

Infinitival Ain't in African American English

Kimberley Baxter and Jonathan Stevenson

The present paper documents the first stage in the creation of an atlas of African American English (AAE) syntax, charting the relative use of the uninflected form of ain't (inf.aint) (1) against didn't (2) in a large Twitter corpus spanning three years from 2012-2015.

- (1) I ain't see that coming (inf.aint)
- (2) I didn't see that coming (inf.didnt)

The semantic near-equivalence of (1) and (2) allows us (with caveats) to consider the two forms as variants of a single variable (Labov et al., 1968; Wolfram and Schilling-Estes, 2016; Fisher, 2018) whereby the relative frequency of inf.aint may be measured against inf.didnt to provide a reliable index of use across and between places. Meanwhile, the scale of the data available via Twitter's API allows for unprecedented resolution at the level of small towns and suburbs.

In line with previous studies that use Twitter data for dialect research (Eisenstein, 2013; Jones, 2015; Stevenson, 2016; Willis, 2020; Strelluf, 2019, 2020), results in many cases appear to follow established dialect 'faultlines' (Eisenstein, 2013, p.1) whilst also highlighting particular hotspots of use. Variation appears to correlate with African American population distribution, supporting the claim that inf.aint is indeed associated with AAE.

The distribution also supports the notion that inf.aint is a relatively recent development in AAE, innovated amongst populations that travelled to the Northern industrial areas of the US during the Great Migration. This is shown in notably higher pockets of use in those regions in the Twitter corpus as well as substantial variation between places. Further, the finding that inf.aint was originally rare in early AAE in the Southern US (Kautzsch, 2000), is partially supported, with on average lower rates on Twitter in many parts of the South Atlantic region that otherwise have high rates of AAE use, such as Atlanta, GA. However, the atlas highlights some localised areas of use that buck this trend as well as sharp boundaries that warrant more detailed investigation. Results show that inf.aint is extremely widespread, with some areas such as Gary, Indiana and Augusta, Georgia exhibiting up to 40% usage in comparison with inf.didnt. Whilst there is semantic equivalence between inf.aint and inf.didnt, equivalence between other AAE forms and their Mainstream American English (MAE) counterparts is only partial and elusive. This is true for forms such as habitual be, and perfective done). The inf.aint data, then, can provide a metric against which the frequency of otherwise less trackable—but more stable—structures may be compared.

The project ultimately aims to produce a functioning syntactic atlas which can serve as a rich resource for further investigation by the academic community. Whilst the initial focus is on Twitter, the atlas will ultimately combine a range of methodological approaches and hopes to foster collaboration between researchers.