BICLCE Bamberg 2019
Workshop 1 – Challenges of computer-mediated communication data: #obstacles #opportunities

Convenors
Marie-Louise Brunner (Trier University of Applied Sciences), Stefan Diemer (Trier University of Applied Sciences), Matt Gee (Birmingham City University) & Andrew Kehoe (Birmingham City University)

List of presenters and workshop schedule.
Presentations have the usual 20+10 format. We scheduled 6 presentations and 30 minutes for introduction and general discussion, for a total of 7 slots and 3:30h workshop duration, not counting breaks

- Introduction (Conveners) (10mins)
- Paper 1: Sofie Decock (Ghent University) & Rebecca Van Herck (Ghent University)
- Paper 2: Ursula Lutzky (WU Vienna) & Matt Gee (Birmingham City University)
- Paper 3: Andrew Kehoe (Birmingham City University) & Matt Gee (Birmingham City University)
- Paper 4: Aatu Liimatta (University of Helsinki)
- Paper 5: Theresa Müller (Saarland University)
- Paper 6: Marie-Louise Brunner (Trier University of Applied Sciences) & Stefan Diemer (Trier University of Applied Sciences)
- Discussion (20mins)

Workshop description and paper abstracts

Workshop description
Computer-mediated communication (CMC) provides an ideal opportunity to study developments in contemporary English as they happen. This workshop aims to provide a discussion forum for researchers who are compiling and analyzing CMC data from a wide range of contexts. We invite research on social media (Facebook, Twitter, Instagram, etc.), blogs, forums, e-commerce, as well as online (email, IM) and video-mediated communication.
In particular, we are interested in research which discusses issues arising from the complex nature of CMC data. The workshop considers questions about data compilation, management, and analysis and invites researchers to approach data in new and flexible ways. We welcome innovative solutions and/or interdisciplinary studies combining fields such as corpus linguistics, sociolinguistics, (multimodal) discourse analysis, and pragmatics.
Recent studies in CMC focus, for example, on regional and social variation (Huang et al. 2016), stance and solidarity (Evans 2016), self-branding (Page 2012), (im)politeness and online aggression (Lambert Graham & Hardaker 2017, Lutzky & Kehoe, 2017), topic and sentiment (Bourlai & Herring 2014, Kehoe & Gee 2012), multimodality, and interaction strategies (Brunner, Diemer & Schmidt 2016, Honey & Herring 2009).

By bringing together researchers working with CMC data we intend to foster an exchange of ideas for dealing with complex and multimodal datasets. Contributions to the workshop should thus illustrate problematic aspects and possible solutions with the help of concrete linguistic case studies. The workshop will close with a round table providing room for a discussion of opportunities and challenges with regard to this new data.

References

Bourlai, E. and Herring, S.C., 2014. Multimodal communication on tumblr: i have so many feels! *Proceedings of the 2014 ACM conference on Web science.* ACM, pp. 171-175.

Brunner, M.-L., Diemer, S., and Schmidt, S., 2016. “It’s always different when you look something from the inside” – Linguistic innovation in a corpus of ELF Skype conversations. *International Journal of Learner Corpus Research* 2(2): 321-348.

Diemer, S., Brunner, M.-L., and Schmidt, S., 2016. Compiling computer-mediated spoken language corpora: Key issues and recommendations. *International Journal of Corpus Linguistics* 21(3): 349-371.

Evans, A., 2016. Stance and Identity in Twitter Hashtags. Language@Internet 12(1).

Honey, C. and Herring, S.C., 2009. Beyond microblogging: Conversation and collaboration via Twitter. *HICSS'09, 42nd Hawaii International Conference on System Sciences* 2009: 1-10.

Huang, Y., Guo, D., Kasakoff, A. and Grieve, J., 2016. Understanding US regional linguistic variation with Twitter data analysis. *Computers, Environment and Urban Systems* 59: 244-255

Kehoe, A. and Gee, M., 2013. eMargin: A Collaborative Textual Annotation Tool. Ariadne 71.

Kehoe, A. and Gee, M., 2012. Reader comments as an aboutness indicator in online texts: introducing the Birmingham Blog Corpus. *Studies in Variation, Contacts and Change in English*, ed. by S. Oksefjell Ebeling, J. Ebeling and H. Hasselgård. Helsinki: Varieng.

Kehoe, A. and Gee, M., 2007. New corpora from the web: making web text more 'text-like' in P. Pahta, I. Taavitsainen, T. Nevalainen and J. Tyrkkö (eds.) *Towards Multimedia in Corpus Studies.* University of Helsinki.

Lambert Graham, S. and Hardaker, C., 2017, (Im)politeness in digital communication. in J. Culpeper, M. Haugh & D. Kádár (eds), *The Palgrave Handbook of Linguistic (Im)politeness.* Palgrave, London, pp. 785-814.

Lutzky, U. and Kehoe, A., 2017. "I apologise for my poor blogging": Searching for Apologies in the Birmingham Blog Corpus. *Corpus Pragmatics* 1(1): 1-20.

Muñoz, A.S., 2016. Attending Multi-Party Videoconference Meetings: The Initial Problem. *Language@Internet* 13(3).

Page, R., 2012. The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse & Communication* 6(2): 181-201.

Paper abstracts

Paper 1:
(De)legitimation strategies in company e-mail replies to customer complaints and their related social media comments by (potential) customers

Rebecca Van Herck & Sofie Decock (Ghent University)

When a service recovery attempt fails, customers turn to social media to confront the company with their inaptitude to solve the complaint, hoping to elicit reactions from the company and other customers. This interaction between (potential) customers and companies is examined from a legitimacy perspective: Which strategies do they use to (de)legitimize themselves or the other party in private and public CMC communication? And how do these strategies contribute to their interactional dynamics? Based on a move analysis, we performed a discursive legitimation analysis of 150 e-mails (company replies to complaints) and 300 related Facebook comments (Reyes, 2011). Results showed that companies seek legitimacy in e-mail replies to customers by balancing interpersonal (e.g. apology) and transactional (e.g. compensation) legitimation strategies. Customers try to delegitimate companies by attacking their moral legitimacy, the pragmatic legitimacy of the decision, and the customer service agent's communication skills. When responding publicly to these attacks, we notice a shift in the companies' approach: they tend to restrict themselves to interpersonal legitimation strategies and are reluctant to be publicly transparent about the transactional side of complaint handling. This seems to silence the public discussion, but its disadvantage is the lack of public evidence of successful complaint handling.

References

Bhatia, V., 1993. *Analyzing genre: Language use in professional settings.* London: Longman.

Bhatia, V., 2008. Genre analysis, ESP and professional practice. *English for Specific Purposes* 27(2): 161-174.

Einwiller, S. and Steilen, S., 2015. Handling complaints on social network sites – An analysis of complaints and complaint responses on Facebook and Twitter pages of large US companies. *Public Relations Review* 41: 195-204.

Reyes, A., 2011. Strategies of legitimization in political discourse: From words to actions. Discourse & Society 22(6): 781-807.

Suchman, M., 1995. Managing legitimacy: Strategic and institutional approaches. *The Academy of Management Review* 20(3): 571-610.

Swales, J. and Najjar, H., 1987. The writing of research article introductions. *Written Communication* 4(2): 175-191.

Upton, T. and Cohen, M. A., 2009. An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies* 11(5): 585-605.

Vaara, E. and Tienari, J. (2008). A Discursive Perspective on Legitimation Strategies in Multinational Corporations. *The Academy of Management Review* 33(4): 985-993.

Van Leeuwen, T. (2007). Legitimation in discourse and communication. *Discourse and Communication* 1(1): 91-112.

Paper 2:

"holy shit that is awesome good for you!" – The study of speech acts in online comments

Ursula Lutzky (WU Vienna) & Matt Gee (Birmingham City University)

This paper explores the possibilities of studying speech acts in large corpora. In corpus pragmatics, the study of speech acts has been mainly concerned with finding ways of identifying and searching for these pragmatic constructions in electronic corpora. This is because most speech acts are defined with regard to their function rather than the specific linguistic form in which they appear in texts. Consequently, previous research has often studied either a closed set of verbs and constructions that typically function as a specific speech act (e.g. sorry or apologise for the speech act of apology, see Deutschmann 2003) or analysed the metacommunicative potential of these expressions, studying both performative and discursive uses of speech acts (see e.g. Jucker and Taavitsainen 2014).

In this study, we elaborate on a new methodological approach that involves studying one particular type of text at a time as well as focusing on a specific position within the text to uncover manifestations of speech acts (Lutzky and Gee 2018). We present the results of a contrastive analysis of online comments drawn from two corpora that were compiled by the Research and Development Unit at Birmingham City University. These include a sub-corpus of 86 million words of comments on blog posts, which were published on the WordPress and Blogger hosting sites between 2000 and 2010 and form part of the Birmingham Blog Corpus (http://www.webcorp.org.uk/blogs), and a second dataset of 280 million words of reader comments which were left on articles on The Guardian website between 2006 and 2010.

The analysis of online comments on blogs and newspaper articles illustrates how the use of specific speech acts may be identified by focusing on the initial position in this type of text. The contrastive perspective allows insights into similarities and differences with regard to the way in which blog and newspaper comments are initiated and to the types of speech acts appearing in these comments. The results show that while blog comments reflect more positive attitudes, with speech acts such as complimenting and thanking occurring frequently at their onset, newspaper comments appear to be more matter of fact and argumentative in nature, with speech acts such as asserting and suggesting appearing at the start of the comments.

References

Deutschmann, M., 2003. *Apologising in British English.* Umeå: Umeå University.

Jucker, A. H. and Taavitsainen, I., 2014. Complimenting in the history of American English. Taavitsainen, I., Jucker, A.H. and Tuominen, J. (eds.). *Diachronic Corpus Pragmatics*. Amsterdam: John Benjamins, pp. 257-276.

Lutzky, U. and Gee, M., 2018. ‘I just found your blog’. The pragmatics of initiating comments on blog posts. *Journal of Pragmatics* 129: 173-184.

Paper 3:

Analysing emojis in context in a corpus of Twitter data

Andrew Kehoe & Matt Gee (Birmingham City University)

Social media has become an increasingly multimodal space. Highfield & Leaver (2016: 47) explore the growing significance of emojis, images, animated GIFs and videos, noting that “visual social media content can highlight affect, political views, reactions, key information, and scenes of importance”. Emojis in particular have received increasing attention from a variety of perspectives (e.g. Dresner & Herring 2010, Ljubešić & Fišer 2016, Miller et al. 2016, Wijeratne et al. 2016, Maíz-Arévalo 2017).

In this paper we carry out a large-scale data-driven corpus pragmatic analysis of emoji use on Twitter. Our corpus contains a random selection of 40 million tweets in English and German downloaded between February and May 2018 using the Twitter API. Table 1 shows the proportion of tweets in our random sample containing multimodal elements.

Language	Total	Images	Videos	Animated GIFs	Emojis
English	38,902,024	8,625,473 (22%)	2,259,007 (6%)	530,617 (1%)	6,381,274 (16%)
German	493,183	68,205 (14%)	8,083 (2%)	3,953 (1%)	65,528 (13%)

Table 1. Frequency of multimodal elements in a random sample of tweets

Our focus in this study is on the 6 million tweets containing at least one emoji (a considerably higher proportion than those containing videos or animated GIFs). Table 2 shows the most frequent emojis found in our random sample, using the official character names as defined in the Unicode standard.

English		**German**	
Emoji	**% of Tweets**	**Emoji**	**% of Tweets**
😂 Face With Tears of Joy	4.42	😂 Face With Tears of Joy	2.49
😭 Loudly Crying Face	1.85	❤️ Red Heart	1.36
❤️ Red Heart	1.51	😍 Smiling Face With Heart-Eyes	0.89
🔥 Fire	1.31	🔥 Fire	0.57
😍 Smiling Face With Heart-Eyes	1.16	👍 Thumbs Up	0.56
💕 Two Hearts	0.61	😊 Smiling Face With Smiling Eyes	0.53
🤣 Rolling on the Floor Laughing	0.51	😉 Winking Face	0.51
👏 Clapping Hands	0.49	🤔 Thinking Face	0.43
🙏 Folded Hands	0.47	😘 Face Blowing a Kiss	0.39
✨ Sparkles	0.46	😭 Loudly Crying Face	0.39

Table 2. Most frequent emojis in a random sample of tweets

Of particular interest to us are those emojis with more than one possible use. In our previous corpus pragmatic studies we used collocational analyses to search for apologies in a large corpus of blog posts and comments (Lutzky & Kehoe 2017) and to examine the use of swearwords in the same corpus (Lutzky & Kehoe 2016), finding collocational analyses to be a useful method for disambiguating different uses of the same word. We now apply the same technique to the disambiguation of emojis. For example, the ‘folded hands’ emoji can be interpreted as thanking, pleading, praying or giving a high five, and we find collocates reflecting all of these uses. We are also able to examine collocation between emojis. We find that in the vast majority of cases emojis tends to co-occur with themselves (e.g. ‘face with tears of joy’ collocates with itself 4.9 million times) but, interestingly, we also find collocation between semantically-related emojis (e.g. ‘smiling face with heart eyes’ and ‘red heart’).
Throughout the paper we use specific examples to illustrate how the words and other emojis collocating with a particular emoji offer clues about its intended meaning and illocutionary force, and we describe some of the challenges we faced, relating to collocational span and the frequent repetition of emojis within single tweets.

References

Dresner, E., Herring, S.C., 2010. Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force. Communication Theory 20, 249–268.

Highfield, T., Leaver, T., 2016. Instagrammatics and digital methods: studying visual social media, from selfies and GIFs to memes and emoji. Communication Research and Practice 2, 47–62.

Ljubešić, N., Fišer, D., 2016. A Global Analysis of Emoji Usage, in: Proceedings of the 10th Web as Corpus Workshop.
Lutzky, U. and Kehoe, A., 2016. “Your blog is (the) shit”: a corpus linguistic approach to the identification of swearing in computer mediated communication. International Journal of Corpus Linguistics 21(2), pp.165-191.
Lutzky, U. and Kehoe, A., 2017. “I apologise for my poor blogging”: Searching for apologies in the Birmingham Blog Corpus. Corpus Pragmatics 1(1), pp.37-56.
Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., Hecht, B., 2016. “Blissfully Happy” or “Ready to Fight”: Varying Interpretations of Emoji, in: Proceedings of the Tenth International AAAI Conference on Web and Social Media. Presented at the ICWSM.
Maíz-Arévalo, C., 2017. Emoticons in transactional and interactional exchanges: social networking chitchat versus working negotiation. ODISEA. Revista de estudios ingleses.
Wijeratne, S., Balasuriya, L., Sheth, A., Doran, D., 2016. EmojiNet: Building a Machine Readable Sense Inventory for Emoji. Proceedings of the International Workshop on Social Informatics.

Paper 4:
Is text length a linguistic variable? Evidence from social media
Aatu Liimatta (University of Helsinki)

When text length within a corpus varies dramatically, normalized word frequencies are an unreliable indicator of linguistic variation. In order to reduce this effect in statistical corpus studies, it is common to simply remove texts shorter than some threshold from the data. For example, studies on Wikipedia often exclude the shortest articles (cf. Hiltunen 2014). In many published corpora, only a sample of fixed length is included from each text to lessen the effect of text length on token frequencies (cf. e.g. Francis and Kučera 1964). Ideal text length has been the subject of discussion for a long time (cf. e.g. Biber 1993).

However, texts have a structure: a short text is not equivalent to a sample of similar length from a longer text. Limiting in any way the length of texts considered in a study means that that shorter texts and their unique features are effectively ignored in studies. This is not so problematic with many of the more "traditional" genres included in corpora, as they tend to be made up of longer texts to begin with. However, if we want to study variation within various CMC genres, and particularly between the smallest units of social media communication, i.e. between single social media comments, the issue becomes more noteworthy, as social media comments typically consist of very short texts or texts of varying length.

In this paper, I will focus on Reddit, the third-largest English-language social media platform; preliminary data indicates that 50% of Reddit comments are under 20 words long. Sharing many similarities with traditional online forum platforms, Reddit contains discussions on a wide variety of topics. Moreover, Reddit effectively does not limit the length of a comment (unlike e.g. Twitter), which makes it possible to compare comments of very different length.

Starting with the hypothesis that text length is determined in part based on the purpose of the text, and so is related to the register of the text, we should be able to observe differences in the distribution of register features such as word length or first-person pronouns and their collocation patterns in texts of different length. These differences will be analyzed using statistical corpus methods and provide evidence for whether or not text length in itself is a relevant register feature in social media comments, a genre which is characterized by highly varying text length. Preliminary results based on Reddit data indicate that the distribution of first-person pronouns correlates to a degree with text length, with shorter comments generally containing more first-person pronouns.

References

Biber, D., 1993. Representativeness in corpus design. Literary and Linguistic Computing, 8(4), 243-257.

Francis, W. N. and Kučera, H., 1964. *A Standard Corpus of Present-Day Edited American English, for use with Digital Computers (Brown)*. Providence, Rhode Island: Brown University.

Hiltunen, T., 2014. Choice of national variety in the English-language Wikipedia. Tyrkkö, J. and Leppänen S. (Eds.), *Texts and Discourses of New Media.* Helsinki: VARIENG. Retrieved from http://www.helsinki.fi/varieng/series/volumes/15/hiltunen/.

Paper 5:

Face-threatening acts and impoliteness on social media – Webcare on Instagram

Theresa Müller, Saarland University

Studies of face-to-face analyses and polls of customer behaviour suggest that while customer behaviour in service interactions varies, positive politeness strategies rather than negative strategies as defined by Brown and Levinson (1978) by companies are prevalent (cf. Lerman 2006, Ryoo 2005) in response to face-threatening acts (cf. Du et al. 2010). This paper contributes to the study of webcare interactions through a study of face-threatening acts and impoliteness in customer interaction on the social medium Instagram. The study is based on data collected from the Instagram page of the fashion company ASOS. The dataset was analyzed both qualitatively and quantitatively.

In a first step, Instagram webcare interactions by the selected company were collected for analysis. Politeness strategies based on both positive politeness and negative politeness were considered in detail, and customer questions initiating webcare interactions were qualitatively analyzed and categorized as potentially face-threatening or not face-threatening. All questions and remarks categorized as face-threatening were considered for further analysis if ASOS responded to them. The company's responses to these potential face-threatening acts were then studied and compared to positive politeness and negative politeness strategies (cf. Brown and Levinson 1978). Potential face-threatening acts by customers, responses by ASOS, as well as replies to the company's response were considered.

A quantitative corpus analysis was performed on the data based upon the qualitative findings of the discourse analysis, counting the total usage of politeness strategies in the data collected. Each webcare response by ASOS was categorized as using positive or negative politeness.

The study illustrates that the usage of positive politeness strategies is a key webcare method on the social medium Instagram in the collected company data when reacting to potential face-threatening acts. This is in line with findings from face-to-face service failure studies (Du et al. 2010, Lee et al. 2013) The findings are then discussed further, also considering the effects of positive and negative politeness on the overall tone of discourse.

References:

Brown, P and Levinson, S., 1978. Universal in Language Usage: Politeness Phenomena. Questions and politeness: Strategies in social interaction, ed. by Goody, E.N. Cambridge

University Press, pp.56-311.
Du, J., Fan, X. and Feng, T., 2010. An experimental investigation of the role of face in service failure and recovery encounters. Journal of Consumer Marketing, 27(7), pp.584-593.
Lee, Y.L., Sparks, B. and Butcher, K., 2013. Service encounters and face loss: Issues of failures, fairness, and context. International Journal of Hospitality Management 34: 384-393.
Lerman, D., 2006. Consumer politeness and complaining behavior. Journal of Services Marketing 20(2): 92-100.
Ryoo, H.K., 2005. Achieving friendly interactions: A study of service encounters between Korean shopkeepers and African-American customers. Discourse & Society, 16(1), pp.79-105.

Paper 6:

"We're always looking for great content" – Customer Interaction via Instagram

Marie-Louise Brunner & Stefan Diemer (Trier University of Applied Sciences)

The paper examines customer communication via Instagram in selected companies. Instagram is a key company advertising and webcare channel, particularly for companies that address an international and young customer base. Our aim is to document existing practice that can help to optimize channel-specific customer communication for an international and plurilingual audience. Starting from a documentation of existing interactional practices in Instagram (Lee & Chau 2018, Mapes 2018, Tuten & Solomon 2017), our approach aims at extending this research to a multilingual, cross-border setting, focusing in particular on how companies can address global audiences and how success in social media communication can be evaluated. For the study, public Instagram interactions of selected internationally active companies (e.g. MyMuesli, Yves Rocher, SAP, Adidas, Bitburger) and self-marketing influencers (e.g. Kylie Cosmetics, Jamie Oliver) were collected, analyzed and classified. Documented communication includes branding, advertising, sales, customer engagement, and complaint management. Successful interactions are used to create a baseline of how Instagram 'best practice' works linguistically, based on likes, reach, and customer engagement as criteria.

A qualitative analysis provides an overview of interactional strategies used in Instagram. We expand and support the qualitative findings with a limited quantitative analysis of general thread sentiment and company tone.

Instagram affordances include features such as hashtag clouds, Instastories, and visual interaction. Results indicate that Instagram communication showcases a distinct company tone and media-specific content as part of strategic posting concepts. Hashtags are used to extend messages and portray social media identity while introducing a personal, authentic and humorous tone. Customer interaction also reflects these affordances and the intended company tone. While the key medium for reaching multilingual audiences is English, companies also establish multiple monolingual channels to reach audiences with specific language backgrounds. Even in monolingual settings, plurilingual customer engagement is achieved by multilingual hashtags, automated translation, and hybrid advertising. Customer feedback and interaction is also frequently plurilingual. Culture-specific images may prompt intercultural negotiations in comments, engaging customers further. Instagram Stories have speedily become an efficient marketing tool, re-establishing traditional time-bound and temporary advertising channels in a low-cost and opt-in social medium setting (Hu et al. 2014).

In general, Instagram provides an ideal platform for valorizing customer testimonials and outsourcing webcare, with customers themselves serving as disseminators and troubleshooters. In sum, the study provides insights into Instagram customer interactions and illustrates how companies create an authentic and effective social media presence for a national and international customer base, how they engage customers and build rapport across borders, and how they maintain their brand identity.

References

Hu, Y, Manikonda, L. and Kambhampati, S., 2014. What we instagram: A first analysis of instagram photo content and user types. Eighth International AAAI conference on weblogs and social media.

Lee, C. and Chau, D., 2018. Language as pride, love, and hate: Archiving emotions through multilingual Instagram hashtags. Discourse, context & media 22 (2018): 21-29.

Mapes, G., 2018. "(De) constructing distinction: Class inequality and elite authenticity in mediatized food discourse." Journal of sociolinguistics.

Tuten, T. L., and Solomon, M.R., 2017. Social media marketing. London: Sage.

Digitising the Linguistic Atlas of Scotland:
Scope and Potential of a new Corpus

John Kirk

The origins of the present work lie with the various attempts by the author to digitise the material of the Linguistic Atlas of Scotland (LAS), vols. 1 and 2 (Mather & Speitel 1975, 1977) which culminated in several publications (Kirk 1994, 1996, 2001; Kirk & Kretzschmar 1992; Kirk & Munroe 1989; Kirk, Munroe & O'Kane 1994) and Master Theses (Cooper 2001, Doherty, 2001, O'Kane 1990, O'Neill 2002, Robinson 2002). With preliminary funding by the University of Vienna, the project has been resurrected in 2019 and begun anew. Since the first endeavours, there has been a revolution in geographical information systems, the interactivity of databases, and automatic cartography, so that the present endeavour has in effect to start afresh.

A second interest in the original first project was to initiate the quantitative analysis of dialect data, inspired by Goebl (1982) (cf. also Viereck 1985, as applied to the data of the Survey of English Dialects) and by work in the USA which culminated in Kretzschmar & Schneider (1996). Burkette & Kretzschmar (2018) have recently shown, on the basis of the item 'parlour', how complexity theory provides a good explanation for dialect survey data, in particular showing "that most linguistic data are, in fact, not normally distributed. The type of distribution we most often encounter for language data is that of the hyperbolic asymptotic curve (A-curve)" (Kretzschmar & Burkette, in prep. p. 5). This hypothesis lends itself to testing by the LAS data as it is well-known that many Scots lexical items are pan-Scottish (i.e. widely distributed throughout the Lowland Scots-speaking areas) whereas many others are quite localised. However, to test this thoroughly, a digitised atlas would be required.

The present paper begins with a critical look at the onomasiological approach displayed in LAS1 and LAS2 with the preference for an ortho-lexical, quasi-phono-lexical categorisations, which excludes oncers, and which avoids physical maps. The present approach seeks to recategorise the data in terms of purely lexical (i.e. etymologically-lexical) types and thus is able in to include the many omitted oncers listed in the atlas appendices. The new atlas uses physical base maps and makes uses of interactive mapping techniques, including displays fir for age and gender. Preliminary examples appear in Hessle (2019).

A further aim is a re-evaluation of the distribution of lexical types per map item; and to account afresh for the origins of these items. Hessle (2019) shows that for many speakers along the Fife littoral, for 'splinter' many informants are recorded as offering both *skelb* and *skelf*. Far from being simply phonological variants, *skelb* is shown to have Gaelic origins whereas *skelf* appears to have come from Middle Dutch. Although traditionally regarded as a Scots-speaking area, Fife was early a stronghold of Gaelic settlement, as shown extensively in the five volumes of *The Place-Names of Fife* (Taylor & Markus 2006-2013). In the medieval period, for reasons of trade, many Flemish immigrants settled along the Fife littoral, as reflected in the Dutch architecture of Culross (cf. Price 2013), and contributing loanwords to the vocabulary of Scots, most strikingly the word for what became Scotland's national sport: *golf* (cf. Murison 1971).

The paper will report ongoing developments by the time of the conference. It will meet all the workshop's requirements by being corpus-based; by triangulating lexical analysis, cartography and socio-cultural history; by taking a quantitative approach (following the lead being shown by Kretzschmar & Burkette); by making a new use of an old resource with new and innovative software; and by drawing comparison with similar lexical dialectological projects (e.g. in the US: cf. Burkette & Kretzschmar 2018).

References

Burkette, Allison & William A. Kretzschmar, Jr. 2018. *Exploring Linguistic Science: Language Use, Complexity and Interaction*. Cambridge: Cambridge University Press.

Cooper, Anne Marie. 2001. *An interactive web-based wordmap generator of information on the occurrence of variations of particular standard items in areas in Northeast Scotland*. Belfast: Queens University M.Sc. dissertation in Computer Science (unpublished).

Doherty, Gavan. 2001. *Word Maps of Northern Ireland*. Belfast: Queens University M.Sc. dissertation in Computer Science (unpublished).

Goebl, Hans. 1982. *Dialektometrie, Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Hessle, Christian. 2019. *Towards a Digital Version of the Linguistic Atlas of Scotland*. Vienna: University of Vienna BA dissertation (unpublished).

Kirk, John M. & George Munroe. 1989. 'A Method for Dialectometry', *Journal of English Linguistics* 22. 97–110. [special issue subtitled *Computer Methods in Dialectology* edited by William. A. Kretzschmar, Edgar Schneider & Ellen Johnson]

Kirk, John M. & William A. Kretzschmar, Jr. 1992. Interactive Linguistic Mapping of Dialect Features, In *Literary and Linguistic Computing* 7. 168–175.

Kirk, John M. 1994. 'East Central Scots: A Computerised Mapping Package'. In Alexander S. Fenton and Donald A. MacDonald (eds.), *Studies in Gaelic and Scots*. Edinburgh: Canongate Academic. 48–68.

Kirk, John M., George Munroe & M.J. Damian O'Kane. 1994. 'Electronic Word Maps'. In Susan Hockey and Nancy Ide (eds.), *Research in Humanities Computing*. Oxford: Oxford University Press. 195–223.

Kirk, John M. 1996. 'Maps: Dialect and Language'. In Ron E. Asher and J. M. Y. Simpson (eds.), *The Encyclopedia of Language and Linguistics*, vol. 5, 2363–2377. Oxford: Pergamon Press.

Kirk, John M. 2001. Maps: Dialect and Language. In Rajend Mesthrie (ed.), *Concise Encyclopedia of Sociolinguistics*. Oxford: Elsevier Science. 350–362.

Kretzschmar, William A., Jr. & Allison Burkette. in preparation. 'Complex Systems and British Isles Survey Data'.

Kretzschmar, William A., Jr. & Edgar W. Schneider. 1996. *Introduction to Quantitative Analysis of Linguistic Survey Data: An Atlas by the Numbers*. Thousand Oaks, CA: Sage Publications.

Mather, James, Y. & H.-H. Speitel. 1975. *The Linguistic Atlas of Scotland: Scots Section*, vol. 1. London: Croom Helm.

Mather, James, Y. & H.-H. Speitel. 1977. *The Linguistic Atlas of Scotland: Scots Section*, vol. 2. London: Croom Helm.

Murison, David D. 1971. The Dutch element in the Vocabulary of Scots. In Adam J. Aitken & Angus McIntosh & Hermann Pálsson (eds.) *Edinburgh Studies in English and Scots*. London: Longman. 159–176.

O'Kane, M. J. Damian. 1990. *A Relational Database for Linguistic Geography*. Belfast: Queens University M.Sc. dissertation in Computer Science (unpublished).

O'Neill, Michael. 2002. *A Web-interactive Word Geography of Northern Ireland*. Belfast: Queens University M.Sc. dissertation in Computer Science (unpublished).

Price, Matthew. 2013. Crowsteps in Fife: The Flemish connection, Part 1. Scotland and the Flemish People Project. University of St. Andrews. https://flemish.wp.st-andrews.ac.uk/2013/12/06/crowsteps-in-fife-the-flemish-connection-part-1/. (30 January, 2019.)

Robinson, Tracy Angela. 2002. *An Interactive Word Atlas of East-Central Scots for the World Wide Web*. Belfast: Queens University M.Sc. dissertation in Computer Science (unpublished).

Taylor, Simon & Gilbert Markus. 2006–2013. *The Place-Names of Fife*, 5 vols. Stamford: Shaun Tyas.

Viereck, Wolfgang. 1985. Linguistic Atlases and Dialectometry: The Survey of English Dialects. In John M. Kirk & Stewart F. Sanderson & John Widdowson (eds.), *Studies in Linguistic Geography: The Dialects of English in Britain and Ireland*. London: Croom Helm. 94–112.

When intuitions (don't) fail: Combining syntax and sociolinguistics in the analysis of Scots

E. Jamieson, Shouchun Chien, Gary Thoms, David Adger, Caroline Heycock & Jennifer Smith

In the examination of non-standard morphosyntactic forms such as those in (1), a perennial problem for sociolinguists is that the 'linguistic features of interest are of such low frequency that quantitative studies are not feasible' (Labov 1996:78).

(1)	a. Aye the floor *needs renewed*.	(needs +ed)
	b. Ken, and if you *div* mess up, which you will.	(*div* for *do*)
	c. The windows *didna* have *nae* glass in them.	(negative concord)

An obvious solution to this problem is to adopt methodologies from syntax, specifically acceptability judgment data gleaned from speaker intuition. However, sociolinguists have been generally reluctant to adopt such a technique, perhaps due to the repeated observation that speakers may 'agree that a certain form is completely unacceptable, yet use it themselves freely in every-day speech' (Labov 1996: 78).

In this paper we present a new digital resource which employs both sociolinguistic methodologies in spoken data and intuitions captured through acceptability judgments in the analysis of non-standard morphosyntactic forms. By having access to both these data types from the same speakers, we have the opportunity to assess the reliability of the judgment data as measured against the spontaneous speech we recorded.

The data come from the Scots Syntax Atlas which sets out to map syntactic structure across time and space in 140 locations across Scotland. We gathered speakers' intuitions through a acceptability judgment questionnaire (e.g. Schütze 1996, Barbiers & Bennis 2007), carried out with over 500 speakers across 200 data points, as well as spontaneous spoken data gathered through sociolinguistic interviews (Labov 1966) between pairs of speakers in each location, c.300 hours in total.

In this analysis, we target the three non-standard morphosyntactic forms in (1), which are said to be differentiated both geographically and socially: the needs +ed form (1a), the auxiliary verb *div* (1b) and negative concord (1c). For each of these morphosyntactic structures, we first map the results from the acceptability judgment data. We then turn to the corpus of spoken data to test whether the patterns of acceptability from the judgment data map on to production data. Our results show that the reliability of judgments is variable-dependent. The needs +ed form (1a) is accepted in most varieties across Scotland and also appears in a diverse range of dialects in the spoken data. The auxiliary *div* (1b) is much more geographically circumscribed in the elicited data, and so too it is in the spoken data, appearing in the north east and the borders only. The results for negative concord (1c) are much more mixed in the judgment data, with no clear pattern of acceptability. However, there are clear geographic and socially distributed patterns of use that arise from the spoken data, suggesting that 'intuitions fail' (Labov 1996) with this particular variable.

These results show that it is not the case that 'any grammatical pattern that is perceived as regional [is] suppressed in introspection' (Labov 1996:98). We discuss these results in the light of constraints on a speaker's ability to tap intuitions about non-standard morphosyntactic forms, specifically how both social and psycholinguistic pressures may operate in governing the reliability of judgment data.

References

Barbiers, Sjef & Hans Bennis. 2007. The Syntactic Atlas of the Dutch Dialects: A discussion of choices in the SAND project. *Nordlyd* 34. 53–72.

Labov, William 1966. *The social stratification of English in New York City.* Washington D.C.: Center for Applied Linguistics.

Labov, William. 1996. When intuitions fail. In L. McNair, K. Singer, L Dolbrin & M. Aucon (eds.), *Papers from the Parasession on Theory and Data in Linguistics. Chicago Linguistic Society (CLS) 32*. 77–106.

Schütze, Carson T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology.* Chicago: University of Chicago Press.

Progressive verb forms in Scottish Standard English: A questionnaire-based approach

Manfred Krug, Ole Schützler & Graeme Trousdale

The use of progressive verb forms with stative verbs (e.g. *liking, wanting, understanding, thinking, doubting*) is a common feature of many varieties of English world-wide (cf. Kortmann & Lunkenheimer 2013). There is ample evidence in the literature that be V-*ing* constructions with stative verbs are also frequent in Scottish varieties of English (along the entire Scots-English continuum; cf. McArthur 1979). For example, Miller (2008: 307) comments on such patterns in Scots, which he views as symptoms of change in progress; further discussions are found in Aitken (1979: 105), Millar (2007: 75) and Beal (1997: 372–373). Smith (2012: 23) classifies the use of be + V-*ing* to stative verbs as a widespread 'covert Scotticism' likely to be used in Scottish Standard English (SSE), and it is also identified as a promising object of investigation by Schützler, Gut & Fuchs (2017).

In our paper, we specifically target raters' intuitions concerning the use of progressive forms with stative verbs in the standard language, i.e. SSE, using the *Bamberg Questionnaire for Lexical and Morpho-Syntactic Variation in English* (Krug & Sell 2013). This has been used in research on several varieties of English (e.g. Krug, Schützler & Werner 2016) and contains a large number of lexical and grammatical items. We will in this paper concentrate on questionnaire results for the following three:

(1) I'm really liking this film.

(2) Are you understanding my point?

(3) What are you wanting?

Informants were asked to rate on a six-point Likert scale (the extreme poles being "no-one" and "everyone") the likelihood of the three items to be used in semi-formal writing and in colloquial speech. The former context was tested by presenting participants with written stimuli, and the latter by presenting twice each of the sentences as audio stimuli spoken by a native speaker. The data for our study are taken from 78 questionnaires representing SSE usage; for comparative purposes, we investigate 32 questionnaires completed by speakers of Southern Standard British English (SSBE) and 52 questionnaires completed by speakers of American English (AmE). This results in a total of *n* = 972 items across the three target sentences and two modes of production. For the quantitative analysis, we treat the Likert-scale as continuous, converting responses into values on a scale of 0–5 (cf. Wray & Bloomer 2012: Ch. 13; Agresti & Finlay 2009: 40). The data are subsequently analysed using a linear mixed-effects regression model with VARIETY, MODE and GENDER as fixed effects and RATER as a random effect.

We find in our data that (i) in SSE all three verbs indeed receive higher ratings than in both AmE and SSBE, that (ii) the gradation in terms of usage ratings found in SSBE (*liking* > *understanding* > *wanting*), as well as the strong dispreference for *want* progressives in SSBE and AmE, are largely absent from SSE, that (iii) gender patterns (e.g. Labov 2001) indicate a norm proscribing against be V-*ing* with stative verbs in (written) SSBE but not so much in the other two varieties, and that (iv) the expected quasi-stylistic differences between speech and writing are also less pronounced in SSE and AmE. We complement our findings with corpus data from ICE-GB, ICE-USA and ICE-SCO (see Schützler, Gut & Fuchs 2017) to reflect on usage-based (particularly constructional) approaches to dialect variation (e.g. Mukherjee and Gries 2009).

In our paper we thus confirm the special status of be V-*ing* constructions in SSE, provide more detailed insights into lexical and socio-stylistic factors that have an effect on their variation, and show that, with regard to this particular feature, the two British varieties of SSE and SSBE are most remote from each other.

References

Agresti, Alan & Barbara Finlay. 2009. *Statistical methods for the social sciences.* Upper Saddle River, NJ: Pearson Education.

Aitken, Adam J. 1979. Scottish speech: a historical view with special reference to the Standard English of Scotland. In Adam J. Aitken & Tom McArthur (eds.), *Languages of Scotland*, 85–118. Edinburgh: Chambers.

Beal, Joan. 1997. Syntax and morphology. In Charles Jones (ed.), *The Edinburgh history of the Scots language*, 335–377. Edinburgh: Edinburgh University Press.

Kortmann, Bernd & Kerstin Lunkenheimer (eds.). 2013. *The Electronic World Atlas of Varieties of English.* Leipzig: Max Planck Institute for Evolutionary Anthropology. http://ewave-atlas.org.

Krug, Manfred & Katrin Sell. 2013. Designing and conducting interviews and questionnaires. In Manfred Krug & Julia Schlüter (eds.), *Research methods in language variation and change*, 69–98. Cambridge: Cambridge University Press.

Krug, Manfred, Ole Schützler & Valentin Werner. 2016. Patterns of linguistic globalization: Integrating typological profiles and questionnaire data. In Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja & Sarah Chevalier (eds.), *New approaches to English linguistics: Building bridges*, 35–66. Amsterdam/Philadelphia: John Benjamins.

Labov, William. 2001. *Principles of linguistic change*, vol. 2: *Social factors*. Oxford: Blackwell.

McArthur, Tom. 1979. The Status of English in and furth of Scotland. In Adam J. Aitken & Tom McArthur (eds.), *Languages of Scotland*, 50–67. Edinburgh: Chambers.

Millar, Robert. 2007. *Northern and Insular Scots*. Edinburgh: Edinburgh University Press.

Miller, Jim E. 2008. Scottish English: Morphology and syntax. In Bernd Kortmann & Clive Upton (eds.), *Varieties of English*, vol. 1: *The British Isles*, 299–327. Berlin: Mouton de Gruyter.

Mukherjee, Joybrato & Stefan Th. Gries. 2009. Collostructional nativisation in New Englishes: Verb-construction associations in the International Corps of English. *English World-wide* 30(1). 27–51.

Schützler, Ole, Ulrike Gut & Robert Fuchs. 2017. New perspectives on Scottish Standard English: Introducing the Scottish component of the International Corpus of English. In Sylvie Hancil & Joan C. Beal (eds.), *Perspectives on Northern Englishes* (Topics in English Linguistics 96), 273–302. Berlin: Mouton de Gruyter.

Smith, Jennifer. 2012. Scots and varieties of Scots. In Bernd Kortmann & Kerstin Lunkenheimer (eds.), *The Mouton World Atlas of Variation in English*, 21–29. Berlin: Mouton de Gruyter.

Wray, Alison & Aileen Bloomer. 2012. *Projects in linguistics and language studies: A practical guide to researching language*, 3rd edn. London: Hodder Arnold.

Variety-specific coding practices within a shared system: Modal verbs expressing strong obligation in Scottish, Southern British and American English

Jenny Herzky, Ole Schützler & Zeyu Li

This paper focuses on modal verbs of strong obligation in Scottish Standard English (SSE), American English (AmE) and (Southern) British English (BrE). In the latter two varieties, the system of modal and semi-modal verbs has been in flux at least since the middle of the twentieth century (Leech 2003; Smith 2003; Millar 2009). Many developments are explicable as 'modal decline' (Leech et al. 2009), i.e. a decrease in the frequencies of core modals partly compensated by an increase in the frequencies of semi-modals. From an onomasiological perspective focused on strong-obligation modality, such developments can be interpreted as changing coding practices for this particular semantic category, namely an increased relative frequency of the verbal predicates *need to* and *have to,* and a concomitant decrease in the frequency of *must* (cf. Millar 2009: 204, 208–209). For Scots (and Scottish English), Miller (2008: 305) makes even stronger claims to the effect that obligation is expressed by the verbs *have to* and *need to*, while *must* hardly expresses strong obligation (ex. 1) but is reserved for the expression of epistemic modality (ex. 2; also cf. Miller & Brown 1982; Millar 2009: 204; Kirk 1987).

(1) Scotland's mountains and wild lands are one of our greatest treasures and **must** be protected. (deontic/obligation; ICE-SCO:rep-062)

(2) I keep thinking about the other tenants in the building and how they **must** be feeling. (epistemic; ICE-SCO:rep-056)

Due to the nature of their data, it is uncertain how results from Miller & Brown (1982), Kirk (1987), Miller (2008) and Millar (2009) can be generalized to SSE. This is what we put to the test in our contribution.

We focus on current SSE usage, comparing new corpus material from ten written genres in the Scottish component of the *International Corpus of English* (ICE-SCO; cf. Schützler, Gut & Fuchs 2017) to corresponding texts from ICE-GB and ICE-USA. The verbs included are *must*, *have to*, *need to*, (*have*) got to, and *need* (cf. Smith 2003). While the main focus is on inter-varietal differences, we also include subject (1st, 2nd and 3rd person) and tense (past, present, future) as predictor variables in the statistical models. We proceed in two steps. First, we inspect the general text frequency of (semi-)modal constructions expressing strong obligation by means of a negative binomial regression model. Second, we investigate which verbs are selected within this semantic domain by means of a multinomial regression model.

We find that modal-verb constructions of strong obligation are most frequent with third-person subjects and least likely with second-person subjects. While this pattern is shared by all three varieties, SSE stands out in exhibiting a particularly high frequency of strong obligation with first-person subjects. We draw on democratisation theory (cf. Farrelly & Seoane 2012; Millar 2009) to account for these findings. In BrE and SSE *have to* and *need to* compensate for the avoidance of *must* in combination with second-person subjects, while AmE uses only *have to* for this end. Furthermore, our data show that *need to* is generally much more frequent in SSE.

Based on our data, we find that none of the modal constructions in SSE are extreme in that they would be unacceptable (or even salient) in other standard varieties (cf. Miller 2008 for comments on Scots). However, we argue that while there is of course a shared system of modal verbs of strong obligation, precise patterns of use differ considerably between SSE, BrE and AmE. The Scottish variety seems to stand out in various respects and is more than "Standard English spoken with a Scottish accent" (Stuart-Smith 2008: 48).

References

Farelly, Michael & Elena Seoane. 2012. Democratization. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 392–401. Oxford: Oxford University Press.

Kirk, John M. 1987. The heteronomy of Scots with Standard English. In Caroline Macafee & Iseabail Macleod (eds.), *The nuttis schell: Essays on the Scots language presented to A. J. Aitken*, 166–181. Aberdeen: Aberdeen University Press.

Leech, Geoffrey. 2003. Modality on the move: The English modal auxiliaries 1961–1992. In Roberta Facchinetti, Manfred Krug & Frank Palmer (eds.), *Modality in contemporary English* (Topics in English Linguistics 44), 223–240. Berlin/New York: Mouton de Gruyter.

Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.

Millar, Neil. 2009. Modal verbs in TIME: Frequency changes 1923–2006. *International Journal of Corpus Linguistics* 14(2). 191–220.

Miller, Jim. 2008. Scottish English: Morphology and syntax. In Bernd Kortmann & Clive Upton (eds.), *Varieties of English*, vol. 1: *The British Isles*, 299–327. Berlin: Mouton de Gruyter.

Miller, Jim & Keith Brown. 1982. Aspects of Scottish English syntax. English *World-wide* 3(1). 3–17.

Schützler, Ole, Ulrike Gut & Robert Fuchs. 2017. New perspectives on Scottish Standard English: Introducing the Scottish component of the International Corpus of English. In Sylvie Hancil & Joan C. Beal (eds.), *Perspectives on Northern Englishes* (Topics in English Linguistics 96), 273–302. Berlin/Boston: De Gruyter Mouton.

Smith, Nicholas. 2003. Changes in the modals and semi-modals of strong obligation and epistemic necessity in recent British English. In Roberta Facchinetti, Manfred Krug & Frank Palmer (eds.), *Modality in contemporary English* (Topics in English Linguistics 44), 241–266. Berlin/New York: Mouton de Gruyter.

Stuart-Smith, Jane. 2008. Scottish English: Phonology. In Bernd Kortmann & Clive Upton (eds.), *Varieties of English*, vol. 1: *The British Isles*, 48–70. Berlin: Mouton de Gruyter.

Rhotics in Scottish Standard English

Philipp Meer, Robert Fuchs, Anika Gerfer, Ulrike Gut & Zeyu Li

The rhotic sound in Scottish Englishes has various phonetic realisations, varying from alveolar trills and taps to approximants. In addition, in coda position, <r> is increasingly not realised in some speaker groups (Stuart-Smith et al. 2007, Dickson & Hall-Lew 2017, Stuart-Smith 2003, Lawson et al. 2008, 2011), although Scottish English is still usually described as “generally rhotic” (Wells 1982: 10, Stuart-Smith 2008: 64). Previous studies have shown that the realisation of onset and coda /r/ in Scottish Englishes varies with socio-economic class, age, gender as well as amount of contact with British English speakers (Stuart-Smith et al. 2007, 2014, Dickson & Hall-Lew 2017, Stuart-Smith 2003, Schützler 2013, 2015). Moreover, it varies with language-internal factors such as stress and phonetic environment (Lawson et al. 2008, Schützler 2010, 2013, Dickson & Hall-Lew 2017), with language-internal factors outweighing the social factors (Jauriberry et al. 2015). While previous studies focussed on urban Scottish Englishes spoken in the Central Belt, this study aims to explore the distribution and realisation of rhotics in the nation-wide standard (Scottish Standard English, SSE), which is still largely unexplored in this respect (but see Jauriberry et al. 2015).

In particular, it is investigated whether an underlying rhotic standard exists for SSE speakers from all over Scotland, whether and where rhotics are realised as trills and taps, which are considered “traditionally Scottish” (Johnston 1997, Stuart-Smith 2008), and what factors influence variation in the realisation and distribution of the SSE rhotic sounds. To this end, formal speeches given in the Scottish Parliament by 49 speakers (aged 14-71) from all over Scotland were orthographically and phonemically annotated, and all rhotics were analysed auditorily by at least two coders (n=5536; overall intercoder agreement: 91.01%). We conducted separate analyses (binary mixed effects logistic regression models) for /r/ in syllable onset, non-linking coda, and linking coda positions.

The results show that SSE is variably rhotic with 54% of all non-linking coda /r/ realised (see Fig. 1). Trills and taps are more frequent in onset position, where they account for 37.1 % of all occurrences. In non-linking coda position, only 6.3 % are realised as trill/tap. The realisation of /r/ varies systematically with:

- preceding and following phonetic context (tokens that occur intervocalically favour trill/tap – independently of whether they occur in onset (64.2%) or linking coda position (59.0%),
- status as single consonant or cluster (although tokens in clusters generally favour rhoticity, singleton non-linking coda /r/ are more likely to be realized as tap/trill)
- and, in onset /r/, also depending on whether they occur in a content or function word (content words favour trill/tap).

While age as a single factor did not have any significant effect, there was a significant interaction of age and gender: middle-aged females were more likely to be rhotic than middle-aged men; similar effects did not occur in the other age groups. Moreover, female speakers, more so than men, favour trills and taps over approximants in linking coda position.
In sum, our results contradict the idea of SSE being “generally [or almost fully] rhotic” but confirm previous reports of increasing occurrence of non-rhoticity, not just by specific speaker groups, but in a formal context. Variation in the realization of rhotics in SSE is foremost affected by language-internal rather than language-external factors. In contrast to previous findings on (non-standard) urban Scottish Englishes, in SSE, females rather than males show a higher proportion of “traditionally Scottish” pronunciations.

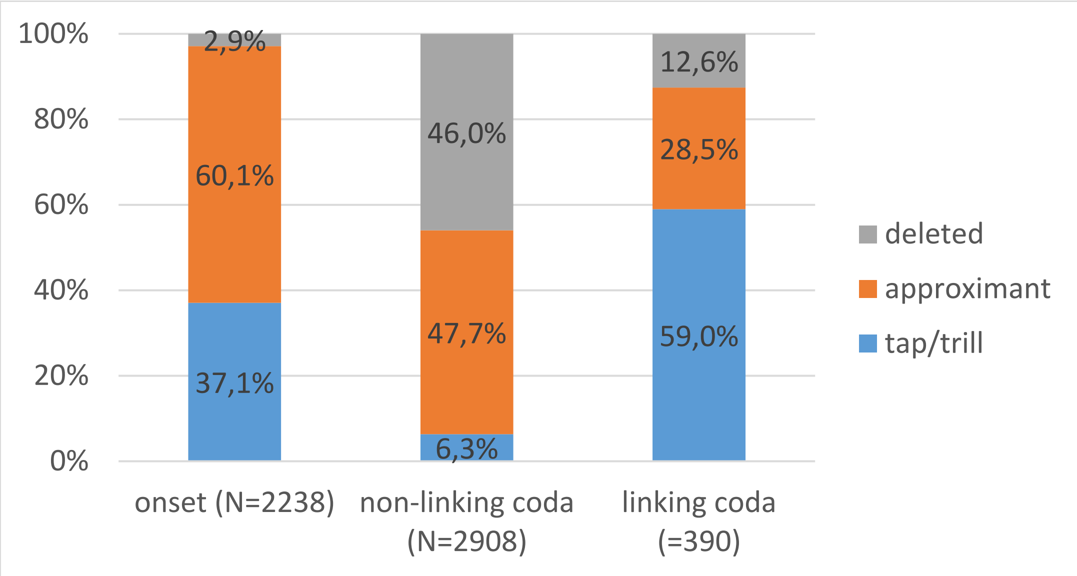

Fig. 1. Overview relative type-token frequencies by context and realisation of /r/ (%).

References

Dickson, Victoria & Lauren Hall-Lew. 2017. Class, gender, and rhoticity: The social stratification of non-prevocalic /r/ in Edinburgh speech. *Journal of English Linguistics* 45(3). 229–259.

Jauriberry, Thomas, Rudolph Sock & Albert Hamm. 2015. Phonetic variation in Standard Scottish English: Rhotics in Dundee. *Proceedings of the 18th International Congress of Phonetic Sciences [ICPhS]*, Glasgow.

Johnston, Paul. 1997. Regional variation. In Charles Jones (ed.), *The Edinburgh history of Scots*, 433–513. Edinburgh: Edinburgh University Press.

Lawson, Eleanor, James M. Scobbie & Jane Stuart-Smith. 2011. The social stratification of tongue shape for postvocalic /r/ in Scottish English. *Journal of Sociolinguistics* 15(2). 256–268.

Lawson, Eleanor, Jane Stuart-Smith & James M. Scobbie. 2008. Articulatory insights into language variation and change: Preliminary findings from an ultrasound study of derhoticisation in Scottish English. *University of Pennsylvania Working Papers in Linguistics* 14(2). 101–110.

Schützler, Ole. 2013. The sociophonology and sociophonetics of Scottish Standard English (r). In Peter Auer, Javier Caro & Götz Kaufmann (eds.), *Language variation – European perspectives IV*, 215–228. Amsterdam: Benjamins.

Schützler, Ole. 2015. *A sociophonetic approach to Scottish Standard English*. Amsterdam: Benjamins.

Stuart-Smith, Jane. 2003. The phonology of modern Urban Scots. In John Corbett, Derrick McClure & Jane Stuart-Smith (eds.), *The Edinburgh companion to Scots*, 110–137. Edinburgh: Edinburgh University Press.

Stuart-Smith, Jane. 2008. Scottish English: Phonology. In Bernd Kortmann & Clive Upton (eds.), *Varieties of English*, vol. 1: *The British Isles*, 48–70. Berlin/New York: Mouton de Gruyter.

Stuart-Smith, Jane, Eleanor Lawson & James M. Scobbie. 2014. Derhoticisation in Scottish English: A sociophonetic journey. In Chiara Celata & Silvia Calamai (eds.), *Advances in Sociophonetics*, 57–94. Amsterdam: Benjamins.

Stuart-Smith, Jane, Claire Timmins & Fiona Tweedie. 2007. Stuart-Smith, Jane, Claire Timmins & Fiona Tweedie. 2007. ‘Talkin’ Jockney’? Variation and change in Glaswegian accent. *Journal of Sociolinguistics* 11(2). 221–261.

Wells, John. 1982. *Accents of English*. Cambridge: Cambridge University Press.

The fate of the Scottish Vowel Length Rule in contemporary Scottish English

Jane Stuart Smith & Rachel MacDonald

The Scottish Vowel Length Rule (SVLR) posits that in word-final stressed syllables, certain Scottish English vowels are short except before /r/, voiced fricatives and morpheme boundaries. (e.g. Aitken, 1981, 2016; Scobbie, Hewlett &Turk, 1999). The current manifestation of the SVLR continues a much earlier process, which is known from historical record to have begun in the 1600s, and appears to be gradually receding. The number of vowels to which it now applies is a matter of debate. It is also closely linked to the definition of 'Scottish English', which now tends to follow e.g. Aitken (1984) and takes Scottish English as a sociolinguistic continuum with Scottish Standard English (SSE) at one end and a local variety of Scots at the other (e.g. Stuart-Smith, 2004). Current descriptions of the SVLR agree that it does not apply to /ɪ/ and /ʌ/, which are short in all environments in all dialects (e.g. for Scots, Aitken, 1981, 2015; for SSE, Scobbie et al, 1999). Describing Scots, Aitken (1981, 2015) excludes an additional four of the 19 vowels featured in his Historical Table of the Scots Vowels and concludes that the SVLR applies to all the remaining vowels in some dialects, and to a number of these vowels in all dialects. More recent research suggests that in SSE the Rule only applies to /i/, /ʉ/, and /ai/ (e.g. Scobbie et al, 1999), although there is new evidence to suggest more vowels may be affected in the North-East (Warren, 2018). Weakening of the SVLR is linked with dialect contact with Anglo-English in Edinburgh (Hewlett, Matthews & Scobbie, 1999), while prosodic factors play a key role in real-time evidence for weakening of the Rule in Glasgow vernacular (Rathcke & Stuart-Smith 2016; Stuart-Smith, José, Rathcke, Macdonald, & Lawson, 2017).

This paper is the first to consider the fate of the SVLR by vowel, dialect, social factors, and time period, from a large-scale analysis of vowel duration from approximately 500 speakers across three SSE corpora (Arthur the Rat, DoubleTalk , How Stable is the Standard) and five regional vernacular Scots corpora (Sounds of the City, SCOTS, One Speaker Two Dialects, Brains in Dialogue, Glaswasian) together covering the main dialect areas across Scotland: the Central Belt (Glasgow and Edinburgh), the South/Borders, the North East and the Highlands & Islands. Large-scale acoustic analysis across multiple datasets, also controlling for prosodic factors such as speech rate and phrase position, is possible using the Integrated Speech Corpus ANalysis (ISCAN) tool (e.g. McAuliffe et al, 2017).

Our dataset allows us to consider the fate of the SVLR for all stressed monophthongs, in SSE and Scots, with respect to the key dimensions of interest. Recordings of both spontaneous and controlled speech from the late 1940s until the mid-2010s in a number of dialects enable inspection over time and across the lexicon (Rathcke & Stuart-Smith, 2016). In addition, the inclusion of gender and ethnicity within the dataset allows for investigation of the role of social factors in the maintenance of the SVLR, including the influence of heritage ethnic background (c.f. Hewlett, Matthews & Scobbie, 1999; Stuart-Smith, Timmins, & Alam, 2011). Using these data, we answer the following questions:

- Which vowels are subject to the SVLR across Scottish dialects?
- How use of the SVLR influenced by social factors, such as gender, social class, and ethnicity?
- How has the implementation of the SVLR changed over time in Scottish English?

References

Abercrombie, David. 1979. The accents of Standard English in Scotland. In Adam J. Aitken & Tom McArthur (eds.), *Languages of Scotland*, 68–84. Edinburgh: Chambers.

Aitken, Adam J. 1981, 2015. The Scottish Vowel-length Rule. In Adam J. Aitken & Caroline Macafee (eds.), *Collected writings on the Scots language*. Scots Language Centre. http://media.scotslanguage.com/library/document/aitken/The_Scottish_Vowel_length_Rul%20e_(1981) (22 January, 2019). Originally published in: Michael Benskin & Michael Louis Samuels (eds.). 1981. *Someny people longages and tonges: Philological essays in Scots and mediaeval English presented to Angus McIntosh*, 131–157. Edinburgh: The Middle English Dialect Project.

Aitken, Adam J. 1984. Scots and English in Scotland. In Peter J. Trudgill (ed.), *Language in the British Isles*, 517–532. Cambridge: Cambridge University Press.

Hewlett, Nigel, Ben Matthews & James M. Scobbie. 1999. Vowel duration in Scottish English speaking children. *Proceedings of the 15th International Congress of Phonetic Sciences [ICPhS]*, San Francisco. 2157–2160.

McAuliffe, Michael, Elias Stengel-Eskin, Michaela Socolof & Morgan Sonderegger. 2017. Polyglot and speech corpus tools: A system for representing, integrating, and querying speech corpora. *Proceedings of Interspeech 2017*, Stockholm.

Rathcke, Tamara V. & Jane H. Stuart-Smith. 2016. On the tail of the Scottish Vowel Length Rule in Glasgow. *Language and Speech* 59(3). 404–430.

Scobbie, James M., Nigel Hewlett & Alice E. Turk. 1999. Standard English in Edinburgh and Glasgow: The Scottish Vowel Length Rule revealed. In Paul Foulkes & Gerard J. Docherty (eds.), *Urban voices: Accent studies in the British Isles*, 230–245. London: Arnold.

Stuart-Smith, Jane. 2004. Scottish English: Phonology. In Bernd Kortmann, Edgar W. Schneider, Kate Burridge, Rajend Mesthrie & Clive Upton (eds.), *A handbook of varieties of English*, vol. 1: *Phonology*, 47–67. Berlin/New York: Mouton de Gruyter.

Stuart-Smith, Jane, Brian José, Tamara Rathcke, Rachel Macdonald & Eleanor Lawson. 2017. Changing sounds in a changing city: An acoustic phonetic investigation of real-time change over a century of Glaswegian. In Chris Montgomery & Emma Moore (eds.), *Language and a sense of place: Studies in language and region*, 38–65. Cambridge: Cambridge University Press.

Stuart-Smith, Jane, Claire Timmins & Farhana Alam. 2011. Hybridity and ethnic accents: A sociophonetic analysis of 'Glaswasian'. In Frans Gregersen, Jeffrey K. Parrott & Pia Quist (eds.), *Language variation – European perspectives III: Selected papers from the 5th International Conference on Language Variation and Change in Europe [ICLaVE] 5)*, Copenhagen, 2009 (Studies in Language Variation 7), 43–58. Amsterdam: Benjamins.

Warren, David. 2018.The Scottish Vowel Length Rule in North East Scotland. PhD Thesis, University of Aberdeen (unpublished).

Language use in the peripheries and centres of megacities: Comparing Hong Kong and Moscow

Carolin Biewer, Elena Dieser, Lisa Lehnen, Ninja Schulz (Würzburg)

Sub-communities in megacities emerge and dissolve as residents move to, from and within the urban space trying to get access to "the symbolic centre of the city" (Saunders 2011: 3) but possibly ending up at the periphery. Some communities are spatially connected, others segregated. Some districts offer traditional living space, others are transitional. Studies in the field of urban sociolinguistics have so far underestimated these urban dynamics and focused on a restricted number of cities located in 'Western nations', particularly where English is the native language of most residents, even though megacities are largely situated in Asia, Africa and South America. This paper tries to bridge this gap by investigating the interrelationship of language and space in Hong Kong and Moscow.

Both these cities are culturally and linguistically diverse. With Cantonese spoken as a mother tongue, and English and Putonghua known already before school by a large number of residents (Bacon-Shone et al. 2015: 18), multilingualism is a widespread phenomenon in Hong Kong. The same holds for Moscow which may appear monolingual at first sight but proves to be much more multifarious: a considerable amount of Muscovites, for instance Tatars, Greeks, Lithuanians, Uzbeks, use more than one language and many Russians who moved to Moscow from other cities or areas speak regional varieties (Baranova et al. 2018, Kitajgorodskaja & Rosanova 2010). Concerning space, Hong Kong has developed into "a wealthy commercial and entrepreneurial powerhouse" since the 1960s, attracting international businesses and immigration (Bolton 2000: 268); Moscow has experienced similar social, political and economic changes. In this context, housing scarcity and redevelopment projects constitute issues that currently affect the dynamics in both cities.

Using a web-based questionnaire, we study sociolinguistic background and accent perception of inhabitants of different districts in both megacities. Moreover, we investigate how residents construct their identities within space by eliciting reports on their living conditions, position in society and aspirations. With this combination of sociolinguistic and discursive methodologies, we aim at making residents' thoughts, emotions and sense of belonging visible, and revealing patterns in the urban dynamics of Hong Kong and Moscow. Taking a comparative approach, the results of this study will also shed light on the interaction between local and global processes that shape language and space today.

References

Bacon-Shone, John, Kingsley Bolton & Kang Kwong Luke. 2015. *Language Use, Proficiency and Attitudes in Hong Kong*. Hong Kong: Social Sciences Research Centre, University of Hong Kong.

Baranova, Vlada V., Yuri B. Koryakov & Yulia V. Mazurova (eds.). 2018. *Acta Linguistica Petropolitana* 14(3). St. Petersburg: Institute for Linguistic Studies, Russian Academy of Sciences.

Bolton, Kinglsey. 2000. The sociolinguistics of Hong Kong and the space for Hong Kong English. *World Englishes* 19(3). 265-285.

Kitajgorodskaja, Margarita V. & Nina N. Rozanova. 2010. Jazykovoe suščestvovanie sovremennogo gorožanina. Moskva: Na materiale Moskvy.

Saunders, Doug. 2011. *Arrival city: How the largest migration in history is reshaping our world.* London: Windmill Books.

Identity conceptions, linguistic repertoires, and language attitudes and use in multilingual Singapore: Developments and recent changes

Sarah Buschfeld (Dortmund), Jakob Leimgruber (Freiburg), Peter Siemund (Hamburg)

Singapore is one of the foremost examples of a multiethnic, multilingual metropolis. Over the years, English has developed a special status in Singapore, though it also coexists with a variety of languages (particularly Malay, Chinese, and Indian languages). Growing numbers of children and young people in Singapore now acquire English from birth, usually in combination with one or more other languages of the country's (and their families') linguistic repertoires.
The present paper investigates the effects of multilingualism on young people's linguistic repertoires and identities, as well as on the development of language attitudes and use. The participants are young Singaporeans of different ethnicities (Chinese, Malay, and Indian), all of whom can be categorized as L1 speakers of English. Participants were divided into three age groups: (1) preschool children (1;4 to 6;9, n = 29); (2) schoolchildren (7;0 to 12;1, n = 8) (Buschfeld fc.); (3) post-secondary students between the ages of 17 and 24 (n = 450) (Leimgruber et al. 2018, Siemund et al. 2014). Data on the participants' linguistic repertoires and language attitudes and use were collected by means of questionnaires.

The results suggest similarities between the groups when it comes to their linguistic repertoires and language attitudes and use (e.g., bilingualism from birth as the predominant acquisition pattern), but also some differences between ethnicities (e.g., differences in the frequency of English usage) and age groups. This is true even within the family and peer groups, where language use is highly correlated with age and shared repertoire, suggesting a gradual apparent-time shift towards English, with non-official varieties relegated largely to use with the older generation.

These results are complemented and largely confirmed by findings on the use of selected linguistic features by the preschool and schoolchildren: realization of subject pronouns, past-tense marking, and vowel quality and quantity (Buschfeld fc.). Alongside prominent quantitative differences between Chinese and Indians, the data also show some homogenization effects, especially at the qualitative level. For example, Indians tend to use the Chinese-derived past-tense marker (verb + finish) the same way their Chinese peers do, even though this structure, not being part of their linguistic repertoires, cannot be transfer-induced.

Overall, our findings suggest that in multilingual and multiethnic Singapore, we find tendencies of both divergence and convergence between the different ethnic groups that can be explained in light of both identity conceptions (ethnic dissociation and consociation due to spatial proximity) and the decades-long evolution and spread of English at the expense of other languages, evidenced in shifting repertoires.

References

Buschfeld, Sarah. Fc. *Children's English in Singapore: Acquisition, Properties, and Use*. Routledge.

Leimgruber, Jakob R. E., Peter Siemund & Laura Terassa. 2018. Singaporean students' language repertoires and attitudes revisited. *World Englishes* 37(2). 282-306.

Siemund, Peter, Monika Schulz & Martin Schweinberger. 2014. Studying the linguistic ecology of Singapore: A comparison of college and university students. *World Englishes* 33(3). 340-362.

Avenues for approaching the urban – the case of Brooklyn, New York

Beatrix Busse, Kristin Berberich, Jennifer Smith (Heidelberg)

Meaning-making processes in large cities pose new analytical challenges not only in terms of their size and density, but also in terms of the many interconnected and simultaneous actions occurring in multiple inter-semiotic forms (Busse & Warnke 2014: 522). In a corpus-based analysis of place-making in several neighborhoods in Brooklyn, New York, we discuss important issues in approaching urban place-making methodologically, highlighting the role of triangulation (Taylor & Marchi 2018) for the study of variation in the urban space. Further, we address the interdisciplinary nature of urban studies and emphasize points of overlap between linguistics, human geography and sociology and discuss possibilities of enhancing the linguistic methodological toolkit for the analysis of cities more broadly. Following Busse & Warnke's (2014) urbanity model, this paper argues for the collection of heterogeneous and novel data sets to analyze language variation as part of multi-semiotic urban place-making processes (Busse 2019). Using data collected through fieldwork, we show how identity formation and negotiation in Brooklyn, New York, contributes to the creation of place(s) in the borough. Drawing on variationist concepts such as enregisterment (Busse 2019, Johnstone 2017), we show how social actors use discursive patterns as a way of signaling belonging in urban spaces (Modan 2007, cf. also Cornips & de Rooij 2018) by combining methods from variational sociolinguistics, linguistic anthropology and discourse analysis.

References

Busse, Beatrix. Fc. 2019. Patterns of discursive urban place-making in Brooklyn, New York. In Viola Wiegand & Michaela Mahlberg (eds.), *Corpus Linguistics, Context and Culture*. Berlin: De Gruyter Mouton.

Busse, Beatrix & Ingo H. Warnke (eds.). 2014. *Place-Making in urbanen Diskursen*. Berlin: De Gruyter Mouton.

Cornips, Leonie & Vincent de Rooij (eds.). 2018. *The Sociolinguistics of Place and Belonging: Perspectives from the margins*. Amsterdam, Philadelphia: John Benjamins.

Johnstone, Barbara. 2017. Characterological Figures and Expressive Style in the Enregisterment of Linguistic Variety. In Chris Montgomery & Emma Moore (eds.), *Language and a Sense of Place: Studies in Language and Region*, 283-300. Cambridge: Cambridge University Press.

Modan, Gabriella G. 2007. *Turf wars: Discourse, diversity, and the politics of place*. Oxford: Blackwell.

Smakman, Dick & Patrick Heinrich (eds.). 2018. *Urban Sociolinguistics: The City as a Linguistic Process and Experience*. London: Routledge.

Taylor, Charlotte & Anna Marchi (eds.). 2018. *Corpus approaches to discourse: A critical review.* London: Routledge.

The sociolinguistic economy of Berlin

Theresa Heyd (Greifswald), Britta Schneider (Frankfurt an der Oder)

In this paper, we report on Anglophone practices in Berlin (Heyd & Schneider in press) and the various ways in which they are connected to and intersect with the larger sociolinguistic economy of Berlin (Heyd, Schneider & von Mengden in press). The history of English in Berlin is on the one hand specific and highly localized, through factors such as the presence of the Allied Forces in post-WWII Berlin and the emergence of an Anglophone creative class in the 1970s. On the other hand, present-day Berlin lends itself to comparison with other cities perceived as transnational, drawing in diverse global workforces ranging from highly mobile elites and creative industry members to precariously employed service workers, as well as refugees and asylum seekers.

Based on data from very different settings, such as the Anglophone New African Diaspora in Berlin on the one hand, and English-coded elite practices of consumption such as Third Wave Coffee and Craft Beer on the other hand, we emphasize that English – and other forms of linguistic practice in Berlin – is first and foremost to be understood through an appreciation of the sociolinguistic economy that shapes the city. Thus, the perspective that we propose here reveals diversity, translocal ties and inter-ethnic relationships that go well beyond established categories of diasporic pockets or ethnic speech communities. The role of contemporary discourses on the instrumental or social value of languages also comes to the fore. We see that the way linguistic resources can be commodified – or not – impacts on their possibility to become visible and on the way their presence is discursively negotiated. Thus, Heller's and Duchêne's work, showing that value is no longer only attributed to languages on grounds of expressing national/ethnic pride and solidarity but can be based on their potential to gain economic profit (Heller & Duchêne 2012), is relevant to interpret the translocal indexical meanings of English and other languages in Berlin.

The importance of economic ideologies in framing indexical values, as it emerges from our data, leads to a critical questioning of the 'city' as an a priori frame of analysis for linguistic variation and change. We argue that traditional assumptions about language and the city – such as its inherent cosmopolitanism and role as a platform for contact, but also the notion of "urban exceptionalism" that posits an essential difference between urban language use and assumed non-urban contexts – need to be reexamined for a better understanding of language in society in the 21st century.

References

Heller, Monica & Alexandre Duchêne. 2012. Pride and profit: changing discourses of language, capital and nation-state. In Monica Heller & Alexandre Duchêne (eds.), *Language and late capitalism*, 1-20. New York: Routledge.

Heyd, Theresa & Britta Schneider. (in press). Anglophone practices in Berlin: from historical evidence to transnational communities. In Ray Hickey (ed.), *English in the German-speaking world*. Cambridge: Cambridge University Press.

Heyd, Theresa, Britta Schneider & Ferdinand von Mengden (eds.). (in press.) *The sociolinguistic economy of Berlin*. Berlin: de Gruyter.

Multilingualism and the role of English in the United Arab Emirates, with views from Singapore and Hong Kong

Peter Siemund (Hamburg), Ahmad Al-Issa (Sharjah), Sharareh Rahbari (Hamburg), Jakob Leimgruber (Freiburg)

A former British protectorate, the UAE has witnessed unprecedented economic and cultural development since its foundation in 1971. Foreign labor and investment play a central role in this development, yielding a sharp imbalance between the local Arab and the foreign population. A share of no less than 85% of highly transient foreigners strongly impacts the local linguistic texture, with many languages competing in the public sphere, creating a highly multilingual landscape. English occupies a special and highly prominent role in this texture, as it is used as a foreign language, a second language, and as a lingua franca. Moreover, it occurs in its standardized varieties (British or American English), but crucially also in several non-standard forms, as foreign labor is recruited from places formerly under British or American influence (India, Pakistan, Philippines, Malaysia, Singapore, Ghana, Kenya, etc.). The UAE thus offers a unique linguistic laboratory in which the interaction and competition of Arabic and English can be studied, in addition to the blending and perhaps levelling of English non-standard varieties. In the present contribution, we present evidence from the ongoing research project *Language Attitudes and Repertoires in the Emirates* (LARES), exploring the tension between English and Arabic, the use of Modern Standard Arabic (MSA) in relation to local forms of Arabic, the prominence of English in specific sectors of society (education, economy, tourism), the increasing use of English as a home language, as well as the presumed emergence of a new variety of English that may tentatively be called 'Gulf English'. We conclude by comparing the situation in the UAE with the similarly high-contact urban areas of Singapore and Hong Kong, city-states that differ in their histories and their local use of English in several ways. Specifically, we argue that while Hong Kong, and Singapore to a lesser degree, remain firmly situated within Kachru's outer circle, the UAE shows signs of moving from the expanding circle inwards towards the outer circle. Moreover, due to the high influx of foreign labor from former British and American colonies (India, Philippines, etc.), the UAE has developed into a melting pot of different varieties of English that we endearingly refer to as 'Milton Keynes II'.

References

Siemund, Peter & Jakob Leimgruber (eds.). (in prep.) *Multilingual Global Cities: Singapore, Hong Kong, and Dubai*. Singapore: Routledge.

English in advertising, signage and graffiti in the Ruhr Metropolis: Occurrences, functions, forms and attitudes

Evelyn Ziegler (Duisburg-Essen)

Occurrences of English in urban public signage have been greatly documented in linguistic-landscape research (Shohamy et al. 2009, Shohamy & Gorter 2010). As most studies deal with issues related to language policy in multilingual societies (including postcolonial contexts) or the emergence of New Englishes, exhaustive empirical studies on the role and use of English in postmodern societies characterized by migration and globalization are missing, especially for Germany.

This paper presents findings from the interdisciplinary research-project *Signs of the Metropolises: Visual multilingualism in the Ruhr Area/Germany*, which investigated the occurrence, regional distribution, function, production and perception of visual multilingualism in representative neighbourhoods of the cities Essen, Dortmund, Bochum and Duisburg (Ziegler et al. 2018). According to the general theme of the conference I will focus on the occurrence and use of English in the Ruhr Metropolis as well as on attitudinal aspects and what they reveal about the prestige of English in the investigated neighbourhoods.

A corpus of 25,504 digital photographs of traffic signs, informative signs (directions, hours of opening), prohibitive signs and warnings, shop signs, signs denoting companies, workshops and businesses, billboards, building names, commemorative signs, artistic signs and graffiti (pieces, tags, stickers) forms the basis of the *Signs-of-the-Metropolises-Database.* The following questions will be addressed:

1. What is the overall visibility of English and how does it interact with other languages?
2. In which types of discourse is English most present (commercial, regulatory, infrastructural, transgressive, artistic and commemorative, cf. Scollon & Scollon 2003)?
3. What forms of English are striking with regard to orthography, grammar, word formation, wordplay and translations?

The results reveal that almost 20% of all language occurrences in the linguistic landscape of the Ruhr Metropolis are English. Whereas English is very present in the commercial discourse and in graffiti it is less visible on institutional signs, which are dominantly German. 92% of these signs are monolingual German signs, only 8% are bilingual signs of which 90% include English.

Furthermore, I will present results from a corpus of 120 on-site interviews which were carried out in a conversational manner in the selected neighbourhoods. The interviews were conducted with 65 men and 55 women between 18 and 80 years of age. The results reveal that English is the second most often perceived language (after Turkish) and that it enjoys a very high status. When asked which languages should appear alongside German on signs in public institutions, more than 40% of informants chose English.

References

Scollon, Ron & Suzie Wong Scollon. 2003. *Discourses in Place: Language in the Material World*. London: Routledge.

Shohamy, Elana, Eliezer Ben-Rafael & Monica Barni (eds.). 2010*. Linguistic Landscape in the City*. Bristol, Buffalo, Toronto: Multilingual Matters.

Shohamy, Elana & Durk Gorter. 2009. *Linguistic Landscape. Expanding the Scenery*. London: Routledge.

Ziegler, Evelyn, Heinz Eickmans, Ulrich Schmitz, Haci-Halil Uslucan, David H. Gehne, Sebastian Kurtenbach, Tirza Mühlan-Meyer & Irmi Wachendorff. 2018. *Metropolenzeichen: Atlas zur visuellen Mehrsprachigkeit der Metropole Ruhr*. Duisburg: UVRR.

BICLCE 2019 Workshop

New Horizons in Cohesion Studies: Widening the Angle

Christina Sanchez-Stockhammer (LMU Munich) and Christoph Schubert (University of Vechta)

Program

(1) Cohesion Studies: Status Quo and Quo Vadis
Christoph Schubert (University of Vechta)

(2) Cohesion through Word Formation: Exploring the Level of Sublexical Cohesive Ties
Christina Sanchez-Stockhammer (LMU Munich)

(3) Where to Draw the Line when the Lines are Blurred: A Computational and Functional Analysis of Cohesion in Comics
Paul Gahman (University of Erlangen-Nuremberg)

(4) Multimodal Cohesion in Online Newspaper Articles about Migration
Helene Schmolz (University of Passau)

(5) Lexical Hypercohesion on YouTube
Sanna-Kaisa Tanskanen (University of Helsinki)

(6) In the Ear of the Beholder? Genre, Framing and Cohesion in Film Trailers
Christian R. Hoffmann (University of Augsburg)

(7) Multimodal Cohesion, Narrative Interpretation and Empirical Foundations
Chiao-I Tseng (University of Bremen)

Abstracts

Cohesion through Word Formation: Exploring the Level of Sublexical Cohesive Ties

Christina Sanchez-Stockhammer (LMU Munich)

In their seminal account of cohesion, Halliday and Hasan (1976: 7) describe various types of semantic "text-forming relations which cannot be accounted for in terms of constituent structure". Their subtypes of cohesion are situated

1) on the level of grammar (reference, substitution, ellipsis),
2) on the level of the lexicon (reiteration and collocation) and
3) on the intermediate level of conjunction.

What Halliday and Hasan (1976) do not discuss at all, by contrast, is the question whether cohesive ties can also be recognised below the level of words. Since their original definition of cohesion does not contradict this possibility, the present paper discusses the potential of the sublexical level as a possible complementation of Halliday and Hasan's (1976) classical model.

The paper explores how word formation processes and the morpho-semantic units they combine can bring about cohesion in texts.

The past decades have seen research on multiple aspects of cohesion, but this topic is still surprisingly unexplored, with few studies e.g. by Dederding, Dressler, Kastovsky, Lipka and Wildgen dating mainly from the 1980s.

The present contribution goes beyond previous research by providing a systematic and comprehensive discussion of word formation and morphology in the light of Halliday and Hasan's (1976) model of cohesion. We will first determine whether it is possible to find equivalents for Halliday and Hasan's original cohesion categories on the sublexical level and then discuss to what extent these sublexical cohesion categories are restricted by level-specific aspects like part of speech, word formation type (e.g. compounding and affixation), headedness and spelling (e.g. with a space or a hyphen; cf. also Sanchez-Stockhammer 2018). Finally, we will explore whether the characteristics of the suggested new level of cohesion (like the partial recurrence described e.g. by de Beaugrande & Dressler 1981) give rise to any completely new types of cohesive tie.

References

Beaugrande, Robert-Alain de & Wolfgang Dressler. 1981. *Introduction to text linguistics*. London: Longman.

Dederding, Hans-Martin. 1983. Wortbildung und Text: Zur Textfunktion (TF) von Nominalkomposita (NK). *Zeitschrift für germanistische Linguistik* 11. 49-64.

Dressler, Wolfgang. 1979. Zum Verhältnis von Wortbildung und Texlinguistik. In János Petöfi (ed.), *Text versus sentence: Basic questions in text linguistics*, 96-106. Hamburg: Buske.

Halliday, Michael A.K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Routledge.

Kastovsky, Dieter. 1981. Word-formation: A functional view. *Folia Linguistica* 16(1-4). 181-198.

Lipka, Leonhard. 1987. Word-formation and text in English and German. In Brigitte Asbach-Schnitker & Johannes Roggenhofer (eds.), *Neuere Forschungen zur Wortbildung und Historiographie der Linguistik. Festgabe für Herbert E. Brekle zum 50. Geburtstag*, 59-67. Tübingen: Narr.

Sanchez-Stockhammer, Christina. 2018. *English compounds and their spelling*. Cambridge. Cambridge University Press.

Wildgen, Wolfgang. 1981. Makroprozesse bei der Verwendung nominaler ad hoc-Komposita im Deutschen. *Folia Linguistica* 16(1-4). 297-344.

Where to Draw the Line when the Lines are Blurred: A Computational and Functional Analysis of Cohesion in Comics

Paul Gahman (University of Erlangen-Nuremberg)

Thanks in part to Halliday and Hasan's (1976) explication of cohesion in text, linguists have long established the relation between cohesive devices employed, text type and register (Brooks and Warren 1972; Louwerse, McCarthy, McNamara & Graesser 2004). In the past few decades, the genre of comics has become the focus of research from both semiotic (Kress & van Leeuwen 1996; Barthes 1997; Groensteen 2007; Kukkonen 2011) and cognitive linguistic perspectives (Eisner 1996; Cohn 2013). However, very little research has been devoted to cohesion in the multimodal context of comics.

The present research addresses this deficit by examining text in English graphic novels and superhero comics. Upon separating spoken from written text found in and outside of speech bubbles, Coh-Metrix is used to establish the texts' cohesive and semantic attributes (e.g. lexical density, connectives, type-token ratio, syntactic complexity, etc.). The most common cohesive devices and functions have been collated to present a lexical repertoire emblematic of comics. Passages with low inter-sentential cohesion have undergone semantic and discourse analyses to determine the causes of breaks in cohesion. Further, semiotic analyses use the images with which the incohesive text appears in order to account for any cohesion deficiencies. These analyses facilitate identifying systematicity in cohesive breaks and may provide insight into whether the text or image is more pivotal in maintaining cohesion in comics. Finally, while it is assumed that images may fill cohesive gaps in either the spoken or written text, onomatopoeia, due to its iconic character, may also act as a cohesive tie in comics. The paper therefore also investigates the role onomatopoeia plays in maintaining cohesion amongst panels.

References

Barthes, Roland. 1977. The rhetoric of the image. In Roland Barthes. *Image, music, text*, 32-51. Trans. Stephen Heath. New York: Hill and Wang.

Brooks, Cleanth & Robert P. Warren. 1972. *Modern rhetoric*. New York: Harcourt Brace Jovanovich.

Cohn, Neil. 2013. *The visual language of comics: Introduction to the structure and cognition of sequential images*. London: Bloomsbury.

Eisner, Will. 1996. *Graphic storytelling and visual narrative*. New York/London: W.W. Norton & Company.

Graesser, Arthur C., Danielle S. McNamara, Max M. Louwerse & Zhiqiang Cai. 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers* 36(2). 193-202.

Groensteen, Thierry. 2007. *The system of comics*. Jackson: University Press of Mississippi.

Halliday, Michael A.K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Kress, Gunther & Theo van Leeuwen. 1996. *Reading images: The grammar of visual design*. London: Routledge.

Kukkonen, Karin. 2011. Comics as a test case for transmedial narratology. *SubStance: Graphic Narratives and Narrative Theory* 40(1). 34-52.

Louwerse, Max M., Philip M. McCarthy, Danielle S. McNamara & Arthur C. Graesser. 2004. Variation in language and cohesion across written and spoken registers. *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. 843-848.

Multimodal Cohesion in Online Newspaper Articles about Migration

Helene Schmolz (University of Passau)

Research in multimodality has grown over the past few years, surely also due to visual elements becoming more important in our society (‘pictorial turn’) and due to digital resources blending previously separated media (e.g. videos and audio files are now commonly embedded in news webpages). Thus, multimodality has also become an interesting field for linguists. However, multimodal critical discourse analysis as well as research combining multimodality, (critical) discourse analysis and corpus linguistics are still rare (Bednarek & Caple 2017; Yalaz & Zapata-Barrero 2018). At the same time, cohesion, i.e. grammatical and lexical links, has played a minor role in the analysis of multimodal discourse, although semantic links, i.e. coherence, have been addressed by many researchers, based mostly on Halliday and Hasan (e.g. Liu and O’Halloran 2009; Subrayan Michael et al. 2013; Crane 2016). Research on cohesion investigating a larger multimodal corpus of newspapers is still missing.

Based on contributions from Machin & Mayr (2012), Caple (2013), and Bateman (2014), the paper intends to examine cohesion in online newspapers articles about migration. In particular, it investigates grammatical and lexical links between image(s) and text (headline, caption, body of an article). To do so, a corpus of online newspaper articles dealing with migration in a wide sense (e.g. immigration, emigration, refugees) has been compiled. The corpus consists of quality newspapers from three English-speaking countries (UK, USA and Australia). It comprises about half a million words and includes over 1,300 images.

The paper will discuss the types of cohesive links occurring most frequently as well as the implications these links carry for the representation of migrants. Additionally, corpus linguistic and qualitative findings will be used to embed multimodal cohesion occurring in these articles in a context.

References

Bateman, John. 2014. *Text and image. A critical introduction to the visual/verbal divide*. New York: Routledge.

Bednarek, Monika & Helen Caple. 2017. *The discourse of news values. How news organisations create newsworthiness*. Oxford: Oxford University Press.

Caple, Helen. 2013. *Photojournalism: A social semiotic approach*. Basingstoke/New York: Palgrave Macmillan.

Crane, Paul A. 2016. Texture in text: A discourse analysis of a news article using Halliday and Hasan's model of cohesion. http://lingueletteratureculturestraniere.uniroma3.it/bacheca/llopriore/wp-content/uploads/sites/38/2016/12/CRANE-on-Cohesion.pdf (30 January 2019).

Liu, Yu & Kay L. O'Halloran. 2009. Intersemiotic texture: Analyzing cohesive devices between language and images. *Social Semiotics* 19(4). 367-388.

Machin, David & Andrea Mayr. 2012. *How to do Critical Discourse Analysis: A multimodal introduction*. Los Angeles: Sage.

Subrayan Michael, Angelina, Chittra Muthusamy, Puspalata C Suppiah, Caroline Joseph & Sulia Masturina binti Che Razali. 2013. Cohesion in news articles: A discourse analysis approach. *International Journal of Applied Linguistics & English Literature* 2(3). 129-133.

Yalaz, Evren & Ricard Zapata-Barrero. 2018. Mapping the qualitative migration research in Europe: An exploratory analysis. In Ricard Zapata-Barrero & Evren Yalaz (eds.), *Qualitative research in European migration*, 9-31. Cham: Springer.

Lexical Hypercohesion on YouTube

Sanna-Kaisa Tanskanen (University of Helsinki)

This paper explores lexical cohesion on a YouTube channel. The aim is to find out how lexical cohesive devices establish links between the video clips and the comments written by users as well as between the comments themselves. Following Schubert (2017), I call cohesion occurring in a multimodal YouTube context *hypercohesion*. This term draws attention to the fact that cohesion in social media exists not only within and between verbal contributions but also across modes.

The YouTube channel analysed in the present study is *The Hydraulic Press Channel* (HPC), which publishes videos of various objects being crushed in a hydraulic press. The objects range from Lego toys and gummy bears to a bowling ball and a Nokia 3310. Created and operated by a Finnish couple, Lauri and Anni Vuohensilta, the channel functions in English: all videos include commentary in English by Lauri Vuohensilta in his very distinctive Finnish accent, and most of the comments are also in English. The channel has been in existence since 2015, and currently it has over two million subscribers and around 300 million total views (Wikipedia, HPC).

The paper argues that lexical cohesive devices used in the comments contribute to coherence by establishing links between the comments themselves and between the video clip and the comments. Both reiteration and collocation relations are analysed, with the former dominating the hypercohesive profile of the multimodal platform. Cohesive ties mostly exist between the clip and the individual comments, as users voice their reactions to what they have seen. At the same time, however, links are created between the comments themselves, for instance through repetitions. Occasionally users create a more dialogical discussion within their polylogue, directly addressing or answering another user. Special attention is paid to cohesive ties between such comments. The findings from the analysis are compared with earlier research (e.g. Hoffmann 2012 on weblogs and Tanskanen 2006 on mailing-list discussions).

References

Hoffmann, Christian R. 2012. *Cohesive profiling: Meaning and interaction in personal weblogs*. Amsterdam & Philadelphia: John Benjamins.

Schubert, Christoph. 2017. Discourse and cohesion. In Christian R. Hoffmann & Wolfram Bublitz (eds.), *Handbooks of pragmatics (HOPS),* Vol. 11: *Pragmatics of social media*, 317-344. Berlin: De Gruyter Mouton.

Tanskanen, Sanna-Kaisa. 2006. *Collaborating towards coherence: Lexical cohesion in English discourse*. Amsterdam & Philadelphia: John Benjamins.

In the Ear of the Beholder? Genre, Framing and Cohesion in Film Trailers

Christian R. Hoffmann (University of Augsburg)

This talk explores the use of recurring patterns of audio-visual cohesion in a specialised genre-balanced corpus of contemporary US-American film trailers. I apply various corpus linguistics tools and methods, e.g. concordancers, GraphColl, keyword analysis, semantic and syntactic taggers (cf. Flowerdew & Mahlberg 2009; Duguid & Partington 2017) to inform more recent analytical frameworks mapping film discourse cohesion in telecinematic discourse (Janney 2012; Tseng 2013; Bateman, Wildfeuer & Hiippala 2017; Schubert 2017; Hoffmann & Kirner-Ludwig forthcoming). More generally, I hope to find some answers to the following two research questions:

(a) How can we use current corpus linguistics tools to best detect *verbal* cohesive patterns (grammatical and lexical cohesion; cf. Halliday & Hasan 1976) within and across larger sets of film trailers?
(b) How can we code and elicit *visual* cohesive patterns in different film trailer genres (e.g. by means of framing, camera movement, mise-en-scene, transitions), and how do these means cluster in expectable patterns around meaningful narrative segments of the trailer?

Drawing on software tools such as Wmatrix, TAACO or Lancsbox, I will first track the relative frequencies and dispersion of (grammatical and lexical) cohesive cues across the trailer corpus. Particular emphasis will be put on the (semantic) categorization of recurring noun and verb phrases and the genre-specific use of collocations. In a second phase, a small number of cinematographic techniques will be coded for the orientation stages of each trailer. The results uncover recurring patterns of the visual techniques, i.e. visual collocations, that directors and editors routinely apply to yield strategic effects in film trailers.

References

Bateman, John, Janina Wildfeuer & Tuomo Hiippala. 2017. *Multimodality: Foundations, research and analysis – a problem-oriented introduction*. Berlin: Walter de Gruyter.

Duguid, Alison & Alan Scott Partington. 2017. Forced lexical primings in transdiscoursive political messaging: How they are produced and how they are received. In Michael Pace-Sigge & Katie J. Patterson (eds.), *Lexical priming: Applications and advances*, 67-92. Amsterdam: John Benjamins.

Flowerdew, John & Michaela Mahlberg (eds.). 2009. *Lexical cohesion and corpus linguistics.* Amsterdam: John Benjamins.

Hoffmann, Christian & Monika Kirner-Ludwig (eds.). Forthcoming. *Telecinematic stylistics*. London: Bloomsbury.
Janney, Richard. 2012. Pragmatics and cinematic discourse. *Lodz Papers in Pragmatics* 8(1). 85-113.
Schubert, Christoph. 2017. Discourse and cohesion. In Christian R. Hoffmann & Wolfram Bublitz (eds.), *Handbooks of Pragmatics (HOPS)*. Vol. 11: *Pragmatics of social media*, 317-344. Berlin: De Gruyter Mouton.
Tseng, Chiao-I. 2013. *Cohesion in film: Tracking film elements*. Houndmills: Palgrave Macmillan.

Multimodal Cohesion, Narrative Interpretation and Empirical Foundations

Chiao-I Tseng (University of Bremen)

This presentation will show how the analytical method of multimodal cohesion is an effective tool for empirically unpacking several significant issues with regard to narrative comprehension and stylistic interpretation in multimodal texts such as film and comics. Multimodal cohesion is a linguistics-informed framework for narrative construction in audiovisual and graphic media (Tseng 2013; Tseng & Bateman 2018; Tseng et al. 2018). It explains how characters, objects and settings are identified and tracked in verbal, visual and audio modalities and what technical details of filmic and graphic composition afford inter-modal tracking.

This talk will first present how to systematically construct cohesive structures of the main identities of narrative characters and settings and how these structures reflect the film viewers' and comic readers' narrative interpretation paths. I will then present two studies, showing the integration of systematic analysis of multimodal cohesion and empirical methods. This combination balances the two main research streams of narrative interpretation: 1. the recipient-based question of how the direction of viewers' attention affects comprehension and 2. the narrative-based issue of how narrative elements in film and comic guide meaning interpretation.

The first study combines cohesion analysis in film and the method of eye-tracking. Based on an analysis of the beginning of Hitchcock's *The Birds* (1963) which combines semiotic and empirical methods, I will address how specific and generic cohesive cues play significant roles in directing viewers' narrative understanding.

The second study unravels cohesion in comics. Using the example of the comic adaptation of Paul Auster's *City of Glass* (1994), I will present how the character-centred cohesive patterns are the main trigger for character immersion in the genre of metafiction. The results of an eye-tracking study support this hypothesis.

In summary, the synthesis of multimodal cohesion analysis with an eye-tracking approach provides powerful empirical evidence for the hypothesis that even if the possible high-levelled narrative interpretations may at first appear boundless, textual artefacts are, in fact, often designed to guide their receivers in specific, intended directions.

References

Tseng, Chiao-I. 2013. *Cohesion in film: Tracking film elements*. Basingstoke: Palgrave Macmillan.
Tseng, Chiao-I & John A. Bateman. 2018. Cohesion in comics and graphic novels: An empirical comparative approach to transmedia adaptation in *City of Glass*. *Adaptation* 11(2). 122-143.
Tseng, Chiao-I, Jochen Laubrock & Jana Pflaeging. 2018. Character developments in comics and graphic novels: A systematic analytical scheme. In Alexander Dunst, Jochen Laubrock & Janina Wildfeuer (eds.), *Empirical comics research: Digital, multimodal, and cognitive methods*, 154-175. London: Routledge.

Language attitude research among mobile speakers from anglophone Africa

Johanna Hartmann (University of Münster)

Mobile speakers of English – particularly those from anglophone Africa – have been covered in relatively few studies in language attitude research (e.g. Githiora 2008; Mutonya 2009). Given that mobility is an essential part of present-day realities and the "sociolinguistics of globalization is perforce a sociolinguistics of mobility" (Blommaert 2010: 28), it would be desirable to increase research interest in such contexts. It is, however, necessary to adapt research designs and methods to the specific demands of migration contexts. It has for instance been pointed out that attitude research among African migrants should turn away from a focus on the former colonizers' languages and varieties (Githiora 2008: 19) and that evaluative adjectives should be adapted to the specific context (Mutonya 2009).

This paper uses experiences from a language attitude study among educated Kenyans and Nigerians in Germany and the UK to address the following questions:

1. Which challenges does a focus on mobile speakers from anglophone Africa pose in language attitude research?
2. Which methods and approaches might be helpful in handling these challenges?

In the present study, an indirect and a direct approach were employed to elicit language attitudes: A verbal guise study with authentic vocal stimuli representing varieties of English from seven countries was combined with semi-structured interviews. Evaluative adjectives for the verbal guise study were elicited in a pilot study.

A mixed-methods approach is generally considered useful in order to obtain comprehensive results (e.g. McKenzie 2010: 52), but it turned out to be of unexpected value for this project because more emphasis could be put on the interviews when it became clear that the research participants were more comfortable with personal interviews than survey-based research. A research design that offers such flexibility is particularly important in migration research (Bose 2012: 288). In-depth interviews also offer researchers an opportunity to establish the necessary trust with the research participants. This is essential in dealing with migrants who might be afraid of giving away sensitive information or of being misrepresented in scientific publications. Furthermore, mobile speakers are not necessarily organized in fixed groups, which makes it difficult to recruit participants for quantitative research. Qualitative research approaches can serve as a door opener to gain access to further participants. In order to successfully analyse language attitudes in mobility contexts, a combination of qualitative and quantitative methods and a flexible research design therefore seem to be called for.

References

Blommaert, Jan. 2010. *The sociolinguistics of globalization.* Cambridge: Cambridge University Press.

Bose, Pablo S. 2012. Mapping movements: Interdisciplinary approaches to migration research. In Carlos Vargas-Silva (ed.), *Handbook of Research Methods in Migration,* 273-294, Cheltenham: Edward Elgar Publishing.

Githiora, Christopher. 2008. *African Immigrants' Attitudes toward African American Language English.* Saarbrücken: VDM Verlag.

McKenzie, Robert. 2010. *The social psychology of English as a global language: Attitudes, awareness and identity in the Japanese context.* Dordrecht: Springer.

Mutonya, Mungai. 2009. *Language attitudes towards varieties of African English.* Saarbrücken: VDM Verlag.

Mapping perceptions in New Englishes

Mirjam Schmalz (University of Zurich)

Language attitude research in New Englishes often fails to employ mixed methods, but rather relies heavily on vocal stimuli. This paper will therefore present an alternative method for the field, which has traditionally been used in perceptual dialectology studies in the US and the UK (see e.g. Montgomery 2015) but is almost always neglected in perceptual studies on New Englishes; namely the mapping of perceptual landscapes. This perceptual mapping is achieved through map drawing tasks undertaken by informants, which are then visualised using GIS tools. In the present paper, this will be illustrated on the basis of first-hand data collected in St. Kitts in 2018 and 2019.

St. Kitts forms part of the island state of St. Kitts and Nevis and with approximately 40,000 inhabitants, it belongs to the smaller islands of the Lesser Antilles of the Caribbean. Within the wider region, it was the first island to be settled by the English in 1623 and was used to explore and settle other islands of the surrounding area (Parkvall 2000: 123). Up to the present day, mobility strongly shapes the island’s population and their language practices, as many Kittitians go abroad for work, tertiary education or because of family ties that connect them with other Caribbean islands or the mainland. Moreover, there is also a considerable number of incoming people, due to the thriving tourist industry and the work this provides on the island. Thus, St. Kitts and its speakers lend themselves well to the study of perceptions towards both local language varieties and overseas varieties of English.

With the data from this island the present paper focuses on the perceptions of Kittitians and attempts to perceptually map parts of the Caribbean from the viewpoint of one of the smaller islands of the wider region. Using ArcGIS to visualise the perceived linguistic variation, a first perceptual map of the area will be presented. Furthermore, data gathered in sociolinguistic interviews and the judging of vocal stimuli will be added to the analysis. This paper consequently proposes that the geographical mapping of perceptions, in combination with other methods, will provide us with a better and more holistic picture of the perceptions and attitudes present in New Englishes speech communities.

References

Montgomery, Chris. 2015. Perceptions of borders and boundaries in the North of England. In: Hickey, Raymond (ed.), *Researching Northern English*, 345-369. Amsterdam: Benjamins.

Parkvall, Mikael. 2000. *Out of Africa: African influences in Atlantic Creoles*. London: Battlebridge Publications.

Research methods in the Indian diaspora: An investigation into folk scrutiny and accent variation

Giuliana Regnoli (University of Naples 'L'Orientale' & University of Heidelberg)

Recent research on the Indian diaspora has raised the question of the extent to which linguistic differences inform our understanding of community structures, attitudinal orientations and identity development (Sharma 2017; Hundt & Sharma 2014). However, little attention has been given to research methods aiming to investigate folk perceptions regarding accent variation in Indian transient communities.

The present work intends to focus on the relationship between language attitudes and folk awareness in a multilingual transient speech community of Indian university students located in Heidelberg, Germany. Drawing on qualitative and quantitative data collected through sociolinguistic interviews, surveys and accent identification tasks, the paper will report on the importance of relying on experimental research designs and mixed-method approaches in order to shed light on the highly complex sociolinguistic situation of transient communities. As a matter of fact, in a community where no previous investigation had ever been carried out - except for the author's master thesis research (Regnoli 2016) - a quantitative or qualitative approach alone seemed rather unrealistic. For this reason, the study used a convergent parallel design (Creswell and Plano Clark 2011) drawing on perceptual dialectology, linguistic anthropology and sociolinguistic research methods in order to test whether students' folk perceptions regarding accent variation had a common ground with analysed linguistic data (Niedzielski & Preston 2003; Duranti 1997).

Reported results suggest that methodology triangulation may be an effective paradigm in folk linguistic and attitude studies for its valid, replicable and reliable results. As far as the present study is concerned, network integration played a crucial role on attitudinal orientation and identity development. Moreover, the transient aspect of the community proved to be a valuable sociolinguistic factor in the fostering of in-group affiliations and distance and that speakers' constant need to renegotiate identity is indexed in their willingness (or not) to shed their regional, linguistic and ethnic identities in deference to their pan-Indian ones depending on the community's fluid sociolinguistic circumstances.

References

Creswell, John & Vicki Plano Clark. 2011. *Designing and conducting mixed methods research*. London: Sage.

Duranti, Alessandro. 1997. *Linguistic anthropology*. Cambridge: Cambridge University Press.

Hundt, Marianne & Devyani Sharma. 2014. Introduction. In Marianne Hundt & Devyani Sharma (eds.), *English in the Indian Diaspora*, 1-8. Amsterdam: Benjamins.

Niedzielski, Nancy & Dennis Preston. 2003. *Folk linguistics*. Berlin: de Gruyter.

Regnoli, Giuliana. 2016. *Indexicality and contextualization: Linguistic, cultural and social stances of Indian English speakers in Heidelberg*. MA dissertation. Dept. of literary, linguistic and comparative studies: University of Naples 'L'Orientale'.

Sharma, Devyani. 2017. World Englishes and sociolinguistic theory. In Markku Filppula, Juhani Klemola & Devyani Sharma (eds.), *The Oxford Handbook of World Englishes*, 232-251. Oxford: Oxford University Press.

Question intonation patterns in educated Nigerian English

Warsa Melles (University of Dortmund) and Folajimi Oyebola (University of Münster)

The aim of this research is to investigate the prosodic marking of questions in Standard Nigerian English. The study focuses on intonation patterns used by educated Nigerian speakers of English when asking 'wh-' questions and 'yes/no' questions. Previous studies on intonation patterns in Nigerian English have suggested that intonation generally reflects the prosodic feature of the speaker's first language (Udofot 1997; Jowitt 2000; Gut 2005). However, there is a lack of systematic investigation of question intonation pattern and its correlation with social factors. In the present study, 101 speakers were selected from the International Corpus of English (ICE) Nigeria based on their sociolinguistic background. Social factors such as gender and ethnicity were considered. The audio recordings were annotated in Praat 5.4.0 (Boersma & Weenink 2015). The Tones and Breaks Indices (ToBI) transcription convention (Silverman et al. 1992) was used for the intonation transcription. Boundary tones (initial & final) were marked, transcribed and classified into fall, rise, fall-rise, rise-fall and level tones. Identified tokens were analysed using a mixed effects logistic regression model with Rbrul (Johnson 2008). Preliminary results show that there are similar question intonation patterns among speakers of Nigerian English, regardless of their ethnic origin or gender. The question type usually influences the final boundary tone- while the 'wh-' questions mostly end with a fall, the 'yes/no' questions end with either a rise or a fall. On the other hand, the initial boundary of the two question types mostly consists of a level tone.

References

Boersma, Paul & David Weenink. 2015. Praat: Doing phonetics by computer (Version 6.0.05) [Computer program]. http://www.praat.org. (26 May, 2019)

Gut, Ulrike. 2005. Nigerian English prosody. *English World Wide* 26(2). 153-177.

Johnson, Daniel Ezra. 2008. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3(1). 359-383.

Jowitt, David. 2000. Patterns of Nigerian English intonation. *English World Wide* 21(1). 63-80.

Silverman, Kim, Mary Beckman, John Pitrelli, Mori Ostendorf, Colin Wightman, Janet Pierrehumbert & Julia Hirschberg. 1992. ToBI: A standard for labelling English prosody. *Proceedings Second International Conference on Spoken Language Processing*, 867-870. Banff: Canada.

Udofot, Inyang. 1997. *The rhythm of spoken Nigerian English.* Doctoral Dissertation. University of Uyo, Nigeria.

Wunder, Eva-Maria, Holger Voorman & Ulrike Gut. 2010. The ICE Nigeria corpus project: Creating an open, rich and accurate corpus. *ICAME Journal* 34. 78–88.

Compiling a “small” corpus: Issues of data and comparability in a corpus of Grenadian English

Ryan Durgasingh (University of Münster)

Chief among the strengths of the ICE suite of corpora is the project’s ongoing emphasis on comparability, both cross-varietally (catering to analyses across national standards) and cross-textually (important to analyses focused on contextually dependent text types). While the geographic spread of national varieties represented in ICE is large, its current lack of any particular regional focus can prove problematic to researchers wishing to explore features which may be sensitive to potential areal properties – not surprising given the continued focus on delineating the corpora according to “L1/ENL varieties or ‘core Englishes’ […] and […] L2/ESL varieties or ‘new Englishes’ […] corresponding to Kachru’s (1985) model of ‘Inner’ and ‘Outer’ Circles of English (Kirk & Nelson 2018: 1). Because of the dearth of material from various small nation states, and a general tendency to treat large, culturally/economically dominant varieties as indicative of their entire regions, it is often difficult to assess whether or not regional varieties may exist in more than a theoretical sense (c.f. Allsopp’s definition of “Standard Caribbean English” (1996: lvi), for example). Clearly, there is a need for more comparable corpora intra-regionally, yet, their compilation poses specific challenges that need to also be addressed. By using the example of the compilation of a “small” (both in size and in terms of the socio-demography of the variety’s location) Grenadian corpus, this paper aims to explore the various issues arising out of spoken data collection and comparability experienced by the present author in order to carry out morphosyntactic analyses of English in the Caribbean using the aforementioned compiled Grenadian data, ICE-JAM, and the as-yet-unreleased ICE-TNT corpora.

The presentation will explore the balancing act between practical concerns (what is feasible given the limitations of a single researcher’s time/resources, and the availability of data in a small nation state), and larger theoretical concerns (how does one gather enough data, given the practical concerns, in order to be representative of a particular variety?). In the case of the latter, the importance of a driving theoretical framework, namely Tannen’s (1982) oral-literate continuum, and the attendant arrangement of text types along this continuum so as to both maintain comparability across varieties, and to be representative of language in typical/important domains of use within the variety, will be discussed.

Also important to the discussion will be the various corpus design specifics of collecting different text types – from exploiting academic/layperson networks, to using social media to gather data. The paper ends with general recommendations of those text types which may be more, or less, difficult for researchers hoping to compile small corpora for comparison with existing ICE corpora.

References

Allsopp, Richard. 1996. *Dictionary of Caribbean English usage*. Oxford: Oxford University Press.

Kirk, John & Gerald Nelson. 2018. The International Corpus of English project: A progress report. *World Englishes* 37(4). 1-20.

Tannen, Deborah. 1982. *Spoken and written language: Exploring orality and literacy*. New Jersey: ABLEX.

ICE corpora, register, and omitted variable bias: A multi-dimensional perspective

Axel Bohmann (University of Freiburg)

The development and continuing growth of the International Corpus of English (ICE) project (Greenbaum & Nelson 1996) has marked a substantial improvement of the empirical base for the description of many New Englishes. Over the past twenty years, this has led to a proliferation of solid, corpus-based studies (e.g. contributions in Seoane & Suárez Gómez 2016; Hundt & Gut 2012). While the importance of the ICE corpora can hardly be overstated, their pioneering role and the amount of research conducted on their basis have turned them into highly normative resources for establishing descriptive facts about individual varieties as well as for drawing cross-varietal comparisons.

The present study takes a critical perspective on this normative status. The focus is specifically on one aspect that has been part of the ICE sampling framework from the outset but has rarely been addressed in explicit statistical modeling: the role of register (Biber & Conrad 2009). The study compares the statistical power of two different extra-linguistic factors in predicting multivariate patterns of linguistic variation in ten ICE corpora: a)the variety a given ICE corpus text represents, and b)the ICE text category label of a text. Results indicate that text category labels are more effective at predicting multidimensional patterns of variation than national variety labels, usually by at least one order of magnitude.

This finding leads to a second problem, namely: omitted variable bias in ICE-based research that does not include register directly in its statistical modeling. In a case study on verb complementation, it is shown that apparent cross-varietal differences diminish in statistical significance once register is systematically accounted for. In other words, cross-varietal differences are heavily mediated by situational and genre context and tend to be exaggerated if register is not directly considered as a predictor of variation.

The results of the study should not be interpreted as a wholesale rejection of ICE-based research so much as a contribution towards methodological refinement. They suggest that, when comparing New Englishes on the basis of ICE data, register relations are an important level of explanation that deserves more explicit attention. Moreover, the multidimensional patterns of variation established on the basis of the ICE corpora can be used to gauge the register differentiation of bigger, messier corpora (e.g. GloWbE; Davies & Fuchs 2015), thus yielding a promising marriage of established, carefully-curated and more recent mega-corpora.

References

Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.

Davies, Mark & Robert Fuchs. 2015. Expanding horizons in the study of World Englishes with the 1.9 billion word Global Web-based English Corpus (GloWbE). *English World-Wide* 36(1). 1-28.

Greenbaum, Sidney & Gerald Nelson. 1996. The International Corpus of English (ICE) project. *World Englishes* 15(1). 3-15.

Hundt, Marianne & Ulrike Gut (eds.). 2012. *Mapping unity and diversity world-wide: Corpus-based studies of New Englishes*. Amsterdam: Benjamins.

Seoane, Elena & Cristina Suárez Gómez (eds.). 2016. *World Englishes: New theoretical and methodological considerations.* Amsterdam: Benjamins.

Converging offline and online ethnographic research: Insights from Belizean and Nigerian data sets

Theresa Heyd (University of Greifswald) and Britta Schneider (University of Frankfurt Oder)

In this presentation, we want to discuss, on the one hand, the relevance of ethnography for establishing a realistic estimation of language practices that avoids a Eurocentric bias in data interpretation, which, in linguistic traditions, tend to conceptualize language as a cognitive, mostly homogenous and bounded entity. This is particularly problematic in socially complex and/or multilingual sociolinguistic settings, to which virtually all settings in which *New Englishes* have emerged belong. In order to understand the concepts of language – in other words, the language ideologies – of speakers, qualitative ethnographic methodology is crucial. Methodological openness can give access unexpected indexical meanings of language use and thus to discourses that, firstly, may be different from Western scientific ideologies of language and that, secondly, may impact on language practice (e.g. in practices that symbolize fixity, belonging, 'correctness' or, on the other hand, creativity or playfulness). Therefore, prolonged field stays, participant observation and the collection of field notes allow for understanding situated behavior and for getting access to insiders' views (see e.g. Geertz 1973; Gobo 2008), which are important to triangulate linguistic data. This also allows for a more critical approach towards the sampling of informants as stays in the field, bringing about diverse social contacts, make the limitations of only documenting language practices of highly educated speakers visible.

At the same time, we argue that ethnographic data collection can no longer be realized as observing offline practices in the field only. Everyday language use cannot be understood as either written and formal, or oral and informal as digital linguistic practice does not exist alongside but converges with and extends situated language use. While linguistic theories so far have been dominantly shaped by an *a priori* assumption of oral face-to-face communication as the 'natural' form of communication, emanating from the cognitive competences of the individual and contrasting with written language (see e.g. Stubbs 1980: 109), these phonocentric language ideologies are generally not tenable in an age of convergent digital practice (see e.g. Blommaert 2013). As converged offline-online linguistic practices have a crucial impact on the constitution of social relationships and on dynamics of mobility, it becomes necessary to include online data into ethnographic observation and to develop new forms of data collection that allow for an understanding of how offline and online language practices interact.

In order to illustrate our claims, we discuss examples from our own *New English* data, among them field notes, interview data, online data and spoken (offline) data from Belize and from the Nigerian digital diaspora.

References

Blommaert, Jan. 2013. Writing as a sociolinguistic object. *Journal of Sociolinguistics* 17(4). 440-459.

Geertz Clifford. 1973. *The interpretation of cultures.* New York: Basic Books.

Gobo Giampietro. 2008. *Doing ethnography.* London: Sage.

Stubbs, Michael. 1980. *Language and literacy: The sociolinguistics of reading and writing*. London: Routledge.