Poster presentations, Language Sciences Symposium 2015

Capturing semantic constraints in implicit language learning

Dimitris Alikaniotis, da352@cam.ac.uk
John Williams, jnw12@cam.ac.uk
Dept. of Theoretical and Applied Linguistics

In distributional semantics, words acquire meaning by exploiting the statistical information that is inherent in their linguistic environment. A common criticism towards such semantic representations is that they lack a rich conceptual structure, questioning the cognitive relevance of such statistical mechanisms during language learning.

Here we show that distributional semantic models provide a good fit to data obtained from implicit language learning experiments on adults (e.g. Leung & Williams, 2014). In these experiments participants are introduced to novel non-words, which co-occur with already known words conditioned on underlying semantic regularities, such as concrete/abstract, animate/inanimate. Participants can implicitly learn such underlying semantic regularities, although whether they do so depends upon the nature of the conceptual distinction involved and their first language.

In the present study, we trained two vector-space models based on the distributional semantics of an English and a Chinese corpus. We used the models' resulting semantic representations as input to a feed-forward neural network, which predicted the novel non-words, discovering the relevant elements of the input representation. Using datasets provided from four behavioural experiments, which used different semantic manipulations, we obtained generalisation gradients that matched closely those of humans, capturing the effects of various conceptual distinctions and crosslinguistic differences.

Crowdsourcing a multilingual speech corpus: recording, transcription and annotation of the CrowdIS corpora

Andrew Caines, apc38@cam.ac.uk
Christian Bentz, cb696@cam.ac.uk
Calbert Graham, crg29@cam.ac.uk
Paula Buttery, pjb48@cam.ac.uk

ALTA Institute and Dept. of Theoretical and Applied Linguistics

We present the 'CrowdIS' corpora – CrowdISeng and CrowdISengdeu – being collected from English native speaker ('eng') and German/English bilinguals ('engdeu') via crowdsourcing platforms (hence 'Crowd') and funded by an INTERSPEECH special session (hence 'IS'). Efforts to collect the corpora are ongoing, and so we describe the collection methodology, our objectives for the corpora, and explain how to stay informed of developments. The corpora will be made freely available to other researchers.

It is well-known that building speech corpora is a time-consuming and expensive process: one estimate puts the cost of transcription at €1 per word, before the cost of any extra annotation (Ballier & Martin 2013). We present a method of collecting speech corpora via crowdsourcing facilities, showing that we can reduce costs considerably by distributing the work among multiple online workers. To collect recordings, we used the Crowdee application

designed for Android operating systems (www.crowdee.de), and for transcription and annotation we uploaded those recordings to CrowdFlower (www.crowdflower.com).

Our primary motivation in designing this project was to obtain a benchmark corpus of English native speakers undertaking tasks similar to those typically contained in learner corpora; and in our case relating to certificates of business English. Hence, a majority (65%) of Crowdee funding was allocated to the recordings needed for CrowdISeng, enabling a maximum of 130 individuals to make contributions.

Participants are required to be resident in the United Kingdom, United States or Canada, and it is a stated requirement of the task that English should be their mother tongue. They are presented with various business-related scenarios (e.g. starting a business, hosting visitors, sports sponsorship), and posed five questions (or 'prompts') about each scenario. In total, the jobs feature twenty prompts and participants are expected to produce approximately 300 seconds (5mins) of speech.

The German/English task designed for the bilingual corpus (CrowdISengdeu) is the same in design as for CrowdISeng, except that participants have to be resident in Germany and need to define themselves as bilinguals with either language as mother tongue. They answer prompts about the same two scenarios in both English and German, and are expected to provide 150 seconds of English and 150 seconds of German. Funds currently allow for a maximum of 50 contributors.

At the time of writing the corpora are 30% complete. Participation and assistance with publicity is very much welcomed, whilst researchers interested in eventually 1 obtaining the corpora may bookmark our reserved Speech and Language Data Repository URLs2.

1 Further information at: http://apc38.user.srcf.net/outreach/#crowd

2 http://sldr.org/ortolang-000913

Building passive constructions with features

Ruyi Dai, rd486@cam.ac.uk
Faculty of Asian and Middle Eastern Studies

My poster presentation will explore the syntactic and semantic properties of Chinese and English passives, which constitute the theoretical linguistic background of my PhD project, investigating the second language acquisition of Chinese *bei* passive constructions by adult English learners. The passive construction is commonly used worldwide, yet languages vary in their ways of forming passives, using the available choices given by Universal Grammar (e.g. Chomsky, 1982). In English, passive constructions are typically formed with the combination of an auxiliary verb *be* and a passive participle with the passive suffix *–en* (or *–ed*). Consider (1a) and its passive form (1b), where *the teacher* is the agent of the action *criticizing* and comes after the verb and the preposition *by*.

(1) a. The teacher criticized Xiaoming.
b. Xiaoming was criticized by the teacher.

By contrast, with less rich inflectional morphology than English, Chinese employs an individual passive marker *bei* to mark passive sentences. For example, (2b) is the passive form of (2a):

(2) a. laoshi piping-le Xiaoming.
teacher criticize-Perf Xiaoming

'The teacher criticized Xiaoming.'
b. Xiaoming bei laoshi piping-le.
Xiaoming BEI teacher criticize-Perf
'Xiaoming was criticized by the teacher.'

The word order of (2b) is different from (1b), because the agent *laoshi* (teacher) occurs before the verb. Different word order and formation strategies are the most obvious differences between English and Chinese passives, but definitely not the only ones.

For instance, features are properties of words, which are considered to be the basic building blocks of syntax (Adger, 2003). Given that features are at the heart of Chomsky's (1995) Minimalism, 'any study of language acquisition done within this framework is now a study of the acquisition of features' (Travis, 2008:23). From this view, parametric linguistic variation is determined by differences in feature inventories, distribution and combination.

I provide a feature-based cross-linguistic analysis of passive constructions. Regarding the syntactic analyses, I adopt Collins' smuggling approach (2005) to English *be* passives, as well as the raising approach to Chinese *bei* direct passives and indirect passives based on Liu (2012) and Huang (1999).

For the semantics, I propose a feature-based account and argue that [telic], [dynamic], [adversative], [indefinite] and [partitive] features differentiate Chinese *bei* passives from English *be* passives. The [telic], [dynamic] and [adversative] features constitute the L2 target feature bundle of Chinese *bei* direct passives. The feature bundle of *bei* indirect passives has additional [indefinite] and [partitive] features.

References

Adger, D. (2003) Core syntax: A minimalist approach (Vol. 33). *Oxford: Oxford University Press.*
Chomsky, N. (1982) Lectures on government and binding. *Dordrecht: Foris.*
Chomsky, N. (1995) The Minimalist Program. *MIT Press.*
Collins, C. (2005) A smuggling approach to the passive in English. *Syntax 8, 81-120.*
Huang, C.-T.J. (1999) Chinese passives in comparative perspective. *Tsinghua Journal of Chinese studies 29(4), 423-509.*
Liu, N. (2012) On the syntax, semantics and typology of non-canonical passives. *Ph. D. dissertation, Beijing Language & Culture University.*
Travis, L. (2008) The role of features in syntactic theory and language variation. *In Liceras, J., Zobl, H. & Goodluck, H. (eds.) The role of formal features in second language acquisition, 22-47.*

Distributional semantic functions

Guy Emerson, gete2@cam.ac.uk
Computer Laboratory

In this poster, I will introduce a novel semantic framework that both draws on recent advances in machine learning and maintains strong links with formal linguistic theory.

Distributional semantics aims to learn semantic representations from large corpora (Harris, 1954; Firth, 1951), and the state of the art is to represent meanings as points in a vector space, whether as embedding vectors (Mikolov et al., 2013) or as count vectors (Turney and Pantel, 2010). While the idea of using a geometric space to model meaning is well established (Osgood, 1952), there is a range of phenomena studied in formal semantics which are not naturally captured by modelling lexical meaning as single points – such as vagueness,

polysemy, hyponymy, compositionality, entailment, and context dependence. Although there have been efforts to resolve these issues, none of them can be considered solved.

Instead, I propose modelling the meaning of a predicate not as a point in a semantic space, but as a function on the space, mapping to values in the range [0; 1]. Each point in the space represents a possible entity, and the function's value at that point represents the probability that a competent speaker would judge the predicate to be applicable to such an entity. For example, in the one-dimensional space of hues, the function for the predicate red would take values near 1 for reddish hues, dropping lower towards orange or purple, then dropping close to 0 towards yellow and blue. Such a probabilisticmodel can naturally model vagueness (Sutton, 2013), polysemy and hyponymy.

I will explain how these semantic functions can be incorporated with Dependency Minimal Recursion Semantics (Copestake, 2009), by using a probabilistic graphical model. One set of nodes represents a set of entities, and another set of nodes represents the applicability of predicates on those entities. The entity nodes are connected according to the DMRS links. In this probabilistic framework, we can perform Bayesian inference, giving us a natural model of context dependence, and allowing evaluation by means of a textual entailment task.

I will also show how this probabilistic model can be implemented using a Restricted Boltzmann Machine, a powerful and flexible tool that has been successful in various applications (Hinton, 2010). The functions can be learnt using a corpus of DMRSs, such as WikiWoods (Flickinger et al., 2010), and an unsupervised learning algorithm that uses Gibbs sampling to deal with the latent entities in each DMRS.

References

Copestake, A. (2009). Slacker semantics: Why superficiality, dependency and avoidance of commitment can be the right way to go. In Proceedings of 12th Conference of the European Chapter of the Association for Computational Linguistics.

Firth, J. R. (1951). Modes of meaning. Essays and Studies (The English Association) 4, 118–149.

Flickinger, D., S. Oepen, and G. Ytrestøl (2010). WikiWoods: Syntacto-semantic annotation for English Wikipedia. In Proceedings of the 7th International Conference on Language Resources and Evaluation.

Harris, Z. S. (1954). Distributional structure. Word 10, 146–162.

Hinton, G. (2010). A practical guide to training restricted boltzmann machines. Technical Report 003, University of Toronto.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013). Efficient estimation of word representations in vector space. In Proceedings of the 1st International Conference on Learning Representations.

Osgood, C. E. (1952). The nature and measurement of meaning. Psychological Bulletin 49(3), 197–237.

Sutton, P. (2013). Vagueness, Communication, and Semantic Information. PhD thesis, King's College London.

Turney, P. D. and P. Pantel (2010). From frequency to meaning: Vector space models of semantics. Journal of Artificial Intelligence Research

---

Automated assessment of non-native speech using spectral features of vowels

Calbert Graham, crg29@cam.ac.uk
Paula Buttery, pjb48@cam.ac.uk
Andrew Caines, apc38@cam.ac.uk
Francis Nolan, fjn1@cam.ac.uk

ALTA Institute and Dept. of Theoretical and Applied Linguistics

The use of human assessors to evaluate spoken language proficiency has proven to be both time-consuming and expensive. This has resulted in considerable interest in developing systems for automating the assessment of non-native speech. There is also need to utilise linguistically meaningful measures that are based on perceptual cues used by humans to judge pronunciation. Previous research on non-native acquisition of vowel systems [1] suggests a strong link between vowel production accuracy and speech intelligibility. It is well known that the acoustic and perceptual identification of vowels rely on formant frequencies. However, formant analysis may not be viable in large-scale corpus research, given the need for manual correction of tracking errors. Spectral analysis techniques have been shown to be a robust alternative to formant tracking [2]. This paper explores the use of one such technique for modelling (SSBE) vowel spectra – the discrete cosine transformation [3] – as a feature in the automatic assessment of non-native English speakers. Mel-scaled DCT coefficients were calculated over a frequency band of 200-4000 Hz. Results show a statistically significant correlation between coefficients and pronunciation scores of speakers and suggest that this technique holds some promise for automated L2 assessment.

References

1. [1] Bohn, O and Flege, J. (1990). Interlingual identification and the role of foreign language experiencein L2 vowel perception. Applied Psycholinguistics, 11, 303-328.
2. [2] Palethorpe, S., Wales, r., Clark, J. E., & Senserrick, t. (1996). Vowel classification in children. Journal of the Acoustical Society of America, 100, 3843–51.
3. [3] Harrington, J. (2010). Phonetic Analysis of Speech Corpora. Wiley-Blackwell, Oxford.

Perceptual aftereffects after adaptation to different speech rates – behavioural and MEG studies

Saskia Helbling, saskia.helbling@mrc-cbu.cam.ac.uk
Helen Blank, helen.blank@mrc-cbu.cam.ac.uk
Leonhard Waschke, leonhard.waschke@gmx.de
Stephen A. Engel, engel@umn.edu
Matthew H. Davis, matt.davis@mrc-cbu.cam.ac.uk

MRC Cognition and Brain Sciences Unit

Neural oscillations in auditory areas entrain to low-frequency amplitude modulation in the speech envelope related to syllable rate and prosody *(Ding & Simon, 2013; Gross et al., 2013; Peelle et al., 2013)*. Previous MEG studies showed that low-frequency oscillations entrain also to the faster rhythm of time-compressed speech as long as the speech remains intelligible *(Ahissar et al., 2001)*. In a novel MEG study we seek to extend these earlier findings and investigate cerebro-acoustic coherence between speech envelopes and low-frequency oscillations during adaptation to fast and slow speech. We further aim to elucidate the neural underpinnings of perceptual aftereffects which can be observed after listening to time-compressed or time-expanded speech. Those aftereffects of slower or faster perceived speech rate are of particular interest as entrained oscillations have been suggested to carry temporal predictions about the incoming speech signal *(Schroeder et al., 2009, Davis and Peelle, 2012, Henry and Obleser, 2012; Ding and Simon, 2014)*.

We will present data from behavioural studies which show that participants tend to hear speech at a normal speech rate as being slower than normal after listening to fast speech and as faster than normal speech after listening to slow speech (Experiment 1). These aftereffects are shown to depend on the duration of the adaptation period and to decay over time after the

adaptation period ends (Experiment 2). After having established the presence of speech rate aftereffects and the time course of their decay, we designed an MEG study (Experiment 3) to investigate changes in cerebro-acoustic coherence during adaptation to different speech rates and during a decay period in which normal speech is heard as fast or slow due to perceptual aftereffects. We hypothesise that the cerebro-acoustic coherence during adaptation will be most pronounced at frequencies corresponding to the respective syllable rate for normal (~4Hz), fast (~6Hz) or slow (~2.4Hz) speech. Furthermore, we expect that coherence during decay periods of normal speech – when there is a mismatch between speech rate expectations based on the previous adaptation period and the actual speech rate – will be lower than for a control condition where subjects are adapted to normal speech. We will present preliminary data on these expected changes in cerebro-acoustic entrainment, using phase-coherence measures and cross-correlation analyses on beamformer source-reconstructed MEG data.

The dynamics of phonological and syntactic processing in spoken sentence comprehension

Hun Choi, hun@csl.psychol.cam.ac.uk
Billi Randall, billi@csl.psychol.cam.ac.uk
Barry J. Devereux, barry@csl.psychol.cam.ac.uk
Lorraine K. Tyler, l.tyler@csl.psychol.cam.ac.uk

Centre for Speech, Language and the Brain, Dept. of Psychology

Speech comprehension engages complex cognitive processes, including the rapid analysis of the lexical properties of words and their on-line integration to produce a representation of the structure and meaning of the utterance. Such rapid analysis and integration are based on actively constraining an upcoming word using a developing representation. To investigate the neural mechanisms of such incremental speech processing, we analysed magnetoencephalography (MEG) signals with representational similarity analysis (RSA; Kriegeskorte et al., 2008) which allows us to compare similarity patterns of brain activation with similarity patterns of psycholinguistic variables to identify neural regions and timings associated with those different language variables.

The main focus of the experiment was the role played by the main verb in constraining the upcoming complement structure for syntactic integration. For this purpose, we varied the likelihood of complement structures following main verbs (e.g. “searched the” vs. “searched for”). This was because the lexically-driven constraints of verbs strongly affect the ease of integrating the complement structures into a developing representation.

To model the dynamic processes of lexical analysis and integration, we computed a phonological model based on phonological overlap and a syntactic surprisal model reflecting the integration of the complement structure based on lexical constraints of the main verb. Based on previous findings (Tyler et al., 2013; Tyler & Marslen-Wilson, 2008), we extracted the data from regions of interest (ROIs) known to be involved in lexico-phonological and syntactic processing -- left Heschl’s gyrus (LHG), posterior superior and middle temporal gyrus (LpSTG; LpMTG), and left pars triangularis (LBA45) as well as their right-hemisphere homologues. From the complement word onset, we correlated the similarity patterns in the data for each ROI with the similarity patterns in our models for every time-point, and statistically tested the strength of this correlation over time (Maris & Oostenveld, 2007).

We found a strong correlation for the phonological model in the bilateral HG and pSTG, peaking around 100-120ms which reflects the lexico-phonological analysis of the complement words. Activity in L-BA45 in response to the surprisal measure was seen from

110ms to 255ms which reflected the degree of difficulty in integrating the complement structure. These results suggest that when starting around the time that bilateral phonology-related activity has peaked in the temporal ROIs, the complement word is integrated into the developing syntactic representation. L-BA45 is involved in this process and reflects the integration process by differentiating the items that are more difficult to integrate from those that are less difficult. The results suggest that the LIFG is sensitive to experience-based expectations about syntactic structure.

High rising terminals as markers of ethnicity in Sydney

Anna Jespersen, abj28@cam.ac.uk
Dept. of Theoretical and Applied Linguistics

One of the most salient sociophonetic markers of Australian English is the high rising terminal, a rising intonation contour used with declarative utterances. High rising terminals are well documented in standard Australian English, but we have very little evidence of their phonetics and phonology in the English spoken by Aboriginal Australians.

This study looked into the frequency of use, phonology and phonetic realisations of high rising terminals in Sydney Aboriginal English. This variety has attracted very little attention from linguists, and the few studies published focus on stereotypically Aboriginal features, and find no or few consistent differences from the standard (Eagleson 1977, 1978, 1984; Malcolm and Koscielecki 1997). However, sociolinguistic work has shown that ethnic minority groups, especially in urban environments, are likely to use linguistically distinct ethnolects (Foulkes and Docherty 2006; Kerswill, Torgersen, Fox and Cheshire 2005).

In order to investigate language use in Aboriginal Sydney, the speech of 22 Sydneysiders, 11Aboriginal and 11 non-Aboriginal, were recorded from local radio shows and transcribed in the ToBI intonational framework. Declarative intonational phrases (IPs) were categorised as having falling, level, rising or complex final f0 contours, and in IPs with rising contours, f0 was measured at pitch accents and boundary tones. Mixed-effects models were used for the analysis.

The two varieties were shown to be very alike in the distributions of declarative, IP-final intonation contours. For the rises, 5 broad types of rises were found, low rises, low- onset high rises, highonset high rises, and two types of fall-rises, low and high. The two varieties again showed very similar patterns in terms of the frequency of use of each rise type. However, in terms of phonetic realisation, the varieties were clearly distinct. Rises produced by non-Aboriginal were on average 2 semitones higher than those produced by Aboriginal speakers, and this difference was both significant overall and with 4 of the 5 rise types. There was no effect of age or gender, which points to some stability.

These results show that Aboriginal speakers in Sydney do not produce more high rising terminals than standard Australian English speakers, or different types of rises. Instead, the difference between the two varieties, while pervasive, was much more subtle. This suggests Aboriginal speakers do have some ethnolectal markers, but points towards a low degree of covert prestige. It also highlights the need for sociolinguistic studies to look at fine-grained variation, and at variables that are not stereotypically associated with the variety.

Rhotics in Singapore Englishes: a story of sociophonetic variation

Geraldine Kwek, gsck2@cam.ac.uk
Dept. of Theoretical and Applied Linguistics

Most phonetic research on segmentals in Singapore English (SgE) has been dedicated to the acoustic features of vowels. Studies on SgE consonants, on the other hand, have been largely phonological and descriptive in nature. This is especially evident in the discussion of rhotics in SgE. The focus has been primarily on rhoticity and, more particularly, the social factors which have attributed to the presence of word-/syllable-final /r/-realisation. The postvocalic /r/ has been shown to be present in SgE, a variety traditionally known to be non-rhotic (Deterding 2007; Poedjosoedarmo 2000a/b). It has been posited to be an influence of social factors, such as exposure to American media and tertiary education in American universities, and has also been associated with higher levels of education and prestige (Tan 2012). These studies, however, have been focused largely on SgE speakers from the major ethnic group (i.e. Chinese) and thus have limitations on being a true representation of SgE.

Moving on from an exclusive study of the /r/ in SgE and taking into account SgE speakers of the other ethnic groups (i.e. Malay, Indian, Eurasian), this study focuses on emergent patterns of phonetic and sociophonetic /r/-variation in this variety of 'New English'. The aims of the study are to provide detailed acoustic and auditory evidence for these variations, and to investigate the extent to which they serve as forms of social indexicality for speakers of SgE.

The approach taken in this study is twofold. An auditory analysis is done to investigate the types of /r/ variants present in SgE and the phonotactics of these different /r/ variants while an acoustic investigation focuses on determining their phonetic features. Coupled with the aim of documenting an unprecedented phonetic and phonological account of the /r/ in SgE, this will, ideally, contribute to the discussion on the phonetic correlates of /r/ in established studies (e.g. Catford 1977; Ladefoged & Maddieson 1996; Lindau 1985; Nolan 1983).

Part of this study also concentrates on sociophonetic implications that arise from the two-part analysis. Preliminary results show four different /r/-realisations in SgE; two of which are more dominant. There is also evidence for ethnically-determined variation in both the realization and phonotactics of /r/ in SgE, suggesting the possibilities of linguistic innovation and sound change. This adds a different dimension to discussions on the sociophonetics of /r/ in SgE and calls for further consideration of previous findings.

Generic and quantificational generalisations

Dimitra Lazaridou-Chatzigoga, dl518@cam.ac.uk
Napoleon Katsos, nk248@cam.ac.uk
Linnaea Stockall, l.stockall@qmul.ac.uk

Dept. of Theoretical and Applied Linguistics

Generic statements such as 'tigers have stripes' attribute properties to kinds, but unlike universally quantified statements ('Every tiger has stripes'), they are tolerant of exceptions (albino tigers). Generic generalisations have long been studied in formal semantics, within which genericity is frequently viewed as a species of quantification. Even though generics have been studied since the seventies (e.g. Carlson 1977) they still remain a rather controversial topic when it comes to deciding how to represent their semantic interpretation, how to build their logical form and how to model their truth conditions (Mari et al. 2013). Within formal semantics, modal and the probabilistic approaches seem most prominent, both of who treat genericity as akin to quantification (Krifka et al. 1995). Little is known though about how children learn the meaning of genericity or how process it, despite the obvious importance of

understanding how we think and speak properties and kinds. An influential proposal suggests that genericity is the default and mode of thinking and that universal quantifiers are initially misinterpreted as generic, an effect that has been called the Generic Overgeneralisation effect (Leslie 2007, 2008, Gelman 2010, Leslie et al. 2001). However this approach is not fully informed by cross-linguistic observations or theories of quantification. To establish a critical evidence-base, we investigated how adult native speakers of two languages with distinct ways of expressing generic statements, English and Greek, understand these statements and how young native speakers of English learn their meaning. In this poster, we will present the results of a recently completed adult study, where we explored whether the GOG effect can be replicated while manipulating different levels of context (neutral, supportive and contradictory). We found that the effect is more context-dependent than in previous studies, and that the data are best explained by other processes, e.g. quantifier domain restriction.

References

Carlson, G. (1977). Reference to Kinds in English. Ph.D. thesis, University of Massachusetts at Amherst.

Gelman, S.A. (2010). Generics as a window onto young children's concepts. In F. J. Pelletier (Ed.), Kinds, things, and stuff: The cognitive side of generics and mass terms. New directions in cognitive science, 100-123. New York: Oxford University Press.

Krifka, M., Pelletier, F., Carlson, G., ter Meulen, A., Chierchia, G. and Link, G. (1995). Genericity: An Introduction. In G. Carlson and F. J. Pelletier (Eds.) The Generic Book. Chicago. Chicago University Press, 1-125.

Leslie, S.J. (2007). Generics and the structure of the mind. Philosophical Perspectives 21, 375-405.

Leslie, S.J. (2008). Generics: Cognition and acquisition. The Philosophical Review, 117(1), 1-49.

Leslie, S.J., Khemlani, S. and Glucksberg, S. (2011). All Ducks Lay Eggs: The Generic Overgeneralization Effect. Journal of Memory and Language 65(1), 15-31.

Mari, A., Beyssade, C. and del Prete, F. (2013). Genericity. Oxford: OUP.

How do we determine the regional origin of speakers based on their pronunciation?

Adrian Leemann, al764@cam.ac.uk
Phonetics Laboratory, Dept. of Theoretical and Applied Linguistics

The speech stream consists of segmental features, individual sounds, and suprasegmental features, speech melody for example. Both these levels carry diagnostic information for accent identification. What current research fails to resolve is the question of the individual contribution of these cues in the identification process. We address this gap by means of an experiment in which segmental and suprasegmental cues are separated from one another. This can be achieved by manipulating the speech signal: the suprasegmental features of dialect X will be morphed onto the segments of dialect Y and vice versa [cf. 1]. These manipulated stimuli are then played to listeners who are asked to judge whether this sentences was from a speaker of dialect X or Y.

In a between-subject design with three conditions (clear speech, duration manipulated, duration + intonation manipulated), we recorded 6 Bern (BE) and 6 Valais (VS) Swiss German speakers. Each speaker read the same 10 sentences. The material was manually labeled in Praat [2]. For the manipulation conditions, the respective suprasegmental features of sentence 01 of speaker BE01, for example, were morphed onto the segments of sentence

01 of speaker VS01 and vice versa [3]. 60 Zurich German listeners (20/condition) participated in the experiment. Listeners heard 120 stimuli of 5-8 seconds each. Following stimulus presentation, subjects indicated whether the sentence heard was BE or VS Swiss German by clicking on the corresponding button on a laptop.

Results indicate that (i) listeners are very sensitive to discriminating between the two accents and (ii) suprasegmental cues seem to occupy only a marginal role when it comes to identifying a speaker's accent.

Research strategies for automatic spoken language assessment

Andrey Malinin, am969@cam.ac.uk
Rogier van Dalen, rcv25@cam.ac.uk
Kate Knill, kate.knill@eng.cam.ac.uk
Yu Wang, yw396@cam.ac.uk
Mark Gales, mjfg100@cam.ac.uk

Dept. of Engineering and ALTA Institute

Automatic spoken language assessment systems are becoming increasingly important to meet the demand for second-language assessment,especially for English. While modern automatic assessment systemsbased on speech delivery features perform well in many cases, there are situations where backing-off to a human grader is necessary. Two scenarios are considered in this work: when the automatic grader is assessing a spoken response which is significantly different from the training data, causing the grader to assign incorrect grade; and when the response is inconsistent with the topic of the question, which cannot be detected by a grader using response delivery based features. Rejection strategies to address both these situations are described. First, using a Gaussian process based grader provides a measure of the uncertainty of its prediction in addition to the prediction itself. This uncertainty measure is sufficiently accurate to indicate which automatic grades should be re-graded by humans. Second, a topic adapted Recurrent Neural Network Language Model (RNNLM) is used to detect whether the spoken response is appropriate for the assessment prompt. Here, an RNNLM is adapted to example responses for each prompt. These adapted models are then used to assess the probability of the response for each prompt. By ranking these probabilities it is possible to detect inappropriate responses to prompts. These approaches are evaluated on the Business Language Testing Service (BULATS) online spoken tests for English.

Off-record indirectness: intuition and practice

Jessica Soltys, jms256@cam.ac.uk
Napoleon Katsos, nk248@cam.ac.uk
Dept. of Theoretical and Applied Linguistics

Off-record indirect speech acts (ORI) are pragmatically ambiguous and encode two meanings – a literal meaning and a functional meaning. The speaker deliberately phrases in the act in such a way that is relevant enough to generate the intended implicature and, if the hearer is amenable to the act, achieves the desired perlocutionary effect. The ambiguity of the ORI act affords the speaker plausible deniability.

Traditional accounts of off-record speech attribute the use of indirectness to politeness-based concerns, specifically the desires to avoid imposition or offense and, thereby, to respect the hearer's negative face wants. Requests for favours are among the prototypical uses of polite ORI (Brown and Levinson 1987).

Recent game-theoretic approaches, such as the Strategic Speaker account, focus on the use of ORI in controversial scenarios such as bribes and sexual propositions (Pinker 2007; Pinker, Nowak and Lee 2008; Lee and Pinker 2010). In these cases, ORI offers an optimal payoff ratio to the speaker – he balances risks (whether legal, financial, psychological, or interpersonal) with anticipated rewards.

The proposed poster will present two empirical studies. The first study is a novel qualitative experiment about trends in the use of ORI among native English speakers. Participants relate both their experiences with ORI and their intuitions about why ORI is used. The data were coded in accordance with the above-mentioned theories (as well as other accounts not discussed here). The majority of responses (51%) were related to politeness-based accounts, while only 7% described Strategic Speaker-based uses of ORI.

The second experiment, a modified version of Lee and Pinker's 2010 study, describes the use of ORI in scenarios depicting favours, bribes, and propositions, respectively. For each speech act type, sample utterances are provided, along with the mean use of ORI, as summarised below.

Speech act	Mean ORI
Bribe	0.78
Proposition	0.92
Favour	0.03

In Experiment 1, participants do not report using ORI in Strategic Speaker-based situations, but frequently produce ORI utterances in these circumstances, of their own accord, under experimental conditions in Experiment 2. Likewise, in Experiment 1, participants frequently recount polite uses of ORI, yet in Experiment 2, they do not produce ORI in the scenario tailored to politeness-based accounts. This discrepancy will be discussed in the proposed poster.

References

Brown, Penelope and Stephen C. Levinson, *Politeness: Some universals in language use* (Cambridge: Cambridge University Press, 1987).
Lee, James J. and Steven Pinker, 'Rationales for indirect speech: The theory of the Strategic Speaker', *Psychological Review*, 117.3 (2010), 785–807.
Pinker, Steven, 'The evolutionary social psychology of off-record indirect speech acts', *Intercultural Pragmatics*, 4.4 (2007), 437–461.
Pinker, Steven, Martin A. Nowak, and James J. Lee, 'The logic of indirect speech', *Proceedings of the National Academy of Sciences*, 105.3 (2008), 833–838

Token-based generalisation is enhanced by overnight consolidation

Viñals, L, lv288@cam.ac.uk
Mirković, J.
Gaskell, M.G.
Davis, M, matt.davis@mrc-cbu.cam.ac.uk

Dept. of Theoretical and Applied Linguistics and MRC Cognition and Brain Sciences Unit

The learning and generalisation of inflectional morphology has been shown to depend on both type and token frequency. Here, we focused on understanding how speakers generalise to novel forms that are phonologically similar to both existing high type frequency regulars and high token frequency irregulars (e.g. *plit* as in *fit-fitted* and *sit-sat*). In these ambiguous cases,

generalisation can be based on type (*plitted*) or token (*plat*) frequency. We were specifically interested in assessing whether a period of overnight consolidation changes the representation of type and token frequency and, in turn, their relative influence on generalisation behaviour.

To address this question, we extended a paradigm developed by Mirković and Gaskell (in prep.) and created an artificial language containing regular and irregular plural affixes that combined with novel nouns referring to the occupation of male and female characters. The plural forms of the novel words were designed such that the majority had regular plural affixes and a minority had irregular plural affixes. That is, regulars had high type frequency (12 items) and irregulars had low type frequency (3 items). Irregulars also had high token frequency (24 presentations/item) relative to regulars (6 presentations/item). One subset of irregulars had an ambiguous phonological cue (e.g. *arb: varb, farb)* that was associated with a high token frequency plural affix in irregulars (*varbesh*) but a high type frequency plural affix in regulars (*farbaff* but also *grollaff, shilnaff*, etc.).

Eighteen participants were trained on these novel plural affixes in 3 sessions over the course of one week. On day 1, participants learned 36 novel masculine and 36 novel feminine singulars. On day 8, participants learned the plural form for 18 words in one gender (consolidated affixes). On day 9, participants learned the plural form for 18 different words in the other gender (unconsolidated affixes) before completing a generalisation test in which novel gendered singular forms were presented. For novel singular forms containing the ambiguous phonological cue, participants were more likely to generalise based on token frequency for consolidated affixes but more likely to generalise on type frequency for unconsolidated affixes.

These results suggest changes in the representation of newly-learned inflected words following overnight consolidation that lead to differential influences of type and token frequency during generalisation. Enhanced token-based generalisation following overnight consolidation is explained within a complementary learning systems framework.

Text readability assessment for second language learners

Menglin Xia, mx223@cam.ac.uk
Computer Laboratory/ALTA Institute

Developing reading ability is an essential part of language acquisition. However, finding the proper reading materials for training and testing second language (L2) learners at a specific level of proficiency is a demanding and time-consuming task for English instructors as well as the readers themselves. An automated text readability assessment system aims to analyze and assess the reading difficulty of texts according to levels of language proficiency. It can be applied to automatic selection of the appropriate reading materials from newspapers or webpages for practice and for language testing, which enables many pedagogical applications by supporting readers in their second language education.

Text readability assessment is to quantify the readability of text through readability measures by assigning a numerical score or other form of estimation that indicates a degree of reading difficulty for a group of readers to the text. Early works on readability were often based on traditional readability formulas computed from surface textual characteristics, such as syllable and word counts. Compared to traditional methods, a machine learning approach enables the investigation of richer text representations derived from computational linguistics, which can provide more accurate and robust analysis of text difficulty.

In my experiments, readability assessment is addressed as a supervised machine learning problem and, in particular, a ranking approach is adopted and compared against a classification method. A pairwise ranking algorithm rankSVM and a standard support vector machine classifier are implemented. A range of linguistic features are explored in the readability assessment experiment and their predictive power for reading difficulty prediction are investigated. The set of features includes traditional readability measures, lexical features, parse tree syntactic features, parsing complexity measures and language modelling (LM) based features. A cross validation process is adopted for evaluation.

Generalisation and consolidation in learning to read words and name objects in an fMRI study

Connor Quinn1,2, Jo Taylor3, Matt Davis1
1. MRC Cognition and Brain Sciences Unit, Cambridge, UK
2. Department of Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK
3. Department of Psychology, Royal Holloway, University of London, Egham, Surrey, UK

Learning to read recruits regions of the brain that would otherwise be involved with other processes, e.g., visual object recognition and language. The extent to which these regions become specialised for reading, rather than other processes is a debated issue, in particular the role of fusiform areas (Price & Devlin, 2011; Dehaene & Cohen, 2011). Teaching participants to read an artificial orthography and to name novel objects offers a useful perspective on this debate.

Taylor et al. (2014) showed additional neural activity in anterior fusiform regions when participants learned to name novel objects and left parietal cortex while learning to read new words written in unfamiliar symbols. These findings might be taken to challenge theories in which fusiform regions are specialised for reading (Cohen et al., 2002) but could also be attributable to these novel items not yet being consolidated in memory. Here we explore the impact of overnight consolidation on activation associated with reading a novel writing system and naming novel objects. We combined the word/object learning task from Taylor et al. (2014) with a design in which participants learn different items on two successive days before scanning on the second day (cf. Davis et al., 2009). Consequently items learned on day 1 should be consolidated at the time of scanning.

Results showed participants could read untrained words by generalising spelling-sound mappings. By using a functional localiser to analyse neural responses to novel written words and novel objects we show that novel items activate the same neural systems as existing objects and words. Furthermore, there was substantial overlap for words and objects in fusiform regions, with more activation for both existing and novel objects compared to both existing and novel words. Analysis of the functional localiser data was extended to include response time, revealing increased activation for objects over words in fusiform gyrus, even after accounting for the increased BOLD response due to slower response times for objects compared to words.

References

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, 125(5), 1054–1069.

Davis, M. H., Di Betta, A. M., Macdonald, M. J., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience*, 21(4), 803–820.

Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, 15(6), 246–253.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2014). Distinct neural specializations for learning to read words and name objects. *Journal of Cognitive Neuroscience*, 26(9), 2128–2154.