Gregorian chant, as an oral musical tradition, was performed by singers that had to memorize thousands of melodies. Each melody has a set of properties, one of which is what mode it belongs to within the modal system. To understand the learning process principles of chants, it may be helpful to decompose melodies into smaller units and analyze their relationship to modality. In this work, we compare Bayesian and neural network unsupervised segmentation methods. We measure their performance on evaluation metrics we design in order to examine the chant's properties with respect to the memorization challenge considering the modality aspects. For this purpose, we have two datasets, one with over thirteen thousand antiphons and the other with over seven thousand responsories. We find the Pitman-Yor process to be a more fitting model than BERT for this particular task, especially the conditional Pitman-Yor process model we proposed to segment each mode independently. We provide several clear arguments that modality and chant segmentation are closely connected. We also dispute the claim by Cornelissen et al. [2020] that the natural segmentation by chant words or syllables is best in terms of mode classification, and we provide a new state-of-the-art performance on the mode classification task.