Duncker & Humblot GmbH

Chapter Title: Report on the Workshop on Forensic Linguistics at the ILLA Relaunch Conference in September 2017
Chapter Author(s): Carole E. Chaski, Victoria Guillén Nieto and Dieter A. Stein

Book Title: Legal Linguistics Beyond Borders: Language and Law in a World of Media, Globalisation and Social Conflicts.
Book Subtitle: Relaunching the International Language and Law Association (ILLA).
Book Editor(s): Friedemann Vogel
Published by: Duncker & Humblot GmbH. (2019)
Stable URL: https://www.jstor.org/stable/j.ctv1q69j9f.20

Reports from the ILLA Relaunch Conference 2017 in Freiburg

Report on the Workshop on Forensic Linguistics at the ILLA Relaunch Conference in September 2017

By *Carole E. Chaski*, Georgetown, *Victoria Guillén Nieto*
and *Dieter A. Stein*, Düsseldorf

I. Introduction

At the relaunch of the International Language and Law Association (ILLA) in September 2017, the forensic linguistics workshop harkened back to the first conference associated with ILLA which was held in 2006 in Dusseldorf at Heinrich Heine University. Under the guidance of Prof. Dr. Dieter A. Stein, the 2006 Conference included the broad spectrum of work in language and law, including forensic linguistics and forensic phonetics. We are pleased to report that forensic linguistics, including phonetics, remained a focus within ILLA in the relaunch.

At the same time, the forensic linguistics workshop looked forward to the broadening and strengthening of the "ILE paradigm in forensic linguistics." This paradigm is associated with the Institute for Linguistic Evidence (ILE). Founded in 1998, ILE was the first non-profit organization devoted to research in forensic linguistics. The ILE paradigm developed from Chaski's Visiting Research Fellowship at the United States Department of Justice, National Institute of Justice from 1995 – 1998. At this time in the United States, the Daubert standard requiring known reliability and error rates for scientific evidence was just beginning to affect forensic science techniques. Chaski developed the paradigm as a way for forensic linguistics to meet the new standard for scientific evidence. The ILE paradigm provides principles so that methods in forensic linguistics can meet legal standards for scientific evidence, be admitted for trial testimony and be practiced as safe investigative techniques. A growing body of research has shown that forensic examiners can be cognitively biased toward confirmation of a desired outcome (e. g. Dror (2012), Kassin et al. (2013a, 2013b), Murrie et al. (2013), Dror et al. (2015)). The ILE paradigm protects the forensic examiner from cognitive bias by proposing that methods, and standard operating procedures for applying methods, be developed outside of any litigation on ground truth data so that the method's results cannot be contaminated by any desired outcome when the method is applied following standard operating procedures.

Succinctly, from the legal perspective, scientific evidence must be scientifically respectable and reliable enough to be worthy of the Court's time. The ILE paradigm proposes that forensic linguistic methods can become fully-admissible scientific evidence when they are:

1) developed independent of any litigation, in the course of normal scientific research;

2) tested for accuracy outside of any litigation;

3) tested for accuracy on "ground truth" data;

4) able to work reliably on "forensically feasible" data;

5) tested for known limits correlated to specific accuracy levels;

6) tested for any errors of individual testing techniques that could cause accumulated error when combined with other techniques;

7) replicable when performed by trained analysts;

8) related to a specific expertise and academic training;

9) related to standard ("generally accepted") techniques within the specific expertise and academic training; and

10) related to uses outside of any litigation in industries or fieldwork in the specific expertise (cf. Chaski 1997, 2001, 2013).

In line with current legal standards, the ILE paradigm puts the focus on the method itself, rather than the status or persuasion of the practitioner. By focusing on the calculation of reliability and error rate, the ILE paradigm opens the door for innovations in forensic linguistics and provides linguists of all ranks a chance to develop methods that may become fully-admissible scientific evidence. The forensic linguistics workshop provided a venue in which various aspects of the ILE paradigm were evident in papers from linguists at all stages in their careers. As we present an overview of the workshop, it will be obvious how current research exemplifies various aspects of the ILE paradigm, with some presentations fully embracing the ILE paradigm for admissible forensic linguistic evidence.

The workshop covered a wide range of topics in forensic linguistics: author identification, speaker identification, forensic linguistic corpus development and text-typing, language crimes, and courtroom interaction.

II. Author Identification

In "Style and Authorship", Carole Chaski presented the argument that identifying style is not the same exercise as identifying authorship. For many forensic linguists, "style" is approached through superficial and highly salient features such as word choice and word frequency. Chaski's argument was based on three things. First, stylometric approaches to authorship have failed to identify authorship based on blind

tests. Second, stylometrists are now providing software to obfuscate authorship by manipulating style. Third, when a stylist attempted to foil a syntactic approach to authorship by editing the style of a text, his stylistic changes were not able to change the syntactic measurements, based on statistical testing of the original and edited version of the text. Therefore, Chaski concluded that style and authorship are not actually equivalent terms, contrasting style as a surface feature of language from authorship as a deep feature of language.

In "Approaches to style in qualitative authorship analysis", Eilika Fobbe presented the results of a stylistic analysis of three short extortion letters by applying Brinker's text linguistic approach to the texts. The paper examined the author's approach to the complex communicational task of extortion and focused on obligatory and optional text patterns, on how they were arranged and linguistically expressed. It was pointed out that by including the text level into stylistic analysis, one could extract additional information about the author's way of dealing with the situation even from very short texts. It was also suggested that this information could possibly be of investigative value.

In "Benchmarking Author Recognition Systems for Forensic Application", Hans van Halteren described how he benchmarked a reasonably good authorship recognition system against the written texts in the British National Corpus. Given the excellent author recognition results in (near) ideal circumstances, Van Halteren doubts anyone would hesitate using the system in court. However, in worse and realistic circumstances, evoked primarily by reducing the amount of training and test material, the recognition quality quickly degraded. Van Halteren strongly suggested that, for a system judgement to be accepted in court, its quality should be tested on material of known origin under the same circumstances that apply in the court case.

In "Demonstration of ALIAS: Automated Linguistic Identification & Assessment System", Carole Chaski showed the ALIAS software analysing forensic linguistic data. First, Chaski presented reasons for automating forensic linguistic methods, including objectivity, algorithmic standardization, lack of confirmation bias, and removal of human fatigue. Second, Chaski demonstrated the components of ALIAS[TM], including TATTLER[TM,] the text analysis system, and the modules for analytical procedures, ALI[TM] for fully automated linguistic analysis, ALEX[TM] for automated linguistic analysis with expert interaction, and ALISTAR[TM] for automated linguistic analysis with scientist-linguist input. Finally, Chaski presented the research paradigm underlying the ALIAS analytical procedures (Chaski 2001, 2005, 2013).

III. Speaker Identification

In "Voice profiling: forensic phonetics applied in the case of the clumsy kidnapper", Gea de Jong-Lendle showed how forensic phonetics can be applied in the case of a kidnapping. In this case the police requested a forensic analysis of three phone calls made to the family of the victim. The caller, assumed to be the same in all three

calls, spoke German fairly fluently but clearly with a foreign accent. De Jong-Lendle demonstrated the phonetic methods to create a profile of the kidnapper based on his voice, his language and other linguistic features, including geographical dialectology.

IV. Spanish Forensic Corpus

In "A Spanish Corpus for Forensic Linguistic Research", Ángela Almela Sánchez-Lafuente, Carole Chaski, Gemma Alcaraz Mármol, Clara Pallejá López, Victoria Guillén Nieto and Arancha García Pinar presented the methods for developing a database of Spanish writing that can be used for forensic linguistic research, including data collection procedures and challenges. Specifically, the main instrument used for data collection has been adapted from Chaski (2001; 2013) and translated into Spanish. This instrument consists of ten tasks, by means of which the subjects are asked to write formal and informal texts about different topics which evoke different registers. To date, 93 undergraduates from University of Alicante, Polytechnic University of Cartagena and University of Castilla-La Mancha have already participated in the study. Further, using the same instrument, text data have been collected from prisoners who fit a specific crime profile. The corpus has four potential uses within forensic linguistics, namely identification, text-typing, inter-textuality, and linguistic profiling, which is in line with some previous analyses of Spanish corpora for forensic purposes (e. g. Almela et al. 2013; 2015).

V. Forensic Computational Linguistics

Gender-based violence is receiving more and more attention from professionals and researchers within the legal, criminological and psychological fields, exploring several aspects related to both the victim and the abuser. However, the linguistic profile of those involved in gender-based violent acts has been hardly explored. In the English-speaking world, several scholars have demonstrated the applicability of linguistic analysis in detecting some subjects' tendency to commit certain types of crime (Drouin et al. (2017); Taylor et al. (2013)). Within this framework, "The Gender-based Abuser: A Proposal for Forensic Linguistic Analysis from a Computational Perspective" by Ángela Almela Sánchez-Lafuente, Carole Chaski, Gemma Alcaraz Mármol, Clara Pallejá López and Pascual Cantos Gómez presented a pilot study of quantitatively and computationally differentiating the language of domestic abusers from a control group. The domestic abusers have been convicted of a violent crime in the domestic context, while control group members have not. The results of analyzing the two groups' linguistic behavior in writing, responding to the same stimuli were presented. Furthermore, results of clustering and classification to determine the statistical reliability of differentiating the language of domestic abusers were presented.

VI. Court Testimonies / Courtroom interaction

Elizabeth Allyn Smith in collaboration with Myriam Raymond-Tremblay presented "Influences of felicitous and non-felicitous presuppositions on belief in French", the first full-scale replication of Loftus (1976) and Loftus and Burns (1982) in a language other than English. The results suggest the possibility of a cultural difference when it comes to the effect of presupposition on long-term memory, though further corroboration is needed. This work contributes to an increasingly strong body of evidence showing that linguistic factors can bias evidence and testimony obtained through questioning techniques. Further, the empirical results undermine the prevailing assumption of some interview techniques that interviewees will automatically challenge questions that presuppose false information. The study indicated the possibility of a high rate of false confessions given the number of presuppositional questions used in interviews.

Carl Vogel and Justine Reverdy presented "Levels of linguistic description in repetition effects associated with 'successful' and 'unsuccessful' dialogue". This research focused on the methodology for quantifying the certainty one can have in rejecting a null hypothesis that interlocutors have not understood each other in dialogue. It was shown that the circumstances brought to court in which the precise question at hand is whether one individual can be said to have understood another, with merely a transcript available as the basis for decision. This method requires calibration on many data sets, and one factor in this calibration relates to the levels of linguistic description involved, just as in authorship attribution research, the linguistic level of tokenization has also been explored extensively. This paper analyzed the levels of repetition (which may be understood in relation to grounding for repetition of "other", and evolving dialogue plan maintenance for repetition of "self") visible at levels of description from sub-lexical to phrasal and levels inclusive of abstract syntax, in short dialogues that external observers are likely to characterize as successful or not with clear intuitions. Analysis was conducted with respect to a data set in which an independent assessment of communication success is available in the form of performance of a collaborative task. The data used was the Human Communication Research Centre (HCRC) Map Task dialogue corpus. One participant communicates a path (Information Giver, IG) on a map to the other (Information Follower, IF). Success in the task was measured in terms of the deviation between the two paths for the map. All participants have four attempts at the task. Factors that were analyzed included partner familiarity, dialogue role, availability of eye contact. It was proved that participants do better as task experience increases. Familiar partners have slightly more success than unfamiliar partners. Where there is significant other-repetition (summing across levels of linguistic description), this yields low deviation scores (more success) for familiar partners in first games but high deviation for unfamiliar partners. Repetition "helps" unfamiliar partners after the first game. Significant repetition (self- and other-repetition) by information givers leads to greater task success. In contrast, lexically based repetition (self- and other-) by information followers cor-

relates with unsuccessful communication; but IF-repeated syntactic structures (self- and other-) does signal successful communication. This contributes to understanding of baselines in repetition effects for "normal" dialogues (Reitter / Moore 2007; Vogel, 2013; Reverdy / Vogel, 2017a, 2017b). An interpretation of the results in legal settings draws to light interrogations: one might reconsider the merits of interrogators repeating themselves or words and phrases of a suspect; one might feel better about suspects repeating themselves and their interrogators.

In "Veracity Assessment of speaker-witnesses in child-abuse cases", Martina Nicklaus and Dieter Stein discussed methodological issues relating to the linguistic underpinning of approaches and procedures in a legal context where verbally given evidence is paramount in establishing veracity, in addition to established psychological evaluation. It would seem that linguistic expert knowledge is called for in the form of modern psycholinguistic theory of the narrative that relates output at each occasion to specific factors at the time of production. Another important issue is establishing a linguistic "baseline": What is a deviation from such a baseline, such as repetition, false starts, tense changes, particles etc., which cannot automatically or schematically be interpreted in a simplistic way as indicators if veracity or otherwise without taking into account the other types of information such as videotapes or recordings, if and hopefully also available. The discussion was based and inspired by work on a concrete case of child abuse as reflected in a narrated interview.

VII. Language Crimes

In "Defamation as a language crime" Victoria Guillén-Nieto (University of Alicante) examined defamation as a language crime within the theoretical framework of impoliteness theory (Culpeper 2011), e. g. affective impoliteness. Guillén Nieto's research is grounded in the analysis of a reference corpus of 150 judgments rendered by High Courts of Justice in Spain between 2013 and 2017. The study demonstrated that impoliteness theory provides the language expert with a socio-pragmatic categorisation of offence, e. g. face intentionality, face attack, and face loss, as well as with a description and explanation of the processes involved in the construction and deconstruction of offence: (a) a bottom-up process through which the victim perceives and constructs the offender's intentionally face-threatening behaviour. And (b) a top-down process through which the court deconstructs and appraises behavior in context, and within a legal culture and system, e. g. Constitution, Civil Law, Penal Law, Penal Code, etc. Findings from this study showed that in Spanish courts not every offence implying face damage is a crime by law, e. g. minor injuries, low intensity insults, degrading expressions, and cross-accusations with no public significance were decriminalised by law in 2015. Moreover, courts in Spain apply weighting techniques in their legal reasonings. More specifically, they clearly favour the prevalence of the Rights to Freedom of Expression and Information over the Right to Honour

when the alleged defamatory text is based on facts and does not include high intensity insults, degrading expressions and false accusations.[1]

VIII. Conclusion

Overall, the workshop demonstrated that forensic linguistics can be a productive and worthwhile field when it focuses on applying reliable methods from mainstream linguistics to forensic questions. The variety and depth of knowledge of linguistics in the presentations are extremely encouraging for the field and we look forward to continued growth of forensic linguistics in the supportive and intellectually-stimulating context of International Language and Law Association.

IX. Literature

Almela, Ángela / Alcaraz-Mármol, Gema / Cantos, Pascual (2015): Analysing deception in a psychopath's speech: A quantitative approach. In: *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada* 31(2), pp. 559–572.

Almela, A. / Valencia-García, Rafael / Cantos, Pascual (2013): Seeing through Deception: A Computational Approach to Deceit Detection in Spanish Written Communication. In: *Linguistic Evidence in Security, Law and Intelligence* 1(1), pp. 3–12.

Chaski, Carole (1997): Who Wrote It? Steps Toward a Science of Authorship Identification. In: *National Institute of Justice Journal.* September 1997. Available online at http://www.ncjrs.orgNCJ184604.

Chaski, Carole (2001): Empirical Evaluations of Language-based Author Identification Techniques. In: *International Journal of Speech, Language and Law* 8(1), pp. 1–66.

---

[1] As an additional study to the one presented in the forensic linguistics workshop, Guillén Nieto presented "Impoliteness theory and the appraisal of intentional defamation in Spanish courts" in the general session of ILLA Conference. Drawing on the same reference corpus of 150 judgements for cases of defamation rendered by High Courts of Justice in Spain between 2013 and 2017, she focused, this time, on the impact of impoliteness strategy (conventional direct formulaic impoliteness vs nonconventional indirect formulaic impoliteness) on the final verdict. After analyzing the impoliteness strategies present in the defamatory texts of the 150 cases together with their corresponding verdicts in the High Courts of Justice (acquittal vs guilty verdict), both qualitatively and quantitatively, e.g. Chi-square test, the null hypothesis was proved, e.g. conventional direct formulaic politeness does not promote guilty verdicts, and nonconventional indirect formulaic impoliteness does not have impact on acquittals. Moreover, the fact that Spanish courts seem to allow more cases in which defamation is encapsulated in conventional direct formulaic impoliteness than those in which the offence is packed in indirect formulaic impoliteness, suggested that the latter may be considered an elusive defamatory strategy since the linguistic evidence may not be strong enough to build a court case of defamation. Other factors rather than the way the meaning is conveyed seem to have more impact in the final verdicts, e.g. the professional status of the target victim, the public significance of the offence, widespread dissemination, reckless contempt of the truth, among others.

Chaski, Carole (2005): Who's At the Keyboard? Recent Results in Authorship Attribution. In: *International Journal of Digital Evidence* 4(1). Available online at http://ijde.org.

Chaski, Carole (2013): Best Practices and Admissibility of Forensic Author Identification. In: *Journal of Law and Policy* 21(2), pp. 332–376.

Culpeper, Jonathan (2011): Impoliteness. Using Language to Cause Offence. Cambridge: Cambridge University Press.

Dror, I. E. (2012): Cognitive bias in forensic science. In: *Science & Technology Yearbook*, pp. 43–45, McGraw-Hill.

Dror, I. E. / McCormack, Bridget M. / Epstein, Jules (2015): Cognitive Bias and Its Impact on expert Witness and the Court. In: *The Judge's Journal*, Vol. 54 No. 4. Available online at https://www.americanbar.org/publications/judges_journal/2015/fall/cognitive_bias_and_its_impact_on_expert_witnesses_and_the_court.html.

Drouin, M. / Boyd, R. L. / Hancock, J. T. / James, A. (2017): Linguistic analysis of chat transcripts from child predator undercover sex stings. In: *The Journal of Forensic Psychiatry & Psychology*. Vol. 28, Issue 4. Available online at https://doi.org/10.1080/14789949.2017.1291707.

Kassin, S. M. / Dror, I. E. / Kukucka, J. (2013): The forensic confirmation bias: Problems, perspectives, and proposed solutions. In: *Journal of Applied Research in Memory and Cognition* 2, pp. 42–52.

Kassin, S. M. / Dror, I. E. / Kukucka, J. (2013): New application of psychology to law: Improving forensic evidence and expert witness contributions. In: *Journal of Applied Research in Memory and Cognition* 2, pp. 78–81.

Kelty, S. F. / Julian, R. / Ross, A. (2013): Dismantling the Justice Silos: Avoiding the pitfalls and reaping the benefits of information-sharing between forensic science, medicine and law. In: *Forensic Science International.*

Loftus, Elizabeth (1976): Unconscious transference in eyewitness identification. In: *Law & Psychology Review* 2, pp. 93–98.

Loftus, E. / Burns T. E. (1982): Mental shock can produce retrograde amnesia. In: *Memory & Cognition* 10, pp. 318–323.

Murrie, D. C. / Boccaccini, M. T. / Guarnera, L. A. /Rufino, K. A. (2013): Are forensic experts biased by the side that retained them? In: *Psychological Science* 24.

Reitter, David / Moore, Joanna (2007): Predicting Success in Dialogue. In: *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 808–815.

Reverdy, Justine / Vogel, Carl (2017): Measuring Synchrony in Task-Based Dialogues. In: *Proceedings of the 18th Annual Conference of the International Speech Communication Association (INTERSPEECH2017)* pp. 1701–1705.

Reverdy, Justine / Vogel, Carl (2017): Linguistic Repetitions, Task-based Experience and a Proxy Measure of Mutual Understanding. In: *Proceedings of the 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2017)*, pp. 395–400.

Taylor, P. J. / Dando, C. J. / Ormerod, T. C. / Ball, L. J. / Jenkins, M. C. / Sandham, A. / Menacere, T. (2013): Detecting insider threats through language change. In: *Law and Human Behavior* 37(4).

Vogel, Carl (2013): Attribution of Mutual Understanding. In: *Journal of Law and Policy* 21(2), pp. 377–420.