New Developments in Quantitative Metrics

XML

Papers

Making Visible the Invisible: Metrical Patterns, Contrafacture and Compilation in a Medieval Castilian Songbook

Gimena del Rio Riande, SECRIT-CONICET; Clara Martínez Cantón, UNED; and Elena González Blanco-García, UNED

Panel Organizer

David J. Birnbaum, University of Pittsburgh, djbpitt@gmail.com

Panel Synopsis

This panel presents new primary research results in the formal study of poetry and poetics that have been made possible by the development and use of innovative digital technologies. The research questions underlying the presentations are varied, as are the linguistic and cultural traditions (early modern and modern Russian, medieval Spanish, and Urdu [in comparison with Hindi/Sanskrit and Persian/Arabic]). What unites the three presentations is (1) a focus on using digital technologies to create humanities knowledge that would not otherwise be possible; (2) the development of innovative methodologies that are able to address those research questions; and (3) the building of new digital tools that make it possible to address new research needs.

Our panel responds to the following specific areas of emphasis noted in the original call for paper:

• *Humanities research enabled through digital media, data mining, software studies, or information design and modeling.* The focus of our panel is on new types of research results that are made possible by the development of original digital tools and methods. In this sense, the research results are foremost, but the projects that produce that research are methodologically innovative, and the research results would not have been attainable without that innovation.

• *Creation and curation of humanities digital resources.* The creation of plain-text poetry archives is relatively straightforward: the text can be generated through OCR or by repurposing digital files originally created to produce print editions. The creation of structured poetry archives is also relatively straightforward: the general hierarchical structure of poetry is represented by pseudo-markup layout in plain-text editions, and is amenable to autotagging with the aid of regular-expression parsing. The creation of poetry archives with metrical and rhyme annotation, however, is difficult because it requires linguistic knowledge, and the presentations on this panel describe new technologies that were developed in order to (1) facilitate the machine-assisted creation of these types of metrically annotated poetic corpora, and (2) undertake original research about formal metrical practice on the basis of large digital corpora.

• *Social, institutional, global, multilingual, and multicultural aspects of digital humanities . . . For the 2015 conference, we particularly welcome contributions that address 'global' aspects of digital humanities including submissions on interdisciplinary work and new developments in the field.* Our panel includes three research reports from three diverse poetic traditions: early modern and modern Russian, medieval Spanish, and Urdu (in comparison with Hindi/Sanskrit and Persian/Arabic). The projects that produced these research results have operated independently, but within their highly varied cultural contexts they address similar types of research questions.

• *Digital humanities in pedagogy and academic curricula.* The projects that contribute to our panel, which were designed to create new research knowledge in the humanities, were developed in many cases with attention to pedagogical and curricular concerns. For example, some portions of the development for these projects was carried out in the context of digital humanities courses, with undergraduate and graduate students making contributions to authentic humanities research as a way of learning to be digital humanists.

2. Making Visible the Invisible: Metrical Patterns, Contrafacture, and Compilation in a Medieval Castilian Songbook

Del Rio Riande, G., Martínez Cantón, C. and González Blanco-García, E.

ReMetCa, Digital Repertoire on the Metrics of the Medieval Castilian Poetry (*Repertorio Métrico Digital de la Poesía Medieval Castellana*) is an online, free-access digital tool designed to undertake simultaneous complex searches on the metrical and rhyming patterns of Medieval Castilian poetry (starting from the late-12th-century epics to the poetry of the 16th-century Castilian *Cancioneros*). ReMetCa is part of the corpus of online digital resources on Medieval Romance poetry, alongside such others as the ones related to Galician-Portuguese (MedDB, The Oxford Cantigas de Santa Maria Database), French (BedTrouveres, Nouveau Naetebus), and Occitan and Catalonian poetry (BedT, Corpus des Trobadours). The research project that sustains ReMetCa aims to integrate traditional studies of philology (especially those pertaining to metrics) with digital humanities, revising and classifying the Castilian corpus in a hybrid digital framework that embeds TEI-Verse module tags in a relational database that works altogether with a controlled vocabulary on Medieval Castilian poetry.

One interesting case of study that illustrates the topic of the panel is our digital approach to the *Cancionero de Baena*, a large songbook containing almost 600 poems transcribed and compiled in the first half of the 15th century by Juan Alfonso de Baena, scribe of the court of King Juan II of Castile. The data retrieved from the tagging and classification of the metrical and rhyming patterns of a large section of Baena's songbook—the one regarding the *antiquiores* or the eldest poets, and those that composed their texts mainly in the second half of the 14th century—yielded interesting results in the area of the Hispanic studies devoted to metrics. On the one hand, we discovered that the *antiquiores* composed a large part of their poems using a pattern almost unknown by their predecessors, the Galician-Portuguese troubadours: the octosyllabic *octava* (eight-lined stanza, lines of eight syllables that may sometimes be isometric or heterometric). This pattern was shaped in a body of four stanzas (*glosa*) with rhymes structured in a singular pattern (*rimas singulares*) and words stressed on the penultimate syllable (*rima grave* or *femenina*). Furthermore, we were able to identify some groups or cycles of poems composed on the basis of the metrical and rhyming imitation or *contrafacture* (→ 4x8@8) (Spanke, 1928; Marshall, 1980; Rossell, 2000).

It was the theoretical organization of the XML markup as a discursive system (Jockers and Flanders, 2013), which we used to shape an ontology framework (available at http://www.purl.org/net/remetca and http://datahub.io/dataset/remetca-ontology), that in practice helped us to move from the descriptive to the connotative dimension, thus making visible the invisible. Apart from using the expected TEI-verse attributes such as @met for our schemes based on the number of syllables of each line and on the number of lines (e.g., 8,8,8,8), and @rhyme for the rhyming structure of the stanzas (e.g., abba), there were some new attributes that did not exist in the TEI-Verse module and that we decided to add to our XML schema: @asonancia, an attribute that indicates the two possible values of the rhyming typology: 'asonante' (which means that only vowel sounds are repeated) and 'consonante' (which means that every sound after the stressed syllable is repeated); @unisonancia, which takes the values of 'unisonante' or 'singular' and shows whether the same rhyming scheme (e.g., abba) is repeated in different stanzas or not; and @isometrismo, which states whether all the stanzas have the same number of syllables (isométrico) or not (heterométrico). It was the joint work of all these attributes that retrieved the metrical and rhyming patterns and led us to those new results. In addition to this, the whole analysis of the corpus resulted in an unexpected fact: Juan Alfonso de Baena may have organized and compiled the texts of the different poets in his songbook guided not only by chronological order but also following metrical patterns and types of stanza.

With the help of our digital tool we will illustrate possible contrafactures, common metrical and rhyming patterns, and cycles of poems in the *antiquiores*' corpus, and give an account of more complex definitions. The examples will also serve as an opportunity to cast our eyes again on the macro-microanalysis (Jockers, 2013; Jockers and Flanders, 2013; Liu, 2014) and data-text (Marche, 2012) debates in the field of literary studies and the concepts of close-distant reading (Moretti, 2013; Latour, 2014) and relate them to a subject of study that is interested in formal patterns (and not as much in content) and acquires new meaning when compared through large corpora: metrics.

References

Asperti, S., Zinelli, F., et al. (n.d.). *BedT, Bibliografia Elettronica dei Trovatori.* www.bedt.it.

Brea, M., et al. (n.d.). *MedDB: Base de datos da Lírica profana Galego-Portuguesa*. http://www.cirp.es/bdo/med/meddb.html.

González-Blanco, E., et al. (n.d.). *ReMetCa: Repertorio Métrico Digital de la Poesía Medieval Castellana.* www.remetca.uned.es.

Jockers, M. L. (2013). *Macroanalysis: Digital Methods and Literary History.* University of Illinois Press.

Jockers, M. L. and Flanders, J. (2013). A Matter of Scale. http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1106&context=englishfacpubs.

Latour, B. (2014). Opening plenary, *Digital Humanities 2014 (DH2014)*, http://dh2014.org/videos/opening-night-bruno-latour/.

Liu, A. (2014). *The Laws of Cool: Knowledge Work and the Culture of Information.* University of Chicago Press, Chicago.

Marche, S. (2012). Literature Is Not Data: Against Digital Humanities. *Los Angeles Review of Books*, http://www.lareviewofbooks.org/article.php?id=1040&fulltext=1.

Marshall, J. H. (1980). Pour l'étude des *contrafacta* dans la poésie des troubadours. *Romania* CI, pp. 289–335.

Moretti, F. (2013). *Distant Reading.* Verso, London.

Parkinson, S., et al. (n.d.). *The Oxford Cantigas de Santa Maria Database.* http://csm.mml.ox.ac.uk/.

Rossell, A. (2000). Intertextualidad e intermelodicidad en la lírica medieval. In Bagola, B. (ed.), *La Lingüística española en la época de los descubrimientos: Actas del Coloquio en honor del profesor Hans-Josef Niederehe, Tréveris 16 a 17 de Junio de 1997.* Hamburg: Helmut Buske, pp. 149–56.

Seláf, L. (n.d.). *Le Nouveau Naetebus. Répertoire des poèmes strophiques non-lyriques en langue française d'avant 1400.* www.nouveaunaetebus.elte.hu.

Spanke, H. (1928). Das öftere Aufreten von Strophenformen und Melodien in der altfranzösichen Lyrik. *Zeitschrift für französichen sprache und Literatur,* 51, pp. 73–117.

3. Using Bioinformatic Algorithms to Analyze the Politics of Form in Modernist Urdu Poetry

Pue, A. S., Teal, T. K. and Brown, C. T.

This paper has two aims. First, it shows how the authors—a humanist and two computational biologists—adapted graph-based algorithms used in genome assembly and multiple sequence analysis to scan the meter of Urdu poetry. Second, applying these techniques to modernist free-verse poetry of the early 1940s, the paper argues that data-rich analysis of poetic meter offers humanistic insights into the politics of literary form.