

---

# Isolating the locus of informational interference during speech-on-speech listening

Sarah Knight\*<sup>1</sup> and Sven Mattys<sup>1</sup>

<sup>1</sup>University of York – United Kingdom

## Abstract

Speech-in-noise research often distinguishes between energetic masking (EM: interference between target and masker at the periphery) and informational masking (IM: interference higher in the auditory pathway). IM can itself be broken down into acoustic and spatial factors on the one hand, and factors related to long-term linguistic knowledge on the other. We use the term "informational interference" (inf-int) to refer to this latter type of IM. A typical manifestation of inf-int is the "masker intelligibility effect", whereby an intelligible masker is more detrimental to target perception than an acoustically-similar but unintelligible masker. However, this effect has usually been demonstrated using maskers constructed from connected speech (e.g. spoken sentences). As a result, it is difficult to determine which specific characteristics of the intelligible masker speech (e.g. syntactic, semantic, phonetic) underlie any observed effects.

In an ongoing series of studies (current N=360), we tested the masker intelligibility effect using word list maskers (i.e. maskers formed by concatenating words produced in isolation by a single talker). These maskers contain lexical and semantic information but lack syntax and sentence-level prosody. They were compared to several manipulated masker versions designed to vary in inf-int (i.e. intelligibility) and EM: 1) speech-modulated noise (SMN) – unintelligible, relatively high levels of EM; 2) time-reversed speech – also unintelligible but with recognisably speech-like sounds, similar levels of EM to natural speech; 3) 8-channel noise-vocoded speech (NVS) – intelligible, interim levels of EM (more opportunities than SMN for "glimpsing" of the target, but fewer opportunities than natural speech); 4) time-reversed noise-vocoded speech – unintelligible, interim levels of EM. Target speech was unmanipulated meaningful sentences, and target and masker talkers were matched for gender.

Results showed that performance was significantly poorer for the SMN masker than for all other maskers, emphasising the substantial role of EM in determining target intelligibility. However, comparisons between the other maskers showed that intelligible word list maskers (including both natural and noise-vocoded versions) were no more detrimental to target intelligibility than acoustically-matched, unintelligible equivalents (time-reversed natural and noise-vocoded versions). In other words, the results showed no evidence for the masker intelligibility effect. This was despite a well-powered design and clear evidence for effects related to EM. This suggests that the locus of inf-int may not be at the lexical-semantic level, and may instead reside in masker characteristics associated with connected speech, such as sentence-level syntax and/or prosody. It is also possible that specific features of our word list maskers, such as rhythmic or lexical predictability, reduced their ability to generate inf-int.

---

\*Speaker

Future iterations of the paradigm will explore this possibility by using word list maskers constructed from a larger number of more phonologically-varied words.

By systematically varying our maskers along several parameters, we have taken steps towards isolating the specific characteristics contributing to inf-int. More generally, these results highlight the difficulty of consistently characterising and empirically quantifying IM. They reveal that even "established" effects may not emerge in certain listening situations, and hence caution against indiscriminate appeals to the concept of IM in speech-in-noise research.