

Song structure, voice identity and digital audio

*Tom Parkinson*¹

¹Royal Holloway, University of London

Tom.Parkinson@rhul.ac.uk

Abstract (max 300 words)

This is a necessarily speculative attempt to link insights from recent experiments into voice identity processing with songwriting in a post-digital context. In the early years of mass adoption of digital music production, the novelty of an absolute volume threshold resulted in the *loudness wars* where music was engineered to be as loud as possible. In the last decade, however, the volume limit of digital audio has been approached as a compositional rather than a technical problem and has manifested new kinds of songwriting. Aside from the volume threshold, there are two other factors at play here: the audible frequency spectrum and vocal clarity. The need to make music that fills the audible spectrum, is loud, and has comprehensible vocals has resulted in linear, rather than simultaneous frequency balancing where, in the most simplistic examples, vocals are present when there is no competition for frequency space with music and loud music happens when there is no singing. In many cases, however, this is a dynamic process and unadorned vocals are present only in the opening section of the song. Across the last decade, the initial sixty seconds of a significant number of commercial pop songs have been characterised by the linear accumulation of acoustic complexity in the frequency space of the voice. It's possible, therefore, that what is at stake is not the clarity of the lyrics but the identity of the voice: once that identity is established, the voice can co-exist with other elements of the music whilst retaining its perceptual coherence. This idea intersects with research on time-based voice identification and proposes further questions about what it means to become familiar with a voice in stylised contexts.

[279 words]