This study investigates the potential contribution of visual information in the audiovisual perception of prosodic contrastive focus in French. Contrastive focus is used to attract the listener’s attention to a specific part of the utterance. Mostly conceived of as auditory/acoustic, it also has visible correlates which have been shown to be perceived. This study aimed at analyzing the interaction between audition and vision for the perception of prosodic focus by using a whispered speech paradigm. It was based on audiovisual recordings from 4 speakers wearing no facial markers. It combined performance assessment to reaction time measurements and confirmed and extended preliminary results obtained on 2 speakers wearing facial markers (needed for a parallel articulatory analysis). The results showed that adding vision to audition for the perception of prosodic focus can not only improve perception performances but can also reduce perceptual cognitive load. A further analysis suggested that the two modalities are integrated for the perception of prosodic focus. Visual only perception was facilitated for whispered speech suggesting an enhancement of visual cues. Moreover, the absence of facial markers did not impair perception. Therefore facial markers do not seem to enhance the relative importance/salience of visual cues.