El proyecto «Poetry standarization and Linked Open Data – POSDATA (GA-679528)», dirigido por Salvador Ros Muñoz, compila quinientos sonetos extraídos del corpus DISCO, que oscilan entre la Edad de Oro y el siglo XIX, así como un importante número de composiciones del siglo XX, incluidas por el estudiante Alberto Barbado González.

Estos poemas, presentados ante los becarios en formato *Excel*, debían ser analizados desde una perspectiva que abandonase los cánones filológicos y que se adentrase en una lectura incardinada en lo emocional. Para ello, lejos, como se ha dicho, de cuestiones hermenéuticas, se facilitaron a los becarios diversos parámetros que medían, mediante una escala numérica, no solo la felicidad o el miedo que una palabra provocaba en ellos, sino la familiaridad, la edad o la capacidad que estas tenían para pergeñar una imagen en la mente del lector. Sin embargo, estas escalas primigenias que se centraban en la emotividad de los vocablos (premisa que tuvo que ser rechazada y que generó algún que otro contratiempo ya que no es los mismo estudiar una palabra aislada que un soneto) fueron remodeladas para facilitar y hacer más plausible la labor de investigación. En primer lugar, las horquillas numéricas que requería cada estado, y que fluctuaban en algunos casos entre el 1-5, 1-9 y 1-7, se estandarizaron en 1 y 5. Así el 1, determinaba el valor «muy poco», el 2 «poco», el 3 «bastante» y el 4 «mucho». De este modo, si la aversión que un soneto nos despertaba era alta se marcaba dicha casilla con 4, mientras que se calificaba con un 1 aquel estado que no generaba un sentimiento aguzado.

Una vez solventado el aspecto de la gradación de los números se procedió a eliminar dos parámetros que resultaban confusos y que no aportaban información relevante para el trabajo. Los estados *Age of Availability* y *Familiarity,* íntimamente entrelazados según la opinión de los anotadores, se suprimieron al considerar no solo la literalidad con la que un leedor de 2 a 11 años (lapso temporal propuesto inicialmente en el estudio) se acerca a lo poético, sino con la firme creencia de que un enorme porcentaje de los sonetos examinados no resultan accesibles durante esos años. Esta teoría resulta completamente factible si tomamos en consideración poemas como «Crece el insano ardor, crece el engaño» (número 254) o «La tirana codicia del hermano» (248), en los que se requiere un conocimiento de las fuentes clásicas, o si estudiamos textos imbricados en la doctrina estoica o en tópicos de la tradición amorosa.

Similar supresión fue llevada a cabo en las 89 condiciones mentales que fueron propuestas, inicialmente, como segunda parte del estudio. Esos estados, que recogían disposiciones como la idealización, la toxicomanía o el acoso, quedaron reducidos a 21 gracias a una muestra inicial que dejaba patente no solo lo reiterativo de los susodichos estados, sino a la conjetura de que una cantidad tan desmedida de los mismos suponía un escollo difícil de abordar en el tiempo establecido. Así, los conceptos finales fueron acometidos teniendo en cuenta no solo la agitación que generaba en el estudioso, sino la impresión que un determinado poeta quiso transmitir en sus versos.

A raíz de estos cambios y como consecuencia, como ya se ha señalado, de las limitaciones temporales la muestra inicial de 500 sonetos fue reducida a 300. Sin embargo, esta disminución no merma el interés de este estudio pues este se perfila como un interesante modelo de futuras investigaciones.