Variability in the prosodic realization of remote past in African American English

Kristine Yu & Alessa Farinella
University of Massachusetts, Amherst

In African American English (AAE), the remote past can be indicated with the so-called "stressed been" aspectual marker (*BIN*), e.g., *Lola BIN married* means 'Lola has been married for a long time and is still married.' *BIN* has been anecdotally described as stressed or having a high tone or pitch contour [1, 2]. The scant extant acoustic work indeed reports that *BIN*—on average—has higher intensity, duration, and f0 than surrounding material [3, 4], as exemplified in the "big" *BIN* f0 peak of Fig. 1a. But [2, 4] also mentioned instances where *BIN* had a "small" f0 peak preceded by a much larger peak, e.g., Fig 1b. Fig. 1a, b are different renditions of the same *BIN* utterance elicited in [4] in the same remote past context from a small, historically-segregated African American community in southwest Louisiana. This paper re-analyzes data from [4] (8 speakers: 5F, 3M, 570 tokens) to look at within-category variability in *BIN* utterances, asking whether multiple acoustic correlates known to signal prosodic prominence co-vary within individual *BIN* utterances in a trading relation or enhancement relationship [5]. In realizations where there is a "big" *BIN* f0 peak, the intensity on *BIN* might be lower (trading) or higher (enhancement) than realizations where there is a "small" *BIN* f0 peak. Similarly, a "smaller" *BIN* f0 peak may be compensated by more f0 range reduction in the post-*BIN* region than for "bigger' peaks.

Because "big" vs. "small" *BIN* variants have been anecdotally described in terms of relative f0 peak height between *BIN* and in the pre-*BIN* region, we computed the 90th percentile f0 value in pre-/on-/post-*BIN* regions to create a 3-point f0 shape profile of the utterance reminiscent of topline parameterizations of f0 contours [8] (Fig. 1). Popular methods using fine-grained f0 shape components (FPCA, GAMMs, time series clustering), which rely on smooth, continuous curve trajectories, were not appropriate for this study, as our utterances weren't well-controlled and included many voiceless intervals where f0 is not well-defined. The choices of the coarse 3-point topline and 90th percentile were robust to segmental perturbations/f0 contour interruptions and extendable to large analysis windows in less controlled, spontaneous speech. The topline parameterization (Fig 2) revealed that some speakers tended to produce "smaller" *BIN* variants, e.g., la04, and others "bigger" variants, e.g., la09, suggesting the possibility of speaker-specificity [10] in *BIN* realization, even when expressing the same remote past meaning. Fig. 3 shows that a "smaller" peak on *BIN* relative to the pre-*BIN* region was not compensated for with greater post-*BIN* f0 range compression. Smaller peaks predicted less post-*BIN* compression (LMM of *BIN*/post-*BIN* f0 diff. ~ *BIN*/pre- *BIN* f0 diff.: $\beta = 0.63$, $t = 9.9$, $p < 2e\text{-}16$), while "bigger" *BIN* peaks predicted more.

We also computed 90th percentile 3-point toplines for intensity and ratios of duration (normalized for speech rate, number of syllables in the region) between *BIN* and the pre- and post-*BIN* regions. Higher intensity and longer duration may contribute to enhancing a higher pitch percept ([6], [7]), but [5] found evidence of a trading relation between f0 peak size and timing of an accent and duration. We found that, while normalized *BIN* duration was overall longer than expected (median ratio 1.72) based on pre-*BIN* syllable durations, it varied freely whether the *BIN* f0 peak was "bigger" or "smaller" relative to the pre-*BIN* region. We did find some initial evidence that "smaller" *BIN* f0 peaks relative to the pre-*BIN* region correlated with "smaller" *BIN* intensity peaks (LMM of *BIN*/pre-*BIN* intensity diff ~ *BIN*/pre-*BIN* log f0 diff: $\beta = 5.5$, $t = 5.8$, $p = 2e\text{-}8$, intercept n.s. different from 0, $p = 0.764$). All together, we found no evidence for trading relations indicative of a prominence target achieved by balancing f0 topline shape components, or of f0 with duration and intensity. The notable phonetic variability from "big" to "small" *BIN* realizations, despite a common remote

past semantic meaning, corroborates [5] in cautioning against relying only on "phonetics-first" approaches to discovering intonational categories, like in f0 contour clustering [10].

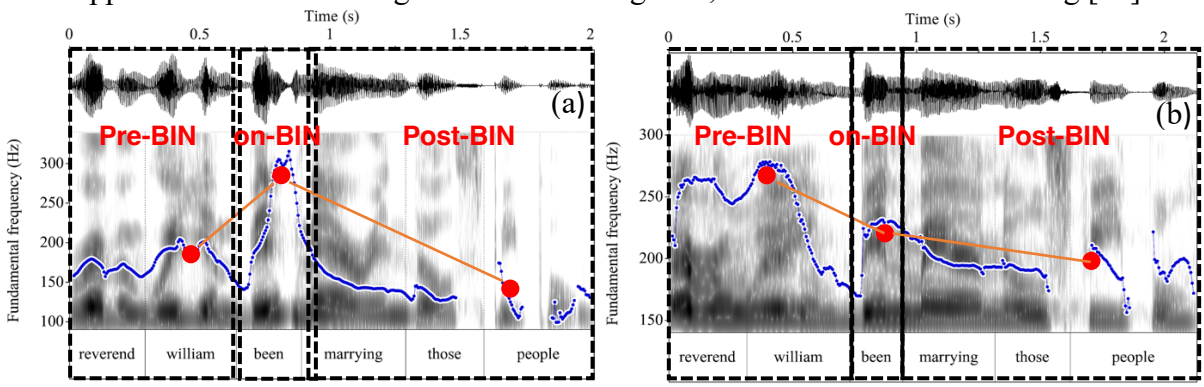

Fig 1. (a) Spkr. la09: "big" *BIN* f0 peak variant. (b) Spkr. la04: "small" *BIN* f0 peak variant. Red dots/lines show topline connecting 90th percentile f0 within each *BIN* region (see Fig. 2)

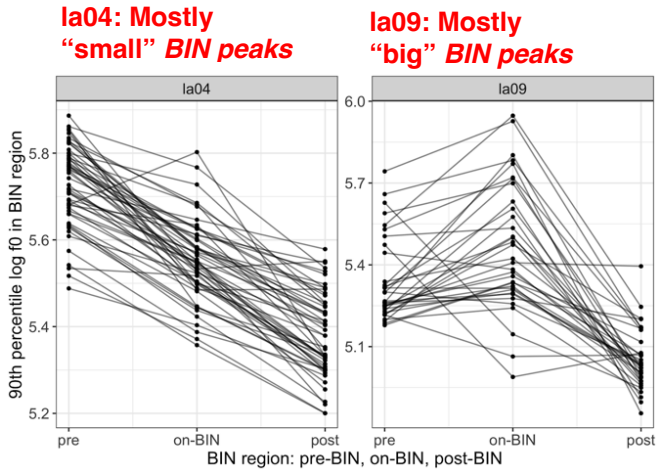

Fig 2. Log F0 topline parameterization of BIN variants shapes for la04 and la09.

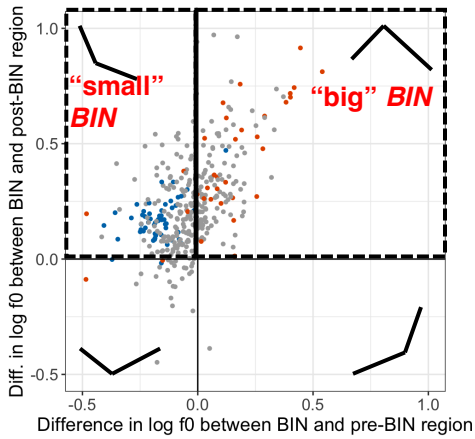

Fig. 3. Topline variation: diff. in log f0 between BIN regions, la04: blue, la09 red.

References

[1] Spears, A. K. (2017). Unstressed been: Past and Present in African American English. American Speech, 92(2): 151-175.

[2] Weldon, T. L. (2019). Race, Class, and Linguistic Camouflage: Remote Past BEEN and the Divergence Debate. In The Routledge Companion to the Work of John R. Rickford.

[3] Beyer, T., et al (2015). "Misintserpretation of African American English BIN by adult speakers of standard American English." Language & Communication 45: 59-69.

[4] Green, L. et al. (2022). "Range in the Use and Realization of BIN in African American English." Language and Speech 65.4: 958-1006.

[5] Arvaniti, A. et al. (to appear). Variability, overlap, & cue trading in intonation.

[6] Oxenham AJ (2012). Pitch perception. J Neurosci 32:13335–13338.

[7] Steffman, J., and S-A Jun (2019). "Perceptual integration of pitch and duration: Prosodic and psychoacoustic influences in speech perception." JASA 146.3: EL251-EL257.

[8] Cooper, W. E., & Sorensen, J. (1981). Fundamental Frequency in Sentence Production.

[9] Cangemi, F., et al.. (2016) Speaker-specific intonational marking of narrow focus in Egyptian Arabic. SP-2016, 335-339.

[10] Kaland, C. (2023). Contour clustering: A field-data-driven approach for documenting and analysing prototypical f0 contours. JIPA, 53(1), 159–188.