The blended nature of legal-lay language in Italian and English:
A corpus-based contrastive analysis

Lucia Busso

Aston Institute for Forensic Linguistics, Aston University

l.busso@aston.ac.uk

The present contribution proposes a contrastive analysis of lexico-grammatical features in English and Italian legal-lay language (henceforth: LLL) – defined here as any legal text type aimed at a non-specialist audience (Tiersma, 1999; Bhatia, 2010). Studies on LLL as an independent genre are still scarce in the linguistic literature (Van Boom et al., 2016; Conklin et al., 2019), which predominantly focus on legal language simplification (*inter alia*: Cortelazzo, 2008; Adler, 2012; Mori, 2019).

The study investigates a specialistic corpus of LLL using collostructional analysis (Römer, 2009; Stefanowitsch, 2013) and a comparative frequency analysis with other specialised corpora. The research has two primary aims: first, to provide an exploratory account of the lexico-grammatical features of LLL and second, to ascertain whether such features can be considered idiosyncratic. Particularly, the research sets off from two interrelated hypotheses: that LLL exhibits idiosyncratic lexico-grammatical characters, different from specialised legal jargon and non-specialised written prose alike; we further hypothesize that LLL will show a 'blended' nature, with mixed characters between these two genres. To analyse both the lexical and the grammatical end of the lexico-grammar continuum (Halliday, 1991; Gabrielatos, 2018), the study takes the constructionist standpoint that language is formed by *constructions*, holistic pairs of form (syntax) and function (semantics) (Goldberg, 2006, 2019). Construction Grammar is in fact increasingly applied in both synchronic and diachronic corpus-based studies (Gries, 2013; Hilpert, 2013). A growing body of literature has also been using constructionist tenets for the analysis of genre (Hoffmann & Bergs, 2018).

The research is conducted on a self-compiled specialised corpus that comprises several textual types ascribable to LLL: CorIELLS (CORpus of Italian and English Legal-lay textS, Busso [accepted]). To analyse grammatical patterns and the lexical items they tend to co-occur with in CorIELLS, collostructional analysis is adopted. This family of quantitative methods finds statistically associated subcategorization preferences (i.e., lexical items) for a given abstract grammatical construction. Simple and covarying collexeme analyses are employed to analyse four grammatical constructions: nominalizations heading prepositional phrases, participial constructions, modal verbs, and passive constructions. These constructions were selected at different abstractedness levels, based on previous research on legal and bureaucratic grammatical features in both Italian and English (Garavelli, 2001; Brunato and Venturi, 2014; Coppolella, 2014; Brunato, 2015; Mori, 2019).

Statistically associated collexemes found with the collostructional analysis are further contrasted to the same structures in different corpora of legal and non-specialised written prose: the specialised legal and written prose subcorpora of the Italian reference corpus CORIS (Rossini-Favretti, 2000) and for English the imaginative subcorpus of the BNC and an ad-hoc created subcorpus of EurLEX-English. Data are analysed with linear mixed-effect-modelling to highlight differences in usage. Results from the analyses of both languages will be further compared and discussed to find differences and similarities in the lexico-grammatical profile of LLL in Italian and English.

The analysis of English is still ongoing, but findings from Italian appear to preliminary validate our hypotheses: subcategorization preferences show a mixture of specifically legal and more colloquial lexical choices. Moreover, constructions are used significantly differently than in both non-specialised prose and legal jargon (Busso, submitted).

References

Bangalore, S. & Joshi, A.K. (1999) Supertagging: An approach to almost parsing. Computational linguistics, 25(2), 237-265.

Adler, M. (2012). *The Plain Language Movement*. Oxford University Press.

Bhatia, V.K. (2010). Specification in Legislative Writing: Accessibility, Transparency, Power and Control. In M. Coulthard, & A. Johnson, (eds) *The Routlege Handbook of Forensic Linguistics* (p.50). Routledge.

Brunato, D. (2015). Operationalizing linguistic complexity from a NLP perspective: The computational assessment of text readability. *Unpublished Phd Thesis University of Siena.*

Brunato, D., and Venturi, G. (2014). Le tecnologie linguistico-computazionali nella misura della leggibilità di testi giuridici [computational linguistic technologies in legal texts readability measure]. *Informatica e diritto*, XL (XXIII), 111-142.

Busso, L. (accepted at the *Società Linguistica Italiana Conference* 2020, postponed to September 2021). *CorIELLS:* A specialised bilingual corpus of lay legal communication.

Busso, L. (submitted). Lexicon and grammar in legal-lay language: a quantitative corpus study on Italian

Coppolella, M. (2014). Modal values of verbal forms in the European Charter for Regional or Minority Languages. A linguistic comparative analysis of the English, Italian and Spanish versions. *Comparative Legilinguistics*, *18*, 24-46.

Conklin, K., Hyde, R., & Parente, F. (2019). Assessing plain and intelligible language in the Consumer Rights Act: a role for reading scores? *Legal Studies,* 1-20.

Cortelazzo, M. (2008). È meglio essere chiari e leggibili [It is best to be clear and readable]. *Etica* 10(3), 37-45.

Gabrielatos, C. (2018). The lexicogrammar of BE interested: description and pedagogy. In *Corpora and Lexis* (pp. 240-276). Brill Rodopi.

Garavelli, B. (2001). Le parole e la giustizia. Divagazioni grammaticali e retoriche su testi giuridici italiani [Words and justice. Grammatical and rhetorical digressions on Italian legal texts]. Einaudi.

Goldberg, AE. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Goldberg, AE. (2019). *Explain Me This*. Princeton University Press.

Gries, S. Th. (2013). Data in construction grammar. In T. Hoffmann and G. Trousdale (Eds.), *The Oxford handbook of Construction Grammar* (pp. 93-108). Oxford University Press.

Halliday, M.A.K. (1991) Corpus studies and probabilistic grammar. In Aijmer, K. & Altenberg, B. (eds.), English Corpus Linguistics: Studies in honour of Jan Svartvik. London: Longman. 30-40.

Hilpert, M. (2013). *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax* (Studies in English Language). Cambridge University Press.

Hoffman, T., & Bergs, A. (2018). A construction grammar approach to genre. *CogniTextes* (18), (open access version https://journals.openedition.org/cognitextes/1032).

Mori, L. (2019). Complessità sintattica e leggibilità. Un monitoraggio linguistico per la valutazione dell'accessibilità dei testi legislativi europei e italiani [Syntactic complexity and readability. Linguistic monitoring for the assessment of the accessibility of European and Italian legislative texts]. *Studi Italiani di Linguistica Teorica e Applicata* [Italian Studies of Theoretical and Applied Linguistics] (48), 627-657.

Römer, U. (2009). The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics*, *7*(1), 140-162.

Rossini-Favretti, R. (2000). Progettazione e costruzione di un corpus di italiano scritto: CORIS/CODIS [Design and construction of a written Italian corpus: CORIS / CODIS]. In R. Rossini-Favretti (ed.), *Linguistica e informatica. Multimedialità, corpora e percorsi di apprendimento*[Linguisticsand Computer science: Multimediality, corpora, and learning](pp. 39-56). Bulzoni. URL:http://corpora.dslo.unibo.it/TCORIS/

Stefanowitsch, A. (2013). Collostructional analysis. InT. Hoffmann & G. Trousdale (eds.), *TheOxford handbook of construction grammar*(pp.1-16). OxfordUniversity Press.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Bodleian Libraries,University of Oxford, on behalf of the BNC Consortium. URL:http://www.natcorp.ox.ac.uk/

Tiersma, P. M. (1999).*Legal language*. University of Chicago Press.

Van Boom,W. H., Desmet, P., & Van Dam, M. (2016). "IfIt's Easy toRead, It's Easy to Claim"—The Effect of the Readability of Insurance Contracts on Consumer Expectations and Conflict Behaviour. *Journal of Consumer Policy*, 39(2), 187–197.