The everyday world was making meaningful sound well before language developed. Self-produced sound exhibits a tight binding between source event and interpretation. These assertions naturally prompt thoughts about the role of these types of sound in the evolution of human auditory capability, especially language. The recent resurgence in theorizing about the origin of language has included other human capabilities as potential generative factors (e.g., gestures), but there has been little discussion about the possible role of capability to process everyday and self-produced sound. This is surprising given that recent neurocognitive research has supported common processing for speech and meaningful sound. Key attributes of language processing are the ability to interpret not just signs but symbols, and the ability to employ generative grammar processing to produce and interpret symbol strings. Association of sound events with causes and sequences, auditory representation of meaning, persistence of meaning under event and sequence variation, and persistence of event interpretations beyond what is currently present (temporal and spatial abstraction) are all fundamental requirements for these language attributes, and it is asserted that these capabilities might have developed to accurately perceive everyday and self-produced sound.