# ABSTRACT

A voice dialogue processing method and apparatus (300) based on a multi-modal feature, and an electronic device. The method comprises: acquiring, in the process of performing dialogue interaction with a user, first voice information that the user currently inputs, wherein the first voice information comprises a silent segment (101); determining, according to text information of the first voice information and historical context information of the first voice information, semantic feature information of the text information (102); determining, according to a voice fragment, which is before the silent segment, in the first voice information, phonetic feature information of the first voice information (103); acquiring temporal feature information of the first voice information (104); and determining, according to the semantic feature information, the phonetic feature information and the temporal feature information, whether the user ends voice input (105).

(Fig. 1)