



Manchester Forum in Linguistics

A conference for early career researchers in linguistics

University of Manchester

28-29 April 2017



www.mfilconf.co.uk

mfilconf@gmail.com

The 2017 Manchester Forum in Linguistics is supported by the University of Manchester's Department of Linguistics and English Language, artsmethods@manchester and the Linguistics Association of Great Britain.

Contents

General information	1
Programme	2
Plenary speakers	4
<i>Origin stories: Evaluating theories for the origins of Canadian raising</i> , Amanda Cardoso	5
<i>Using quantitative syntactic data to assign dates of composition to Old English texts</i> , Aaron Ecay	6
<i>Social media and the new frontiers of forensic authorship profiling</i> , Andrea Nini	7
<i>Language structure from social practice</i> , Péter Rácz	8
Talks	9
<i>Velar nasal plus in the north of (ing)land</i> , George Bailey	10
<i>Expletives in Icelandic: A diachronic study</i> , Hannah Booth	11
<i>Minimal effort and the relative order of double topics in Mandarin Chinese</i> , Sherry Yong Chen	12
<i>Implementing reproducibility in phonetic research: A computational workflow</i> , Stefano Coretta	13
<i>The syntax-pragmatics interface: Unifying contrastiveness: Evidence from North-Eastern Italian dialects (NEIDs)?</i> , Simone De Cia	14
<i>How many or how much? – Counting and measuring in Chinese</i> , Christopher Hicks	15
<i>English influence of verb-noun collocations in Quebec French</i> , Sarah-Kay Hurst	16
<i>Agree manifestations in Arabic</i> , Marwan Jarrah	17
<i>Verb third in Kiezdeutsch: The development of an (optional) subject position?</i> , Benjamin Lowell Sluckin	18
<i>Voice onset time as an index of sexuality</i> , Bradley Mackay	19
<i>Labelling in ECM and its implications for phase theory</i> , Manabu Mizuguchi	20
<i>Linguo-cognitive aspect of lexico-semantic group of nouns denoting “lodging” in English in comparison with the Spanish and Armenian languages</i> , Lilit Movsesyan	22
<i>What does a suspicious review look like? Developing linguistic heuristics for detecting fake online reviews</i> , Olumide Popoola	23
<i>Differential recipient marking and the typology of argument marking: New evidence from Romance languages</i> , Hagay Schurr	24
<i>German psych verbs and the stage-level/individual-level distinction</i> , Livia Sommer	25
<i>On first conjunct agreement in Spanish</i> , Bo-Fan Syue	26
<i>Experimental insights on the Romanian ethical dative</i> , Alina Tigau	27
<i>Fronting in the T[ɹ:]n: Evidence of /u/-fronting in North East England</i> , Jasmine Warburton	28
<i>In the /i/ of the beholder: Looking at vowels with ultrasound tongue imaging</i> , Fabienne Westerberg	29

Posters	30
<i>Surprises on the hunt for causes</i> , Bernat Castro López	31
<i>Proficiency and working memory in L2 sentence processing: Evidence from eye-movement data</i> , Zhe Gao, Yanhui Zhang & Meng Jiang	32
<i>A proposal for a shared classification of code-switching and code-blending</i> , Eleonora Marocchini	33
<i>Camfranglais as a transnational practice – being Cameroonian in a ‘superdiverse’ context</i> , Constance Mbassi Manga	34
<i>Intuitive and reflective inference in the interpretation of metaphor</i> , Ramona Pistol	35
<i>Constructional change in the lifetime: it-cleft foci as a case study for change in individuals</i> , William Standing	36

General information

This booklet contains both the conference programme and the abstracts for all oral and poster presentations to be given at the 2017 edition of the Manchester Forum in Linguistics (mFiL 2017). Abstracts are divided into sections according to whether they are plenaries, talks or posters and arranged in alphabetical order by the surname of the lead author.

Other information about registration, accommodation and finding one's way around Manchester and to the conference venue can be found on the conference website: www.mfilconf.co.uk. Any updates regarding the conference should also be communicated by the organisers on social media: www.twitter.com/mfilconf and www.facebook.com/mfilconf. Should the need arise, more crucial information will also be circulated by e-mail.

All members of the organising committee will be wearing purple name badges. If you have any questions or problems, feel free to speak to one of them.

Talks will take place in either the Conference Room (C1.18) or Seminar Room 2 (CG58) in the Ellen Wilkinson Building's Graduate Centre at the University of Manchester on Oxford Road. The poster session as well as lunch and coffee breaks will be held in the atrium of the same building. With the exception of plenary speakers, who have each been allocated a one hour slot, speakers will have a total of 30 minutes: 20 or 25 minutes for the talk itself and 10 or 5 minutes for questions, this being the speaker's choice.

If you require Wi-Fi, eduroam is available in the Graduate Centre; however, for those coming from institutions that are not part of the eduroam network, if you ask a member of the organising committee, they will be able to provide you with a visitor username and password that will enable you to use the university Wi-Fi.

Organising committee:

- Stephen Nichols (lead organiser)
- George Bailey
- Fernanda Barrientos
- Hannah Booth
- Stefano Coretta
- Simone De Cia
- Deepthi Gopal
- Christopher Hicks
- Henri Kauhanen
- Donald Morrison
- Sascha Stollhans

Last update: 26 April 2017

Saturday 29 April

- 10:00-11:00 **Plenary III:** Dr Aaron Ecay, *Using quantitative syntactic data to assign dates of composition to Old English texts* (Conference Room)
General linguistics (Conference Room) **Syntax II** (Seminar Room)
Chair: Stephen Nichols Chair: Christopher Hicks
- 11:00-11:30 Olumide Popoola, *What does a suspicious review look like? Developing linguistic heuristics for detecting fake online reviews* Hannah Booth, *Expletives in Icelandic: A diachronic study*
- 11:30-12:00 Sarah-Kay Hurst, *English influence of verb-noun collocations in Quebec French* Manabu Mizuguchi, *Labelling in ECM and its implications for phase theory*
- 12:00-13:30 **Lunch & coffee** (Atrium)
Syntax III (Conference Room) **Phonology II** (Seminar Room)
Chair: Simone De Cia Chair: George Bailey
- 13:30-14:00 Benjamin Lowell Sluckin, *Verb third in Kiezdeutsch: The development of an (optional) subject position?* Jasmine Warburton, *Fronting in the T[ɹ:]n: Evidence of /u/-fronting in North East England*
- 14:00-14:30 Christopher Hicks, *How many or how much? – Counting and measuring in Chinese* Bradley Mackay, *Voice onset time as an index of sexuality*
- 14:30-15:30 **Careers panel** (Conference Room)
- 15:30-16:00 **Coffee** (Atrium)
- 16:00-17:00 **Plenary IV:** Dr Amanda Cardoso, *Origin stories: Evaluating theories for the origins of Canadian raising* (Conference Room)
- 17:00-17:30 **Closing remarks** (Conference Room)

Plenary speakers

Origin stories: Evaluating theories for the origins of Canadian raising

Amanda Cardoso, University of Leeds

Varieties of English across the world, such as Canadian English (Chambers 1973), have developed PRICE and often MOUTH nucleus raising before voiceless consonants, so that the vowels in the words *tight* and *tide*, and *lout* and *loud* are produced differently but are perceived as the same. Nucleus raising of PRICE and MOUTH in different varieties are often independent innovations, and this observation has led to a range of proposals to account for the origins and development of these types of patterns. However, it is often difficult to empirically evaluate these proposals as little or no recorded data exist for the time period before nucleus raising was present. The present talk examines a recently described instance of PRICE and MOUTH nucleus raising in Liverpool English (Knowles 1973, Cardoso 2015), which is used to evaluate three of these proposed theories.

Trudgill (1986) suggests that PRICE nucleus raising in varieties of Canadian English originated from new-dialect formation. A new dialect emerges from the intensive contact among several different source varieties, where small adjustments in individuals' speech occur over several generations (Trudgill 2004). Liverpool English, a relatively recent case of new-dialect formation, emerged in the mid-19th century as a result of extensive and prolonged immigration mostly from the surrounding areas of Lancashire and from Ireland (Cardoso 2015). Therefore, if PRICE and MOUTH nucleus raising in Liverpool English is the result of new-dialect formation, speakers born shortly after the dialect was formed would be predicted to exhibit PRICE nucleus raising.

Moreton & Thomas (2007) propose that PRICE nucleus raising is the result of different co-articulatory pressures from the following consonant, which are reanalysed in future generations. Their findings suggest that the offglide of PRICE before voiceless stops is peripheralised, which translates into a diphthong that is composed of a short nucleus and a long offglide. Subsequent generations reanalyse the short nucleus as a raising of the nucleus. Therefore, if PRICE nucleus raising in Liverpool English is the result of phonetic co-articulatory pressures, it would be predicted that speakers of Liverpool English would begin with a difference in the ratio of nucleus to offglide before voiceless consonants compared to voiced consonants, but not nucleus raising. In this proposal, PRICE nucleus raising should emerge gradually over time. Moreton & Thomas (2007) do not include MOUTH nucleus raising in their analysis, but do suggest that a similar proposal should hold for MOUTH nucleus raising.

Gregg (1973) proposes that Canadian Raising is the result of historical reflexes of PRICE and MOUTH vowels. The following voiceless obstruents inhibit PRICE and MOUTH nuclei from lowering. In other words, PRICE and MOUTH before voiceless obstruents are stuck in a previous historical reflex of the target vowels, while PRICE and MOUTH in other contexts continue to later stages of the target vowels. Therefore, if PRICE and MOUTH nucleus raising is the result of historical non-lowering, it would be predicted that speakers of Liverpool English would begin with all the nucleus of PRICE and MOUTH the same in all environments and a gradual lowering of the nucleus over time would occur, except before voiceless consonants.

The current investigation provides a dynamic acoustic analysis of the PRICE vowel in the Origins of Liverpool English Corpus (Watson & Clark 2017), a dataset with speakers born shortly after new-dialect formation in Liverpool, and a recently collected corpus (Cardoso 2015). These datasets are used to evaluate the predictions of the three proposals for the origins of these processes.

My findings suggest that PRICE and MOUTH nucleus raising is not present in the oldest speakers of the current investigation. Therefore, PRICE and MOUTH nucleus raising in Liverpool English is likely not the result of new-dialect formation. Despite the absence of nucleus raising for the older speakers, I did find a difference in the inflection points for PRICE before voiceless consonants and voiced consonants. Furthermore, nucleus raising before voiceless consonants and nucleus lowering before voiced consonants is found to emerge gradually in apparent time, which results in a clear difference in the height of the nucleus of PRICE before voiceless consonants compared to voiced consonants for young speakers. These results provide some support for the proposal by Moreton & Thomas (2007) and possibly for Gregg (1973).

Using quantitative syntactic data to assign dates of composition to Old English texts

*Aaron Ecaj, University of York**

In this talk I will explore an ongoing project to use quantitative syntactic data to assign dates of composition to Old English texts. One of the primary texts of interest in this regard is the poem *Beowulf*: though it survives only in a manuscript from ca. 1000, most linguists and scholars from other disciplines hold the opinion that the text is actually much older, from the 8th or early 9th century. After surveying evidence from a variety of syntactic features of OE, we observe that syntactic evidence is in line with this view (and lends further weight against proposals in the literature that *Beowulf* has a later date of composition). Turning to more abstract questions, I will aim to illustrate how a traditional philological question like the dating of texts can be recast first in terms of “digital humanities” and then in terms of statistical algorithms. Projects which span these domains of scholarship have the potential to enrich the field and provide an opportunity for young researchers to develop their own perspective on venerable research questions.

*Presenting joint work with Susan Pintzuk and Richard Zimmermann.

Social media and the new frontiers of forensic authorship profiling

Andrea Nini, University of Manchester

Authorship profiling could be defined as the analysis of the language of an anonymous text with the aim of extracting clues that can narrow down the list of suspect authors. Within a forensic linguistic context, this type of analysis is applied in ‘single text problems’, that is, cases of disputed authorship of malicious communication for which a small comparison set of suspect writers is missing (Grant 2008). These cases often involve ransom demands or threatening messages in high-risk situations. Despite the importance of providing intelligence quickly and reliably for cases involving profiling, forensic linguistics has made very little progress in developing objective, reliable and fast methods to tackle these problems. In the majority of the successful cases, forensic linguists have employed *ad hoc* strategies coming from sociolinguistics and dialectology involving extensive manual analyses. Although research in computer science has revealed that it is possible to apply reliable algorithms to profile the anonymous author of a text in terms of their basic demographic details such as gender or age (e.g. Schler et al. 2006, Argamon et al. 2009), the application of these algorithms is not always possible for forensic malicious texts, both because of their register characteristics and their short length. The main problem for forensic authorship profiling is the absence of resources containing sociolinguistic or dialectal information on the use of linguistic structures that can be easily queried automatically. Until a solution to this problem is found, a reliable and general method of profiling the author of an anonymous malicious text will be impossible to develop.

In the present talk, after explaining the state of the art of forensic authorship profiling using real-life cases as examples, it will be proposed that the new frontier of forensic authorship profiling lies in the mining of social media. For the first time in history, linguists can benefit from a massive amount of data often rich in geographical and social details in the form of social media texts published online, such as twitter posts, Facebook posts, forum posts or blogs. By mining the cyberspace for this type of data it is possible to perform on-the-fly and specialised *ad hoc* sociolinguistic studies or dialect surveys that can be key in extracting clues from questioned malicious texts. The problem of the absence of resources can thus be solved using methods to extract sociolinguistic information from large amounts of social media data. The example that will be presented in this talk is the use of geotagged tweets to deduce the geographical origins of the authors of malicious texts using a corpus of about 9 billion words including 890 million tweets collected in 2014. It will be demonstrated that it is possible to extract the dialectal signal of salient lexical items in a text and from that to deduce the most likely geographical origin of an anonymous author.

The talk will outline the progress made so far and suggest new directions. It will be proposed that, besides dialectal information, it is also possible to infer richer socio-demographic details about the anonymous author just using spatial information with the addition of census data. Moreover, it will be pointed out that future research will have to address the problem of the selection of features, which is currently too heavily dependent on the analyst’s expertise.

Language structure from social practice

Péter RÁCZ, University of Bristol

Our language is shaped by the world we inhabit. Some of our cues are environmental: tigers have stripes, green fruit is probably not ripe, a sudden noise means trouble. Most of them, however, are social. Human history is a history of communities and shared habits. Rituals and practices have a tremendous importance as exponents of accepted conventions. The influence of social practice, one way or another, have been well documented for individual linguistic traits such as category formation or grammaticalisation. Diachronic linguistics and variationist sociolinguistics have contributed to a broader view of language structure as a result of language change driven by individual practices in the community. Social practice is now also recognised as an important predictor in speech production and processing.

Using examples from my research, I give an overview of the ways in which social practice affects language structure on various levels. First, I use typological data collected with Fiona Jordan to look at the effect of marriage rules and social complexity on variation across kinship term systems. Second, based on work with Joe Blythe, I show how practices in name avoidance shape grammar in Murrinhpatha, an Australian Aboriginal language. Then I move on to the effect of our perception of the interlocutor, from my lab work with Jen Hay and Christoph Bartneck in New Zealand. Finally, I illustrate how our knowledge of social contexts shapes lexical processing in an online experimental task, citing work with Jen Hay, Janet Pierrehumbert and Clay Beckner.

These studies show the pervasive effect of social practice on language structure, from semantic categories to grammatical structures to low-level lexical processing. By drawing a connection across these domains, I advocate a view of language as a fundamentally social process, framed by the everyday practice of the individuals who use it.

Talks

Velar nasal plus in the north of (ing)land

George Bailey, University of Manchester

Despite the breadth of work investigating (ing) across the English-speaking world (see Hazen 2006), the situation in northern dialect regions of England where competition exists between *three* variants, [m]~[ɪŋ]~[ɪŋg], is comparatively understudied. This paper highlights the ways in which (ing) in the north west of England departs from the previously attested patterns, and is the first thorough investigation of the non-coalesced [ɪŋg] form (termed ‘velar nasal plus’ by Wells 1982) that is exclusive to these varieties. This study also covers the behaviour of word-final velar nasals in stressed syllables (e.g. [sɪŋg]~[sɪŋ] ‘sing’), marking the first time that this has been studied under the variationist paradigm.

Through the analysis of spontaneous and elicited speech from twenty-four sociolinguistic interviews conducted in the Manchester area (which yield over 5,200 tokens), this study first reveals that grammatical category plays no role in constraining (ing) variation in these varieties. This is particularly striking given the strength of this effect elsewhere, both in the US (Labov 2001) and elsewhere in the UK (e.g. Tagliamonte 2004 in York). It likely stems from the overall predominance of *-in* in conversational speech, where rates are consistently high across all speakers and remain so even in many contexts that typically disfavour this variant.

In stressed syllables, by contrast, language-internal factors play a crucial role in constraining variation; it is best modelled by positing a post-nasal /g/-deletion rule that shows sensitivity to the lifecycle of phonological processes and its synchronic implications (Bermúdez-Otero & Trousdale 2012). Social factors play no role here; its rate of application is predicted almost entirely by assuming cyclic application across stem-, word-, and phrase-level phonology such that the surface rate of deletion is a function of how many chances the rule has had to apply across this stratified phonological system (see Guy 1991 for a similar approach to t/d-deletion). There is also preliminary evidence to suggest a diachronic change towards inhibition of the deletion rule in phrase-final position; the exact reasons for this development are unknown, but it could be seen as a parallel change to other ‘velar fortition’ processes that target the same position, e.g. ejection of phrase-final velars, also found to be increasing over time (McCarthy & Stuart-Smith 2013). Mixed-effects logistic regression confirms the importance of these internal constraints, as well as the absence of any social stratification.

The results here have implications for the homogeneity of sociolinguistic variables across communities, and touch upon important topics in phonological theory. They also suggest that, despite being subject to variationist study for almost sixty years dating back to Fischer’s (1958) seminal work in New England, there are still new insights to be gained from the study of (ing) variation.

Expletives in Icelandic: A diachronic study

Hannah Booth, University of Manchester

This paper presents results from an ongoing corpus-based investigation of the diachronic evolution of the expletive *það* in Icelandic. Modern Icelandic exhibits a particularly diverse range of constructions with *það* (e.g. existentials, weather predicates, impersonal passives, extraposition). However, in Old Icelandic (1150-1350), expletive *það* is notably absent, as in the following attested constructions:

- (1) _____ Var fátt manna heima. (existential)
 ØEXPL was few men.GEN at-home
 ‘There were few men at home.’ (IcePaHC, 1350.FINNBOGI.NAR-SAG,655.1696)
- (2) _____ Líður nú af veturinn og _____ vorar. (weather predicate)
 ØEXPL passes now of winter.DEF and ØEXPL becomes-spring
 ‘The winter now passes and it becomes spring.’ (IcePaHC, 1350.BANDAMENNM.NAR-SAG,.147)

The historical circumstances whereby Icelandic developed overt expletives has been only scarcely investigated (Eythórssón & Sigurðardóttir 2016, Faarlund 1990, Rögnvaldsson 2002) and previous studies *það* to be a relatively recent development. I discuss data for overt expletives in 3 construction types in the IcePaHC corpus (Wallenberg et al. 2011) over 5 time periods and compare them to ‘null’ expletives (contexts where modern Icelandic would require an overt expletive, but where *það* is absent in the data, e.g. 1-2):

Time period	Extraposition type		Presentational type		Impersonal type	
	Main clauses	Embedded clauses	Main clauses	Embedded clauses	Main clauses	Embedded clauses
1150-1350	94%	100%	19%	33%	10%	9%
1351-1550	87%	100%	24%	25%	20%	9%
1551-1750	85%	89%	26%	50%	23%	16%
1751-1900	97%	94%	72%	100%	33%	17%
1901-2008	93%	75%	96%	86%	74%	17%

Table 1: Proportion of overt expletives (versus null) 1150-2008

This data allows the following conclusions to be drawn:

1. The proportion of overt expletives in the extraposition type is already very high in the earliest period (1150-1350) and remains relatively stable over time.
2. The proportion of overt expletives in the presentational and impersonal types 1150-1350 is much lower, but they do still occur. This challenges the standard view that Old Icelandic had no overt expletive aside from in the extraposition type (Faarlund 1990:66; Rögnvaldsson 2002).
3. The presentational type shows a dramatic increase over time, particularly from 1751 onwards.
4. The impersonal type also shows a dramatic increase at least in main clauses, but the development of overt expletives lags here behind that observed in the presentational type.
5. Across all construction types in almost all periods, the preference for overt expletives appears to be less strong in embedded clauses than in main clauses.
6. Even by the latest period (1901-2008), overt expletives are still not fully obligatory in any of the construction types.

More generally, I also discuss how the historical data can shed new light on the possible motivations for expletive *það* in Icelandic. In sum, this investigation challenges a number of standard assumptions in previous work on expletives in Icelandic.

Minimal effort and the relative order of double topics in Mandarin Chinese

Sherry Yong Chen, University of Oxford

This paper presents new data to argue for syntactic constraints on the relative order of nominal topics in double topicalisation constructions in Mandarin Chinese (henceforth MC)

MC topics are derived via either base generation or movement (Badan & Del Gobbo 2011, Huang 1982, Li 2000, Shyu 1995) and always appear in clause-initial position. Being a topic-prominent language, MC is a window into many typologically interesting constructions. Of particular interest here is the so-called double topicalisation construction, where two nominal topics are found in clause-initial position. What kind of surface structure can capture the various internal organisations of double topicalisation constructions?

To answer this question, I take the Minimalist Program as a broad theoretical framework and adopt the Topic Phrase (TP) analysis (Gasde & Paul 1996), which treats TP as a functional category with maximal projection above an IP or a recursive TP, with nominal topics occupying the specifier position of TP and the Topic head can be either empty or occupied by topic markers. In my analysis, I first show that in double topicalisation constructions, the base-generated topic must precede the moved topic:

- (1) a. [Hua], (meigui-hua)_i ta zui xihuan ______i
flower rose-flower 3SG most like
'Flowers, roses he likes them the best.'
b. *[Meigui-hua]_i, (hua) ta zui xihuan ______i
rose-flower flower 3SG most like
*'Roses, he likes flowers best.'

This constraint on the relative order of base-generated topics and moved topics may be explained by the locality constraint on movement, which follows the idea of minimal effort that places shortness requirements on movement operations (Li 2000, Hornstein, Nunes & Grohmann 2005).

Secondly, when both topics are derived via movement, they must be in a 'nested' dependency relation with their gaps in the comment clause:

- (2) a. [Zhejiang shi]_i, (lixiansheng)_j wo gaosu guo ______j ______i
this.CL matter Mr Li I tell PERF
'I have told Mr Li about this matter.'
b. *[Lixiansheng]_i, (zhejiang shi)_j wo gaosu guo ______i ______j
Mr Li this.CL matter I tell PERF
*'I have told Mr Li about this matter.'

This may be explained by Nested Dependency Constraint (Fodor 1978, Kaplan & Bresnan) or Path Containment Constraint (Pesetsky 1982); be it a requirement of the grammar or of the parsing mechanism of language processing, the constraint for double dependencies to be nested is readily motivated and will be used to guide how the two gaps are to be filled, such that these dependencies can be formed or parsed in a speedy and accurate manner, with minimal effort required from either the grammar or the parser.

These findings challenge the view that MC topics are merely constrained via a semantic "aboutness" relation with the comment clause (cf. Xu & Langendoen 1985). They suggest that topicalisation is subject to syntactic constraints that govern the relative order of double topics in MC, which fall under the broad notion of minimal effort. MC has a rigid word order but little inflectional morphology, which sometimes makes it difficult to pin down the grammatical function of a constituent based on traditional diagnostic tests. Nevertheless, the kind of syntactic constraints described above allow MC speakers to interpret double topicalisation constructions in a relatively effortless manner. The grammar that underlies MC topicalisation is designed as such, so that minimal effort is required from the speaker to resolve potential ambiguities despite the lack of overt morphological cues.

Implementing reproducibility in phonetic research: A computational workflow

Stefano Coretta, University of Manchester

This paper deals with the concept of reproducibility and describes a workflow for the development and dissemination of reproducible research in phonetics. A piece of research is reproducible when, along with its results, the data and the computational operations that produced those results are made available to other researchers (Fomel & Claerbout 2009). Reproducibility is particularly helpful in those cases when replicability is not possible or difficult (Peng 2009), which is typically the case in phonetics. Even if they did not explicitly talk about reproducibility, Bird & Simons (2003) argued that linguistic descriptions should be accountable, i.e. they should be accompanied by the corpus upon which the description is based. Thieberger (2004) further stated that descriptions should be directly linked to the linguistic dataset that produced them. While these approaches, which are the first step towards reproducibility, are increasingly becoming a standard in descriptive morphology and syntax, this is less so the case in phonetic research.

Normally, the computational workflow of phonetic research is constituted by the following phases: (a) developing the source code and producing the relative documentation, (b) running the code on the acquired and derived data, and finally (c) writing up the results in a publishable form (article/book, website, ...). In this talk, I will first review the principles of reproducibility in general (Sandve et al. 2013) and how they can be applied in linguistics (Maxwell 2013). I will then introduce a collection of tools (some already existent and some developed by the author) that support the phonetician throughout this workflow with the aim of generating reproducible research. Since most phonetic research is conducted using Praat (Boersma & Weenink 2016), the computational tools discussed here are meant to be used in conjunction with this specific software.

Concerning phase (a) of the workflow, it is common to write scripts using the Praat scripting language to automate repetitive tasks in Praat, like formant or pitch extraction, duration measurements, and voice reports. Syntax highlighting and automatic word completion (autocomplete) are useful features which help the researcher in the writing process. Scripting in Praat can also benefit from literate programming (Knuth 1984), which is the basis of reproducible research (Peng 2015). In literate programming, a source file contains both the code of the programme (or script) and its documentation. Literate programming for Praat is made possible by the software `lmt` (literate markdown tangler) and `noweb`. Using either software, the process called “tangling” extracts (tangles) the code from the source file and creates a script containing the tangle code (which can then be run with the appropriate software, in this case Praat). The programme `pandoc`, instead, allows “weaving” the source code, which is the process of producing a human-readable documentation file. The advantage of using literate programming is that a script can be included in the research write-up and easily reproduced by subsequent users. To date, the Atom editor (<https://atom.io>) supports plain and literate Praat highlighting and autocomplete, thanks to the package `languagepraat` (<https://atom.io/packages/language-praat>).

Moving to phase (b), the processes of writing, tangling, weaving and running Praat scripts can be automated using the R language (R Core Team 2015) and the package `speakr` (which is under development). Regarding phase (c), the package `knitr` (Xie 2014) already adds support for reproducible research in R. Thus, `speakr` and `knitr` can be used in combination to facilitate the use of Praat scripts and documentation of the research. Finally, data, source code and documentation (which includes the results write-up) can be made available to the research community using online services like GitHub (<https://github.com>), the Open Science Framework (<https://osf.io>), or the dedicated Tromsø Repository of Language and Linguistics (<https://opendata.uit.no/dataverse/trolling>), among others. To illustrate the proposed workflow, I will present some practical applications of this method, partly based on the work which I am currently undertaking for my PhD project.

The syntax-pragmatics interface: Unifying contrastiveness: Evidence from North-Eastern Italian dialects (NEIDs)?

Simone De Cia, University of Manchester

In Lamonat and Sovramontino, the NEIDs under investigation, contrastive elements, be they topics or foci, are limited to one occurrence per sentence and must be sentence-initial. The paper aims to (i) shed more light on the variation in the encoding of pragmatic contrastiveness in Romance and (ii) attempt a unitary syntactic analysis of contrastive foci and contrastive topics in the two dialects, namely Lamonat and Sovramontino.

Kiss (1998) argues that an element is contrastive when it specifically refers to an information unit that is contrary to the corresponding interlocutor's presuppositions. A contrastive element can also involve an explicit choice among a finite set of alternatives where the presupposed contrasting information unit has not been previously mentioned in the discourse. Contrastiveness is traditionally considered as a property of foci, more specifically a subtype of narrow focus. López (1999, 2009), nevertheless, shows that in some Iberian varieties of Romance, contrastiveness is not only limited to foci, but can be also borne by left-dislocated, topical constituents. Typologically, Erteschik-Shir (2007) highlights the common syntactic-pragmatic properties of contrastive foci and contrastive topics. She notes that, in many languages (i) sentence initial topics represent so-called contrastive or switch topics and consist of a specific type of topic that has some focus properties; and (ii) sentence initial foci represent restrictive or contrastive foci consisting of a specific type of focus that has some topic-like properties. Lamonat and Sovramontino obey this typological trend and exhibit a sharp distinction between contrastive and non-contrastive elements.

In Lamonat and Sovramontino, all contrastive constituents must surface at the leftmost position of the sentence regardless of whether they constitute foci or topics. By adopting Rizzi's (1997) split-CP, I investigate which left peripheral position contrastive elements occupy in these two varieties; most importantly, I critically evaluate whether or not this position can be regarded as the same for contrastive topics and contrastive foci. To this end, I look at root interrogatives, D-linked *wh*-elements and high adverbs. I propose that in Lamonat and Sovramontino both contrastive foci and topics target the same leftmost position in the left periphery, which is fully activate through an [EDGE] feature every time a contrastive element enters the derivation. This position is beyond ForceP, as shown in the following labelled bracketing: [XP [ForceP [TopicP [FocP [FinP [TP]]]]]].

XP dominates ForceP as (i) contrastive elements are not affected by the sentence type or mood; (ii) contrastiveness is only licensed at the sentence leftmost position; and (iii) despite being topic-like, contrastive elements are not recursive. This position closely resembles FrameP in Romance which hosts 'frame setting' topics such as HTs and circumstantial adverbs (Benincà 2001). Poletto (2000, 2002) and Benincà & Poletto (2004) argue that this 'scene setting position' is located structurally higher than Force. The pragmatics of 'frame setting' elements is fully compatible with contrastiveness as speakers pick a specific element that, implicitly or explicitly, contrasts with the other elements that they could have picked as frame-setters. The paper investigates the similarities and differences between XP and FrameP, assessing whether or not these two positions can be conflated or are in fact one and the same.

Not only does assuming the existence of a single syntactic position for contrastive elements help us better account for non-canonical constituent orders in Romance dialects such as the ones in question, but also typologically aligns NEIDs with other languages, like Japanese and Finnish, which mark contrastiveness through different strategies. In the case of Lamonat and Sovramontino, but also in Finnish, this is done at the syntax level, whereas Japanese resorts to affix marking (see Shimojo 2011 for Japanese and Vilkuña 1995 for Finnish). To conclude, contrastiveness, to some extent, has the property of superseding topic or focus status: our research goes towards a unitary treatment of this syntactic-pragmatic aspect of language in the Romance varieties under investigation.

How many or how much? – Counting and measuring in Chinese

Christopher Hicks, University of Manchester

Some nouns refer to entities which naturally divide into discrete countable units, while others refer to substances which do not. This is standardly termed the count-mass distinction. In languages like English, the first type of noun – the count noun – can combine directly with a numeral when being counted, i.e. *two books*, whereas the second type – the mass noun – must appear with a measure word when being counted, i.e. *two grains of sand* and not *two sands*. In other languages, like Mandarin Chinese, nouns are said to never combine directly with a numeral to be counted; a small element called a classifier is necessary. This has led some to claim that all nouns in Chinese are mass nouns (Chierchia 1995 and Krifka 1995). However, there is a wealth of data that suggests the count-mass distinction is alive and well in, and relevant to, the grammar of Mandarin Chinese. Working within the principles of the Minimalist Program (Chomsky 2000, 2001) but with a minimally specified Universal Grammar, this talk will present some of this data and put forward its consequences for the structure of the nominal phrase both in Chinese and cross-linguistically.

Firstly, and perhaps most importantly, data I have collected shows that a classifier is not always necessary when counting entities in Mandarin Chinese – it seems that it is possible to drop the classifier in certain contexts, primarily in spoken, colloquial Chinese. Based on the types of classifiers that seem optional, I propose this as one of the clearest pieces of evidence for the count-mass distinction in Chinese, linking it to the function of classifiers and locating this phenomenon in the typological space with “non-classifier” languages.

I also contrast how there is a class of measure words in English – usually nouns in their own right – for use with mass nouns, whilst count nouns can appear on their own, with the prevalent view for Chinese: that there is a single grammatical category combining both counting classifiers and measure words. Based on the differences in behaviour between “counters” and “measurers”, I investigate whether there is sufficient evidence for the analysis of classifiers as a single grammatical category, or if there should be two subcategories, or if, like English, measure words should be treated more in line with nouns.

Following from the categorisation of these elements, many linguists also propose that counting and measuring involve different structural configurations. Taking classifiers and measure words as a single category might require a structural contrast, but how do you ensure that classifiers and measure words from this single category only appear in semantically appropriate configurations? On the other hand, if they belong to different categories, then is there the motivation and evidence for different structures? How does this compare to the structures proposed for counting and measuring in English? Using the data, I show that there is motivation for different structures, linking my conclusions to the availability of optional counting classifiers and corresponding structures in other languages.

My findings have implications for various aspects of current syntactic theory. Firstly, a detailed look at counting and measuring structures in Chinese, which has overt elements for both functions, sheds light on how structure may align with function both internal to a single language, and cross-linguistically through comparison with English. Furthermore, this extends to a consideration of grammatical Number, which is present in English but arguably less so, if at all, in Chinese. The count-mass distinction, and count nouns more specifically, link conceptually to the notion of Number, and my analysis considers whether this link might also exist on a structural level. An investigation into issues such as these, at the interface of semantics and syntax, might have far-reaching consequences for theories of syntax, the acquisition of grammar, and language typology.

English influence of verb-noun collocations in Quebec French

Sarah-Kay Hurst, Indiana University Bloomington

This study evaluates ten verb-noun collocations in Quebec French (QF), using journalistic usage and lexicographic representation to evaluate the integration of innovative variants which seem to be calqued from English. For example, in standard QF, the expression ‘to make sense’ is rendered by the expression *avoir du sens* (literally ‘to have sense’). However, *faire du sens* (literally ‘to make sense’) is also attested and this change in the collocate appears to be due to contact with English. Integration of the innovative variants is evaluated in this study by journalistic usage (six online QF newspapers) and lexicographic treatment (seven QF dictionaries).

Collocations, which occupy an importance place in phraseology, second language acquisition, corpus linguistics and lexicography (Hausmann 1979, Cowie 1981, Cop 1988, Bossé-Andrieux & Marsechal 1998, Howarth 1998, Mel’čuk 1998, Siepmann 2005, Verlinde, Binon & Selva 2006), are sometimes defined in terms of statistical co-occurrence as two lexical units that occur together frequently (cf. Hausmann & Blumenthal 2006). Traditional, bipartite collocations are considered to consist of a base, which is selected independently, depending on the communicative desires of the speaker, and a collocate which is chosen as a function of the base (cf. Hausmann 1979, Howarth 1998). For example, for the verb-noun collocation *make sense*, the nominal base *sense* selects the verbal collocate *make*. Since the collocate depends on the base, the choice of the collocate may differ between languages (as is the case between English and standard QF) and may be susceptible to change.

Instances in which a traditional verb-noun collocation alternates with an innovative variant that seems to be calqued from English may also be couched within a discussion of semantic Anglicisms. Semantic Anglicisms or semantic calques are an especially covert type of borrowing in which a French word takes on an additional meaning due to imitative contact with English. The celebrated example is *réaliser* which used to only mean ‘to achieve’, but has come to also mean ‘to realize’ (in the sense of ‘to become aware of’ or ‘to come to understand’) due to contact with English *realize*. Semantic Anglicisms have been addressed in QF (Darbelnet 1976, Lagneux 1988, Paqout 1997), as well as other dialects of French (Picone 1996, Rifelj 1996, Wise 1997, Brasseur 2004, Bernard 2006).

The ten collocations considered in this study are (alphabetically by innovative variant, traditional verb given in parentheses): *faire* (*avoir*) *du sens* ‘**make** sense’, *passer* (*voter*) *une loi* ‘**pass** a law’, *passer* (*adopter*) *une résolution* ‘**pass** a resolution’, *paver* (*ouvrir*) *la voie* ‘**pave** the way’, *placer* (*faire*) *un appel* ‘**place** a call’, *placer* (*passer*) *une commande* ‘**place** an order’, *prendre* (*suivre*) *un cours* ‘**take** a class’, *prendre* (*faire*) *une marche* ‘**take** a walk’, *prendre* (*faire*) *un voyage* ‘**take** a trip’ and *sauver* (*gagner*) *du temps* ‘**save** time’.

The prevalence of the innovative variants in newspapers is measured in terms of the number of tokens as well as the percentage of the total number of tokens for the base, thus comparing the traditional and innovative variants. In order to procure approximate indications of frequency, the most prevalent forms of the ten collocations were searched in each of the six newspapers: the infinitive, the past participle, and all the forms of the present indicative and imperfect indicative. One hundred tokens for each collocation are also examined in more detail – fifty for the base with the traditional collocate and fifty with the innovative collocate (where possible). These results are analyzed by register, indicated by the type of writing, whether the tokens occurred (1) in the main body of an article, (2) in a blog, letter, opinion or editorial, (3) in a comment or (4) in a direct quote. These results are also divided by date.

Lexicographic integration is measured in terms of whether or not the innovative variants are included in the dictionaries and whether or not they are marked as Anglicisms or nonstandard in some way. Six of the ten collocations are found to be fairly well integrated according to journalistic usage and lexicographic treatment. This study thus considers an issue at the crossroads of phraseology, language contact and lexicography.

Agree manifestations in Arabic

Marwan Jarrah, Newcastle University

Data from Arabic varieties reveals that Agree relations (established between a probe and a goal; cf. Chomsky 2000, 2001) must have a phonetic manifestation whose specification depends on certain factors, including the ability of the goal to get assigned a morphological Case. When an Agree relation is established between a case-assigning probe and an overt goal, the phonetic realization of the Agree operation is achieved through case-marking the goal by the probe, as the case of the complementizer *?inn* in Modern Standard Arabic (MSA):

- (1) ?inna ?al-walad-a qara?-a ?ad-dars-a
Comp DEF-boy-ACC read.3MS-IND DEF-lesson-ACC
'The boy read the lesson.'

Following Mohammad (1990, 2000), Ouhalla (1997), Aoun et al. (2010), among many others, the subject *?alwalada* above is assigned Accusative case by *?inna*. Note here that *?inna* is not attached with any agreement suffix, unlike the cases when the subject is a pro. Consider the following sentence:

- (2) ?inna-uh qara?-a ?ad-dars-a
Comp-3SM read.3MS-ND DEF-lesson-ACC
'He read the lesson.'

Within the traditional Arabic grammar view (cf. Ibn Al-Anbari 1961) and some modern proposals (e.g. Mohammad 1990, 2000 and Bemamoun 1993), the weak form suffixed to *?inn* in (2) is a phonetic form of the pro which enters an Agree relation with *?inna*. For these authors, the pro must be realized when it is assigned Accusative Case, a condition which is barely attested beyond (Modern Standard) Arabic grammar. Alternatively and following Shlonsky (1997), I argue that the weak form attached into *?inn* in (2) is an inflectional suffix produced as an outcome of an Agree operation. I argue that *?inn* assigns the pro subject Accusative Case, which is not realized given nullness of the pro. The manifestation of an agree relation between *?inn* and the pro must be therefore achieved through a different strategy, namely via a PF reflex of valuation of $u\phi$ -features of *?inn*. This implies that such a reflex is deleted when *?inn* assigns morphological Case to the subject (as in (1)). This reasoning helps us understand a host of other related phenomena across Arabic varieties. For instance, in Jordanian Arabic, the complementizer *?inn* shows always overt agreement with the following element.

- (3) ?if-jab ?i?taraf ?inn-hum ?ixwat-uh z^harab-u l-walad
DEF-young man confess.3SM.PAST that-3PM brothers-his hit-3PM DEF-boy
'The young man confessed that his brothers hit the boy.'

?inn agrees with the preverbal subject *?ixwatuh* by virtue of the suffix *-hum* attached to the former. This observation is readily accounted for pursuing the approach that a manifestation of an agree relation must hold. Given that spoken Arabic has no case-marking system (see, e.g. Brustad 2000:27, Aoun et al. 2010:15), recourse is made to the PF reflex of valuation of $u\phi$ -features of *?inn* to secure a PF manifestation of the probe-goal relation between *?inn* and the subject in (3). Additionally, this approach accounts for word order asymmetries in MSA and lack of them in Arabic vernaculars in a systematic way. Impoverished agreement is resulted when T^0 assigns morphological case to a subject, the case we only find in VSO in MSA (see, Soltan 2007). This research contributes to the ongoing debate cross-linguistically on the interplay between case and agreement. Both are ways of securing a PF record of agree relations whose output is ruled by economy conditions (cf. Miyagawa 2010). In Arabic, there should be only one record, hence the systematic alternation between case and agreement. The latter is only used when the former fails short of securing the PF record, either because the goal is not overt or because the probe is not a case-assigner.

Voice onset time as an index of sexuality

Bradley Mackay, Newcastle University

While sexuality as an extra-linguistic variable has been studied in a variety of socio-phonetic contexts, the effect that sexuality has on conditioning voice onset time (VOT) remains an area which would benefit from further research. Studies which have been carried out (for English L1 speakers see: Crist 1997, Smyth & Rogers 2008, Smyth, Jacobs & Rogers 2003) have found mean VOT of voiceless stop consonants to be longer for homosexual males than it is for their heterosexual counterparts. Of the three stops, velar /k/ has been found to condition the longest VOT, and bilabial /p/ the shortest. Similarly, the following vowel context has been found to effect VOT, with tense vowels conditioning longer VOT than lax vowels. The three broad research questions this paper addressed were:

(1) *Is there significant difference in VOT measurements between self-defining heterosexual and homosexual males?* (2) *What effects do place of articulation and following vowel context have on VOT?* (3) *Is there a link between VOT and perceived masculinity?*

To approach these questions, respondents were separated into two sub-groups; five self-defining homosexual males and five self-defining heterosexual males. Each respondent was asked to repeat a list of CVC words with /p/ /t/ /k/ in word-initial position. These preceded a variety of different vowel contexts and were repeated five times, yielding a total of 1,050 tokens which were measured manually in Praat. Respondents were then asked to rate themselves on a scale of 1-10 with 1 being the least masculine-sounding, and 10 being the most. A t-test was then used to calculate significance figures.

Mean VOT was found to be longer for the homosexual respondents than for the heterosexual group. This difference was found to be significant ($p=0.02$). Tense vowels were found to condition longer VOT for both groups than lax vowels, supporting findings in previous studies (see Yao 2007, Cho & Ladefoged 1999). Place of articulation was found to effect VOT, with a general pattern of bilabial < alveolar < velar, which supports the findings of Byrd (1993), Hitchcock & Koenig (2014) and Cho & Ladefoged (1999). While a very general pattern was found between masculinity scores and VOT, with respondents scoring themselves low on the masculinity scale generally producing higher VOT scores, the relationship between self-defined masculinity and VOT was not found to be significant.

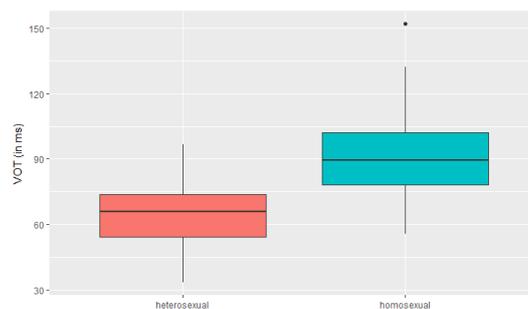


Figure 1: Mean VOT scores for heterosexual and homosexual respondents

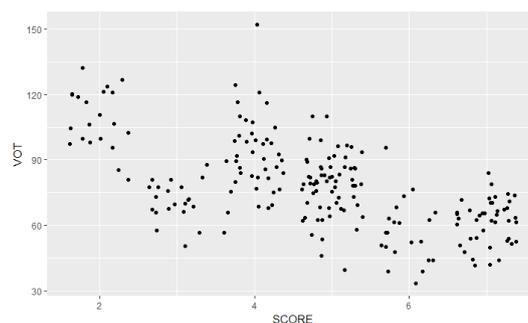


Figure 2: Mean VOT scores plotted against self-defined masculinity scores

Labelling in ECM and its implications for phase theory

Manabu Mizuguchi, Toyo University

[1] Problem

In minimalism, syntactic objects or sets are yielded by Merge in its simplest form and sets created are labelled by the Labelling Algorithm (LA) for interpretation at the interfaces (Chomsky 2013, 2015). This paper considers labelling in ECM and explore its implications. It has been argued that raising of the ECM subject out of the embedded clause into the matrix clause is optional and it can terminate in SPEC-T in the ECM complement (e.g. Lasnik 2003).

- (1) I believe [α the student [to [be [t intelligent]]]]

If the ECM subject does not raise out, a labelling problem will arise with α , the set of the form XP-YP. It has been argued that XP-YP is labelled by LA as far as X and Y agree; the subject, however, does not agree with T in the ECM complement, which is not finite but non-finite T.

[2] Proposal

I argue that the ECM complement is structured in the same way as its finite counterpart(=(2)), proposing that ECM predicates also select C (Bošković 2007, Rooryck 2002).

- (2) a. I believe [that [α the student [T $_{\phi}$ -is [t intelligent]]]]
b. I believe [C [α the student [T $_{\phi}$ [be [t intelligent]]]]]

This proposal solves the problem of labelling of α , explaining why *the student* can halt in the SPEC of T in the ECM complement: since C is merged, the head of *the student* (i.e. *n*) can agree with T in ϕ , which is inherited onto T from C as in the *that*-clause complement. In (2b) as well as in (2a), α can be labelled $\langle \phi, \phi \rangle$ thanks to the agreeing heads (Chomsky 2013, 2015).

The proposal, though it can solve the problem of labelling, leaves the problem of Case valuation. If Case is valued by ϕ -feature agreement(Chomsky 2000 et seq.), then agreement with T in the ECM complement will value the Case of *the student* as nominative just as in (2a). This is empirically wrong as the ECM subject receives accusative Case.

- (3) a. I believe him/*he to be intelligent. b. I believe he/*him is intelligent.

I argue that Case valuation is not achieved solely by ϕ -features but by ϕ -features plus tense (Epstein et al. 2012, Pesetsky & Torrego 2001). The argument that ϕ is not enough is supported by the fact that not only T but also R is engaged in ϕ -feature agreement. If only ϕ is relevant to valuation of Case, Case values assigned by ϕ -feature agreement would be the same. In the ECM complement, since T does not bear tense, ϕ -feature agreement of the ECM subject with the embedded T does not lead to its Case valuation; the subject does not get de-activated and it can multiply agree with the matrix R in ϕ to have its Case correctly valued as accusative.

[3] Implications

Labelling in ECM has implications for phase theory. First, non-phase heads do not inherit all of their properties from phase heads but do have properties of their own (contra Chomsky 2013, 2015): in the case of T, it inherits ϕ from C, a phase head, but bears \pm tense as its inherent property. This implication is in line with the view that lexical items are made up of features, which amounts to saying that lexical items without inherent features cannot exist.

Secondly, non-phase heads play an important role in the construction of phasehood. Note that in (2b), the matrix R agrees with the ECM subject across C and Case is valued as accusative.

- (4) I believe- v^* [R $_{\phi}$ [C [α the student [to [be [t intelligent]]]]]] AGREE(R $_{\phi}$, *the student*)

This suggests that C does not constitute a phase by itself. Given the proposed analysis, I argue that “phase” or phasehood is not intrinsically but contextually determined: it arises due to conspiracy of C (i.e. ϕ) and T (i.e. tense) and there is no notion of “phase head” (likewise, there is no notion of non-phase head). Since tense is missing in (4), unlike in (2a), C-T does not constitute a phase, allowing agreement across the C boundary.

The proposal that phase is defined by ϕ + tense predicts the following four patterns:

- (5) a. $\checkmark\phi, \checkmark$ tense b. $\checkmark\phi, \text{tense}$ c. ϕ, \checkmark tense d. ϕ, tense

In (5), only (5a) constitutes a phase. I argue that (5a-d) are empirically illustrated by (6):

- (6) a. [β The student will read the book]
 b. I believe [β C [the student [T ϕ [be [t intelligent]]]]](=(2b))
 c. [β [That the student is a genius] is obvious] (sentence with the clausal subject)
 d. The student seems [β to be happy] (subject raising)

As for (6c), Mizuguchi (2016) argues that ϕ is de-activated, only tense available in β , and claims that β does not constitute a phase. Given this argument, it argues for (5c), hence contextual determination of phasehood. In (6d), unlike in (6b), the student cannot halt in SPEC-T in β .

- (7) *It seems [β [α the student [to be happy]]] (Epstein et al. 2014)

This follows if unlike in the ECM complement, ϕ is missing from C and labelling of α fails in β . The nP can move out of β in (6d), which shows that β does not form a phase.

If ϕ + tense defines phasehood, phase amounts to a syntactic domain where Case valuation occurs as Case is valued by ϕ + tense. This outcome is empirically supported by scope relations in Japanese discussed by Takahashi (2010), who argues for this proposal with the v^*P phase.

[4] Conclusions

In this paper, I argue, discussing labelling in ECM, (i) that the ECM complement is CP, with ϕ inherited onto T, and (ii) that phasehood is determined contextually by ϕ + tense. If the argument is correct, it argues that agreement plays a vital role in derivation as it defines phase together with tense, and that agreement (or its equivalents) is a universal property.

Linguo-cognitive aspect of lexico-semantic group of nouns denoting “lodging” in English in comparison with the Spanish and Armenian languages

Lilit Movsesyan, Yerevan Brusov State University of Languages and Social Sciences

One of the current trends in modern linguistics is the study of language categorization. In this respect of special importance is semantic categorization of certain phenomena of reality.

In the present research the peculiarities of linguo-cognitive and linguo-cultural relationship are revealed through the study of the lexico-semantic group of nouns denoting “lodging” (hereinafter LSG).

The object of research are nouns (314 nouns in English, 291 nouns in Spanish and 330 lexical units in Armenian) literally or figuratively denoting “lodging” by their first four definitions recorded in dictionaries or revealed through corpus data analysis. Hence, through the multifaceted linguistic analysis of the above named nouns, the research aims to identify the linguo-cognitive aspect of LSG of nouns denoting “lodging” in English in comparison with Spanish and Armenian. The methodological toolkit involves descriptive, contrastive, semantic componential analysis, free association and statistical methods.

To achieve the aim of the research the following objectives are set forth: to *identify* English LSG “lodging” through the thorough consideration of linguo-cognitive, psycholinguistic and corpus approaches and relevant methods in comparison with the Spanish and Armenian languages to *create* linguo-cultural prototype images of the current LSG by means of psycholinguistic experiment of free associations to *discover* innovative methods of language categorization in compliance with the semantic prototype theory by E. Rosch to *reveal* linguo-cognitive peculiarities of LSG “lodging” through the identification of *FrameNet*, *WordNet*, *Sketch Engine* electronic projects and national corpora research possibilities to *categorize* LSG “lodging” as a cognitive frame.

There exist numerous methods of revealing semantic relations of words. One of these is the method of associative experiment which is conducted in different linguistic disciplines. The creation of associations much depends on our past experience and knowledge. Associations may be viewed not just as names of objects or events of reality but units reflecting the surrounding world in the human’s consciousness which contains specific cultural and national features. The associative experiment carried out on such prototypes as *house*, *casa* [kasa], *տն* [tun] helped us reveal the cognitive perceptions of the same notion by three different nations and underline the differences based on the man’s life experience, way of thinking, beliefs, perceptions and also on his settlement.

Along with the development of language corpora, different electronic projects have emerged as a source and tool for multifaceted linguistic investigations. The detailed analysis of their functional features on the material of the LSG under study helped us treat these electronic projects as modern reflections of the classical model of language categorization and revealed their implicit role in categorizing the world around.

The current research also highlights the apparent role of cognitive models in denoting both concrete and abstract phenomena and ensuring their understanding via frames, schemes, scenes, domains, scripts, etc. Frame with its two defined types (communicational and cognitive or conceptual) is given in this research as a unit of knowledge reflecting our perception of the world through different types of lodging.

What does a suspicious review look like? Developing linguistic heuristics for detecting fake online reviews

Olumide Popoola, Aston University

Online reviews are a key part of the consumer decision-making process. €23 billion per year of UK consumer spending is influenced by online reviews, with 82% of adults reading reviews before purchasing products or services and 78% fully trusting what they read. Globally, online reviews have a similar level of impact (Competition and Markets Authority 2015). Yet the proportion of reviews that are fake – i.e. written or commissioned by businesses without disclosure of commercial interest – is estimated at up to 18% (ibid.).

While review websites and businesses such as Amazon.com and Yelp, whose business models rely heavily on review content, have taken a legal and algorithmic approach to combating fake reviews, it has been left to consumer regulatory bodies to educate consumers to ‘tell-tale’ signs of deceptive reviews. Here the focus has been on non-textual elements of reviews such as username, profile creation date and ‘burstiness’ i.e. a rapid increase in quantity of reviews over a short time. This emphasis reflects academic research that has focused on network analysis and development of machine learning NLP tools rather than linguistic characteristics.

This presentation reports on analysis of the DeRev corpus of fake and authentic Amazon book reviews (Fornaciari & Poesio 2014) that utilises Rhetorical Structure Theory (Mann & Thompson 1992) to identify text-structural differences between fake and authentic reviews at three levels of linguistic analysis. At the level of discourse strategy it finds that fake reviews are more likely to focus on plot narration at the expense of actual book evaluation. At the level of coherence it finds that whereas authentic reviews use *Concession* and *Contrast* relations to recommend a book whilst acknowledging its limitations, fake reviews tend to offer unequivocal endorsement. Significant differences are also found in sentence structure, with fake reviews having less embedding and utilising fewer clauses per sentence.

These findings are used to produce heuristics that can help consumers flag online review content as suspicious. It also provides a rationale for feature selection in future computational/forensic linguistic research in this area.

Differential recipient marking and the typology of argument marking: New evidence from Romance languages

Hagay Schurr, Hebrew University of Jerusalem

In Differential Argument Marking (DAM), a subset of free (pro)nominal arguments may or must differ in morpho-syntactic coding from other arguments associated with the same semantic role or syntactic function (e.g. by adpositional marking). In Romance languages, DAM is most commonly recognized indifferent object marking (DOM) (Bossong 1991, Stark 2005, Iemmolo 2010), but recently identified in the marking of indirect objects associated with the semantic role of Goal (DGM) (Luraghi 2011, Fábregas 2015). However, differential Recipient marking (DRM) was only briefly mentioned in Spanish (Company Company 2003, Malchukov 2008).

Based on novel evidence from diachronic and web corpus studies in Spanish and French (11th-16th centuries), I demonstrate that French developed a different DRM pattern and that both patterns are typologically opposite (i.e. unlikely to co-occur); Spanish marks highly individuated Recipients (1), while French marks unindividuated ones, where *individuation* is understood as a cluster scale based on Hopper & Thompson (1980).

(1) Spanish

Si se=vea (Caroline) pues, la=envío contigo
if se(.RFL)=go.PRSSUBJ.3SG then 3SGF.ACC=send.PRS.1SG COM-2SG

‘If she goes (Caroline), so I send her over to you.’ (COM = comitative)

Contrasting with previous synchronic studies in DRM (cf. Malchukov 2008, Kittilä 2006, 2011), I propose grammaticalization paths for Spanish comitative-marked DRM (*con* ‘with’) and for French apudlocative-marked DRM (*chez* ‘at’). The grammaticalization paths share common properties with the recognized grammaticalization of “regular” dative markers (Creissels 2009) (2):

(2) *adessive* → *apudlative* (intransitive clauses) → *apudlative* (ditransitive clauses).

The abstract grammaticalization path in (2) reads such that DRM markers first mark location in place (‘be at/with’), then acquire dynamic motion towards Goal (‘go to’) and only then they mark the Recipient in caused motion or transmission events. The proposed grammaticalization path for Spanish DRM, based on the comitative preposition *con* (‘with’) has far-reaching implications: It counters previous research (Croft 1991, Stroh, Urdze & Stolz 2008) in that:

- (i) comitative markers can and do grammaticalize into Recipient markers
- (ii) (comitative-marking is not limited to so-called “dative-associated” semantic roles in Sinitic varieties (Chappell, Peyraube & Wu 2011) and
- (iii) combined with insights from previous research on DOM and accusative marking (Bickel & Witzlack-Makarevich 2008, Grossman 2013), this research offers a number of predictions concerning the distribution of both types of DRM patterns and the emergence of new dative and accusative argument markers crosslinguistically.

An exploratory sample at my disposal supports my predictions and illustrates that what is currently considered a rare and unexpected grammaticalization of Comitative-to-Recipient marker in Spanish may be more common than previously assumed and shed light not only on DRM but on differential and non-differential argument structure more generally.

German psych verbs and the stage-level/individual-level distinction

Livia Sommer, Humboldt-Universität zu Berlin

The differences in argument structure of psych(ological) predicates among languages is an ongoing puzzle. The thematic role EXPERIENCER is either the grammatical subject (1a) or object (1b) of the psych expression, but without any obvious semantic reason.

- (1) a. Subject-experiencer (SE)

Den Großvater bedauert die Veränderungen.
the.NOM.SG grandfather.NOM.SG bemoan.3.PL the.ACC.SG changes.ACC.SG
EXPERIENCER THEME

The grandfather bemoans the changes.

- b. Object-experiencer (OE)

Die Veränderungen frustrieren den Großvater.
the.NOM.PL changes.NOM.PL frustrate.3.PL the.ACC.SG grandfather.ACC.SG
EXPERIENCER THEME

The grandfather is frustrated by the changes.

As proposed by Pylkkänen (2000), in Finnish stative psych predicates one of the two structural types conveys the semantic properties of Individual-level (IL) predications and the other structural type goes along with stage-level (SL) predications, a distinction which goes back to Carlson (1978) and Milsark (1974, 1977) and further developed by Kratzer (1995) and Chierchia (1995), among others. The most recent approach on these lines is provided by Fábregas & Marín (2015), who argue that Spanish stative subject-experiencer (SE) verbs correspond to IL predicates and Spanish object-experiencer (OE) verbs to SL predicates.

I argue, however, that the SL/IL-distinction appears to be reflected by another structural phenomenon in German psych predications: the explicit realization of the THEME argument as either definite or bare. Whereas bare plural THEME arguments like in (2a) assign a generic primary reading to the predication and provide it with an IL interpretation, a THEME realized by a definite article (2b) gives rise to an existential reading and a SL predication. This accounts for both SE (very few exceptions) and OE verbs.

- (2) a. *Den Großvater frustrieren Veränderungen.*
the.ACC.SG grandfather.NOM.SG frustrate.3.PL changes.NOM.PL

The grandfather is frightened by changes.

- b. *Den Großvater frustrieren die Veränderungen.*
the.ACC.SG grandfather.NOM.SG frustrate.3.PL the.NOM.PL changes.NOM.PL

The grandfather is frightened by the changes.

In a compositional semantics approach those effects could be accounted for as ‘Object effects’ (Husband 2010, 2012). The aspectual properties of the object are transmitted to the verb via event mapping (Kratzer 2004). By incorporating the interaction between verbal and nominal semantics into the analysis, we can shed new light on the discussion of how to modulate a syntactic representation of SE and OE verbs. This way, additional evidence can be provided to discuss the main structural analyses proposed by Belletti & Rizzi (1988) on the one hand and Pesetsky (1995) and Arad (1998) on the other.

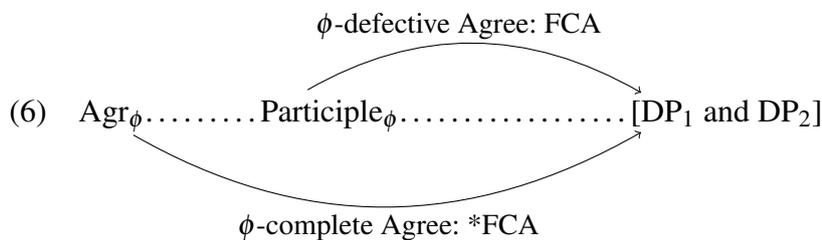
On first conjunct agreement in Spanish

Bo-Fan Syue, National Taiwan Normal University

This study examines first conjunct agreement (henceforth FCA) phenomena in Spanish. As shown in (1-5) below, Spanish allows FCA in existentials, transitives in which a participle agrees with the object and non-finites, but not in passives and reflexive passives (the plural expression ‘together’ is included to rule out an ellipsis analysis (see Camacho 2003 and Marušič, Nevins & Saksida 2007)

- (1) Hay **colgada** una lámpara y un candelabro juntos en el techo
have **hung.FEM** a.FEM lamp.FEM and a.MAS chandelier.MAS together in the ceiling
‘There is a lamp and a chandelier hung together in the ceiling’
- (2) Tenemos **colgada** una lámpara y un candelabro juntos ...
have.1PL ‘We have a lamp and a chandelier hung together...’
- (3) Al estar **colgada** una lámpara y un candelabro juntos ...
because to.be ‘Because a lamp and a chandelier are hung together...’
- (4) *Fue **colgada** una lámpara y un candelabro juntos ...
was.3SG ‘A lamp and a chandelier were hung together...’
- (5) *Se mantiene **colgada** una lámpara y un candelabro juntos ...
se keep.3SG ‘A lamp and a chandelier were hung together...’

In this study, I follow Soltan’s (2007) approach to FCA in appealing to acyclic adjunction, where FCA occurs when the AGREE operation takes place prior to the adjunction of DP₂, whereas if adjunction of DP₂ happens before the AGREE operation, no FCA is observed. Notice that this analysis only applies to post-verbal DPs. Through examining the data above, I propose that ϕ -complete Agree between the verb in the agreement projection and the first conjunct blocks FCA; however, ϕ -defective Agree between the participle and the first conjunct allows FCA. This is sketched in (6):



Besides, FCA is not possible if there is a ϕ -complete probe in the same phase. In addition to the proposal, I would like to point out a difference between A-movement and A'-movement with respect to the possibility for the DP₂ to adjoin. The DP₂ cannot adjoin to an A-moved phrase, but it can adjoin to an A'-moved phrase (cf. Lebeaux 1988, Bošković 2004).

Although FCA has been discussed in Spanish, with the exception of existential constructions (Camacho 2003, Demonte & Pérez-Jiménez 2012), no analysis had been proposed involving Chomsky’s Agree and phase theory.

Experimental insights on the Romanian ethical dative

Alina Tigau, Universität zu Köln

The literature on Romanian dative clitics differentiates between core (subcategorized) (CD) and non-core datives (NCDs), also known as high datives (HDs). Among the latter, one distinguishes between datives of (inalienable) possession, benefactive/adversative, affected/experiential and ethical datives.

This paper aims at providing new insights into the syntactic characteristics of Romanian ethical dative clitics, which we distinguish from other types of NCDs (in line with Joutiteau & Rezac 2008 a.o.), and at paving the way towards an appropriate analysis of these elements. The proposed account rests on newly obtained experimental data¹ uncovering the special status and behaviour of EDs.

Indeed, there exist a number of properties setting EDs aside from other NCDs: EDs have been labelled as ‘non-actantial’ datives, since they are not part of the valency of the verb but have an expressive function, grounding ‘the event structure in relation to the speech participants’ (Delbecque & Lamiroy 1996). As such, they do not affect the truth conditions of the sentence in which they occur, unlike the other NCDs, but merely invoke the addressee or the speaker as a witness or as a vaguely affected party.

Furthermore, EDs allow multiple clitic clusters. This is not possible with other instances of NCDs. Note, moreover, that the multiple clusters of EDs may only contain a 1st and a 2nd person pronoun. When a third person dative clitic pronoun appears in the cluster, it may be interpreted as an argumentative, a possessive, beneficiary or affected dative. Lastly, unlike other NCDs, EDs do not allow a full corresponding DP pronominal.

The properties uncovered above urge one to draw several conclusions: (a) co-occurrence with other NCDs points to the existence of distinct hosting projections (b) the fact that EDs anchors the Speaker/Addressee may prompt one to conclude that EDs merge within the CP, as it has been assumed for German (Abraham 1972). There is, however, strong evidence that EDs are actually merged lower, i.e. their occurrence in embedded clauses (infinitives included) and in questions. We are thus led to conclude that EDs merge within TP, in a special projection where from they may be bound by abstract Speaker/Addressee operators from within CP (Sigurðsson 2004) and thus be interpreted as denoting the Speaker/Addressee. Thus, in line with Baker (2008), we propose person to be a derivative notion, the result of operator-variable agreement. Furthermore, in line with Harbour (2006), Michelioudakis (2012) we claim that ED clitics are specified as [+participant, ±author] and that these specifications amount to the interpretable, lexically valued features probed for by the Speaker/Addressee operators in CP. Moreover, given their featural make up, EDs merge into an ApplP carrying an uninterpretable [+participant] feature, situated below T but above the HighApplP hosting other NCDs.

The analysis makes several correct predictions: the co-occurrence between EDs and other NCDs, the fact that EDs may never surface as fully fledged DPs, their anchoring the Speaker/Addressee to the event denoted by the verb etc.

¹The data brought into discussion represent the result of an experiment unfolded on 80 Romanian native speakers. The experiment was conducted within a formal design framework and verified the syntactic behaviour of ED clitics and other NCDs with respect to multiple clustering and co-occurrence with other datives (both full DPs and clitics).

Fronting in the T[ɜː]n: Evidence of /u/-fronting in North East England

Jasmine Warburton, Newcastle University

One current sound change in progress, the fronting of the back vowel /u/, has been found to effect the majority of varieties of English. This /u/-fronting phenomenon has been found in many British dialects, including RP (Harrington et al. 2008), Carlisle (Jansen 2012), Derby (Sóskuthy et al. 2015) and Manchester (Hughes et al. 2011). Previous research on varieties within North East England has indicated that while Teesside speakers show signs of a more centralised [ɜː] (Pearce 2009, Beal et al. 2012), those in Tyneside are ‘less prone’ to fronting (Watt & Allen 2003). The present study aims to be one of the first to investigate differences in /u/ realizations across the North East, asking whether speakers in the region can be distinguished based on the frontness of their /u/ vowels.

Using interviews obtained from the *Diachronic Electronic Corpus of Tyneside English* (DECTE), 21 speakers (11 Males, 10 Females) from across the region were selected for analysis. The study includes speakers from Teesside, County Durham, Wearside, Tyneside and Northumberland. These selected interviews were then automatically aligned using the *Forced Alignment and Vowel Extraction* suite (FAVE) in order to obtain the formant values for all tokens of /u/.

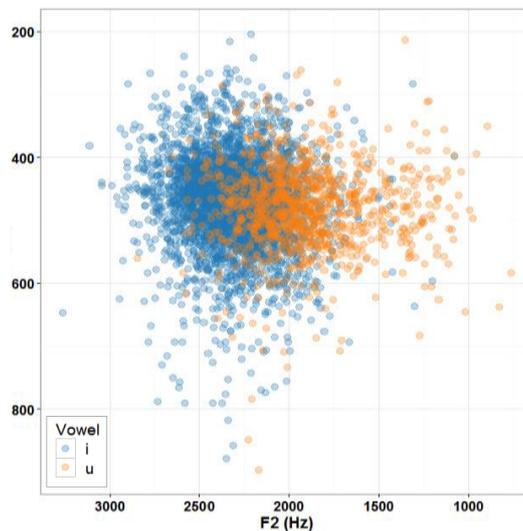


Figure 1: Formant values of /u/ and /i/ for all speakers (blue circles /i/, orange circles /u/)

Results from the present study indicate that, in contrast to previous claims, the whole of the North East is participating in the /u/-fronting sound change. While all speakers exhibited evidence of a fronted /u/, the results point to those in the southern parts of the region being more advanced in this sound change. Notably, when looking at the distance between each speakers’ mean F2 values for the /u/ and /i/ vowels, those from Teesside, Wearside and County Durham had a shorter distance between these vowels, suggesting greater fronting of /u/. Furthermore, the Teesside females showed a similarly fronted /u/ in both coronal and non-coronal contexts; this collapsing across phonetic environments is typically indicative of speakers who are more advanced in a sound change (Turton & Baranowski 2013). In terms of linguistic diffusion and the spread of sound changes, these results are interesting as the innovative fronted /u/ variants appear to be diffusing from the south of the region and not from Tyneside, the urban centre of the North East.

In the /i/ of the beholder: Looking at vowels with ultrasound tongue imaging

Fabienne Westerberg, University of Glasgow

Most phonetic vowel studies rely on acoustic data, since articulatory data has historically been difficult to collect and analyse. But acoustic analysis is not always able to answer questions about vowel articulation, as it is possible for speakers to create the same acoustic output in a number of different ways (Stevens 1972). This phenomenon is known as articulatory trade-off, and poses a particular problem with unusual vowels whose articulation might be unclear. As suitable technology is becoming increasingly available, there is a growing interest in articulatory vowel study, but there are also many challenges as new methods struggle to find their feet. In this paper, I discuss how articulatory methods can contribute to our understanding of vowel production, and address some of the issues faced when working with new technology and new types of data.

A practical example of a vowel whose articulation is disputed is the Swedish *Viby-i*. This is an /i/ vowel with an unusual ‘thick’, ‘buzzing’ quality, used by some speakers instead of the standard [i] vowel. It occurs sporadically in rural dialects across Central Sweden, but is also characteristic of urban speech in Stockholm and Gothenburg (Engstrand et al. 1998, Björsten & Engstrand 1999). In rural environments it is often stigmatised, but in urban speech it is associated with high status and “young, trendy speakers” (Bruce 2010). The articulation of *Viby-i* has not been established, as there seem to be several ways of achieving this vowel (Björsten & Engstrand 1999). To date, there is only one articulatory study of *Viby-i*, pointing towards variation between speakers from different locations (Schötz et al. 2014). However, it is also possible for articulatory variation to be linked to other factors, such as social class (Lawson et al. 2013), which might be relevant to *Viby-i* given its social function.

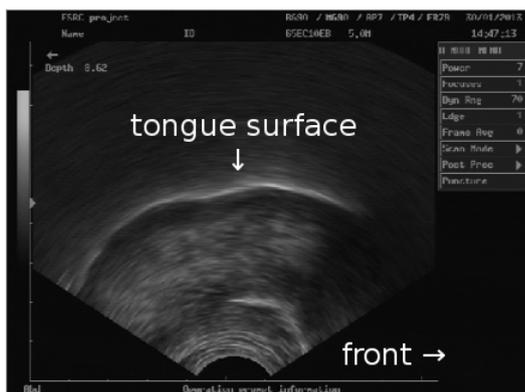


Figure 1: Ultrasound image of the tongue (speaker facing right)

However, it is also limited in that it can only display a two-dimensional image of the tongue surface, which does not include other articulators unless they make contact with the tongue. This makes it difficult to estimate the size and shape of the acoustic chamber that creates the vowel sound. UTI also raises several methodological problems, such as: How do we normalise between speakers whose vocal tracts are inevitably different? How should the tongue contours be quantified and compared? And how stable can we assume vowel production to be in general?

Understanding vowel production has important implications for several linguistic fields, including child speech development, speech therapy, sociolinguistics, and acoustic phonetics. For this purpose, articulatory methods such as UTI are able to provide invaluable insights which cannot be inferred from acoustic information alone. But limitations in the technology itself, and the complexity of the data, mean that there is still a long way to go until we have a full understanding of how vowel articulation works.

Posters

Surprises on the hunt for causes

Bernat Castro López, *Universitat Autònoma de Barcelona*

Bromberger (1992), analyzing “scientific” questions, shows that why-questions (contrary to some intuition) do not have to contain surprising propositions. However, he also acknowledges that this tends to happen, even if it seems probable that “scientific” questions (characterized by the absence of agentivity) are the least likely to involve such an attitude.

Therefore, it seems convenient to formulate as a research question in which cases questions-for-causes are separated from surprise and in which the opposite tends to happen: the specialization of different causal wh-elements co-existing with “why”, and restricting its use, needs to be considered. It seems necessary to take into account the concept of “mirativity” (DeLancey 2012), following recent studies that show the existence of markers of “unexpectedness” in European languages (Rigau 2012, Tirado 2016). This is compatible with a formal approach such as the one adopted here (for which judgments of acceptability and corpora are used).

It seems sound to postulate that different languages have causal wh-elements conveying mirativity (and presupposition, arguably (Fitzpatrick 2005)): “how come”, Spanish “cómo es que” (Colombian “cómo así”), Catalan “com és que” (Balearic “com així”), Italian “come mai”. All of them seem left-peripherally generated (Shlonsky & Soare 2011). May the formation of these structures provide some insight into the ways of expressing of mirativity? Some appear to incorporate a deictic of textual character (“así”, “així”); others appear to have their origin in the “és que” (“it is that”) copulative structure, which introduces presupposed explanations (versus efficient causes) that usually require some “attitudinal” inference on the part of the hearer (Declerck 1992). In any case, data shows that contrasts arise between these wh-elements and “why” (and equivalents), at least in “scientific” questions (1) and questions about states (2) (Spanish examples).

- (1) a. —¿Por qué llueve? {¿Porque el agua se condensa? / # No lo habían dicho.}
—Why does it rain? {Is it because water condenses? / It was not predicted.}
b. —¿Cómo es que llueve? {# ¿Porque el agua se condensa? / No lo habían dicho.}
—How come it rains? {*Idem*}
- (2) # —¿Por qué sabes estonio? / —¿Cómo es que sabes estonio?

Literally: why (you) know Estonian? / how is that (you) know Estonian?

In fact, in some languages there appear to be two co-existing wh-elements conveying mirativity, like “com és que” and “què ho fa que” in Catalan. The second one is interesting because (unexpectedly) it is only synonymous of the first one when the verb of the contained proposition is in indicative mood (3a). If it is in subjunctive mood (3b), there is a question for a non-agentive cause expressed in a nominal (subject-like) form, so as an argumental cause. “Què ho fa que”-subjunctive is appropriate for “scientific” questions, which are typically unconnected with mirativity and, crucially, also lack agentivity.

- (3) a. Què ho fa que l’aigua es congela? {És raríssim! / # La pèrdua de calor?}
what it_{accus} does that the’water itself freezes_{ind}? {is weird! / the loss of heat?}
b. Què ho fa que l’aigua es congeli? {# És raríssim! / La pèrdua de calor?}
what it_{accus} does that the’water itself freezes_{subj}? {*Idem*}

In order to improve the knowledge about the traditionally attested specificity of “why” it is important to understand the limits of its use, and, according to what is argued here, there is a factor, unexpectedness, which cross-linguistically constrains it.

Proficiency and working memory in L2 sentence processing: Evidence from eye-movement data

Zhe Gao*, Yanhui Zhang* & Meng Jiang†, *Chinese University of Hong Kong, †Sichuan International Studies University

For years, the issue of whether sentence processing in first language (L1) and second language (L2) is fundamentally different remains debating. Some propose (Cunnings 2016) that advanced L2 learners process syntactic complexities essentially similarly as native speakers, whereas others (Clahsen & Felser 2006) argue that L2 sentence processing is fundamentally lexical-driven. The current study aims to explore two critical factors contributing to the L2 processing difficulties, including working memory (WM) capacity and L2 proficiency. Extensive studies have demonstrated that WM capacity critically affects the processing of L1 syntactic dependency. The current study investigates the roles of WM in the processing of French relative clauses (RCs) by L2 learners with L1 Mandarin Chinese and its interaction with L2 proficiency.

The RC structures of French and Chinese are distinctively different. French is head-initial whereas Chinese is head-final. Unlike Chinese, French RCs contain morphological cues: *qui* marks subjective RCs; *que* marks objective RCs. Moreover, French objective RCs have a derivation caused by inversion where the RC subject moves behind the RC verb. Studies on French native speakers have shown the most WM demands for the processing of French objective RCs followed by the derivations of objective RCs, and then by subjective RCs. On the contrary, Chinese subjective RCs are more complex than objective RCs. We posit that L2 learners may overcome the bottleneck from Chinese when their French proficiency reach an advanced level.

The study inspects the processing of L2 French RCs by using the eye-tracking protocol. The subjects were 21 advanced L2 French learners and 29 intermediate learners. They had homogeneous background and French learning experience. Besides the experiment, the subjects took a memory test, a French proficiency test and a norming test. In the eye-tracking experiment, they read 24 series of French sentences with RCs and 52 filler sentences. The experiment tracked the first fixation duration, gaze duration, go-past time, first-pass regression ratio and skipping ratio of each critical word. In the data analysis, the processing of French RCs by the two learner groups was compared with those by French native speakers from other eye-tracking studies. The studies revealed that speakers had processing difficulty if the reading durations was the longest, or the regression ratio was the highest among the three clause types.

The study has three major findings: First, the two groups of learners can process the French RCs like native speakers in that the processing difficulty loci are similar for L2 learners and native speakers. Meanwhile, in the advanced group, there are significant interactions between the WM capacity and the gaze durations at the objective RC verbs. The advanced learners may be more sensitive to the RC structure that requires high WM demands. Second, L2 proficiency level influences the speed of L2 processing. Advanced learners have shorter reading times and lower probability of regression than the intermediate learners. Third, different from native speakers, L2 learners are more likely to have regressive saccades. The learners meet the most processing difficulties in the derivations of objective RCs even if the short form of the complementizer and the auxiliary verb (i.e. “qu’a/qu’ont”) indicates a strong lexical hint of the RC structure.

Findings of the study demonstrate supporting evidence for Cunnings’s theory that advanced L2 process sentences similarly as native speakers but they are more sensitive to the interferences within RCs. Lexical cues only do not facilitate the processing of complex syntactic structures. With the improvement of the L2 proficiency, learners can process French RCs similarly as native speakers.

A proposal for a shared classification of code-switching and code-blending

Eleonora Marocchini, Istituto Universitario di Studi Superiori di Pavia

Code-blending is defined as a phenomenon occurring in bimodal bilingual communication where signs in a sign language and words in a vocal language are simultaneously produced (Emmorrey et al. 2005). It seems like bimodal bilinguals tend to prefer code-blending to code-switching (simple alternation) between their two languages, as, in several studies on the phenomenon, 94-95% of the language mixing cases were simultaneous (Petitto et al. 2001, Emmorrey et al. 2005). Subjects of the studies taken into account were all bimodal bilingual children who could be defined or defined themselves as “CODAs” (i.e. Children of Deaf Adults), who acquired a spoken language and a signed language almost at the same time, like unimodal bilingual children acquire their spoken languages.

Since code-switching study has a much longer history, more has been done in order to define terms for a unified classification of its types. After reviewing works by several authors (Alfonzetti, Berruto, Myers-Scotton, Poplack, Romaine and Sankoff) the three main types of code-switching identified by Romaine (1989), who substantially follows Poplack (1980) can be assumed as a summary taxonomy, based on the extent of the switching. This taxonomy consists of: “extrasentential switching” (also called “tag-switching” elsewhere), involving interjections, fillers and discursive markers; “intrasentential switching” (or “code-mixing”), involving words within a sentence or a clause; “intersentential switching” (or “code-switching”) occurring outside the clause level.

Code-blending taxonomies are more various and difficult to unify, as every researcher focuses on different aspects of blending and therefore divides the types on the basis of how independent the strings in the two languages are (Branchini 2012), what order the words occurring in each of the two languages follow (Donati et al. 2009) or the number of signs involved (identifying the type of blending as either “single-sign” or “multi-sign code-blends”, as in Emmorrey et al. 2008).

This paper is to be intended in the spirit of proposal for a taxonomy of code-blending that could highlight the important similarities between the two phenomena, using a shared vocabulary for the classification of the types, and underline the extent of the blending, provided that, in the case of code-blending, what is “extrasentential”, “intrasentential” and “intersentential” is not the border between the two languages, but the boundaries of their simultaneity. We can therefore define:

- (1) “extrasentential blending” what Branchini (2012) calls “non-independent dominant blending”, as signs are simultaneously produced without adding any significant meaning to the sentence in the vocal language (this type of blending is usually composed of what Emmorrey calls “single signs”);
- (2) “intrasentential blending” what Branchini defines as “(non-independent) blended blending”, where the simultaneous production would not be understandable by monolinguals in any of the two languages (just like intrasentential switching), because the constituents are divided between the two codes;
- (3) “intersentential blending” what Branchini labels “independent blending” as the two clauses or sentences are independent from one another and understandable for monolinguals too (just like intersentential switching, although in the case of the blending it is possible to actually say the whole sentence or period in both languages simultaneously, whilst intersentential switching only allows monolinguals to understand the part of the sentence which is said in their language); this third type can be divided into three subtypes according to the classification suggested by Donati et al. (2009), depending on the order followed by the constituents, that can be: (a) that of the sign language for both sentences, (b) that of the vocal language for both sentences, (c) that of the language in which the two sentences are produced, for each of them.

Camfranglais as a transnational practice – being Cameroonian in a ‘superdiverse’ context

Constance Mbassi Manga, Lancaster University

Urban hybrid linguistic registers have been extensively studied and characterised as “youth languages” over the last 30 years (Boyer 1997, Goudaillier 1998, Billiez et al. 2003, Rampton 1995, 1999, Harris 2006 ...). Camfranglais, hereafter CFrA, is no exception. A mixture of Cameroonian and Western national languages, it emerged among the urban youth in the 1980s in Yaoundé, the capital of Cameroon, a West/Central African country with 2 official languages, French and English, and over 250 vernacular languages and ethnic groups. CFrA is often characterised as a non-systematic register that is difficult to identify, or a low variety of Cameroonian French (De Féral 2006:202, 212, Njoh Kome 2009:6), and traditional approaches to studying non-standard registers have failed to circumscribe it. Moreover, over the last thirty years, it has spread beyond the youth and beyond Cameroon; it is not the “transient age-specific phenomenon” (Rampton 2011:278, about ‘Southallian’) it was thought to be, and it is now spoken by adult diasporic Cameroonians.

My ethnographic PhD study elicits an understanding of CFrA from adult speakers living in Paris, London and 2 cities in the US, and exchanges from a Facebook-based blog. This data raises questions about their relationship with CFrA, particularly with reference to its emblematic status among speakers’ linguistic repertoires. Initial findings suggest that most respondents present their use of their vernacular and French/English as beyond their control and/or as a social or professional obligation. In contrast, they present CFrA as a language they chose to use and they make quite strong claims linking the register with their sense of belonging and identity.

What lies behind this erasure of agency when discussing their use of former colonial languages and vernaculars? How far does CFrA qualify as a symbol of users’ affiliation to an alleged transnational Cameroonian community or identity, enacted through the use of CFrA in these Western contexts? This paper will discuss these issues.

Intuitive and reflective inference in the interpretation of metaphor

Ramona Pistol, Middlesex University

This paper argues for an adjustment to Carston's (2010) view of metaphorical interpretation as involving two distinct processing routes. Theoretical accounts of metaphor are often divided into those which see it as arising in language (e.g. Sperber & Wilson 2008), those which see it as a matter of cognition (e.g. Lakoff & Johnson 1999), and those which see it as involving a combination of these (e.g. Steen 2011). Additionally, there is the idea that novel metaphorical expressions require different comprehension processes from the ones activated in more conventionalised cases (e.g. Gentner & Bowdle 2001). To address a perceived shortcoming in the relevance-theoretic account of metaphors relating to the distinction between novel, particularly highly creative and extended metaphors, and more conventionalised metaphors, Carston (2010) suggests two processing routes. The first is a local and rapid, intuitive, process, involving concept adjustment processes, for understanding simple metaphors such as "John is a shark". The latter involves a slower, more reflective, global appraisal of the literal meaning of the entire metaphorical expression, and applies to extended and poetic metaphors, such as the one from Carl Sandburg's poem *The Fog*: "The fog comes/ On little cat feet. / It sits looking/ over harbour and city". This paper argues for a refinement of Carston's model in which the local, intuitive, concept adjustment processing route and the more reflective route are not seen as wholly distinct. Rather, local adjustment processes are seen as sometimes embedded within more reflective inferential processes within which interpretations involving local adjustment are metarepresented. The paper develops understanding of this reflective processing route from the perspective of Mercier and Sperber's (2009, 2011) massively modular view of the human mind, which sees reflective inference as metarepresentational. The argument is developed with reference to examples from naturally occurring everyday discourse and from literary texts.

Constructional change in the lifetime: it-cleft foci as a case study for change in individuals

William Standing, Universiteit Antwerpen

While grammatical changes have primarily been examined for language as an abstract system, they are essentially the result of changes in the mental grammars of actual language users. In constructionist approaches to change, grammatical items are modelled as schematic nodes in a hierarchical and dynamic associative network, which implies *inter alia* that constructions and the links between them are subject to change through analogical attraction, both in terms of formal and semantic make-up, but also in terms of frequency (Hilpert 2014). While such changes have been attested at the community level, it is still unclear to what extent similar processes are at work in language change in individuals. We tackle this key problem through analysis of individual-level change in the case of it-clefts, and the ramifications for related constructions, information structure and communicative function.

Data come from a sample of the EMMA-corpus (Early-Modern Multiloquent Authors), a new large-scale longitudinal corpus comprising the writings of 50 individuals from the 17th-century London-based elite, all of whom were prolific over a prolonged period of time during their adult lives. Features of the it-cleft construction undergo substantial change during the period in question with the rise of new possible foci in clefts, such as the time adverbial in (1), as well as new performative uses of the construction which exhibit backgrounding of subject matter and opinion-as-fact framing functions (demonstrated in (2)).

- (1) “It is not yet that the general Rule fails, because of this Exception. . .”
(L’Estrange, 1680, *The Answer to the appeal expounded*)
- (2) Aman. Pray be so just then to me, to believe, ’tis with a World of Innocency I wo’d enquire, Whether you think those Women we call Women of Reputation, do really ’scape all other Men
(1696 VANBR-E3-P1, 43.108 from Patten, 2012)

The growth of permissible focus types in the it-cleft seems to proceed systematically with adverbials of means and reason attested in the construction prior to those of time and place (Patten 2012). I present corpus evidence that tentatively suggests that the same development is also manifested at the level of the adult individual. Loss of various functions of word order variation during the rigidification of the SV(O) word order may have contributed to the growth of acceptable types of adverbial foci as other strategies, such as inversion, became less viable. It is explored how this shift, which takes place in the syntactic system at large, interacts with more functional motivations specific to clefts. Information structural licensing strategies, such as contrastive focus, are increasingly used to frame marked foci types and increase their rates of acceptability. A possible motivation underlying such extensions is the desire for higher expressivity (e.g. Haspelmath 1999). The interplay of factors allows for new foci to become a productive element of the construction. The analysis of the interaction between clefts, information structure and the grammar at large will add to our understanding of inter-construction (paradigmatic) relations in a constructionist perspective.