Does memory for individual talkers help word segmentation?

Rachel Smith
Department of Linguistics, University of Cambridge, UK
rhs20@cam.ac.uk

Experiments show that episodic memory for spoken words affects performance in many language tasks. But research has normally focused on isolated words and has not specified which acoustic-phonetic properties of the signal are encoded episodically. The present work investigates whether episodic memory is also implicated in word segmentation. If it is, then knowing a talker’s voice should critically involve knowledge of the fine phonetic means by which that talker realises linguistic distinctions around word boundaries. A word-monitoring experiment manipulated talker familiarity (established in a 1½-hour training session) and allophonic detail at a word boundary. Materials were phonemically identical sentence pairs differing in placement of one word boundary, with a concomitant grammatical distinction (e.g. But Pat sawed them / But Pat’s awed them). Participants heard spliced versions of the sentences, spoken by one of 2 familiar talkers or 4 unfamiliar talkers, native speakers of Southern British English. The last part of each version (e.g. awed them) was spliced to an earlier part excised from a different token of the same sentence (Matched condition; But Pat’s) or the alternative sentence (Mismatched condition; But Pat s[awed]). Both parts of each sentence were from the same talker. In a balanced design, participants heard one member of each sentence pair (either But Pat sawed them or But Pat’s awed them) in one condition, Matched or Mismatched.

For each target word (e.g. awed), participants saw it written as they listened to up to seven sentences systematically designed to resemble the critical sentence in different ways (e.g. So Pat’s auctioned them all off; We were sorry to have bored them), spoken by a selection of the familiar and unfamiliar talkers. Participants pressed a button when they heard the target word. Dependent variables were accuracy and reaction time (RT). When the listener knew the talker’s production patterns, allophonic match was expected to be facilitative, and mismatch was expected to be disruptive. In contrast, allophonic match/mismatch was expected to have no effect when the talker was unfamiliar. These predictions were supported for words that lacked an obstruent onset (e.g. awed, rain) but not for words with obstruent onsets (e.g. sawed, strain). For words without obstruent onsets, familiarity and allophonic match increased response accuracy independently, but interacted in RT: RT to familiar talkers was faster than to unfamiliar talkers when the allophones matched, and slower than to unfamiliar talkers when the allophones mismatched, while allophonic match/mismatch did not affect RTs to unfamiliar talkers. For words with obstruent onsets, effects were absent or too weak to achieve statistical significance.

These results for words without obstruent onsets are consistent with the claim that memory for speech and language is neurally organised such that the same sensory information feeds segmental, prosodic, speaker-identity-related and attitudinal percepts. More work is needed to elucidate why responses to words with onset obstruents behave differently. [Supported by the A.H.R.B.]