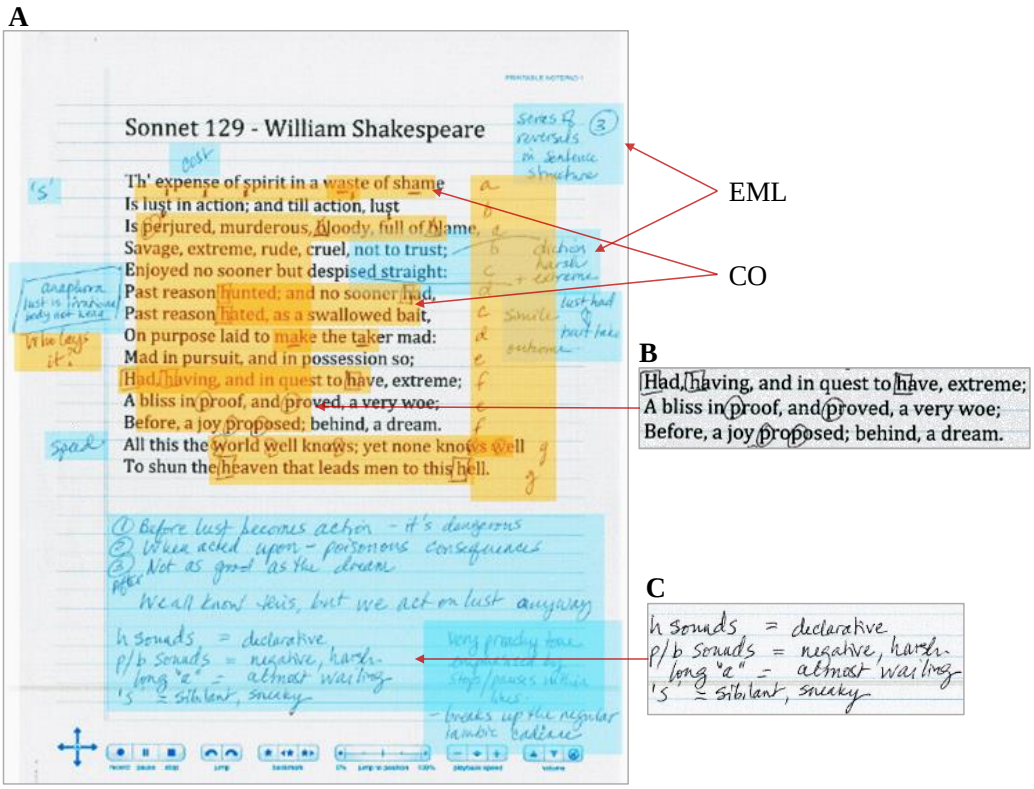


Fig. 2. Annotations made during close reading of poetry. (A) shows *cognitive purpose* codes assigned to annotation units for one of the participants' annotations during the study, as visualized by our coding tool. Orange and blue bounding boxes represent annotation units categorised as CO and EML respectively. (B) is an example of an annotation unit, identifying the repetition of sounds, categorised as CO. (C) shows an example of an annotation unit, noting observations about repetitions of sound across the poem, coded as EML.

in terms of what linguistic features they were identifying, the annotation forms used to mark these features were idiosyncratic, leading to annotation form-function ambiguity within and between participants. Categorization of the annotations based on the external cognition framework, which is more high level and abstract, permitted the intent of the annotations to be discerned consistently across all participants despite this ambiguity. In addition, the coarse categorisation by this code set enabled us to generalize highly specific actions performed by the participants to the more abstract cognitive tasks of hypothesis generation and hypothesis verification.

Codes based on the *cognitive purpose*, served by the process of annotation in text comprehension, included: computational offloading (CO), externalizing to reduce memory load (EML), both computational offloading and externalizing to reduce memory load (EML+CO) and ambiguous (A). CO and EML codes have been derived, in the context of annotations and close reading, based on the main cognitive benefits of using external representations identified by Preece et al. [2015]. CO represents the act of thinking through the content for identifying patterns of interest by annotating whereas EML refers to the act of tracking and assimilating intermediate hypotheses formed based