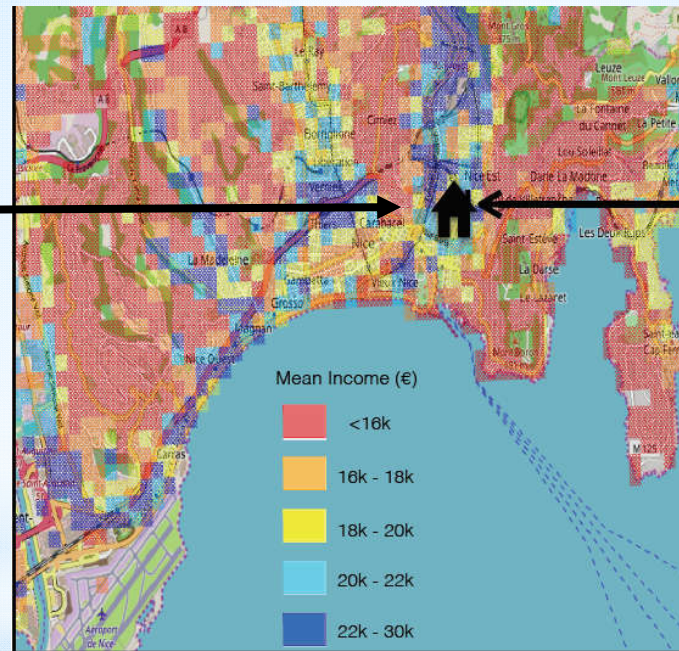


# How to infer socioeconomic status (SES) of Twitter users?

Matching INSEE data and 109 000 geolocated users of Twitter

## INSEE Data

Mean yearly income is 18 to 20 thousands of euros in this square area



## Twitter Data

The geolocated tweets of user #878291414 are most frequently sent from this area

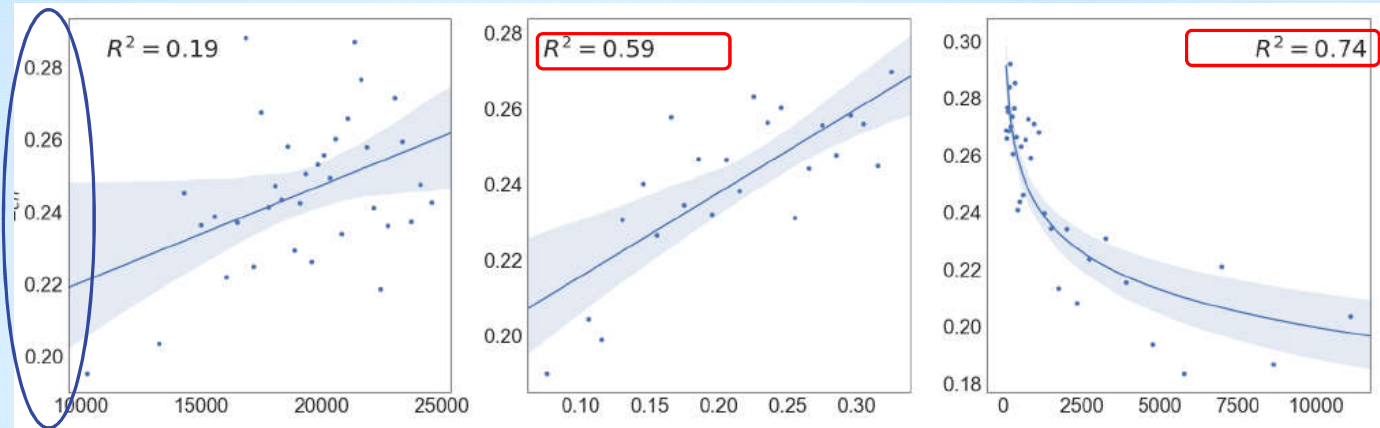
User	Lat	Lon
720624511	0.222689	47.990762
441320265	-52.983982	4.913242
982677272	2.9831989621	48.4007454896
2296639897	2.4912035419	48.9458390389
878291414	2.360299	47.116211
360954534	5.863804	49.303493
190311280	7.023194	43.552704
1498428090	2.289129	48.673666
289953393	2.386827	48.865732
482001789	4.9224447	46.8626553

## Working hypotheses

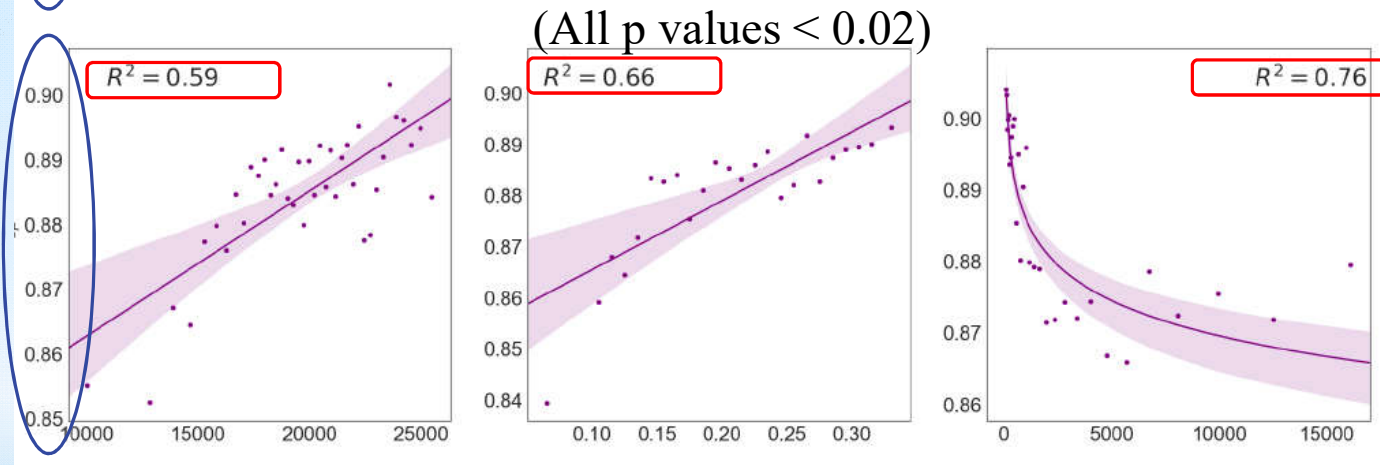
- User #878291414 lives in this area
- The yearly income of user #878291414 is 18 to 20 thousands euros

# Result 1 – Correlation between language and SES

Rate of standard  
*ne* realization



Rate of standard  
-s/-x plural endings



Income

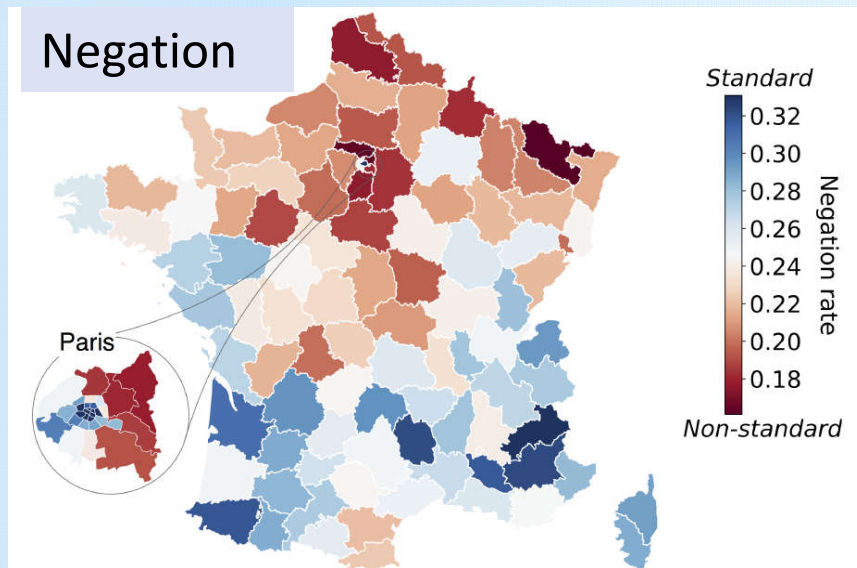
Rate of  
homeowners

Density of  
population

What's new for sociolinguistics ? New research questions...

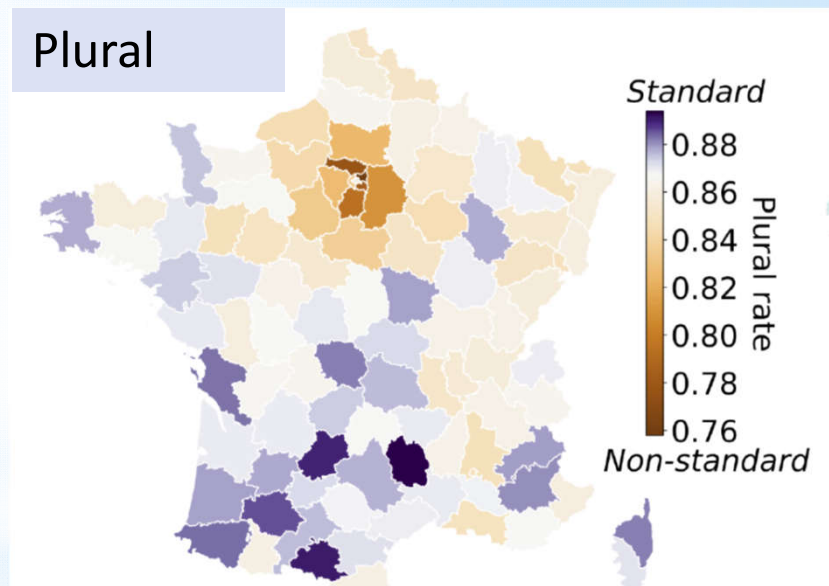
- Correlations are surprisingly strong
- Correlation appears with very small range of sociolinguistic variation
- Different correlation values with different SES proxies

## Result 2 – Spatial correlation : a North-South gradient of variation ?



- Departments of South are more standard than departments of North
- Difference is gradient

Same trends for the plural endings



What's new for sociolinguistics ? The mystery gets deeper...

- No clear dialectal explanation (-s-x plural is **not** a dialectal spoken cue)
- Are twitter users the same in the North and the South : different rates of men/women, higher/lower SES; different network structures ?

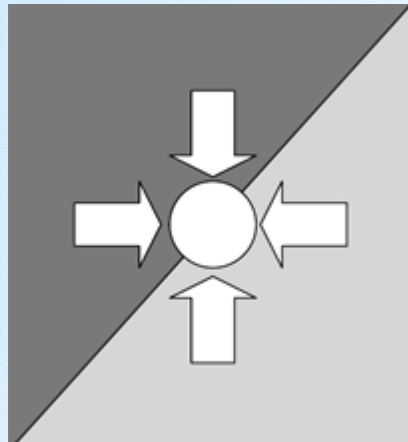
4

To conclude  
Risks and benefits of  
interdisciplinary  
research



# The (power)relationships between disciplines

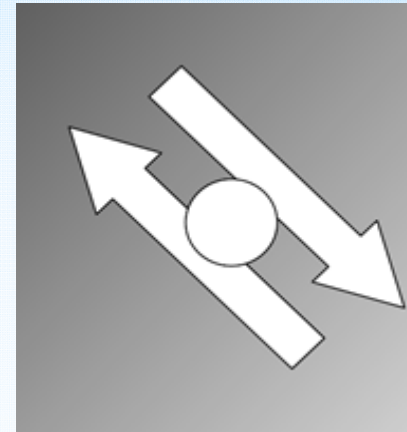
## Case 1 – Balanced relationships between disciplines



### Coordination

Disciplines coordinate for a better understanding of the same object

Methods and theories of each one remain identical



### Fusion

Methods and theories influence each other

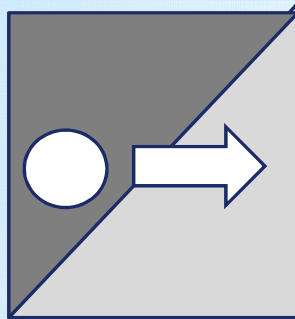
Toward the creation of a new area

Degree of integration between the disciplines

# The (power)relationships between disciplines

Case 2 – **Unbalanced** relationships between disciplines

Degree of **dominance** between the disciplines



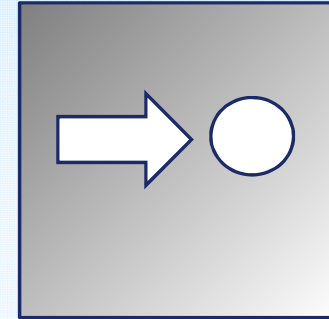
Subcontracting

**Discipline A uses the services of discipline B**

Object of study specified by discipline A

Potential risk when social science interacts with ICTS and Life or Physical sciences

Example of risky question : Social science helps Life science to deal with ethical issues of brain imaging data



Annexation

**Discipline A partially replaces discipline B**

Methods of A apply to objects of B

In the case of **Computational social science** : Researchers from Statistical physics and Computer science analyse « social data ».

There is a risk that the object of social science is annexed by computational social science

# The (power)relationships between disciplines

Example of annexation ?

## SOCIAL PHYSICS

SOCIAL PHYSICS ABOUT BLOG VIDEOS PAPERS NEW & EVENTS CONNECTION SCIENCE CLASS

See also, Cho, 2009, *Ourselves and Our Interactions: The Ultimate Physics Problem?* *Science* 425, 406-408

Social Physics is a new way of understanding human behavior based on analysis of Big Data.

The contributors to the Social Physics are a set of researchers who are connected through their association with the Human Dynamics Lab at MIT.

Contributors include:

<http://socialphysics.media.mit.edu/about/>

Not “Physics” but another way of doing Sociology because the objects of study are not matter, space and time!

My experience

If sociolinguists do not explain, computational scientists ignore the conceptual framework of Sociolinguistics  
Debate within the Sosweet project : “incorrect language” and “nonstandard language” are not identical concepts !

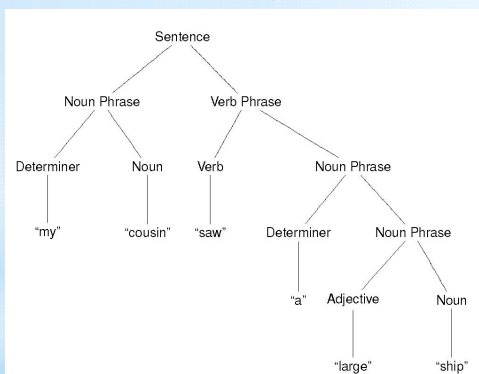


# However interdisciplinary research will help Language science facing a basic challenge

Bender, EM. and Good, J.. 2010. A Grand Challenge for Linguistics: Scaling Up and Integrating Models.  
White paper contributed to NSF's SBE 2020 initiative.  
<https://faculty.washington.edu/ebender/papers/GrandChallenge.pdf>

## 1/ The complexity challenge : integrating the perspectives

Building theories that integrate the different aspects of language: the structural, the social and the cognitive sides at the collective and individual levels



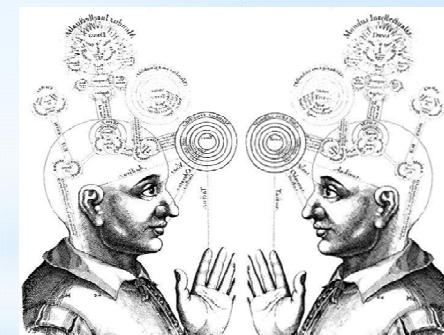
### Structural side

Phonology,  
grammar, and the  
lexicon



### Social side

Language usage and  
interaction



### Cognitive side

Brain and cognitive  
devices underpinning  
language usage and  
interaction

# However interdisciplinary research will help Language science facing two basic challenges

## 2/ The data challenge

Meeting the complexity challenge involves collecting and processing large sets of thick and diverse data

### **Datasets allowing to extract information about**

- The structural organization underpinning the linguistic knowledge
- The social usage of language
- The cognitive cues of the speakers

### **Need for computational methods**

- For assisting linguists in analyzing large-scale data
- For building and testing models accounting for language complexity

# Conclusion : The role of sociolinguistics in interdisciplinarity

1. Language science needs collaboration with...

- Subfields of Social science
- Subfields of Cognitive science
- Subfields of Computational science

2. A way of avoiding unfair subcontracting and annexation is to ensure that each discipline meets its own challenges

3. The inherent interdisciplinarity of sociolinguistics and its numerous connections with other fields place it in a strategic position for facing the challenge of integrating the linguistic, cognitive and social aspects of language both across groups and within individuals.

Contact : [jean-pierre.chevrot@univ-grenoble-alpes.fr](mailto:jean-pierre.chevrot@univ-grenoble-alpes.fr)



THANK  
YOU

The image shows the words "THANK YOU" painted in large, bold, yellow capital letters on a dark asphalt surface. The paint is cracked and shows signs of wear. The text is arranged in two lines: "THANK" on the top line and "YOU" on the bottom line. The background is a textured asphalt pavement with some dark stains and a concrete curb visible on the left side.