Michał Mrugalski
Humboldt-Universität zu Berlin

Provocation and Pre-Diction. Entropy and Surprisal in Terrorism Narratives of the Tsarist Empire 1862-1917 (Russian and Polish Language Texts)

1. Literary Historical and Theoretical Assumptions: Terrorist Realism (TR) as a Narrative Mode

The conspiracy in the name of the oppressed and the assassination of a figure of power as a provocation to universal revolution prevailed in Europe and America in the nineteenth and early twentieth centuries. But nowhere else but in the Tsarist Empire did the development of the political strategy of terrorism coincide with the emergence of an enormous number of narrative texts in almost all the languages of the Empire, in genres ranging from anonymous leaflets to the masterpieces of Dostoevsky, L. Tolstoy, Chekhov, or Przybyszewski and Brzozowski (Mogil'ner 1999, Dieze 2016; Frank 2017; Patyk 2017).

Regardless of their support or opposition to actual violence, these texts have specific formal characteristics due to their thematization of provocative acts of violence and speech; I propose the term Terrorist Realism (TR) as a way of describing this narrative mode.

The identification and description of TR allows, first, a cultural-historical reinterpretation of the development of fictional discourse between 1860 and 1914, not only in Central and Eastern Europe, since TR had a global impact. Second, the description of TR provides new insights into the narrative construction or reproduction of consciousness (cf. Schmid 2017) and its objective and subjective correlates: estrangement (*Verfremdung*, *ostranenie*) and cognitive effort. Third, this historical narrative mode has implications for how terrorism is conceptualized and experienced today. Fourth, its modeling provides an opportunity to coordinate narrative economy with the reader's energetic effort in an attempt to create a cognitive poetics that would be both quantitative and capable of not terribly vulgarizing works of modern art.

Both actual terrorism and TR can be accommodated within a broader framework of energy economics that 1) can be captured quantitatively, 2) functions according to the logic of provocation.

To demonstrate this quantitative analysis of provocation, I propose a postclassical narratological model of TR that combines the classical communicationist approach with postclassical cognitivism ("enactive inference").[1]

Historians describe acts of terror as acts of communication designed to provoke specific responses to surprising violence (Hilbrenner 2022). Literary works, themselves acts of communication on many levels (between protagonists, the narrator and his counterpart, the author and the reader), incorporate these specific violent "utterances" (along with the provocative verbal utterances of terrorists and secret police) into their diegetic worlds and narrative strategies. As for the postclassical level of the model, enactive inference is consistent with TR because it sees cognition as a function of the organism's survival. In TR, where constant threat dominates and provocation undermines all certainties, the subject's cognitive acts, as in enactive inference's description of the functioning of organisms, amount to manic predictions that create reality for the subject (the world and the subject provoke each other), including the subject's interventions in the worlds. The subject parses and hierarchizes the elements of the world (perceived as signs) in terms of their potentially threatening nature and manages its energetic resources according to Hamiltonian machanics (this is the assumption of enactive inference). Predictive world-making, approximating Bayesian probability, takes place in bodily and emotionally charged exchange with the environment. Because of the predictive nature of this worldmaking, which boils down to managing energy to survive, the quantitative analysis of literary cognition can be based on entropy and surprise, measures adapted from statistical mechanics in information theory. There was a precedent: Leon Winiarski was a thinker associated with the terrorist organization The People's Will (via their

[1] As a synthesis of active inference and enactivism, Ramstead 2020; Downey 2017; Venter 2021; Parr 2022.

Polish allies Proletaryat). In 1888-1889, Winiarski formulated quantitative aesthetic theories in terms of energy transfer (effort and entropy) and Hamiltonian mechanics.

2. Operationalizing the Model

The narrative model can be related to well-established methods of corpus linguistics, since both communicationism and enactive inference on the one hand, and many "schools" of linguistics on the other, have embraced information theory's adaptation of the notions of entropy and its correlate, surprise, from statistical mechanics.

Entropy and surprise are correlated with the effort the subject invests in decoding messages. Due to the finite energy resources of speakers, recipients show a preference for a uniform distribution of information (Fenk 1980). Increased information density (measured by surprisal or infomativeness) slows perception so that the perceiver's effort remains constant.

The Russian formalist Viktor Shklovskii (Shklovskii [1917] 2015) points to the temporal prolongation of the perceptual process as the ultimate goal of aesthetic experience, driven by the making-difficult device. More directly, however, I refer to Leon Winnicki's reduction of aesthetics to the equations of Lagrangian mechanics.

The narratological model of TR can be operationalized on the basis of three hypotheses, inspired by the achievements of computational linguistics[2]. All three hypotheses presuppose a distinction between online surprisal (as opposed to average surprisal) and Shannon entropy.[3] Surprisal means the predictability of a word in a sequence given previous words/context [$W_i = -\log_2 P(w_i \mid w_1 \ldots w_{i-1})$], while Shannon entropy is the measure of uncertainty about what is being communicated with respect to the diegetic world as a narrative sequence unfolds proposition after proposition: $H(x \in X) = -\sum P(x) * \log 2\, P(x)$.

1: The narratives of TR tend to expose higher-than-average linguistic measures of entropy and, initially, surprise; there is a correspondence between the subject matter (deadly provocation) and the linguistic features of the texts.

2: The evolution of TR over time should be characterized on the linguistic level by the maintenance of a high degree of entropy (related to the unstable, dangerous composition of the diegetic world) and a decreasing degree of surprise (the conventionalization of a discourse).

3. As for entropy, the uncertainty/confusion in relation to the diegetic world (large semantic figures) can depend either on a homonymy of words or on the hierarchy of details, clues or perceptions in terms of their importance for survival. Both cases appear paradigmatically in A.C. Dolye's "The Speckled Band", where a band of Gypsies is substituted for a strip of material (a metaphor for a snake), while an unimportant detail contained in a relative clause turns out to be critical to the solution. The former type of uncertainty is more related to semantics and metaphor, the latter to syntax and metonymy.

Hypothesis testing involves building your own custom tools, because off-the-shelf solutions, few as they are (https://onlinetexttools.com/calculate-text-entropy), simply calculate the type-token ratio for a text against the general background of an LLM. They cannot account for online surprisal or entropy within a storyline or genre history. In addition, a reference corpus is compiled containing the most representative realizations of the narrative mode TR identified in the scientific literature. The reference corpus is confronted with the archive (nineteenth-century narrative prose). The construction of the tools and corpora has a threefold objective:

1. To train a classifier that would find similar texts in the archive of nineteenth-century prose, thus measuring the impact of the mode.

2. [Related to hypotheses 1 and 2] To study the language of TR on the basis of surprise ("informativeness") and Shannon entropy, as well as auxiliary measures that help to detect

[2] Brouwer et al. 2021; Degaetano-Ortlieb 2019; Degaetano-Ortlieb & Piper 2019; Degaetano-Ortlieb & Teich 2022; Levy 2015; Lowder et al. 2018; Sayeed et al. 2015; Venhuizen et al. 2019; Venhuizen 2019a; Venhuizen et al. 2022

[3] … although average surprisal and entropy are the same for non-skewed probability distributions.

differences between information distributions, such as cross-entropy or KLD, called “relative entropy”.

2.1. To measure with KLD the diachronic differentiation between the reference corpus and the archive (if the difference increases over time, we witness a consolidation of a mode).

2.2. To test whether phrasal standardization occurs, which implies that surprisal should decrease over time.

3. [Related to hypothesis 3] To measure, within particular texts, entropy and online surprisal, which are correlated with the construction of diegetic worlds and the parsing of discourse, respectively. Likewise, experiments with word embeddings should shed light on the complicated relationship between surprisal and entropy on the one hand, and the paralogisms of the modernist theory of surprise or estrangement as signposts of artfulness and eventfulness on the other. While surprisal (as a measure of discourse) is proportional to formalist making-difficult (zatrudnenie) and estrangement (*Verfremdung*, *ostranenie*), entropy seems to function in a different, though intimately related, way. Both decreasing and increasing entropy in relation to the diegetic world involve effort relative to the absolute value of the change. Both gaining and losing confidence in the world is laborious. The greatest difficulty arises when the subject oscillates between certainty and uncertainty, as in the case of provocation by bomb and word.

References:

Brouwer H. et al. (2021): Neurobehavioral Correlates of Surprisal in Language Comprehension: A Neurocomputational Model. In: Frontiers Psychology. Section: Language Sciences 12.

Degaetano-Ortlieb S. (2019): Hybridization effects in literary texts. In: Proceedings of the 10th International Corpus Linguistics Conference, Cardiff.

Degaetano-Ortlieb S., A. Piper (2019): The Scientization of Literary Study. In: Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, 18–28.

Degaetano-Ortlieb S., E. Teich (2022): Toward an optimal code for communication: The case of scientific English. In: Corpus Linguistics and Linguistic Theory 18(1), 175-207.

Dietze C. (2016): Die Erfindung des Terrorismus in Europa, Russland und den USA 1858-1866. Hamburg: Hamburger Edition.

Downey A. (2017): Radical Sensorimotor Enactivism & Predictive Processing – Providing a Conceptual Framework for the Scientific Study of Conscious Perception. In: T. Metzinger & W. Wiese (Eds.). Philosophy and Predictive Processing: 11. Frankfurt am Main: MIND Group.

Fenk A, G. Fenk (1980): Konstanz im Kurzzeitgedachtnis – Konstanz im sprachlichen Informationsfluß? In: Zeitschrift für Experimentelle und Angewandte Psychologie 27(3), 400–414.

Frank M. (2017): *The Cultural Imaginary of Terrorism in Public Discourse, Literature, and Film. Narrating Terror. London: Routledge.*

Hilbrenner A. (2022): Gewalt als Sprache der Straße. Terrorismus und die Suche nach emotionaler Gemeinschaft im Russischen Reich vor 1917. Stuttgart: Steiner.

Levy R. (2015): Memory and surprisal in human sentence comprehension, https://www.mit.edu/~rplevy/papers/levy-2013-memory-and-surprisal-corrected.pdf

Lowder M et al. (2018): Lexical Predictability During Natural Reading: Effects of Surprisal and Entropy Reduction. In: Cognitive Science 42 (4): 1166-1183.

Mogilner M. (1999): *Mifologia podpolnogo cheloveka*' radikal'nyi mikrokosm v Rossii nachala XX veka kak predmet semioticheskogo analiza. Moskva: NLO.

Parr T., G. Pezzulo, K. Friston (2022): Active Inference: the Free Energy Principle in Mind, Brain, and Behavior. Cambridge: MIT Press.

Patyk L (2017): Written in Blood: Revolutionary Terrorism and Russian Literary Culture, 1861-1881. Madison: UW Press.

Ramstead M., M. Kirchhoff, K. Friston (2020). A tale of two densities: active inference is enactive inference. In: Adaptive Behavior 28(4), 225–239.

Sayeed A., S. Fischer, V. Demberg (2015): Vector-space calculation of semantic surprisal for predicting word pronunciation duration. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, vol. 1, 763–773.

Schmid W. (2017): Mentale Ereignisse. Bewusstseinsveränderungen in europäischen Erzählwerken vom Mittelalter bis zur Moderne. Berlin: De Gruyter.

Shlovskii V. ([1917] 2015). Art, as Device. In: Poetics Today 36, 151-174.

Venhuizen N., M. Crocker, H. Brouwer (2019): Expectation-based comprehension: Modeling the interaction of world knowledge and linguistic experience. In: *Discourse Processes 56*(3), 229–255.
Venhuizen N., M. Crocker, H. Brouwer (2019a): Semantic Entropy in Language Comprehension. In: *Entropy 21*.
Venhuizen N. et al. (2022): Distributional formal semantics. In: Information and Computation 287.
Venter, Elmarie. 2021. Toward an Embodied, Embedded Predictive Processing Account. In: Frontiers of Psychology 12.
Winiarski, Léon ([1888-1889]1967): Essais sur la mécanique sociale, ed. by Giovanni Busino. Genève: Librairie Droz.