

THE DEPARTMENTS OF INFORMATICS &
PPLS PRESENT

THE 74th LANGUAGE LUNCH @ EDINBURGH

6-8 PM (UTC) | 8/3/2022
LIVE ON ZOOM



For more details visit:
<https://blogs.ed.ac.uk/language/>

SCHEDULE

6-8 PM (UTC) | 8/3/2022
LIVE ON ZOOM



1800 | SESSION #1

1810 | CHIHIRO TAGUCHI

Mermaid Construction as Raising with a Nominal Predicate

1825 | YAO FU

Scaling Structured Inference with Randomization

1840 | JENNY STEELE

The negotiation of identity in English for Academic Purposes:
Understanding students' experiences in a Scottish university context

1855 | MEET THE SPEAKERS #1

1905 | SESSION #2

1910 | ZHAOXI YAN

The style-shifting of /s/ Among Men Beauty vloggers on YouTube

1925 | JAVIER DE TORRES

Grammatical Aspect in Spanish: A summary of the state of the art

1940 | SHANGMIN GUO

Expressivity of Emergent Languages is a Trade-off between
Contextual Complexity and Unpredictability

1955 | MEET THE SPEAKERS #2

MERMAID CONSTRUCTION AS RAISING WITH A NOMINAL PREDICATE

Mermaid construction (MMC) refers to a construction that is typically a combination of a verb and a noun, of which the noun is grammaticalized to carry an abstract semantic function such as modality, aspect, etc. Sentence (1) is an example of the Japanese MMC. Tsunoda (2020) argues that the underlined parts in (1) constitute a compound predicate, and therefore the MMC is monoclausal. Although Tsunoda (2020) provides a cross-linguistic descriptive study of MMCs, it lacks elaboration on the theoretical analysis of MMC. In this paper, I claim that nouns triggering MMC should be analysed as raising nominal predicates.

- (1) tomodachi=ga nihon=ni ik-u yotei=da.
friend=NOM Japan=DAT go-NPST plan=COP
Literal: “(My) friend is a plan that goes to Japan.”;
Free translation: “(My) friend plans to go to Japan.”

Raising is a functional control where the subject argument of the subordinate clause is raised to the subject position of the matrix clause. I argue that the noun ‘yotei’ in (1) is a predicate that triggers raising. The evidence for this claim is given in the comparison of (2) and (3). Japanese allows for relativization of an oblique NP, so (2) is taken to be a grammatical sentence. In the MMC in (3), however, a nominative NP cannot fill the slot retaining the same meaning, as the slot is already filled by the raised NP.

- (2) [kore=ga [haha=ga sakana=wo yak-u]CP nioi=da]IP.
this=NOM mother=NOM fish=ACC grill-NPST smell=COP
“This is the smell which (comes from where) my mother grills fish.”
(3) *[kore=ga tomodachi=ga [t nihon=ni ik-u]VP yotei=da]IP.
this=NOM friend=NOM Japan=DAT go-NPST plan=COP

CHIHIRO TAGUCHI | PPLS



CHIHIRO'S INTERESTS INCLUDE SYNTAX, TYPOLOGY, AND COMPUTATIONAL LINGUISTICS. CHIHIRO IS ALSO A BIG FAN OF LANGUAGE LEARNING AND HAS RECENTLY BEEN LEARNING KOREAN THROUGH NETFLIX! HE IS CURRENTLY PURSUING AN MSCR IN LINGUISTICS AT EDINBURGH, AND HOLDS AN MA ENGINEERING FROM THE NARA INSTITUTE OF SCIENCE AND TECHNOLOGY.

SCALING STRUCTURED INFERENCE WITH RANDOMIZATION

Discrete structured models for language have enjoyed great progress recently. However, existing dynamic programming (DP)-based inference typically works with a small number of states (usually fewer than hundreds), which severely limits the model capacity. In this work, we propose a family of randomized dynamic programming (RDP) algorithms for scaling structured models to tens of thousands of latent states. Our method is widely applicable to classical DP-based inference (partition, marginal, reparameterization, entropy) and different graph structures (chains, trees, and more general hypergraphs). It is also compatible with automatic differentiation so can be integrated with neural networks seamlessly and learned with gradient-based optimizers. Our core technique is randomization, which is to approximate the sum-product by restricting and reweighting DP on a small subset of nodes, leading to orders of magnitudes computation reduction. We further achieve low bias and variance with Rao-Blackwellization and importance sampling. Experiments on different inferences over different graphs demonstrate the accuracy and efficiency of our methods. Furthermore, when using RDP to train a scaled structured VAE, it achieves better test likelihood and successfully prevents posterior collapse.

Keywords: Structured Prediction, Dynamic Programming, Approximate Inference, Randomization.



YAO FU | INFORMATICS

YAO FU IS AN INFORMATICS PHD STUDENT AT THE UNIVERSITY OF EDINBURGH WORKING WITH PROFESSOR MIRELLA LAPATA. PREVIOUSLY, HE FINISHED HIS BS AT PEKING UNIVERSITY AND MS AT COLUMBIA UNIVERSITY. HE WAS ALSO A VISITING RESEARCH STUDENT AT CORNELL UNIVERSITY BEFORE EDINBURGH. HIS RESEARCH TOPIC IS DEEP GENERATIVE MODELS FOR STRUCTURED PREDICTION AND LANGUAGE GENERATION.

THE NEGOTIATION OF IDENTITY IN ENGLISH FOR ACADEMIC PURPOSES: UNDERSTANDING STUDENTS' EXPERIENCES IN A SCOTTISH UNIVERSITY CONTEXT

Increasing globalisation and student mobility have led to more diversity within Scottish universities in recent years. Since not all applicants from overseas meet the requisite language proficiency scores, most Higher Education institutions run pre-sessional programmes in English for Academic Purposes (EAP). These aim to strengthen learners' knowledge of the academic discourse needed for disciplinary study. Although EAP students are a crucial source of income for universities, and enhance their global standing, they have been overlooked in the literature. This talk will explain the gap in current research and outline the methodology of a planned case study into their identities and language practices.

My project is situated within a framework which views language as social practice and identity as a potential 'site of struggle' (De Costa & Norton, 2016, p.95). Drawing on positioning theory (Davies & Harré, 1990) and Bourdieu's (1977) notion of cultural capital, I will investigate participants' orientation to the new academic discourse community. Non-native users of English may experience negative labelling (e.g., Maringe & Jenkins, 2015), with detrimental consequences for their confidence and sense of belonging. Instead of assuming that EAP students' struggles are the result of individual failings, social context and power relations need to be considered.

I will employ various instruments to collect data, including a survey of pre-sessional EAP students, semi-structured interviews (with students and instructors) and classroom observations. As I am not beginning with any hypotheses, data will be analysed inductively, based on emerging themes. Unlike previous studies conducted in the UK, I will track focal participants (target n=8) from a pre-sessional programme on to their degrees to ascertain how identities change in different contexts. My case study will go beyond a superficial level to gain insight into the real experiences of participants, with key implications for university policies on internationalisation and inclusion.

JENNY STEELE | PPLS



JENNY GRADUATED WITH A DEGREE IN ENGLISH FROM ST ANDREWS, AFTER WHICH SHE SPENT A YEAR IN JAPAN WITH THE JET PROGRAMME, COMPLETED A PGCE, & TAUGHT IN SECONDARY SCHOOLS.

SHE THEN PURSUED AN MSc IN TESOL AT THE UNIVERSITY OF EDINBURGH, FOLLOWED BY EAP TUTORING POSITIONS AT HERIOT-WATT & STRATHCLYDE. SHE IS CURRENTLY A 2ND YEAR PHD CANDIDATE EXPLORING THE IDENTITIES & LANGUAGE PRACTICES OF INTERNATIONAL STUDENTS IN SCOTLAND.

THE STYLE-SHIFTING OF /s/ AMONG MEN BEAUTY VLOGGERS ON YOUTUBE

This paper studies the interaction between sociophonetic variant /s/ in American English and the visual transformation of men beauty vloggers on YouTube. Sibilant fricative /s/ in English is a well-studied index of gender identity. It is found that the frontness of /s/ can index the level of femininity (e.g., Schwartz, 1968; Campbell-Kibler, 2011; Zimman, 2017). Men beauty vloggers are men who apply makeup to themselves in their vlogs in order to teach makeup skills or to recommend beauty products. In their makeup tutorial videos, they usually transform from relatively masculine looks to rather feminine looks. Since /s/ and visual transformation are both embodied variables and are often integrated into the process of self-representation, it is expected that the visual transformation would correlate with changes in their productions of /s/. This paper selects 15 sample videos of 3 vloggers and collects the Center of Gravity of their /s/. Having observed the change of Center of Gravity of /s/ through the sample videos, both intra-speaker style shifts and inter-speaker differences are found. The /s/ of the three sample vloggers have shown three different changing patterns: increasing, decreasing and stable. The results suggests that the sociolinguistic variable /s/ does interact with visual signs in the process of character building. Besides gender identity, /s/ is also related to contextualized personae among the community of men beauty vloggers.

Keywords: Sociophonetics; Language Variation; Gender; Persona.



ZHAOXI YAN | PPLS

ZHAOXI YAN IS A SECOND-YEAR PHD STUDENT AT THE UNIVERSITY OF EDINBURGH. HER RESEARCH INTERESTS LIE IN SOCIOLINGUISTICS; SOCIOPHONETICS; LANGUAGE VARIATION AND CHANGE. THE PAPER TO BE PRESENTED WAS HER MASTER'S THESIS WHICH SHE PROPOSED TO EXPAND AS HER PHD PROJECT.

GRAMMATICAL ASPECT IN SPANISH: A SUMMARY OF THE STATE OF THE ART

When studying Spanish verbs, people often focus on its tense. However, it has a lesser known trait which is just as relevant: aspect. Aspect is a non-deictic category which encodes the way in which the action denoted by the verb takes place. There are two kinds of it: lexical and grammatical. The latter, which is the one I'll be focusing on, is contained in each tense and, on the most basic level, it answers the question: "Did the speaker finish the action or not?". Two tenses may have the same time, but different aspect. Preterite and Imperfect tenses both occur in the past, but, in one case, the event is over, and in the other, it is not. For example:

- (1) a. Ayer escribí un cuento.
Yesterday write-PST.PFV.ISG a short story
'Yesterday you wrote up a short story'
- b. Ayer escribía un cuento.
Yesterday write-PST.IPFV.ISG a short story
'Yesterday you were writing a short story'

Any Spanish speaker can confirm that, in the first sentence, the speaker finished the short story, whereas, in the second one, we don't know if he did. This is not a pragmatic inference of any kind: this information is part of the verb.

In this short talk, I'll briefly describe how grammatical aspect works in Spanish. First, I'll do a basic description of it like the one above, and then I'll provide a few examples of how it interacts with syntactic constituents and lexical aspect, so the audience sees how important this category is in the Spanish language.

Keywords: Spanish; Grammatical Aspect; Perfective; Imperfective

JAVIER DE TORRES | PPLS



JAVIER WANTED TO BECOME A WRITER, BUT INSTEAD DECIDED TO PURSUE A DEGREE IN SPANISH PHILOLOGY AT THE AUTONOMOUS UNIVERSITY OF MADRID, WHERE HE STUDIED LANGUAGE & LINGUISTICS.

ONE DAY, A FRIEND TOOK HIM TO A TALK BY ANTONIO MORENO SANDOVAL, A RENOWNED SPANISH EXPERT IN COMPUTATIONAL LINGUISTICS. LONG STORY SHORT, HE IS NOW READING FOR AN MSC IN SPEECH AND LANGUAGE PROCESSING AT THE UNIVERSITY OF EDINBURGH.

EXPRESSIVITY OF EMERGENT LANGUAGES IS A TRADE-OFF BETWEEN CONTEXTUAL COMPLEXITY AND UNPREDICTABILITY

Shangmin Guo, Kenny Smith, Simon Kirby, Stefano V. Albrecht - University of Edinburgh; Yi Ren - University of British Columbia; Kory Mathewson - DeepMind

Researchers are using deep learning models to explore the emergence of language in various language games, where simulated agents interact and develop an emergent language to solve a task. We focus on the factors which determine the expressivity of emergent languages, which reflects the amount of information about input spaces those languages are capable of encoding. We measure the expressivity of emergent languages based on their generalisation performance across different games, and demonstrate that the expressivity of emergent languages is a trade-off between the complexity and unpredictability of the context those languages are used in. Another novel contribution of this work is the discovery of message type collapse. We also show that using the contrastive loss proposed by Chen et al. (2020) can alleviate this problem, compared with the standard referential loss used by the existing works.

Keywords: Emergent Language; Expressivity.



SHANGMIN GUO | INFORMATICS

SHANGMIN IS A SECOND-YEAR PHD STUDENT IN THE CDT-NLP PROGRAMME. HIS PHD RESEARCH FOCUSES ON HOW GENERALISATION OF MACHINE LEARNING SYSTEMS CAN BE IMPROVED BY REPEATED LEARNING, INTERACTING AND TRANSMITTING, WHICH ARE REFERRED TO AS DEEP ITERATED LEARNING, AND HOW MULTIPLE AGENTS ARE CAPABLE OF COMPLETING COMPLEX TASKS THROUGH THIS LEARNING MECHANISM.

ORGANIZING COMMITTEE

ANNA LAOIDE-KEMP | PPLS
GEORGIA CARTER | INFORMATICS
SYDELLE DE SOUZA | PPLS
NIAMH HARRIS | PPLS
JOYEE LOHANI | PPLS
ABIGAIL LA LIBERTE | PPLS

ACKNOWLEDGEMENTS

THIS EVENT WAS MADE POSSIBLE BY THE GENEROSITY OF THE CDT IN NLP AND
THE SCHOOLS OF PPLS AND INFORMATICS



THE UNIVERSITY of EDINBURGH
UKRI Centre for Doctoral Training
in Natural Language Processing



THE UNIVERSITY of EDINBURGH
informatics



THE UNIVERSITY of EDINBURGH
School of Philosophy, Psychology
& Language Sciences