CAMBRIDGE
UNIVERSITY PRESS

Quotation in earlier and contemporary Australian Aboriginal English

Celeste Rodríguez Louro[1], Glenys Dale Collard[2], Madeleine Clews[3], and Matt Hunt Gardner[4]

[1]The University of Western Australia, Australia, [2]The University of Western Australia, Australia, and Mallee Aboriginal Corporation, Australia, [3]The University of Western Australia, Australia and [4]University of Oxford, UK
Corresponding author: Celeste Rodríguez Louro. Email: celeste.rodriguezlouro@uwa.edu.au

Abstract

We examine constructed dialogue in a longitudinal corpus of Australian Aboriginal English (AE) spoken in Perth, Australia. We conduct a variationist analysis of naturalistic data from forty-six L1 speakers of AE born 1907–2005. We ask, regarding the use of quotative frames, whether AE has changed in line with settler colonial Englishes. We examine whether a division of labor exists in the use of quotative frames, and whether the rise of first-person-marked internal thought reporting attested in settler colonial Englishes is present in AE.

Our statistical modeling shows functional partitioning in how quotative frames are used, with AE speakers strongly encoding direct speech across time. We find that the rise of first-person-marked internal thought reporting has not been systemic in AE. Despite *be like*'s incursion after 1983, the underlying system of AE has not changed. The cultural prerogative to encode speech remains strong despite sustained contact with non-First Nations Australia.

Keywords: Australian Aboriginal English; quotation; direct speech; stability; change

Sociolinguists have long been interested in the actuation problem: why some linguistic varieties undergo specific structural change while others do not (Weinreich, Labov, & Herzog, 1968:112). Here, we engage with this question by examining constructed dialogue, the reproduction of speech, thought, gesture, and sound effects (Tannen, 1986) in a longitudinal corpus of Australian Aboriginal English (AE) spoken in Perth, Western Australia. AE is a post-invasion contact variety, or group of varieties, "spoken by the majority of Aboriginal people throughout Australia" (Eades, 2014:417). These varieties differ in terms of their placement on a continuum ranging from acrolectal to basilectal features (Rodríguez Louro & Collard, 2021a). Acrolectal AE varieties are relatively close to standardized Australian English (AusE); basilectal AE varieties are closer to Kriol, an English-lexified creole spoken in Northern Australia (Schultze-Berndt, Meakins, & Angelo, 2013). Here, we consider "standardized AusE" to be the codified variety used in mainstream institutions, including schools and courtrooms. A longstanding question has been whether AE should be viewed

as a single ethnolect or whether one should speak of "Aboriginal Englishes" (Dickson, 2020:136). We focus on an acrolectal variety of AE spoken in Perth, Nyungar country, Southwest Western Australia. We use the label "AE" to refer to "one major variety of Aboriginal English, which embraces a number of regional varieties" (Malcolm, 2018:124).[1]

A pressing issue for L1 speakers of AE is that of contact with standardized AusE. The AE speakers in our corpus have all worked, lived, and/or attended school alongside speakers of AusE. They have had "more than superficial contact" with AusE speakers, with increased "potential for transfer on all levels of linguistic structure" (Meyerhoff, 2009:298). Additionally, First Nations youth are avid social media and digital technology users (Rice, Haynes, Royce, & Thompson, 2016:3). While specific linguistic material (such as *be like*) may not have spread in this way (Tagliamonte, 2014), these media are likely to impose a normative influence that is not necessarily ethnolect-specific.

The potential for linguistic transfer is heightened in global cities such as Perth. Globalization has recently been linked to the ingress of linguistic variants readily taken up by metropolitan youth in mainstream communities. Quotative *be like* is one such variant, and the quotative system of "settler colonial Englishes" (Denis & D'Arcy, 2018; we adopt the term throughout) has changed drastically since its emergence. For example, in standardized AusE, *be like* has ousted *say* as the most widely used quotative frame and it has been documented among speakers born as early as 1967 (Rodríguez Louro, 2013). Additionally, the expansion of *be like* in settler colonial Englishes has been linked to the rise of psychological drama in narrative, with speakers encoding their thought as monologic performances that, starting in the late 1960s, led to the rise of reporting first-person inner dialogues, "the very niche of *be like*" (Tagliamonte, D'Arcy, & Rodríguez Louro, 2016:839-840). In fact, D'Arcy (2020) has argued that the very reason *be like* became so widespread in settler colonial Englishes is that the entire quotative system was already in flux before *be like*'s emergence. Globalization may have helped *be like* diffuse across settler colonial communities (Buchstaller & D'Arcy, 2009; Tagliamonte et al., 2016), but storytelling in those communities already favored thought and attitude reporting in narrative. *Be like* arrived to fill an existing functional niche and subsequently spread to other kinds of constructed dialogue reporting (see Gardner, Denis, Brook, & Tagliamonte, 2021).

This scenario offers an ideal testing ground. Storytelling "is an essential element of Aboriginal cultural life" (Malcolm, 2018:10), and quotation occupies a premium place in the complicating action of a narrative (Rodríguez Louro & Ritz, 2014:560). Given sustained contact between AE and AusE, has the quotative system of AE speakers in Nyungar country changed in line with the global shifts in settler colonial Englishes, including AusE? What do any similarities/differences tell us about what is systematic and unique in AE? Additionally, in the quotative system of earlier and contemporary AE, which patterns indicate continuity, and which ones signal change?

Variationist research into First Nations languages (including contact varieties like AE) is still emergent vis-à-vis variationist studies elsewhere (Stanford, 2016). Apart from Dickson and Durantin (2019), Dixon (2017), Mailhammer (2021:149–156), and Clews, Rodríguez Louro, and Collard (2021), English-lexified postcontact varieties in Australia have received no attention from the variationist literature. Drawing on Rodríguez Louro and Collard's (2021b) decolonial variationist fieldwork methods,

this paper offers a variationist analysis of longitudinal AE as used in Nyungar country. Our study is important because it shows how a global change may enter an indigenized variety, and because research generalizations are much improved with the participation of diverse populations.

This paper is organized as follows. We first review research into English quotation. We then introduce the data and methods used in our study, followed by our distributional results and statistical models. We finally present our findings vis-à-vis existing trends in the extant literature, underscoring the entrenched practice of direct speech encoding in AE.

Background

In Australia, using First Nations languages is key to signaling connection to one's country, and First Nations storytellers have been known to show association to their country—ancestral land, oceans, waterways, and skies—by switching the language used to fit with where in First Nations country the characters are located (Rumsey, 1993). Importantly, constructed dialogue in spoken language—the reproduction of speech, thought, gesture, and sound effects—has traditionally been used in First Nations Australia to establish connection to country (Moll, 2011:247).

In English, constructed dialogue is achieved syntactically through quotative frames—quotative verbs such as *say, think, go, be like*, etc., as well as the null form (Ø) or "zero" quotative (Mathis & Yule, 1994)—that allow narrators to "take on the voices of characters by shifts of pitch, amplitude, voice quality, prosody, and pacing" (Tannen, 1986:320). Here, we analyze these frames in main clauses extracted from our corpus.

(1) I **said**, "Uncle's ere not far, I know them smells." (TB/m/82/1908)[2]
(2) They coming toward me and I'**m thinking**, "Oh well I've got to get up and get in this trailer." (VB/w/57/1962)
(3) And the people were, oh you know, they'**re going**, "What's the good of it, what's the good of it?" (NH/m/74/1926)
(4) This lady crying, crying crying like old Aboriginal, **Ø** "Ngaa, ngaa, ngaa [Nyungar crying sound]." (SC/w/88/1932)
(5) And she'**s like**, "I didn't want to go in the car with them, Auntie." (JMC/w/34/1985)

Although quotative system changes have operated since at least the mid-eighteenth century (D'Arcy, 2017), the overall distribution of quotative frames in settler colonial Englishes has changed significantly since the 1970s, characterized by the rapid uptake of quotative *be like*, which allows the speaker to suggest that the quoted content is merely "a kind of thing" the person was saying or thinking (Tannen, 1986:321). Synchronic research on this change is prolific (Ferrara & Bell, 1995; Gardner et al., 2021; Rodríguez Louro, 2013; Tagliamonte et al., 2016; Tagliamonte & Hudson, 1999; Winter, 2002). Studies have also focused on how constructed dialogue patterns across discourse types or genres (Rodríguez Louro, Richard, & Bharadwaj, 2020) and on diachronic trends in settler colonial Englishes (Buchstaller, 2011; D'Arcy, 2012, 2018; Rodríguez Louro, 2016).

The rise of *be like* has been observed in ethnolectal varieties of American English, where its high frequency and correlation with reported thought were strikingly similar between white and racialized speakers (Bayley & Santa Ana, 2004; Cukor-Avila, 2012; Fought, 2003; Hansen-Thomas, 2008; Kohn, 2008; Sánchez & Charity, 1999). Comparatively little is known about quotation patterns in postcontact English varieties beyond African American Language. For Trinidadian English, Deuber, Canan Hänsel, and Westphal (2021:450) found that 30% of the quotative verbs extracted from their corpus of conversational interactions were instances of quotative *be like*, which is favored by young women, in line with global trends. Similarly, Höhn (2012:288) concluded that *be like* "has successfully spread to Jamaican English."

The literature is scant regarding how quotation is used in indigenized English-lexified varieties in the southern hemisphere. A notable exception is D'Arcy's (2010) study of quotation in Māori English. D'Arcy (2010) found that "zero" is frequent and that it encodes mimetic speech (speech that includes gesture, facial expressions, and sound effects), patterning in opposition to mainstream New Zealand English.

In her discussion of the system-internal rise of *be like* in settler Englishes, D'Arcy (2020) has recently established a distinction between the rise of *be like* as a system-internal change (from below) as opposed to an imposition from an external system (a change from above). D'Arcy argued that, longitudinally, the rise of *be like* as an encoder of first-person-marked thought and attitude is less salient than underlying and socioculturally relevant systemic shifts in how narrative is formed, including the "lionization of self-revelation" (Ferrara & Bell, 1995:283) characteristic of American English. In sum, D'Arcy's argument established that the reason *be like* has become so widespread is due to a change in the way that stories are told. While erstwhile stories included mostly direct speech in the third person, modern storytelling is replete with first person-marked thought and attitude. What remains to be established is whether this shift in how stories are told applies across cultures, and in post-contact indigenized varieties. Here, we explore whether some of these changes hold in AE.

Our investigation includes rigorous statistical analysis as well as questions about the deep-seated nature of reported speech or thought, its purpose in discourse, and the cultural forces that shape and drive it.

Data and methods

We merge synchronic and diachronic datasets to offer a broad longitudinal view of AE direct quotation. Our corpus consists of two sources of acrolectal AE data collected in metropolitan Perth: (1) a synchronic corpus compiled in 2019/2020; and (2) oral histories recorded between 1977 and 2013 (Clews, 2020), including a collection of first-hand narratives told by Nyungar Elder Tom Bennell, recorded in 1988 and compiled in the collection titled *Kura* (Collard, 1993). These materials amount to 758,556 words of unscripted speech data from twenty-eight women and twenty men born between 1907-2005 who speak AE as their L1. It is likely that some of the speakers in our corpus also speak standardized AusE, and most do not speak Nyungar. However, we did not seek out this information during our data collection

in 2019/2020, nor was it available for the oral histories. Tom Bennell, whom Collard recorded in 1988 for *Kura*, spoke Nyungar and AE but did not speak standardized AusE. Our data represent an authentic speaking act (*yarning*) that can be viewed as typical of AE. The details of our sample are shown in Table 1.[3]

While the diachronic corpus was sourced from oral histories and first-hand narratives compiled in *Kura* (Collard, 1993), our synchronic corpus was secured under First Nations leadership, using yarning sessions led by Collard in 2019/2020 (see Rodríguez Louro & Collard, 2021b). Yarning is a First Nations cultural form of storytelling and conversation. It has been described by Terszak (2008:90) as "a process of […] passing on history and knowledge." The diachronic and synchronic materials have in common the presence of casual dialogic data, including storytelling.

To ensure comparability with the extant literature, our data extraction and coding protocols draw on existing practice. For example, in line with variationist studies of English quotation (e.g., Buchstaller & D'Arcy, 2009; Gardner et al., 2021; Rodríguez Louro, 2013; Tagliamonte et al., 2016), the envelope of variation was functionally circumscribed to include all uses of direct quotation, internal thought and nonlexicalized sounds and gestures by self or others. All instances aligning with this criterion were extracted, including those introduced by an overt verb of quotation (e.g., *say*, *think*, *go*, *be like*), and those prefaced by "zero."

Each quotative token was coded for speaker year of birth, corpus/time of interview, and speaker gender (woman/man). The linguistic factors considered include grammatical person of the quotative verb's subject, tense/temporal reference and aspect of the quotative verb, and content of the quotation.

Speaker gender has been shown to have variable influence in the use of direct quotation, with significant effects noted for quotative *be like*. However, these findings are inconsistent across varieties. For example, while standardized American English-speaking women favor *be like* (Blyth, Recktenwald, & Wang, 1990), this is not the case in other settler colonial Englishes (Buchstaller, 2014:109).

Speaker year of birth and time of interview were included to assess diachronic change. Previous research on settler colonial Englishes has established that people born in the 1960s constitute the first generation of *be like* users (Tagliamonte et al., 2016:838). Across studies, *be like* is increasingly favored by those born in the 1980s and later at the expense of more established quotative frames such as *say*

Table 1. Longitudinal corpus of Australian Aboriginal English (AE) used in this study

Corpus	Collection years	Years of birth	Speaker *N*	Token *N*
Diachronic Corpus of AE (DCAE)	1977, 1980, 1984, 1988, 1989, 1990, 1996, 2000, 2001, 2007, 2012, 2013	1907–1955	17	1257
Synchronic Corpus of AE (SCAE)	2019	1931–2005	31	1147
Total			48	2404

and *think* (Denis, Gardner, Brook, & Tagliamonte, 2019; Rodríguez Louro, 2013), suggesting a change in progress. Real-time studies have verified this change and shown *be like*'s dramatic rising over the last sixty years (Gardner et al., 2021; Labov, 2018).

Regarding grammatical person, while *say* and *go* correlate with third-person subjects across English varieties (Blyth et al., 1990; Tagliamonte & Hudson, 1999) as in (6), *be like* tends to co-occur with first-person subjects (e.g., Romaine & Lange, 1991) as in (7).

(6) And Didi, she **says**, "Oh, I've got to tell you something." (SDC/w/58/1961)
(7) We **were like**, "No, stuff it, we'll walk back home." (AC/m/16/2003)

Across Englishes, and in AusE, quotative *be like* has been shown to correlate strongly with quoted thought and attitudes (Rodríguez Louro, 2013:53). *Be like* has been argued to begin as a thought-introducer or as a form allowing the speaker to remain uncommitted regarding whether what is being encoded is speech or thought (Schourup, 1983:35). The use of quotative frames, especially *be like*, to encode thought has been linked to the rise of a narrative style characterized by psychological drama and first-person monologic performance (e.g., Tagliamonte et al., 2016:839–840). *Say* is usually strongly correlated with third-person speech (Blyth et al., 1990; Tagliamonte & D'Arcy, 2007), and quotative *go* is generally favored with third-person subjects and nonlexicalized sounds (nonvocal sounds such as hand clapping without the person also uttering the word 'clap clap') (Blyth et al., 1990).

For this analysis, quotative verbs including "zero" were coded as clearly introducing speech (8) or thought (9), where the speaker describes their own thought process in the first person, as being indeterminate between introducing speech or thought (10), or as introducing mimicked speech and/or sound effects (as in postposed *go* in [11], which—as Clews et al. (2021) found—is common in AE). Physical gesture was not considered as we lacked video recordings of the diachronic materials. The "indeterminate" category has generally not been used in the extant research, but we deem it important to identify instances that clearly encode speech or thought from those that may ambiguously encode either of these.

(8)   They **was like**, "That means 'fuck'." (CAMJ/w/18/2001)
(9)   I **was like**, "You've got to be joking." (GC/w/60/1958)
(10) I turned on the lights and I **was like**, "What the fuck?" (MJW/w/19/2000)
(11) Anyway you just listen, then "Phoo" the wind it **went**. (SCC/w/87/1932)

The present tense is a persistent structural constraint on *be like* (Blyth et al., 1990). Following D'Arcy (2004:331), recent research has refined definitions, establishing a distinction between surface morphology and temporal reference (Tagliamonte & D'Arcy, 2007). The pertinent contrast is between forms that are morphologically and referentially present and forms that are morphologically present but refer to a past situation (i.e., the *historical present*). We coded each quotative token for one of the following categories: *present*, present temporal reference with present tense

morphology, as in (12); *past*, past temporal reference with past tense morphology, as in (13); *historical present*, past temporal reference with present tense morphology, (14); and *other*, which included future, as in (15), or modals, present perfect, and cases where tense morphology is unexpressed, as in (16).

(12) Many times, I just **say**, "Look I'm from Perth." I **say**, "Look, I don't go out to the country. I only go there for funerals and that. That's all." (CC/m/36/ 1983)

(13) He **said**, "Anyone know what a *yongka* [Nyungar for 'kangaroo'] is?" (HH/ m/59/1960)

(14) And then I looked back and I **think**, "Ooh, they are really gone." (SDC/w/58/ 1961)

(15) Being two Nyungars and like you've got a darker parent than the other, you know, men, because they have that real issue of when the kids are born, they're born fair and they'**ll say**, "That's not my kid." (VB/w/62/1957)

(16) And her dad just **say**, "Go away, girl. Go away. You don't want to drink this here." (JC/w/74/1945)

Aspect/mood conveys the nature of quoted content's placement within its temporal frame—its duration, whether spoken or thought utterances are once-off or, instead, represent ongoing, repeated, habitual action or a fixed state of mind or being. Aspect/mood has been observed in mainstream English quotation to have shifted over time from relatively simple forms (e.g., past tense with perfective aspect) to increasingly complex constructions. D'Arcy (2012:354), for example, noted a complete absence of any nonperfective aspect in the quotation of the earliest-born speakers of New Zealand English, with subsequent generations starting to use "a range of habitual collocations (*would say*, *used to say*, *used to go around saying*)" (see also Gardner et al., 2021). Given the distribution of simple and complex constructions in our data, we coded aspect as either progressive (17) or nonprogressive (12-16) above.

(17) The teacher asked me what my name was, and I **kept saying**, "Koori, Koori." (CB/m/75/1932)

We now examine these sociolinguistic predictors to address whether the shift in how stories are told in settler colonial Englishes is also apparent in AE.

Results

The longue durée of AE quotation

Our findings offer a continuous link between present-day AE quotation and the early twentieth century. We begin by considering the overall distribution of the most frequently used quotative verbs in our two corpora, shown in Table 2.

Overall, 69% of direct quotation is expressed through *say*. "Zero" follows at 15%, while the remaining 16% of the quotative verbs in the sample include "other"[4] quotatives such as *tell*, *ask*, and *sing out*, as well as *think*, *go*, and *be like*. *Be like* is only

Table 2. Overall distribution of quotative verbs in the Diachronic Corpus of Australian Aboriginal English (DCAE) and Synchronic Corpus of Australian Aboriginal English (SCAE)

Quotative verb	DCAE (b. 1907–1955)		SCAE (b. 1931–2005)		All Speakers	
	%	N	%	N	%	N
say	76	956	62	711	69	1,666
"zero"	15.5	195	14	161	15	356
other	4	50	4.6	53	4.5	103
think	3.7	47	4.9	56	4.5	103
go	0.7	9	7.7	89	4	99
be like	0	0	6.7	77	3	77
Total	52	1,257	48	1,147	100	2,404

attested in the SCAE, where it occurs in 6.7% of the data. As Figure 1 shows, this is because, aside from one notable exception,[5] *be like* is only used among speakers born after 1980 and only the SCAE includes speakers born this recently. As Table 2 shows, the use of both *be like* and *go* in the SCAE corresponds with a decline in *say*; additionally, the relative incidence of *think*, "zero," and other quotatives is consistent across both corpora. In the remainder of our analysis, we merge real and apparent time and consider the DCAE and SCAE as a single longitudinal corpus.

The distribution of individual quotative frames, including "zero," across speaker decade of birth is shown in Figure 1. Not only is *say* the favored quotative verb across time, importantly, it also does not show the systematic "trajectory […] of decline" (D'Arcy, 2020:81) attested in mainstream Englishes, including standardized AusE (Rodríguez Louro, 2013:60). The trajectory of *say* in AE is wave-like, recurrently declining and increasing as other more semantically rich, "other" quotatives (though
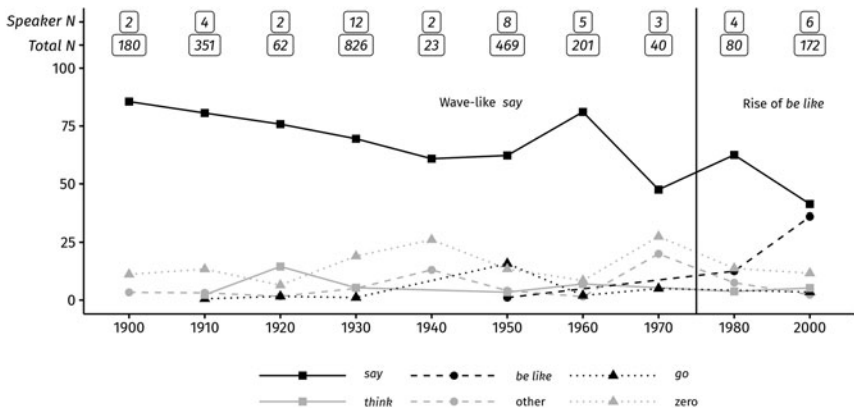

Figure 1. Relative frequency of quotative verbs over time in our Australian Aboriginal English (AE) longitudinal corpus (speakers born 1907–2005).

not *be like*) are used (e.g., among speakers born in the 1940s and especially the 1970s). Equally notable is the rise and fall of the "zero" form in conjunction with the rise and fall of *say* among those born in the 1930s, 1940s, and 1970s: When *say* recedes, the "zero" quotative expands in use. Nonetheless, *say* remains the majority quotative verb across the generations and is still used about 52% of the time among the youngest speakers (born 1980–2005). Conversely, Rodríguez Louro's (2016:147) analysis of mainstream AusE quotatives attested a decline in the use of *say* from 70% (speakers born 1920–1930) to a mere 11% among young speakers (born 1985–2000).

Aside from one notable outlier (see note 5), quotative *be like* is only attested in our data among the youngest speakers (born 1980–2005, right of the vertical line, Figure 1), and it is the second most widely used form (after *say*) among speakers born after 2000 (36%). Tokens of *be like* stem from interactions with a variety of interlocutors, from community elders to peers. Though not explored here, how diverse audiences impact the frequency of *be like* use in AE is an intended stream of future inquiry. Importantly, though the well attested cross-over between *say* and *be like* (D'Arcy, 2020:82; Tagliamonte et al., 2016:832-834) is absent in our AE data, the resolute upward trend in Figure 1 suggests a change in progress, and that, all things being equal, *be like* will eventually overtake *say*.

To further investigate diachronic trends in our data, we employ a Conditional Inference Tree (CIT) analysis, which can "expose the quantitative structure of a data set, pinpointing fine-grained distinctions among predictors" (Tagliamonte et al., 2016:832).[6] The data are partitioned into sections "based on which data points are most different from each other" (Schnell & Barth, 2018:64). The partitioned trees indicate where there are statistically significant differences between predictor levels if all other input predictors are simultaneously considered (Hothorn, Hornik, & Zeileis, 2006; Schweinberger, 2022). To enhance Figure 2's readability, *say* and *be like* are plotted individually while "zero," *go*, *think*, and "other" quotatives (Table 2) are grouped together and labeled "various." Input predictors are speaker year of birth (YOB) and gender (woman, man). While speaker gender is nonexplanatory, speaker year of birth is highly predictive of quotative verb selection. A CIT that additionally included corpus (DCAE versus SCAE) as a predictor did not produce an informatively different tree. The first significant partition in Figure 2 is between those born in 1983 and earlier, and those born post-1983 ($p < 0.001$). These two groups are selected as different because speakers born after 1983 use *be like*, while their predecessors do not (though see note 5). The second significant partition is among those born in 1911 and before, and those born after 1911. Speakers born before 1911 show a remarkable preference for *say*. Those born after 1911 continue to prefer *say*, but this use declines considerably, and direct reported material is encoded via "various" quotative verbs, including "zero," quotative *go*, and a host of other minority forms such as *ask*, *sing out*, and *yell*. This CIT analysis confirms that birth year 1983 marks a critical diachronic boundary in our data.

Multivariate analysis

So far we have shown that, although fluctuating, AE does not show the overall decline that has led to the recent shrinking of *say*'s usage frequency in mainstream varieties.
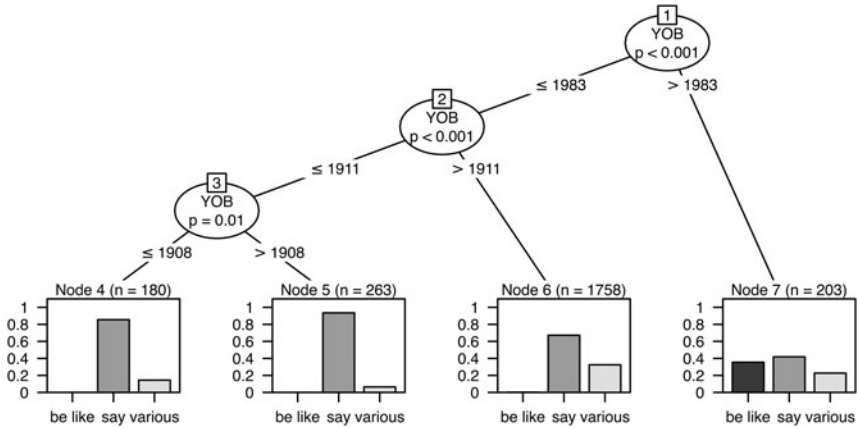
Figure 2. Conditional inference recursive partitioning tree for the realization of different quotative verbs with YEAR OF BIRTH and GENDER as input predictors ($n = 2404$).

We have also established that our AE sample is significantly partitioned in two: those born in or before 1983 and those born after 1983. We now report the results on mixed-effects logistic regression modeling of the linguistic predictors of GRAMMATICAL PERSON OF THE SUBJECT, CONTENT OF THE QUOTATION, TENSE/TEMPORAL REFERENCE, and ASPECT on quotative choice. Some pruning of the data was in order. Quotative *be like* clauses featuring nonreferential subjects (e.g., *It's like, "What are you following me for?"* JMC/w/34/1985), where *it* does not encode an individual or an entity, were excluded from the statistical analysis ($n = 5$) (see Tagliamonte & D'Arcy, 2004, 2007). Additionally, as the following regression analyses examine the potential effects of tense/temporal reference and subject grammatical person on quotative choice, quotative frames with no overt quotative verb were removed from the data. Further, for speakers born in and before 1983, the five *be like* tokens produced by the outlier mentioned in note 5 were excluded.

Table 3 presents the mixed-effects regression model of *say* among speakers born in and before 1983.[7] There are 1,860 quotative tokens from this cohort, of which 85% are *say*. The Akaike Information Criterion (AIC) is an estimator of the prediction error of this model. A lower AIC is a better fit (more explanatory) than a higher AIC. The AIC of a model built with just the random effect SPEAKER is 1,478, which is significantly higher (less explanatory) than that of the model presented ($\chi^2 = 659$, df = 9, $p < 0.001$). The marginal and conditional R^2 values show the proportion of the data explained by the fixed effects predictors (GENDER, PERSON, CONTENT, TENSE/TEMPORAL REFERENCE, and ASPECT) and the fixed effects predictors plus the random effect (SPEAKER) respectively. For this model, the fixed effect predictors are assumed to have a constant relationship with the response variable (choice of *say*) across all observations. But by setting SPEAKER as a random effect, we allow for that fixed relationship to vary from person to person (accounting for individuals having a differing overall likelihood of using *say*). The mean rate of *say* use per speaker is 82% ($\pm 17\%$). The Condition Number (κ) and Variance Inflation Factor (VIF) both assess (multi)

Table 3. Mixed-effects logistic regression testing the fixed effects of Gender, Grammatical Person of the Subject, Content of the Quotation, Tense/Temporal Reference, and Aspect, and a random intercept of Speaker on the realization of quotative *say* in Australian Aboriginal English (AE) among speakers born in and before 1983

AE, speakers born ≤ 1983							
Observations: 1,860 (overall frequency of *say* 85%, *n* = 1581)							
AIC: 847, Marginal R^2 0.58; Conditional R^2 0.67, κ 4.6, VIF ≤ 1.28							
					Observations		
Fixed Effects	Coefficient	SE	z	Sig. Level	Total N	*say*	%
Intercept (all reference values)	3.69	0.37	10.06	***			
Gender (versus men)					635	562	89
women	−0.07	0.37	−0.20		1,228	1,019	83
Person (versus second)[†]					730	652	89
third	−0.17	0.25	−0.69		313	261	83
first	−0.02	0.23	−0.08		820	668	81
Content (versus direct speech)[†]					1,737	1,573	91
other	−3.33	0.58	−5.67	***	25	7	28
thought	−8.10	0.35	−7.39	***	101	1	1
Tense/Temporal Ref. (versus simple past)[†]					1,385	1,291	93
simple present	−1.86	0.35	−5.34	***	90	60	67
other	−2.16	0.26	−8.40	***	245	156	64
historical present	−2.55	0.31	−9.12	***	143	74	52
Aspect (versus non-progressive)					1,735	1,540	89

(*Continued*)

Table 3. (*Continued.*)

AE, speakers born ≤ 1983							

Observations: 1,860 (overall frequency of *say* 85%, *n* = 1581)

AIC: 847, Marginal R^2 0.58; Conditional R^2 0.67, κ 4.6, VIF ≤ 1.28

					Observations		
Fixed Effects	Coefficient	SE	z	Sig. Level	Total N	say	%
progressive	−1.21	0.31	−3.1	***	128	41	32
Random Effects	Variance				Group n		
SPEAKER (Intercept) [††]	0.81				35		

Treatment contrast coding. Model fit by maximum likelihood (Laplace Approximation). Model converges with BOBYQA optimizer with <20,000 iterations. Coefficients reported in log-odds.
Significance level: ∗ $p < 0.05$; ∗∗ $p < 0.01$; ∗∗∗ $p < 0.001$
[†] Simultaneous tests for General Linear Hypothesis
first versus third = 0: Coef. 0.19, SE 0.27, z 0.71, Sig. Level
thought versus other = 0: Coef. −4.77, SE 1.22, z −3.88, Sig. Level ∗∗
other versus simple present = 0: Coef. −0.29, SE 0.35, z −0.82, Sig. Level
historical present versus simple present = 0: Coef. −0.67, SE 0.37, z −1.79, Sig. Level
historical present versus other = 0: Coef. −0.39, SE 0.28, z −1.38, Sig. Level
[††] mean by speaker = 82%, SD = 17%

collinearity in the data (that the differences in *say* choice explained by GENDER, for instance, may be influenced by differences in PERSON). The Condition Number here is less than six, indicating no collinearity (Baayen, 2008:182). Likewise, the VIFs—the proportion of variance of one fixed effect associated with another fixed effect—is ≤1.28 for all predictors, below the conservative tolerance of four and indicating little to no collinearity (e.g., O'Brien, 2007). In Table 3 the model is built with treatment contrast coding whereby the coefficient of the Intercept represents the likelihood (expressed in log-odds) of *say* when each of the fixed-effect predictors (factor group) are set to a reference value (factor): here, GENDER = *men*, PERSON = *second*, CONTENT = *direct speech*, TENSE/TEMPORAL REFERENCE = *simple past*, and ASPECT = *nonprogressive*. These reference values each have the highest frequency of *say* (shown by the Observations columns distributions), though their choice is arbitrary. The coefficients for each nonreference value represent the change in likelihood of *say* if that parameter is switched. Positive coefficients indicate an increased likelihood of *say*, while negative coefficients indicate a decreased likelihood of *say*. Based on the standard error (SE) and *z*-score, whether this difference in likelihood is significant is calculated and represented by asterisks. While women use less *say* than men, this difference is not significant. Likewise, while *say* is used more frequently with second-person subjects, compared to third- or first-person subjects, neither difference is significant. A simultaneous test of the General Linear Hypothesis indicates that the difference between third- and first-person subjects is also nonsignificant. In other words, for the use of *say*, there is neither a significant GENDER nor PERSON effect. The difference between direct speech and other quotation content types and between direct speech and thought is, however, significant. So too is the difference between other quotation content types and thought. Of the one hundred tokens of quoted thought, only one occurred with *say*. There is a significant difference between simple past and all remaining TENSE/TEMPORAL REFERENCE categories, but not a significant difference among the remaining TENSE/TEMPORAL REFERENCE categories. Finally, *say* is significantly more likely in nonprogressive contexts. To summarize, *say* is favored when the simple past is used, when using nonprogressive aspect, and/or when quoting direct speech, while other quotative verbs are favored elsewhere.

As Figure 3a shows, the seven speakers born after 1983 (two men, five women) have a variable quotative system dominated by *say* (women) and "zero" (men).
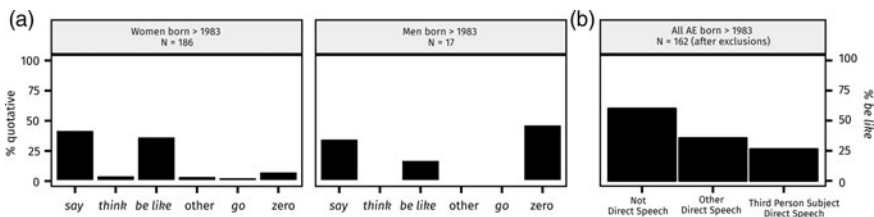

Figure 3. Frequency of different quotatives in Australian Aboriginal English (AE). Panel (a) shows the frequency of quotatives among AE speaking women and men born after 1983. Panel (b) shows the frequency of *be like* across three broad categories of tokens among all AE speakers born after 1983.

However, *be like* is making inroads, especially among young women in line with the general sociolinguistic observation that young women lead in the use of innovative variants (Labov, 2001; though see Denis et al., 2019). For this group, we attempted to construct an identical model as Table 3 with *be like* as the application value. For this analysis, the frame *it's like* must be excluded as it is singular to *be like*, and "zero" tokens must be excluded as they have no subjects and cannot be marked for tense/aspect. Given this, there were only 162 remaining quotative tokens from this cohort. The complex model, as in Table 3, did not converge with so few tokens (the model did not predict the variation in the data well). Further, there is moderate collinearity (regression models assume little to no collinearity) as there are combinations of GENDER, PERSON, CONTENT, TENSE/TEMPORAL REFERENCE, and ASPECT with few or no tokens, or categoricity. For this reason, we simplified our analysis and categorized our data from younger speakers into three broad categories: tokens that are direct speech with third-person subjects, all other direct speech tokens, and all nondirect speech tokens. The frequency of *be like* within these three categories is shown in Figure 3b. Notably, *be like*'s frequency is lowest for tokens of direct speech with third-person subjects, exactly the combination of content type and grammatical person found to resist *be like* in other communities, including mainstream AusE (Rodríguez Louro, 2013). Unfortunately, even a regression model built on the >1983 data using this three-way distinction (AIC = 235) as the only fixed effect and SPEAKER as a random effect does not explain the variability any better than a model that just includes the random effect of SPEAKER (AIC = 236; $\chi^2 = 5.72$, $p = 0.057$). So, while the rate of *be like* for these three broad categories is on-trend for the spread and rise of *be like* worldwide, we cannot validate it statistically with the current dataset.

What's going on with go?

There are only twelve speakers of forty-six who employ *go*. However, they include speakers born in 1911 ($n = 2/256$, 0.78%) and in nearly every decade in the data, including speakers born in 2001 ($n = 6/103$, 6%), though speakers born in the 1950s employ the form the most frequently ($n = 75/406$, 18%). Below we model the use of *go* relative to other overt quotative frames. Note that pruning away "zero" contexts results in forty-three rather than forty-six speakers. The model here is significantly better at explaining the variation in the data than the null model with just the random effect of SPEAKER (AIC = 630; $\chi^2 = 242$, df = 9, $p < 0.001$). The Condition Number (κ) here is 24 indicating medium, though not harmful, multicollinearity (Baayen, 2008:182), which is corroborated by the max VIF of 1.76.

The model indicates that GENDER and PERSON are not significant predictors of *go*. *Go* is significantly favored with quotation that is neither direct speech nor thought (though the difference between thought and direct speech is not significant), which includes nonlexicalized sounds and gestures. Given *say*'s strong association with reporting direct speech ($n = 1657/1887$, 88%) and *think*'s strong association with expressing internal thought ($n = 101/114$, 89%), both *be like* and *go* find their niche expressing other types of quotation.

Zeroing in on "zero"

As shown in Figure 1, there is robust "zero" in the overall dataset ($n = 356/2404$, 15%). This motivated an analysis of "zero" versus overt quotation. Figure 4, a CIT built with GENDER, YEAR OF BIRTH, and CONTENT as predictors, reveals that men use significantly more "zero" quotative frames than women, and that, among both groups, "zero" quotative frames are more likely for content that is not thought (the domain of *think*) or direct speech (the domain of *say*), such as nonlexical sounds (18), similar to *go*.

(18) So he made this sound all the night, Ø "Mm, mm, mm, mm, mm, mm," and he stopped. (RB/m/77/1930)

These trends are verified by a mixed-effects regression analysis, presented in Table 5. TENSE/TEMPORAL REFERENCE, PERSON, and ASPECT were not included as fixed effects in our analysis, as "zero" quotatives neither carry tense/aspect marking nor have overt grammatical subjects. While the regression model is not a great fit to the data (it only accounts for a fifth of the variation, Conditional $R^2 = 0.20$), it performs significantly better than the null model with only the random effect of SPEAKER (AIC = 1,898; $\chi^2 = 34$, df = 3, $p < 0.001$).

Our analysis does not explore in depth the role of mimesis and sound effects on the use of specific quotative frames, including "zero." However, the favoring effect of "zero" quotatives to introduce nonlexicalized sounds in AE is in line with D'Arcy's (2010) findings for Māori English, where "zero" quotatives encoded mimetic speech, patterning in opposition to mainstream New Zealand English.

Our results motivate the question of how quotatives may have been employed in Nyungar, the traditional First Nations language of Southwest Western Australia. The question arises as to whether traditional ways of telling stories may have contributed to the patterns noted above. A modern grammar of Nyungar does not currently exist (though see Douglas, 1976), but an examination of narrative texts in the stories *Mamamg* 'Whale' (Scott, 2011a), *Noongar Mambara Baktij* 'Man Fights Spirit' (Scott, 2011b), and *Noorn* 'Snake' (Scott, 2011c), provides insights into what
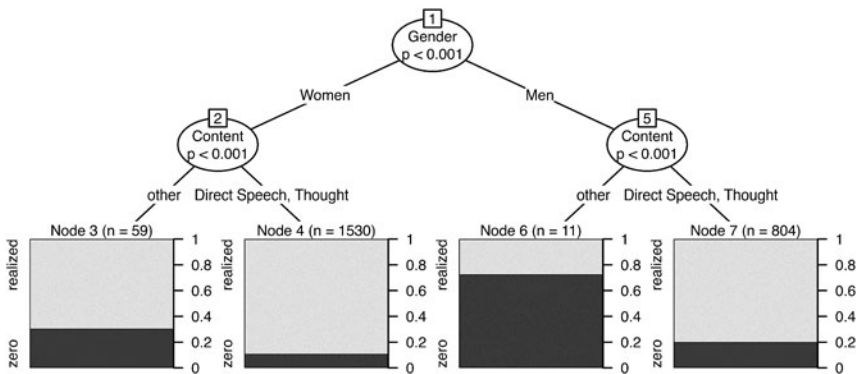

Figure 4. Conditional inference recursive partitioning tree for "zero" versus realized quotative verbs with YEAR OF BIRTH, GENDER, and CONTENT OF THE QUOTATION as input predictors ($n = 2404$).

Celeste Rodríguez Louro *et al.*

Table 4. Mixed-effects logistic regression testing the fixed effects of Gender, Grammatical Person of the Subject, Content of the Quotation, Tense/Temporal Reference, and Aspect, and a random intercept of Speaker on the realization of quotative *go* in Australian Aboriginal English (AE)

AE, all speakers							
Observations: 2,048 (overall frequency of *go* 5%, *n* = 98)							
AIC: 405, Marginal R^2 0.28; Conditional R^2 0.76, κ 24, VIF ≤1.76							
						Observations	
Fixed Effects	Coefficient	SE	z	Sig. Level	Total N	go	%
INTERCEPT (all reference values)	−3.51	1.06	3.29	***			
GENDER (versus women)					1,403	90	6
men	−0.86	0.81	−1.06		645	8	1
PERSON (versus third)†					337	18	6
second	0.23	0.40	0.30		788	42	5
first	−0.37	0.44	−0.82		923	38	4
CONTENT (versus other) †					44	20	45
direct speech	−6.34	0.85	−7.46	***	1,889	73	4
thought	−7.22	1.07	−6.78	***	115	5	4
TENSE/TEMPORAL REF. (versus historical present) †					165	43	26
simple present	−1.07	0.47	−2.25	**	109	17	16
other	−1.79	0.38	−4.72	***	292	27	9
simple past	−4.22	0.51	−8.32	***	1,482	11	1
ASPECT (versus progressive)					171	33	19

(*Continued*)

Table 4. (*Continued.*)

AE, all speakers							
Observations: 2,048 (overall frequency of *go* 5%, *n* = 98)							
AIC: 405, Marginal R^2 0.28; Conditional R^2 0.76, κ 24, VIF ≤1.76							
					Observations		
Fixed Effects	Coefficient	SE	*z*	Sig. Level	Total *N*	*go*	%
non-progressive	−0.90	0.40	−2.22	∗	1,877	65	3
Random Effects	Variance				Group *n*		
SPEAKER (Intercept) [††]	6.39				43		

Treatment contrast coding. Model fit by maximum likelihood (Laplace Approximation). Model converges with BOBYQA optimizer with <20,000 iterations. Coefficients reported in log-odds.
Significance level: ∗ $p < 0.05$; ∗∗ $p < 0.01$; ∗∗∗ $p < 0.001$
[†]Simultaneous tests for General Linear Hypothesis
first versus second = 0: Coef. −0.60, SE 0.37, *z* −1.62, Sig. Level
direct speech versus thought = 0: Coef. 0.88, SE 0.64, *z* 1.38, Sig. Level
other versus simple present = 0: Coef. −0.72, SE 0.48, *z* −1.49, Sig. Level
simple past versus simple present = 0: Coef. −3.15, SE 0.59, *z* −5.37, Sig. Level ∗∗∗
simple past versus other = 0: Coef. −2.43, SE 0.54, *z* −4.52, Sig. Level ∗∗∗
[††]mean by speaker = 2.8%, SD = 14%

Table 5. Mixed-Effects logistic regression testing the fixed effects of Gender, Content of the quotation, and a random intercept of Speaker on the realization of a null quotative in Australian Aboriginal English (AE)

AE, all speakers							
Observations: 2,404 (overall frequency of a "zero" quotative 15%, n = 353)							
AIC: 1,870, Marginal R^2 0.05; Conditional R^2 0.20, κ 13.9, VIF ≤1.00							
					Observations		
Fixed Effects	Coefficient	SE	z	Sig. Level	Total N	"zero"	%
Intercept (all observations)	0.07	0.35	0.19				
Gender (versus men)					815	170	21
women	−0.83	0.27	−3.01	**	1,589	186	12
Content (versus other)†					70	20	29
thought	−1.98	0.23	−4.62	***	126	5	4
direct speech	−1.41	0.29	−4.91	***	2,208	74	3
Random Effects	Variance				Group n		
Speaker (Intercept)††	0.60				44		

Treatment contrast coding. Model fit by maximum likelihood (Laplace Approximation). Model converges with BOBYQA optimizer with <20,000 iterations. Coefficients reported in log-odds.
Significance level: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$
† Simultaneous tests for General Linear Hypothesis
direct speech versus thought = 0: Coef. −0.56, SE 0.33, z −1.71, Sig. Level
†† mean by speaker = 2.2%, SD = 13%

quotation was like in Nyungar. These stories constitute approximately 12,300 words of Nyungar. They contain eighteen quotative frames, eight instances of the verb *wangk* 'say,' as shown in (19), and ten instances of quotation introduced without an overt quotative verb, that is, a "zero" quotative, as in (20). All eighteen quotative frames introduce speech. No quotative-introduced thought encoding is recorded in the stories. The translations into English, and the glosses, were provided by the author in the original publications. We have bolded items and added "Ø" in the gloss to represent "zero." In (20) the "stranger" sees the snake in the presence of other people who warn him to "look out."

(19) *Balaap* **wangk**, '*Nidja demangka birt   nyoondok koorliny.*
    They   **say**   this   old-people track you   going
    'Hey, they said, "Old people come this way, but not for a long time."' (Scott, 2011b:6)

(20) *Waam   baal weyarn.   Ø 'Aliwa!'*
    Stranger he   frightened.   Ø 'Look out!'
    'The stranger was frightened when he saw the snake.' (Scott, 2011c:28)

The Nyungar data depicted in (19) and (20), although limited, align with our finding that, in AE as spoken in Nyungar country, the encoding of speech in storytelling (rather than thought) remains strong.

Discussion

Based on the individual analysis of *say*, *be like*, *go*, *think*, and the "zero" quotatives in our dataset, we argue that there is functional partitioning: different quotatives are marshalled for specific quotative jobs. *Say* is the default choice for expressing direct speech and occurs most often with past tense morphology. *Think*, on the other hand, is specific to expressing internal thought. Other quotative frames, especially *go* and "zero," are selected when what is being quoted is a nonlexical sound. *Go* is especially likely when speakers use the nonpast or progressive aspect for narrating events. While a robust statistical analysis of *be like* was not possible, we have shown that *be like* is incipient for both nonlexical quoting as well as nonthird-person subjects. *Be like* is finding its niche in AE in the exact contexts where *say* and *think* do not dominate, while aligning with the global tendency to use *be like* to quote one's own words and thoughts (Tagliamonte et al., 2016).

The longitudinal perspective adopted here also offers a glimpse of the AE quotative system before *be like*. The literature has established that the quotative system of settler colonial Englishes was already in flux before *be like* emerged. Specifically, D'Arcy (2020) argued that the rise of internal thought reporting in narrative is independent of the rise of *be like* per se and that the quotative system had already reconfigured before *be like*'s appearance, facilitating its selection as the preferred introducer of a speaker's thoughts and attitudes. In settler colonial Englishes, the drift toward *be like* is now understood as related to the increased expression of internal dialogue in narrative (see D'Arcy, 2012; Rodríguez Louro, 2016).

In our Results section, we determined that *say* dominates across generations of AE speakers where it does not show the steady decline noted in settler colonial Englishes, including mainstream AusE (Rodríguez Louro, 2013:60). This trajectory is also evident when considering the overall proportion of quoted direct speech and thought. Viewed longitudinally, the AE quotative system shows conservativism, encoding direct speech to near categorical levels over time, as shown in Figure 5.

While settler colonial English varieties such as mainstream Australian, New Zealand, and Canadian English display "gradual and longitudinal" changes whereby the overall usage frequency of internal dialogue/thought represented as quotation rises "regularly and monotonically" (D'Arcy, 2020:88), AE shows stability across the generations, with quotative frames favored in the expression of direct speech rather than internal thought. The quintessential function of quotation in acrolectal AE as used in Nyungar country is that of direct speech encoding.[8]

The conservative patterns illustrated in Figure 5 show that the "longitudinal, systemic change" that preceded *be like* "by at least a full century" (D'Arcy, 2020:90) in settler colonial Englishes is not attested in AE. AE speakers have historically used quotative frames to introduce direct speech. The increase of internal thought and attitude reporting in narrative, which paved the way for *be like*'s encroachment in various settler colonial Englishes, has been absent in AE. Not only is AE's *be like* a change
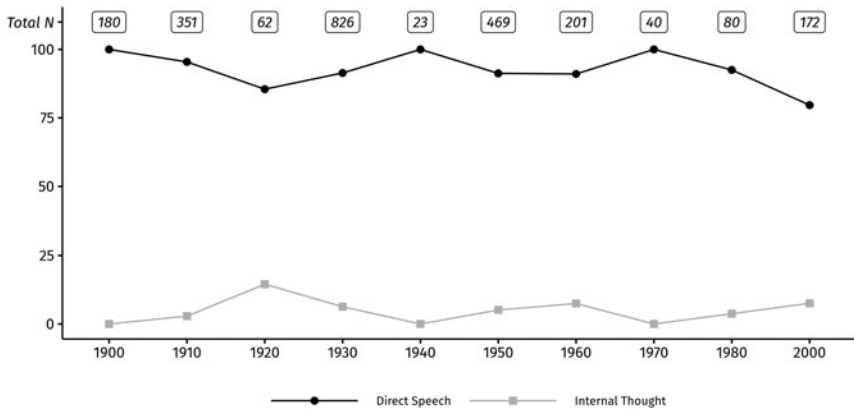
Figure 5. Overall proportion of quoted speech and thought in the Australian Aboriginal English (AE) corpus by speaker decade of birth.

from above (external to the system), the changes motivating its rise (internal thought and attitude reporting in the first person) have not been systemic in AE. The result of these processes is a quotative (*be like*) that, while on the rise, has started off on different footing in AE.

Conclusion

We have shown that, in synchrony, AE quotation is characterized by functional partitioning such that different quotative frames, including *say, be like, go, think*, and "zero," correspond to different pragmatic functions. We have also established, through the analysis of a longitudinal corpus of AE, that AE characteristically utilizes quotative devices to introduce direct speech. Crucially, this trajectory does not entirely diminish with the passing of time.

These findings indicate that AE has adopted the linguistic changes attested in mainstream Englishes (i.e., the use of *be like* as a quotative verb), but the longitudinal trajectory of this development is not entirely akin to that of mainstream Englishes, including AusE. We leave it to future research to determine whether, in the context of direct quotation in yarning, *be like* might herald deep-seated cultural change.

Competing interests. The author declares none.

Notes

1. This position is also in line with Newmark, Walker, and Stanford's (2016:655) findings of similarity across disparate North American Indigenous Englishes.

2. The parenthetical information at the end of our examples shows speaker initials, gender, age at data collection, and year of birth. Any names included in the examples are pseudonyms.

3. The data collected for this project were secured with the express permission of the speakers. The research received ethics clearance through The University of Western Australia (RA/4/20/4977).

4. The "other" category includes *tell* ($n = 44$), *ask* ($n = 18$), *sing (out)* ($n = 13$), *shout (out)* ($n = 8$), *call* ($n = 4$), *(get to) be* ($n = 4$), *cry* ($n = 2$), *reckon* ($n = 2$), *yell (out)* ($n = 2$), *feel* ($n = 1$), *question* ($n = 1$), *scream* ($n = 1$), *threaten* ($n = 1$), *warn* ($n = 1$), *wonder* ($n = 1$).

5. Our outlier is sixty-one-year-old GC, a grandmother of twenty-eight. Her use of *be like* is reminiscent of Tagliamonte and Denis' (2014:122; footnote 42) fifty-five-year-old "hip" grandmother "Leah Thompson" who, like GC, "is highly involved in the care of her grandchildren." AY, a man born in 1955, also uses one instance of *it's like*; however, as existential subjects are unique to *be like*, this token is not included in the regression analysis.

6. The Conditional Inference Tree presented as Figure 2 was created in *R* (R Core Team, 2023, version 4.3.0) using the *partykit* package (Hothorn & Zeileis, 2015; Hothorn et al., 2006).

7. The models presented in Tables 4 and 5 were built in *R* (R Core Team, 2023, version 4.3.0, using the following packages: *lme4* (Bates, Maechler, Bolker, & Walker, 2015); *multcomp* (Hothorn, Bretz, & Westfall, 2008); *MuMIn* (Bartón, 2019); *JGmermod* (Grafmiller, 2018); and *performance* (Lüdecke, Ben-Shachar, Patil, Waggoner, & Makowski, 2021).

8. The AE speakers in our corpus were of similar demographic characteristics (e.g., most had completed some primary school). No current analysis of how social class impacts the use of quotation exists. However, we refer the reader to Howard's (2012) thesis where, for nine working-class men and women who speak AusE as their L1, thought reporting represents 21%, while 79% of the quotative frames are used to encode speech.

References

Baayen, R. Harald. (2008). *Analyzing linguistic data. A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Bartón, Kamil. (2019). MuMIn: Multi-Modal Inference. R package version 1.43.15.

Bates, Douglas M., Maechler, Martin, Bolker, Ben, & Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48. https://doi.org/10.48550/arXiv.1406.5823

Bayley, Robert, & Santa Ana, Otto. (2004). Chicano English: Morphology and syntax. In B. Kormann, K. Burridge, R. Mesthrie, E. Schneider, & C. Upton (eds.), *A handbook of varieties of English, vol. 2: Morphology and syntax*. Berlin: Mouton de Gruyter. 374–390.

Blyth, Carl, Recktenwald, Sigrid, & Wang, Jenny. (1990). I'm like, 'Say what?!': A new quotative in American oral narrative. *American Speech* 65: 215–227. https://doi.org/10.2307/455910

Buchstaller, Isabelle. (2011). Quotations across the generations: A multivariate analysis of speech and thought introducers across 5 decades of Tyneside speech. *Corpus Linguistics and Linguistic Theory* 7 (I): 59–92. https://doi.org/10.1515/cllt.2011.004

Buchstaller, Isabelle. (2014). *Quotatives: New trends and sociolinguistic implications.* Malden, MA: Wiley-Blackwell.

Buchstaller, Isabelle, & D'Arcy, Alexandra. (2009). Localized globalization: A multi-local, multivariate investigation of quotative *be like. Journal of Sociolinguistics* 13: 291–331. https://doi.org/10.1111/j.1467-9841.2009.00412.x

Clews, Madeleine (2020). *Mining the archives: A study of the diachrony of quotation in Australian Aboriginal English.* University of Western Australia: Honors Thesis.

Clews, Madeleine, Rodríguez Louro, Celeste, & Collard, Glenys. (2021). Indigenizing SAY in Australian Aboriginal English. *Australian Journal of Linguistics* 41(4): 1–22. https://doi.org/10.1080/07268602.2021.2007046

Collard, Glenys. (1993). *Kura, stories by Tom Bennell*. Bunbury, WA: Excelsior.

R Core Team. (2023). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Cukor-Avila, Patricia. (2012). Some structural consequences of diffusion. *Language in Society* 41(5): 615–640. https://doi.org/10.1017/S004740451200067X

D'Arcy, Alexandra. (2004). Contextualizing St. John's youth English within the Canadian quotative system. *Journal of English Linguistics* 32: 323–345.

D'Arcy, Alexandra. (2010). Quoting ethnicity: Constructing dialogue in Aotearoa/New Zealand. *Journal of Sociolinguistics* 14: 60–88. https://doi.org/10.1111/j.1467-9841.2009.00437.x

D'Arcy, Alexandra. (2012). The diachrony of quotation: Evidence from New Zealand English. *Language Variation and Change* 24: 343–369. https://doi.org/10.1017/S0954394512000166

D'Arcy, Alexandra. (2017). *Discourse-pragmatic variation in context: 800 years of LIKE*. Amsterdam: John Benjamins.

D'Arcy, Alexandra. (2018). Linguists be like: "Where did it come from?" *New Ways of Analyzing Variation* 47. New York: New York University.

D'Arcy, Alexandra. (2020). Reconfiguring quotation over time and the system-internal rise of *be like*. In P. J. Grund & T. Walker (eds.), *Speech representation in the history of English: Topics and approaches*. Oxford: Oxford University Press. 73–101.

Denis, Derek, & D'Arcy, Alexandra. (2018). Settler colonial Englishes are different from postcolonial Englishes. *American Speech* 93(1): 3–31. https://doi.org/10.1215/00031283-6904065

Denis, Derek, Gardner, Matt Hunt, Brook, Marisa, & Tagliamonte, Sali. (2019). Peaks and arrowheads of vernacular organization. *Language Variation and Change* 31: 43–67. https://doi.org/10.1017/S095439451900005X

Deuber, Dagmar, Canan Hänsel, Eva, & Westphal, Michael. (2021). Quotative *be like* in Trinidadian English. *World Englishes* 40: 436–458. https://doi.org/10.1111.weng.12456

Dickson, Gregory. (2020). Aboriginal English(es). In L. Willoughby & H. Manns (eds.), *Australian English reimagined: Structure, features and developments*. London: Routledge. 134–154.

Dickson, Gregory, & Durantin, Gautier. (2019). Variation in the reflexive in Australian Kriol. *Asia-Pacific Language Variation* 5: 171–207. https://doi.org/10.1075/aplv.00005.dic

Dixon, Sally. (2017). *Alyawarr children's variable present temporal reference expression in two, closely-related languages of Central Australia*. Australian National University: PhD thesis.

Douglas, Wilfrid. (1976). *The Aboriginal languages of southwest of Australia*. Canberra: Australian Institute of Advanced Studies (AIAS).

Eades, Diana. (2014). Aboriginal English. In H. Koch & R. Nordlinger (eds.), *The languages and linguistics of Australia: A comprehensive guide*. Berlin: Walter de Gruyter. 417–447.

Ferrara, Kathleen, & Bell, Barbara. (1995). Sociolinguistic variation and discourse function of constructed dialogue introducers: The case of be + like. *American Speech* 70: 265–290. https://doi.org/10.2307/455900

Fought, Carmen. (2003). *Chicano English in context*. New York: Palgrave MacMillan.

Gardner, Matt Hunt, Denis, Derek, Brook, Marisa, & Tagliamonte, Sali. (2021). *Be like* and the constant rate effect. From the bottom to the top of the S-curve. *English Language and Linguistics* 25(2): 281–324. https://doi.org/10.1017/S1360674320000076

Grafmiller, Jason. (2018). *JGmermod*: Custom functions for mixed-effects regression models. R package version 0.2.0.

Hansen-Thomas, Holly. (2008). An investigation of innovative quotatives in adolescent Chicana English in Texas. *Intercultural Pragmatics* 5: 19–39. https://doi.org/10.1515/IP.2008.002

Höhn, Nicole. (2012). And they were all like 'What's going on?' New quotatives in Jamaican and Irish English. In M. Hundt & U. Gut (eds.), *Mapping unity and diversity world-wide: Corpus-based studies of new Englishes*. Amsterdam: John Benjamins. 263–290.

Hothorn, Torsten, Bretz, Frank, & Westfall, Peter. (2008). Simultaneous inference in general parametric models. *Biometrical Journal* 50(3): 346–363. https://doi.org/10.1002/bimj.200810425

Hothorn, Torsten, Hornik, Kurt, & Zeileis, Achim (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3): 651–674. https://doi.org/10.1198/106186006X133933

Hothorn, Torsten, & Zeileis, Achim. (2015). partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research* 16(118): 3905–3909.

Howard, Rachel. (2012). *'Story gets better': Performed narratives and quotatives across social classes in Australian English speech*. The University of Western Australia: Honors thesis.

Kohn, Mary Elizabeth. (2008). *Latino English in North Carolina: A comparison of emerging communities*. Master's thesis: North Carolina State University.

Labov, William. (2001). *Principles of linguistic change: Vol. II*: *Social factors*. Oxford: Blackwell.

Labov, William. (2018). The role of the Avant Garde in linguistic diffusion. *Language Variation and Change*. 30(1): 1–21. https://doi.org/10.1017/S0954394518000042

Lüdecke, Daniel, Ben-Shachar, Mattan S., Patil, Indrajeet, Waggoner, Philip, & Makowski, Dominique. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software* 6(60): 3139.

Mailhammer, Robert. (2021). *English on Croker Island: The synchronic and diachronic dynamics of contact and variation*. Berlin: De Gruyter.

Malcolm, Ian. (2018). *Australian Aboriginal English: Change and continuity in an adopted language*. Berlin: Mouton De Gruyter.

Mathis, Terrie, & Yule, George. (1994). Zero quotatives. *Discourse Processes* 18(1): 63–76. https://doi.org/10.1080/01638539409544884

Meyerhoff, Miriam. (2009). Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change* 21: 297–317. https://doi.org/10.1017/S0954394509990196

Moll, Andrea. (2011). Natural or unnatural? Linguistic deep level structures in Aboriginal English. A case study of New South Wales Aboriginal English. In J. Alber & H. Rüdiger (eds.), *Unnatural narratives: Unnatural narratology*. Berlin: De Gruyter. 246–268.

Newmark, Kalina, Walker, Nacole, & Stanford, James. (2016). 'The rez accent knows no borders': Native American ethnic identity expressed through English prosody. *Language in Society* 45: 633–664. https://doi.org/10.1017/S0047404516000592

O'Brien, Robert M. (2007). A caution regarding rules of thumb for Variance Inflation Factors. *Quality & Quantity* 41: 673–690. https://doi.org/10.1007/s11135-006-9018-6

Rice, Emma S., Haynes, Emma, Royce, Paul, & Thompson, Sandra C. (2016). Social media and digital technology use among Indigenous young people in Australia: A literature review. *International Journal for Equity in Health* 15(81): 1–16. https://doi.org/10.1186/s12939-016-0366-0

Rodríguez Louro, Celeste. (2013). Quotatives down under: *Be like* in cross-generational Australian English speech. *English World-Wide* 34(1): 48–76. https://doi.org/10.1075/eww.34.1.03rod

Rodríguez Louro, Celeste. (2016). Quotatives across time: West Australian English then and now. In H. Pichler (ed.), *Discourse-pragmatic variation and change in English: New methods and insights*. Cambridge: Cambridge University Press. 139–159.

Rodríguez Louro, Celeste, & Collard, Glenys. (2021a). Australian Aboriginal English: Linguistic and sociolinguistic perspectives. *Language and Linguistics Compass*: 1–12. https://doi.org/10.1111/lnc3.12412

Rodríguez Louro, Celeste, & Collard, Glenys. (2021b). Working together: Sociolinguistic research in urban Aboriginal Australia. *Journal of Sociolinguistics* 25(5): 785–807. https://doi.org/10.1111/josl.12487

Rodríguez Louro, Celeste, Richard, Sophie, & Bharadwaj, Sana. (2020). Another story: *Be like* across discourse types. *English World-Wide* 41(3): 325–351. https://doi.org/10.1075/eww.00053.rod

Rodríguez Louro, Celeste, & Ritz, Marie-Eve. (2014). Stories down under: Tense variation at the heart of Australian English narratives. *Australian Journal of Linguistics* 34(4): 540–556. https://doi.org/10.1080/07268602.2014.929080

Romaine, Suzanne, & Lange, Deborah. (1991). The use of *like* as a marker of reported speech and thought. *American Speech* 66: 227–279. https://doi.org/10.2307/455799

Rumsey, Alan. (1993). Language and territoriality in Aboriginal Australia. In M. Walsh & C. Yallop (eds.), *Language and culture in Aboriginal Australia*. Canberra: Aboriginal Studies Press. 191–206.

Sánchez, Tara, & Charity, Anne. (1999). Use of *be like* and other verbs of quotation in a predominantly African-American community. *Paper presented at the 28th meeting on New Ways of Analyzing Variation in English (NWAVE 28)*. Toronto, October 14-17.

Schnell, Stefan, & Barth, Danielle. (2018). Discourse motivations for pronominal and zero objects across registers in Vera'a. *Language Variation and Change* 30: 51–81. https://doi.org/10.1017/S0954394518000054

Schourup, Lawrence C. (1983). *Common discourse particles in English conversation*. The Ohio State University: PhD dissertation.

Schultze-Berndt, Eva, Meakins, Felicity, & Angelo, Denise. (2013). Kriol. In S. Michaelis, P. Maurer, M. Haspelmath, & M. Huber (eds.), *The survey of pidgin and creole languages*. Volume 1. Oxford: Oxford University Press. 241–251.

Schweinberger, Martin. (2022). *Tree-based models in R*. May 21, 2022. Brisbane: The University of Queensland, School of Languages and Cultures. https://slcladal.github.io/tree.html.

Scott, Kim. (2011a). *Mamang*. Perth: UWA Publishing.

Scott, Kim. (2011b). *Noongar Mambara Bakitj*. Perth: UWA Publishing.

Scott, Kim. (2011c). *Noorn*. Perth: UWA Publishing.

Stanford, James. (2016). A call for more diverse sources of data: Variationist approaches in non-English contexts. *Journal of Sociolinguistics* 20(4): 525–541. https://doi.org/10.1111/josl.12190

Tagliamonte, Sali. (2014). Situating media influence in sociolinguistic context. *Journal of Sociolinguistics* 18 (2): 135–286. https://doi.org/10.1111/josl.12074

Tagliamonte, Sali, & D'Arcy, Alexandra. (2004). *He's like, she's like*: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8(4): 493–514. https://doi.org/10.1111/j.1467-9841.2004.00271.x

Tagliamonte, Sali, & D'Arcy, Alexandra. (2007). Frequency and variation in the community grammar: Tracking a new change through the generations. *Language Variation and Change* 19(2): 119–217. https://doi.org/10.1017/S095439450707007X

Tagliamonte, Sali, D'Arcy, Alexandra, & Rodríguez Louro, Celeste. (2016). Outliers, impact and rationalization in linguistic change. *Language* 92(4): 824–849. http://www.jstor.org/stable/44164126

Tagliamonte, Sali, & Denis, Derek. (2014). Expanding the transmission/diffusion dichotomy: Evidence from Canada. *Language* 90(1): 90–136. https://www.jstor.org/stable/24672021

Tagliamonte, Sali, & Hudson, Rachael. (1999). *Be like* et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics* 3(2): 147–172. https://doi.org/10.1111/1467-9481.00070

Tannen, Deborah. (1986). Introducing constructed dialogue in Greek and American conversational and literary narrative. In F. Coulmas (ed.), *Direct and indirect speech*. Berlin: Mouton de Gruyter. 311–332.

Terszak, Mary. (2008). *Orphaned by the color of my skin: A stolen generation story*. Maleny, Queensland: Verdant House.

Weinreich, Uriel, Labov, William, & Herzog, Martin. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics*. Austin: University of Texas Press. 95–188.

Winter, Joanne. (2002). Discourse quotatives in Australian English: Adolescents performing voices. *Australian Journal of Linguistics* 22(1): 5–21. https://doi.org/10.1080/07268600120122535