This paper begins by reviewing the speech perception literature to predict cues that would and would not be expected to survive energetic masking of various types. The focus is especially (but not exclusively) on spectrotemporal cues to stops in the vicinity of the segment boundary in CV syllables. The second part of the paper discusses influences that can restrict the generality of research findings from isolated CV syllables. This includes ways in which CV syllables change in different phonetic contexts and styles of speech, contributions of the visual modality, and other uses of top-down information, such as phonotactic, lexical, semantic and syntactic probability. The paper concludes by asking whether-and how-simple measures such as CV intelligibility can be used to reflect intelligibility of speech in real-life communicative situations.