CAMBRIDGE
UNIVERSITY PRESS

Subject dislocation in Ontario English: Insights from sociolinguistic typology

Sali A. Tagliamonte and Bridget L. Jankowski

University of Toronto, Canada
Corresponding author: Sali A. Tagliamonte. Email: sali.tagliamonte@utoronto.ca

Abstract

Subject dislocation (SD) is common across languages. In French, it is a vernacular norm. In English, it is comparatively rare. This article examines English SD in a unique contrastive situation in Ontario, Canada: two communities where SD is a community norm, one where individuals speak both English and French (Kapuskasing), and the other where the population speaks English only (Parry Sound). Dislocated subjects are produced by the same underlying linguistic mechanisms in both places, with parallel constraints by type of subject and intervening material, suggesting a typological universal. However, SD is age-graded in Kapuskasing, regardless of heritage language. In Parry Sound, it is obsolescent, in steady decline over the twentieth century. We conclude that while typological trends are underlain by universal cognitive processes, locally embedded sociocultural influences are the source of differentiation.

Keywords: Canadian English; Ontario English; language contact; sociolinguistic typology; obsolescence; age-grading

The repeated expression of a subject, a type of topicalization, is a common feature across the world's languages. In contemporary varieties of French, it is a well-known vernacular feature that has been studied extensively, labeled *subject doubling, subject dislocation* (henceforth SD) or *left dislocation*,[1] depending on syntactic criteria (Auger, 1998; Nadasdi, 1995a; Thibault, 1983; Zahler, 2014). In contrast, in English, doubled expressions of the subject are rare. While they have been present since Old English (Traugott, 2007), the literature suggests this phenomenon has never made up more than 1% of all subjects: compare Old English (.7%) to Switchboard Corpus (.6%) (Liberman, 2008). Although doubled subjects have declined through the history of written varieties of English (Pérez Guerra & Tizón-Couto, 2004), it is reported in contemporary varieties of spoken English (Godfrey & Holliman, 1993). While the feature is sometimes mentioned for English dialects (e.g., Southard & Muller, 1998), even in places where it is attested, such as in the southern United States, it is considered rare.

Tagliamonte and Jankowski (2019) reported high levels of SD[2] in the English spoken in Kapuskasing, Ontario, where locals have Anglophone and Francophone[3] heritage language background, and linked this usage to high levels of French-English bilingualism. A quantitative comparative analysis contrasting English-speaking Anglophones and English-speaking Francophones revealed a trend in apparent-time toward increasing SD, and parallelism in its linguistic patterns in the youngest age groups. SD was used more among young people of both heritage language cohorts, and their usage patterns conformed to a subject constraint well-known in studies of French (Tagliamonte & Jankowski, 2019). They explained these findings as grammatical alignment (e.g., Chandler, 2007:117-124, 306; Cheshire, 1998:139-140) based in part on a rising proportion of heritage Francophones in the community over the last decades of the twentieth century as well as simultaneous increasing bilingualism of residents regardless of heritage language, suggesting that the origin of SD in the Anglophones was influence from French. However, SD is found in many languages as well as being documented throughout the history of English and in contemporary English dialects. These facts present a puzzle: Is the explanation for SD found in language-specific influence or is it some more universal process? Further study was required to elucidate the nature of this feature in Ontario.

Emerging findings from analyses of linguistic features in other bilingual situations suggest that the local embedding of social factors, linguistic constraints, and typological considerations are key to understanding cross-linguistic contrasts. Quantitative studies that examine the structure of variation expose the importance of speaker choices in spontaneous interactions. For example, use of Spanish and Portuguese on the border of Northern Uruguay (Carvalho, 2016:217) demonstrates that use of third-person singular pronouns in the two languages has "not converged into a single variable third-person singular pronoun system." A comparison of English/Spanish subject omission in contemporary varieties (Torres Cacoullos & Travis, 2019) demonstrates parallels in the probabilistic constraints, even in two languages where overall rates are known to differ. These findings champion the comparative variationist enterprise in exposing language-particular versus cross-language patterns and suggest that these methods applied to language contact situations will offer crucial insights to understanding language universals.

In this study, we build on the earlier research of the Kapuskasing bilingual situation (Tagliamonte & Jankowski, 2019) by conducting a comparative study with SD in another community: Parry Sound. We also embellish the comparative enterprise by adding in considerations of sociolinguistic typology (Trudgill, 2011). In this case, we compare distinct communities based on language contact, one highly bilingual, Kapuskasing, where the majority of the population are Francophone, versus another, Parry Sound, where the majority of the population are Anglophone and trace their ancestry to the United Kingdom. These two communities offer antithetic scenarios in Trudgill's typology of community types: they are both small with tight social networks. However, Kapuskasing is high for language contact (type 6) and Parry Sound is low (type 1) (Trudgill, 2011:148). This is an ideal situation for testing the origin of SD in each community and for contributing insight into SD more generally.

Dislocation, left dislocation, subject doubling

Many constructions can be considered "dislocation." In this study, we focus on constructions on the left periphery in which a subject is repeated, resulting in a lexical NP or strong pronoun adjacent to a subject pronoun that further explains or identifies the subject, as in (1).

1. So **Jeff, he** understood horses perfect.
   (male, 75. b. 1905, Parry Sound, Anglophone)

In most cases, the NP subject comes first; however, the first position can also be a strong pronoun, as in (2).

2. My brother went overseas and **me, I** come home.
   (male, 91, b. 1925, Kapuskasing, Francophone)

Many studies of this phenomenon distinguish between "doubling" and "dislocation" (e.g., Culbertson, 2010). In some languages and varieties closely related to Standard and Colloquial French, such as Pied Noir (Roberge, 1990), Picard (Auger, 2003), and Montréal French (Nagy, Blondeau, & Auger, 2003), it is possible for quantified DPs to be doubled, whereas in Standard and Colloquial European French, it is not (e.g., *Personne (*il) ne parle* 'Nobody speaks'). The inability to double a quantified DP has yielded a syntactic analysis of left dislocation for a sentence such as *John il parle* 'John speaks,' placing the DP outside the clause and "the clitic as a resumptive element in canonical subject position" (Culbertson, 2010:105). If quantified DPs can be doubled, as they can in some varieties, then the structure is true "subject doubling" and the clitic is "an agreement marker with the DP in canonical subject position" (Culbertson, 2010:105, based on Rizzi, 1986). Given this analysis and the complete absence of doubling with quantified subjects in our data, it seems clear that the phenomenon in our colloquial English data is SD.[4]

Situating the linguistic variable

SD is a well-known characteristic of topic prominent languages (e.g., Li & Thompson, 1976). English and French are considered subject-prominent languages. However, such languages also have ways of producing topic-comment structures, including adverbial expressions, passive voice, clefting, and fronting. Dislocation is only one of these strategies. A key fact for our study is that repeated subjects are frequent and robust in French but not in English.

In Canada, French and English are official languages; the presence of French varies by region. In the province of Québec, French is the majority language and double expressions of the subject are a prominent feature (Nadasdi, 1995a). In Ontario, French is a minority language in most places. In certain pockets, however, Francophones represent large proportions of the population (Mougeon, Beniak, & Valois, 1985), and their French also has the characteristic doubling of subjects. In contrast, SD is not a prominent feature in English generally nor in the English dialects of Ontario, though

rare attestations and anecdotal reporting attest to its presence in rural varieties and nonstandard registers. In the next section, we review some of the patterns of repeated expressions of the subject in varieties where it has been studied.

Doubling and dislocation in French. In European varieties of French, subject doubling and dislocation are reported to have several linguistic patterns depending on variety and community. Nagy et al. noted that Sankoff posited doubling and dislocation to "represent different stages in the process of grammaticalization [such that] 'The [third person] clitic would become a morphological component of the verb and would acquire quasi-obligatory status, even when a noun phrase with the same referent is present in the clause'" (Nagy et al., 2003:78, citing Sankoff, 1982:85).

Blanquart (2012:25-26) cited De Cat's claims (De Cat, 2007:220-242) that approximately a quarter of clauses in spoken French contained a dislocation, mostly definite noun phrases.[5] Her explanation for these constructions was emphasis or topic shift, used when speakers are expressing opinions, contrast, or turn-taking (Blanquart, 2012:41). Doubling has been especially associated with northern French dialects in the Picard region of France,[6] where it was reported to be near obligatory (Coveney, 2005:103-105), held to convey social meaning, and had an age-grading pattern where middle-age speakers between the ages of 24-37 had the lowest rates. It has also been considered to have "particular significance as a badge of Picard identity" in the Picard French dialect, being "obligatory not only for NPs headed by a noun […], but even for pronouns such as the Picard equivalents of *tout le monde* and *chacun*" (Coveney, 2005:103). It has been a highly stigmatized feature, as evident in the attitudes of traditional grammarians and authors of second language teaching materials until the mid-twentieth century. Subject doubling in French "was already being criticized in the seventeenth century for the superfluous subject pronoun" (Coveney, 2005:97). In Swiss French, Fonseca-Greber (2000) reported that subject doubling occurred 77% of the time, even with indefinite subjects, as in (3):

3. ***Une omelette*** *elle est comme ça.*
   'An omelette she is like that.'

In contrast, Zahler (2014:868-869) found lower rates of doubling in spoken Parisian French (22%) compared to other varieties. She suggested that this was due to the formality of her data and the status of Parisian French as the standard variety.

In Canadian French, referring to this phenomenon as subject doubling, Sankoff (1982) reported an overall rate of 55% in Montréal, while Auger and Villeneuve (2010) reported an overall rate of 45% in Saguenay. In Ontario (Pembroke, North Bay, Cornwall, and Hawkesbury), (Nadasdi, 1995a:7, Table 1) reported 27% overall. Nagy et al. (2003:88-90) reported 37% subject doubling among Anglophones speaking French in Montréal. Those with the highest exposure and integration into Canadian French culture produced the most, suggesting parallel behavior with native Francophones.

Doubling and dislocation in English. In English, SD has the social meaning of rurality, though to our knowledge it is not overtly commented on by traditional grammarians or dialectologists. A few descriptions of Southern and Appalachian US dialects

(e.g., Southard & Muller, 1998; Wolfram & Christian, 1976) mention SD. Yet it is salient enough to be used by authors and comedians to represent nonstandard, rural dialects. Examples (4-6) are from Mark Twain's *Huckleberry Finn* (Southard & Muller, 1998).[7]

4. **The door, it** slammed to.

5. **The widow, she** cried over me.

6. **Jim, he** grumbled a little.

   In some rural Ontario locations with no significant French influence, there is notable SD, and the same nonstandard, rural associations. For example, Canadian comedian Don Harron's signature character, Charlie Farquharson, the personification of a rural Ontario farmer from Parry Sound, Ontario, appeared for over half a century on radio and television, including on the long-running (1969-1992) North American country music sketch comedy series *Hee Haw*. When performing as Farquharson, Harron regularly used SD among his cadre of lexical, syntactic, and phonological vernacular speech features for comedic effect, as in (7-8).[8]

7. **Angolf, he** was checking the top of the tank for the thickness of the ice when he fell through.

8. Well **Bill, he** thunk it over, and he thinks he'll settle for just ten years.

   The point of departure for this study is to probe the linguistic patterns of SD in Ontario English in two situations. The characteristics of Kapuskasing and Parry Sound offer a singular opportunity to compare communities with maximally different language contact situations, where there is salient use of SD.

Circumscribing the variable context

Following variationist sociolinguistic methods, we circumscribe the variable context of SD using the principle of accountability (Labov, 1982:30). We excluded false starts, such as sentence fragments that did not contain a predicate, and where the two co-referential elements for the subject were unclear. Although Nagy et al. (2003:82) included all strong third-person subjects, both pronouns and noun phrases, we restrict the present analysis to noun phrase subjects only. This is because Tagliamonte and Jankowski (2019:4-5) found that noun phrases exhibited a rate of 10% SD ($n = 6828$) in Kapuskasing, whereas subjects outside of first- and third-person contexts, as in the second-person in (9), represented little more than 1% of the data.

9. **Yourself, you're** born and raised in Toronto?
   (male, 33, b. 1982, Kapuskasing, Anglophone)

The data

SD is used by individuals of all ages in both Kapuskasing and Parry Sound in animate referents, both proper names and noun phrases, as in (10-11), as well as with inanimate nouns, as in (12).

10.   **Alice Cooper, he** plays Sudbury all the time.
       (female, 24, b. 1992, Kapuskasing, Anglophone)

11.   … and then **my dad, he** was over the blacksmith shop too …
       (female, 92, b. 1899, Parry Sound, Anglophone)

12.   … **the onions, they** last us about two, three months …
       (male, 18, b. 1998, Kapuskasing, Francophone)

The presence of SD in Parry Sound and in both heritage language groups in Kapuskasing raises important questions: Is the linguistic and social patterning of SD parallel or different, and if so, how? By conducting a consistent analysis in Kapuskasing and Parry Sound, we aim to assess whether it is the same phenomenon or distinct. The results will permit us to understand whether SD is a feature of English or whether it developed in Kapuskasing English due to contact with French, where SD is a characteristic feature. More broadly, we hope to contribute to an understanding of the general process of dislocation in language more generally. Simply put, is influence from French the reason for the relative abundance of SDs in Kapuskasing, or is SD in Ontario English more generally the result of a pervasive phenomenon of topicalization in vernacular speech?

Data and method

The data come from a long-term research program, the Ontario Dialects Project (ODP) (Tagliamonte, 2013-2018).[9] We focus on the two communities, shown in Figure 1, with several Canadian cities indicated for reference: Montréal, Ottawa, Toronto, and Thunder Bay.

Kapuskasing. Kapuskasing has a population of 8,292 (Statistics Canada, 2017) and is 850 km (500 miles) from Toronto, making it relatively isolated from the urbanized south. Kapuskasing was founded in the 1920s with an economy driven by pulp and paper, which remains the dominant industry (Dunlap, 2017; Town of Kapuskasing, 2010). The community has a strong religious tradition, a bent toward outdoor activities, and close-knit social networks (Tagliamonte & Jankowski, 2019:4).

In 2016 the population was over 65% primary Francophone speakers, and over 70% of residents report knowledge of both English and French (Statistics Canada, 2017). This population mix is unusual; although French is one of Canada's official languages, it is by far the minority language in the province of Ontario at only 4%. In Kapuskasing, however, Anglophones and Francophones have been living together in long-term contact since the early 1900s. The high rate of reported bilingualism as compared to the rest of Ontario (and Canada) is unique, offering the opportunity for studying the impact of potential French influence on English in a language contact situation.

Parry Sound. Parry Sound is 225 km (140 miles) from Toronto. In comparison to Kapuskasing, Parry Sound is relatively "old," having been incorporated in 1887.

Figure 1. Map showing Kapuskasing and Parry Sound (Adapted from: https://commons.wikimedia.org/wiki/File:Ontario_Locator_Map.svg).

It is a small town with a population of 6,321 (Statistics Canada, 2017). The area is predominantly Anglophone (93%), with ancestral roots from the British Isles and a few reports of other languages. While this area also had an early economy of pulp and paper, tourism has become its most prominent industry in the late twentieth century (see also Jankowski, Needle, & Tagliamonte, 2022).

The Kapuskasing and Parry Sound corpora comprise conversational interviews conducted using sociolinguistic methods and collected during community-based fieldwork in 2016 and 2018 (Labov, 1972; Tagliamonte, 2006). The sample design is shown in Table 1.[10]

The focal dimension of comparison is the heritage language group—Anglophone or Francophone in Kapuskasing versus Anglophone only in Parry Sound. While there are many other differences between the two towns, they both share the attribute of having salient use of SD. This is particularly germane for assessing the nature of this phenomenon in the Ontario context.

Table 1. Kapuskasing and Parry Sound corpora

Community	Date of collection	# speakers				Birth years	Total words
		Anglophone		Francophone			
		Female	Male	Female	Male		
Parry Sound	1974–2018	23	26	0	0	1885–2001	557,320
Kapuskasing	2016	12	13	11	10	1925–2001	518,713

Coding and analysis

To assess the patterns of variation, we coded linguistic and social factors previously reported to condition SD use in French and English.

Linguistic factors

Many linguistic factors constraining SD have been attested in the literature (see Zahler, 2014). However, in this study we focus on the main factors found to be significant.

Type of subject. The nature of the subject noun phrase figures in every study and can be defined in many different and overlapping dimensions including animacy, definiteness, and specificity (Auger, 1998; Nadasdi, 1995b). However, in studies that focus on third-person subjects only, subject type is most commonly categorized according to kind (e.g., proper versus common, and animacy; Auger & Villeneuve, 2010; Nagy et al., 2003). Across studies, full noun phrase subjects are found to have the most SD, and proper nouns favor it more than common nouns (e.g., Nagy et al., 2003:88). Tagliamonte and Jankowski (2019) showed that this constraint was operational in Kapuskasing English: SD was more frequent with human subjects than nonhuman subjects. This study categorized third-person noun phrases by type and animacy: proper names (13), other humans (14), collectives of humans (15), organizations (16), animals (17), places (18), and things (19).

13. And so old **Missus Vander, she** really liked us …
    (male, 85, b. 1933, Parry Sound, Anglophone)

14. Irma, **my mother-in-law, she** plays cards every Thursday …
    (female, 62, b. 1956, Parry Sound, Anglophone)

15. A lot of **families, they** did their own entertaining
    (female, 93, b. 1924, Parry Sound, Anglophone)

16. **The big companies, they** only took pine and hemlock …
    (male, 86, b. 1895, Parry Sound, Anglophone)

17. **The one rooster, he** was a bad son of a gun …
    (male, 75, b. 1943, Parry Sound, Anglophone)

18. Yeah, **Barcelona, Paris, they,** like **they**'re amazing.
    (female, 38, b. 1980, Parry Sound, Anglophone)

19. And **this stuff** he put on, why, **it** just pretty near drove you nuts.
    (male, 75, b. 1905, Parry Sound, Anglophone)

Since dislocation by type of subject is reported across studies, this factor will offer key insights into different levels of contrast, whether at the community, dialect level, or by heritage language in a French-English contact setting.

Intervening material. Another pattern consistently reported is the effect of material intervening after the subject noun phrase and before the verb. Intervening material has been found to favor SD because of the additional processing load it requires (Auger & Villeneuve, 2010; Nagy et al., 2003:88-89; Zahler, 2014:363). Intervening material can derive from spoken language phenomena, including hesitations, pauses, and other disfluencies. Intervening material can also have linguistic substance, as for instance adverbs, clauses, and other periphrastic constructions can occur between a subject noun and verb (20), as well as the nonlexicalized particles that we transcribe as *ah, uh*, and *um*, as in (21).

20. **The Tower Hill *here*, that**'s a big thing.
    (male, 53, b 1964, Parry Sound, Anglophone)

21. **My younger brother *um* he** did something to the skidoo.
    (female, 59, b. 1957, Kapuskasing, Francophone)

While cognitive processing can be expected to apply uniformly across communities and social groups, it is unknown whether different sources of intervening material will have the same influence on SD.

Social factors

Social patterns have been reported for SD, but unlike the reported linguistic influences, their influence tends to be inconsistent across studies. To test the possible influence of social factors, we included year of birth, perceived gender, education level, occupation (blue versus white collar versus student) and for Kapuskasing, mother tongue of the different heritage language groups. The community-level patterns for these factors will enable us to theorize about the social meaning of SD in each location, and its origin in French and/or English.

Change in progress

A key pattern that emerged in Tagliamonte and Jankowski (2019) was a generational trend, with a significant change arising in the youngest individuals in Kapuskasing. Anglophones and Francophones exhibited a similar correlation of SD with proper names, but there was a marked difference between the oldest and youngest generation such that older Anglophones patterned distinctly from older Francophones. Tagliamonte and Jankowski hypothesized that the parallelism of the type of subject constraint among the youngest generation only was due to an increasingly common vernacular between Francophones and Anglophones. However, without a comparative perspective from an English-only community, such an interpretation could not be conclusive.

If the apparent-time trend in Kapuskasing is indeed the result of alignment, we would *not* expect the same shift in apparent-time among the youngest generation in Parry Sound where there is no language contact. Instead, if SD is a vernacular feature of English in general, we would expect older Anglophones in both Kapuskasing and Parry Sound to have parallel frequency and patterns of SD—and possibly the

older Kapuskasing Francophones as well, since they are, after all, skilled speakers of English. Further, if SD in Kapuskasing is indeed diverging from English patterns under influence from French among the youngest generation, we would expect the younger Kapuskasing and younger Parry Sound individuals to differ with respect to their usage patterns—those in Kapuskasing patterning like French, those in Parry Sound patterning with the rest of the community.

Age-grading

In contrast to an interpretation of change in progress, previous studies of SD in French have reported an age-graded pattern. Coveney found that middle-aged individuals speaking the regional variety of French in Picardie had the lowest rates, which he suggests is the result of "stable age-grading" (Coveney, 2005:105). Tagliamonte and Jankowski (2019) found indications of age-grading in Kapuskasing English as well: middle-aged speakers used SD less than older and younger speakers. The U-shaped curve typical of age-grading is said to be an indication that the phenomenon is a vernacular norm of the speech community, whereby middle-aged individuals suppress nonstandard uses due to participation in the linguistic marketplace of the community (Sankoff, 1978). If SD is age-graded, we would expect the middle-aged individuals of each heritage language cohort to have the lowest rates.

We now turn to the analysis of SD in Kapuskasing and Parry Sound, employing variationist techniques and the comparative method (Poplack & Tagliamonte, 2001:Chapter 5; Tagliamonte, 2002), and situating the study in the context of sociolinguistic typology (Trudgill, 2011).

Results

We begin with distributional analyses and then employ mixed effects statistical modeling in *R* (version 4.2.0; R Core Team, 2022) to determine the patterning of social and linguistic factors, their significance, and relative importance to the production of left-dislocated subjects.

Distributional analysis

A first step in the comparative endeavor is to evaluate use of SD by community, taking individual usage into account. Table 2 shows the proportion of SD in Kapuskasing's two heritage language groups and Parry Sound's monolingual English speakers.[11]

Table 2. Proportion of SD with NP subjects by cohort

	% dislocation	n	Total n
Kapuskasing Anglos	10.1%	400	4050
Kapuskasing Francos	10.4%	289	2791
Parry Sound	5.4%	329	5994

Table 2 shows that Kapuskasing has twice the proportion of SD as Parry Sound, but there is no difference between Kapuskasing Anglophones and Francophones. Note too that the rate is less than one-third the rate reported in studies of Canadian French.

Figure 2 shows proportion of SD by community in a boxplot, which visualizes individual variance across each cohort of heritage language speakers, Parry Sound (PS), Kapuskasing Francophones (KAP.Francos) and Kapuskasing Anglophones (KAP.Anglos). Figure 2 exposes the much lower frequency of SD in Parry Sound, which is below not only the means (horizontal black lines), but also the dispersion for both Kapuskasing cohorts.

The black dots in the Parry Sound plot on the far left of Figure 2 are individuals with higher frequency, suggesting a stylistic feature.[12] The two cohorts in Kapuskasing are more cohesive within groups, overlapping and without outliers. This view indicates a contrast in the community norm for SD over and above individual or heritage language group. The difference, both in frequency and dispersion by cohort in Parry Sound, also argues that although SD is present, it may be produced by a different mechanism than in Kapuskasing.

We begin with distributions to understand the social and linguistic patterns. Then we turn to conditional inference trees to visually model these multiplex comparisons. As the complex interactions in the data become apparent, and the focal patterns of contrast and cohesion emerge, we conduct random forests and mixed effects modeling to confirm whether the relevant patterns are statistically significant, and to what degree, when all of them are considered simultaneously. For discussion of how the combination of these techniques further explanatory adequacy, see Tagliamonte and Baayen (2012) and Gries (2018).

Social patterns

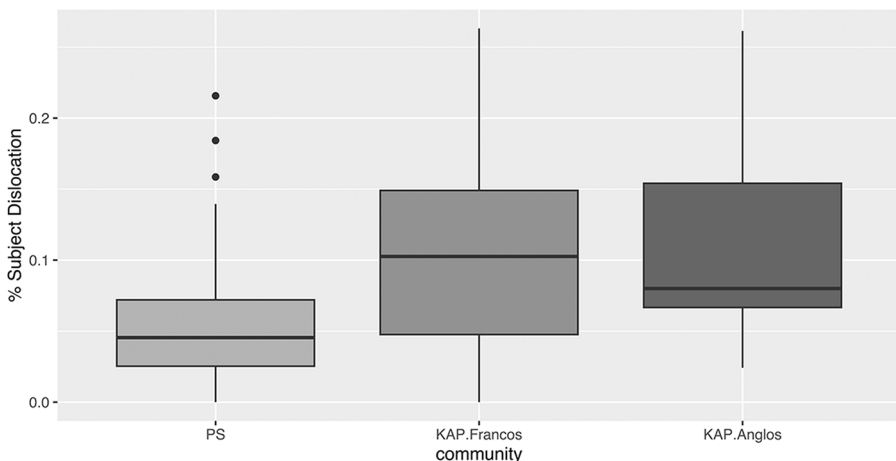As outlined earlier, previous studies of SD report notable social patterning.

Figure 2. Proportion of SD by cohort.

Gender. An important factor in many studies is the difference between men and women (modeled as a binary factor). However, the constraint ranking is inconsistent across studies. In Canadian French in Saguenay, young women favor SD (e.g., Auger & Villeneuve, 2010). In Parisian French, older women are the highest users (Zahler, 2014:368, Table 6). In Montreal, a French-dominant and highly bilingual city, Anglophones speaking French have "questionable significance" of gender due to interactions with other factors such as acquisition and integration into the French milieu (Nagy et al., 2003:94). Figure 3 shows the effect of gender in the Kapuskasing and Parry Sound data.

Figure 3 shows that in Parry Sound men have more SD. In Kapuskasing, men have more SD among the Francophones, but among the Anglophones women have more. However, the differences here are modest, and this comparison does not take any other factors into account.
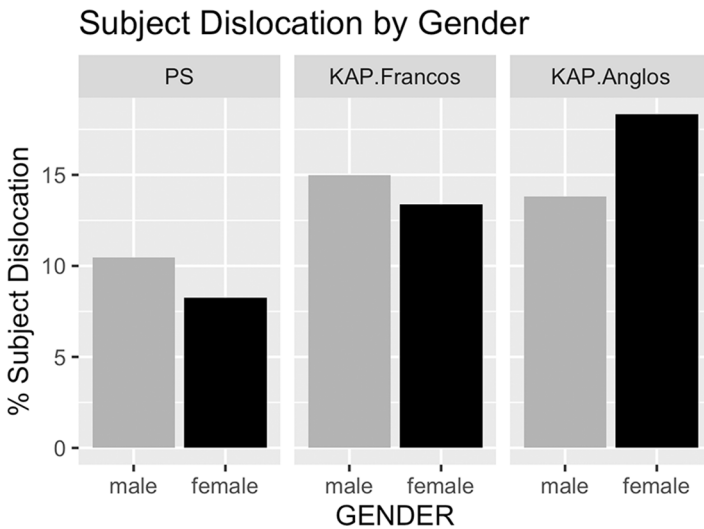

Figure 3. Proportion of SD by gender and cohort.

Occupation and education. Some studies of SD have reported higher usage with less educated speakers, at least among Anglophones speaking Canadian French (Nagy et al., 2003:94), where formality and more standard use appear to dampen the frequency of this nonstandard usage. Figure 4 assesses the influence of occupation, treated as a ternary contrast of blue collar (B), white collar (W), and students (S). Figure 5 assesses education, treated as a binary contrast between those with more than high school education (Y) or high school education or less (N).

Figure 4 reveals that the effect of occupation also varies by community. Among Kapuskasing Anglophones there is little difference between blue-collar workers, white-collar workers, or students. Among Kapuskasing Francophones and in Parry Sound, SD is higher among blue collar workers. Note too that students, who are all younger, have the highest use in Kapuskasing, while there are no students in Parry Sound who use
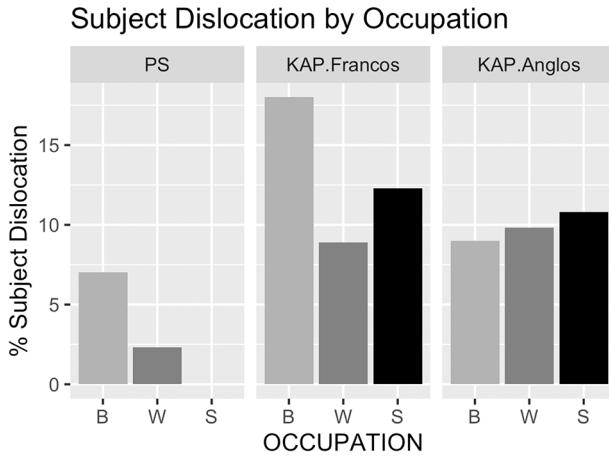
Subject Dislocation by Occupation

Figure 4. Proportion of SD by occupation and cohort.
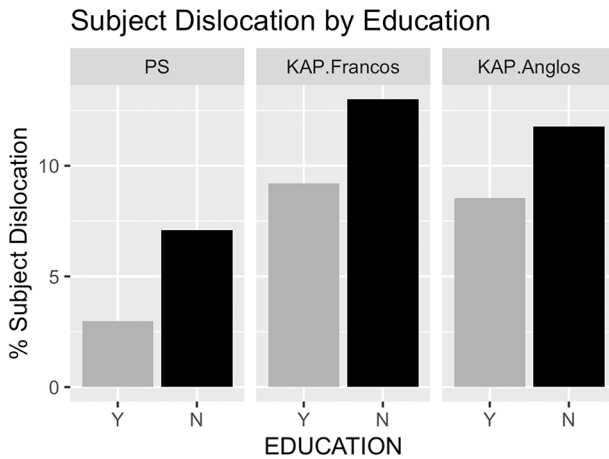
Subject Dislocation by Education

Figure 5. Proportion of SD by education and cohort.

dislocation. Figure 5 shows that the effect of education is consistent across all groups in each community—less educated individuals use more SD.

Date of birth. Of all the social patterns reported for SD, the most widely cited is age of the speaker, presented here according to date of birth binned into decades. Figure 6 shows the proportion in Parry Sound, the top panel compared to Kapuskasing Anglophones and Francophones stacked underneath. This visualization also provides the number of tokens in each binned age group. Recall from Table 2 that the Parry Sound materials offer a timeframe back to the late 1800s due to legacy materials. However, in Kapuskasing, we do not have this longitudinal perspective.
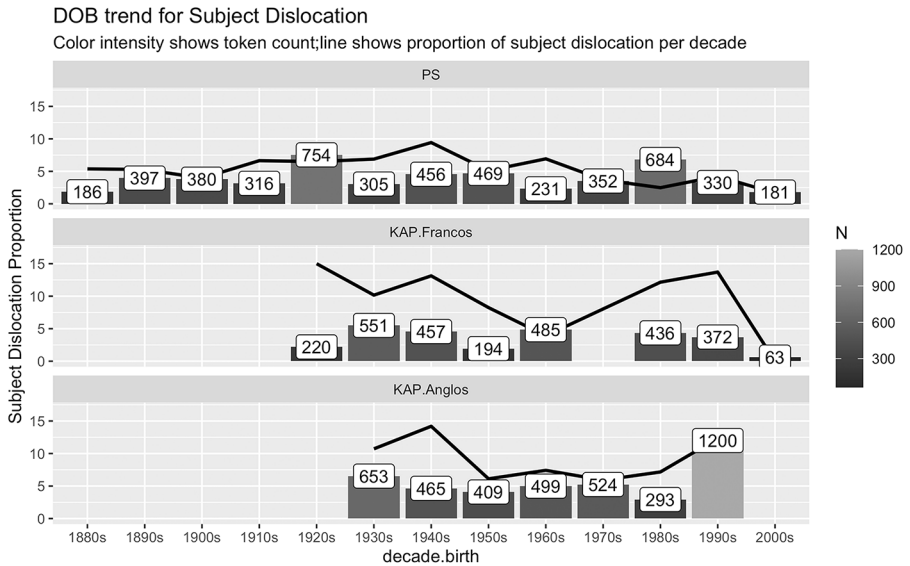
DOB trend for Subject Dislocation
Color intensity shows token count;line shows proportion of subject dislocation per decade

Figure 6. Proportion of SD by decade of birth and cohort.

Figure 6 reveals that in Parry Sound SD is never more than 10% of the data, with a relatively linear decline in use from individuals born from the early to late twentieth century. In contrast, Kapuskasing exhibits an age-graded pattern such that middle-aged individuals use SD least while older and younger speakers use it more. Moreover, this pattern is shared for both Kapuskasing heritage groups.

These social patterns expose a qualitative difference in the social embedding of SD in these communities. In Parry Sound, SD is not only low frequency but also receding and more frequently used by male speakers overall. In Kapuskasing, there are parallels between community groups and a classic age-graded pattern. We now turn to the internal linguistic constraints.

Linguistic patterns

At least two internal patterns have been reported for SD: type of noun subject and intervening material. Most reports document a preference for SD with proper nouns over other human subjects, animals, or inanimate nouns.

Type of subject. Figure 7 plots the use of SD in Kapuskasing Anglophones (KAP.Anglos) and Francophones (KAP.Francos) and in Parry Sound (PS) by type of subject and decade of birth. We display a five-way categorization for type of subject, distinguishing proper names, animate subjects (humans and animals), collectives (of animate subjects), inanimates (organizations and places), and things (concrete and abstract).

Figure 7 shows that proper names (black bars) stand out as having the most SD across cohorts but not in Parry Sound after 1960. In Parry Sound, there are notably

high proportions among those born in the 1940s and 1950s; however, this is the result of heightened use among certain individuals.[13] In other contexts, use of SD is scant, despite many possible contexts in the data. In Kapuskasing, SD is relatively high for Francophones born in the 1920s. Middle-aged Francophones exhibit a dip in usage, but among Francophones born in the 1980s and 1990s, rates are higher, consistent with the overall distribution in Figure 6. Further, Anglophones and Francophones mirror each other. Figure 7 adds the additional perspective that proper names are not only a privileged context for SD in Kapuskasing but that this is consistent across dates of birth. The frequency with other animates (mostly humans) and collectives (e.g., *families* in [15]), is lower and more erratic, likely due to small Ns in some decades. An increase with inanimate nouns among Kapuskasing Francophones in recent decades is apparent (specifically organizations and places, as in *companies* in [16] and *Barcelona* in [18], but the token counts here are very small [$n = 24$]). The category "things," as in (19) *this stuff*, rarely has SD, but in this case this category is very frequent in the data ($n = 3089$).
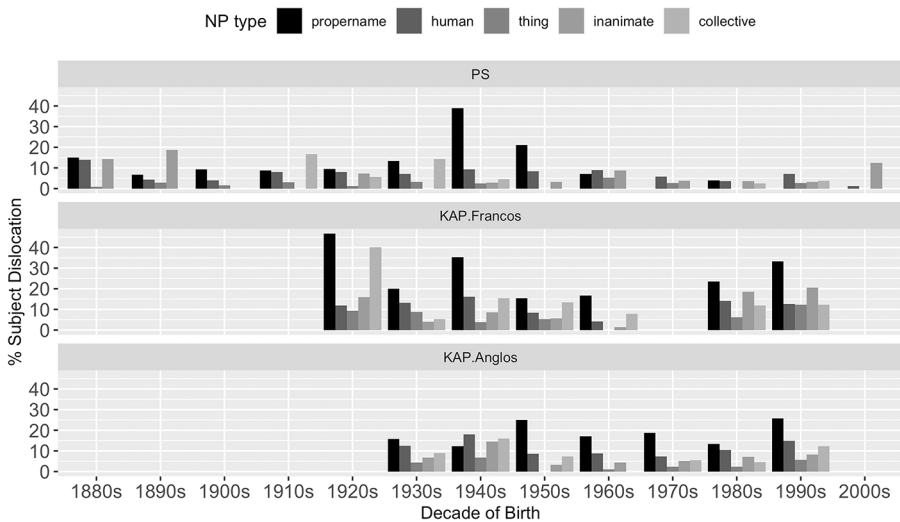

Figure 7. Proportion of SD by type of subject, decade of birth, and cohort.

Intervening material. Figure 8 shows the proportion of SD by the type of intervening material, cohort, and decade of birth. SD is present across all contexts. The main context in all cohorts is with nonlexical hesitations (black bars). This effect is strong in all decades of birth.

Intervening linguistic material, whether adverbs or other clausal elements, is also a dominating context. In both Kapuskasing Francophones and Anglophones, SD is more likely in these contexts. Partitive constructions, like *one of them*, have low frequency of SD, although Kapuskasing shows a small proportion of instances in this context in earlier decades. Across all communities and cohorts, where there is no intervening material (i.e., NONE), SD rarely occurs.

In summary, distributions of the empirical data provide evidence of contrasting and parallel patterns by nature of the intervening material. In all groups, SD arises from structural disruption, whether adverbs or clauses or hesitations. However, a key contrast is that SD in contexts of nonlexical hesitations is the dominating effect in Parry Sound, while in Kapuskasing, SD occurs in both hesitations and intervening linguistic material.
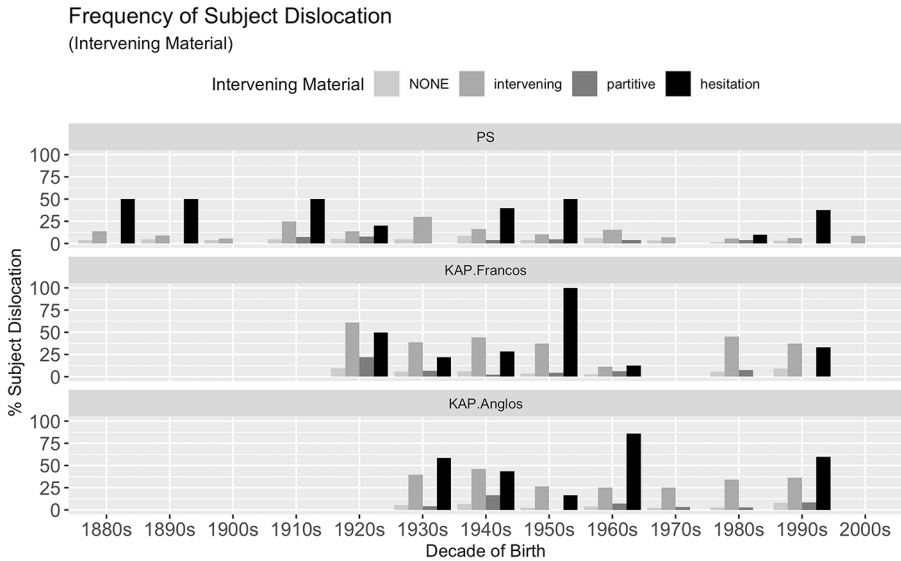

Figure 8. Proportion of SD by intervening material, cohort, and decade of birth.

Statistical modeling

We now turn to statistical modeling to assess the significance of factors when they are treated simultaneously. To uncover the importance of predictors and how they work together, we employ random forests and conditional inference trees (Strobl, Malley, & Tutz, 2009). Then, we use mixed effects generalized linear modeling to evaluate the significance of patterns and interactions (e.g., Baayen, 2008). Interpretation of statistical modeling results compared across cohorts and generations (e.g., Tagliamonte, 2002) helps us situate this feature socially, linguistically, and with respect to language contact, sociolinguistic typology, and language universals.

Random forest

We begin with a random forest analysis (Tagliamonte, 2012) of all the data, which enables us to determine the relative importance of factors. Community is most important, consistent with the frequency differences shown in Table 2 and Figure 2. We therefore partitioned the data in the visualization of Figure 9 by heritage language cohort and community to expose the contrasts and similarities in strength of other
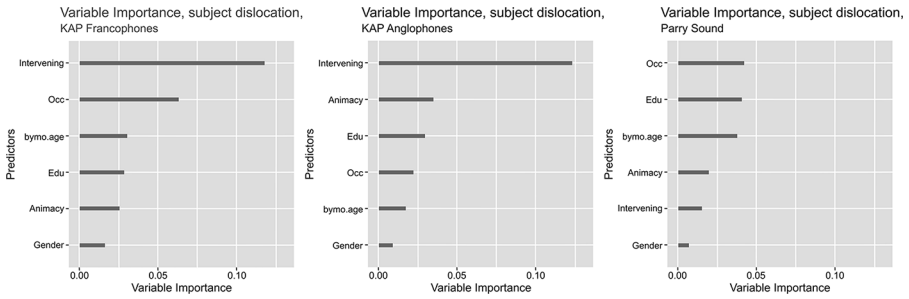
Figure 9. Random forest analysis by cohort and community.

factors. In these models, date of birth, education, and occupation are binned into the divisions that mirror the contrasting patterns in Figures 4-6.[14]

The random forest plots expose a conspicuous distinction between Kapuskasing and Parry Sound. In Parry Sound, the social factors of occupation, education, and speaker age are most important, followed by the internal factors of animacy and intervening material, and gender is the least important. In Kapuskasing, the presence of intervening material is the most important factor in both heritage language cohorts while social factors play a lesser role. The effect of animacy also differs markedly by cohort: among Anglophones, intervening material is the second-ranked factor, while for Kapuskasing Francophones it ranks among the lowest.

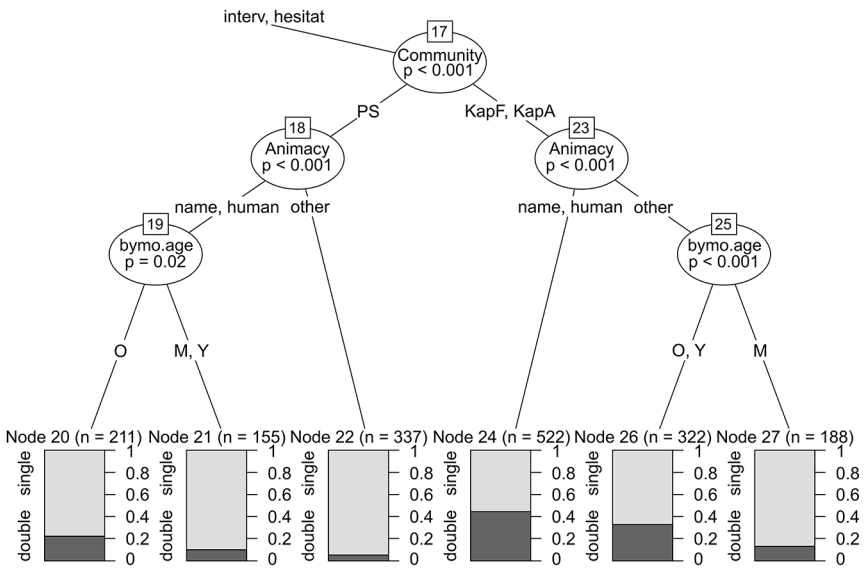### Conditional inference tree(s)

The distributional analyses in Figures 7 and 8 indicated the internal ranking by subject animacy and type of intervening material diverge across cohorts. The next step is to assess the internal structure of each variable using a conditional inference tree analysis, which provides a visualization of the complex interactions within the data.

The conditional inference trees in Figures 10a and 10b present a model that includes community, the two contextual factors (subject type and intervening material), and speaker age, binned as in Figure 9. To avoid small splits in the data, which would be compromised by low numbers, we set the algorithm for a minimal bucket of one hundred tokens.[15] The dark bars indicate the presence of SD. Because the tree is wide and splits in two at the top node, we present the results in two parts. Figure 10a shows the contexts with no intervening material between subject noun phrases and the verb (none). Figure 10b shows the contexts containing intervening material whether intervening linguistic material (interv) or hesitations (hesit).

The dominant effect, over and above that of community in the composite tree, was presence of intervening material, with a main split at the top of the tree (labeled "Between"), indicating a significant main difference in the number of dislocated subjects, depending on whether there is intervening material between the subject and verb. Contexts of intervening material are pivotal. "NP + (**pronoun**) + verb" sequences [none], as in "this guy (**he**) went," contrast with "NP + *intervening material* + (**pronoun**) + verb" sequences [interv, hesitat]: "this guy *at the time/uh uhm* (**he**)
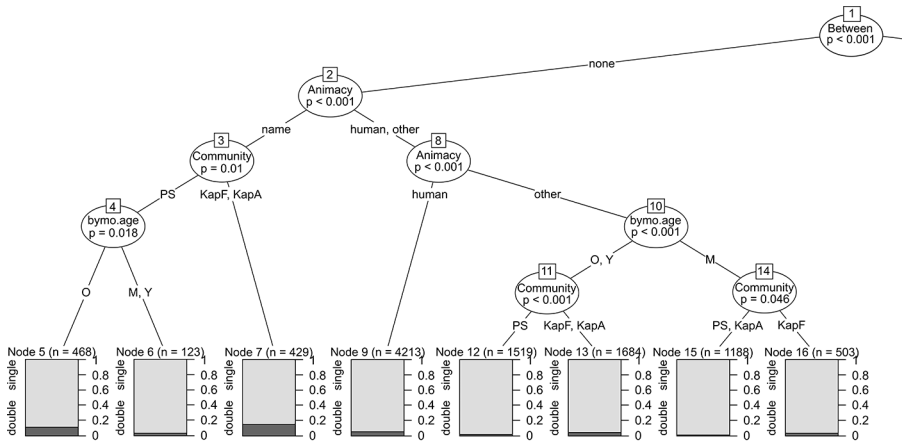
Figure 10. Panel (a). Conditional inference tree plotting linguistic and social factors with no intervening material. Panel (b). Conditional inference tree plotting linguistic and social factors with intervening material.

went." Comparing the frequency in each main branch makes it evident that SD rarely occurs in the "NONE" condition (Figure 10a). In contrast, SD occurs moderately but systematically in the condition "interv, hesitat" (Figure 10b).

In Figure 10a, patterns of animacy are most significant with a split at node 2 between proper names (proper) versus humans and other nouns, and another split at node 6

between human and other subjects, revealing the animacy hierarchy *proper name > human > other*. In Figure 10b, community differences are clearly the main contrast. Kapuskasing splits to the left and Parry Sound splits to the right. The same animacy hierarchy is found in each community, namely *proper name > human > other*; the difference is only in the frequency of SD. The last splits are due to the age cohort. In Parry Sound, older people "O" contrast with middle-aged and younger people, while in Kapuskasing older and younger people pattern together in contrast to the middle-aged group. This is consistent with the distributional patterns in Figure 6.

In summary, Figures 10a and 10b reveal that SD predominately arises in contexts where there is a proper name and distance between the subject and the verb. These contextual influences are the same in all cohorts, but to different degrees. Further, the evidence from apparent-time (speaker age) continues to indicate two distinct trajectories. In Parry Sound, SD is obsolescent while in Kapuskasing it is age-graded.

Modeling

The next step is to subject the data to linear mixed effects modeling (Bates, Maechler, & Bolker, 2011; Tagliamonte, 2012:144-152) using the *lme(4)* package in *R* (version 4.2.0; R Core Team, 2022) in order to further explicate the findings of the distributional, random forest, and conditional inference tree analyses. The model presented in Table 3 includes the main predictors of community, age group, intervening material, and type of subject, with individual as a random effect. In earlier models, education and gender returned nonsignificant results, so these were pruned for the final model. Also included are interaction terms for community and the linguistic factors of animacy, community, and age group, both indicated in the distributional data. The stars indicate whether the factors are significant and to what degree, as per the significance codes listed at the bottom of the table. The model is based on ninety-five individuals and 12,835 tokens.[16] Baseline levels are listed first followed by predicted levels. Percent is the proportion of SD for each level, and *n* gives the number of observations for each level.

The model in Table 3 confirms the trends that were previously visualized in the distributional charts, random forests, and conditional inference trees. Community, age group/date of birth, animacy, intervening material, and occupation are significant, consistent with Figures 9 and 10. Proper names are highly significantly contrasted with other human and nonhuman noun phrase subjects. Subjects adjacent to pronouns are significantly different from those with intervening material (linguistic material or hesitations). Blue-collar workers are significantly different from white-collar workers and students. Finally, the interaction terms demonstrate there is little to no significant interaction across language cohorts with respect to animacy, but significant interactions by age group between Parry Sound and *both* Kapuskasing Anglophones and Francophones.

Summary of results

A key consideration in this study has been to evaluate the use of SD in two Ontario towns with a view to sociolinguistic typology and social characteristics of the communities, including population size, network types, and degree of language contact

Table 3. Mixed-effects logistic regression for the use of SD

Effect	Percent	n	Log-odds	SE	z	p > /z/
Intercept			−2.13	0.17	−12.32	<2e-16***
Community						
Parry Sound	5.5	5994				
Kapuskasing Anglos	10.0	4050	0.98	0.33	3.01	0.003**
Kapuskasing Francos	10.4	2791	1.69	0.40	4.29	1.8e−05***
Age Group/Date of birth						
1936–1956	8.5	6464				
1957–1983	6.1	3745	−0.11	0.30	−0.36	0.717
1984–2001	9.2	2626	−1.15	0.39	−2.97	0.003**
Animacy						
Proper name	14.3	1369				
Human	9.7	5255	−0.90	0.16	−5.54	3.1e−08***
Other	5.0	6211	−1.81	0.18	−10.17	<2e−16***
Intervening material						
Nothing	5.0	11,001				
Adverb, etc.	25.1	1687	2.08	0.08	26.69	<2e−16***
Hesitation	32.7	147	2.15	0.20	10.78	<2e−16***
Occupation						
Blue collar	7.9	5515				
White collar	7.2	5887	−0.79	0.20	−3.90	9.7e−05***
Student	1.1	1443	−1.18	0.49	−2.42	0.015*
Interactions:						
Community/Animacy						
Kapuskasing Francos:Human	2.7	1211	−0.56	0.31	−1.82	0.069
Kapuskasing Anglos:Human	4.0	1719	0.13	0.23	0.57	0.568
Kapuskasing Francos:other	2.0	1473	−0.06	0.32	−0.17	0.862
Kapuskasing Anglos:other	1.9	1950	0.13	0.25	0.51	0.613
Age Group/Community						
1957–1983:Kapuskasing Francos	1.9	1028	−0.43	0.48	−0.90	0.369
1984–2001:Kapuskasing Francos	2.6	533	1.64	0.62	2.63	0.009**
1957–1983:Kapuskasing Anglos	2.9	1634	−0.40	0.44	−0.90	0.366
1984–2001:Kapuskasing Anglos	5.9	1296	1.85	0.64	2.91	0.004**

Significance codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' ', 1

to understand language variation and change. Each analysis conducted has pointed to distinct patterns for Parry Sound as opposed to Kapuskasing. While there are some differences between Anglophones and Francophones in Kapuskasing, most patterns are largely parallel. The contexts that are significantly different provide clues

for understanding the differences between Parry Sound and Kapuskasing. In particular, the interaction between community and intervening material confirms that the contrast between adjacent full NP subject + subject pronoun in Parry Sound is significantly different from Kapuskasing Francophones and Anglophones.

SD is reported in many languages and follows the implicational scale of the animacy hierarchy: *proper names > humans > other subjects*. Well studied in French, SD has been only cursorily reported in dialects of English in the United States and United Kingdom. This study has allowed us to espy two contrasting community types among people speaking English, one in a context of robust language contact between English and French (Kapuskasing) and one in a context of longitudinal monolingualism and low contact (Parry Sound). In both communities, individuals use SD; however, the linguistic and social distinctions by cohort differ, offering new explanations of this phenomenon from the perspective of sociolinguistic typology.

When all linguistic and social factors are modeled together and both communities and cohorts are included in the same analysis as independent variables, the presence of intervening material is the strongest predictor, with a community-level distinction depending on whether SD occurs in the context of intervening linguistic material versus nonlexicalized hesitations. There is also an additional divide by community: in Kapuskasing, SD has an age-graded pattern. In Parry Sound, it is obsolescent. In essence, SD in Kapuskasing and Parry Sound is a systematically conditioned feature of informal, vernacular usage but with distinct sociolinguistic patterns.

In Kapuskasing, where there is language contact, linguistic factors were the most prominent conditioning factor with a significant influence of speaker date of birth and language heritage. Dislocation with proper names stands out across generations as well as human subjects more generally. Notably, the Anglophone speakers in Kapuskasing preserve the Francophone norm of heightened use of SD with *all* human subjects. In Parry Sound, SD is present at low frequencies with only a mild tendency of use with proper names.

Discussion and conclusions

An earlier study of SD in Kapuskasing (Tagliamonte & Jankowski, 2019) suggested that the pattern of change and subsequent parallelism between the young Anglophones and francophones was the result of alignment facilitated by local sociolinguistic conditions. The concept of "alignment" in language can be traced to Goffman (1967:128). In more recent work, it is explained as the use of a linguistic feature to bring together more than one cultural value at the same time—"conceptual alignment" (Chandler, 2007:117-124, 306)—or "syntactic parallelism" (Cheshire 1998:139) to signal congruous discourse between speakers in an interaction. Such "syntactic harmony […] typically occurs when the conversation is proceeding harmoniously, with speakers cooperating to produce felicitous discourse" (Cheshire, 1998:139-140). The status of SD as a more widespread vernacular feature of English is equally compatible with this explanation. However, further research on the nature of SD in English is necessary to understand if and how it is used in other English-speaking communities. As far as Kapuskasing English is concerned, French linguistic features are evident in other areas of grammar, such as the discourse-pragmatic *there* as in "We lost a lot of jobs *there*" (Tagliamonte & Jankowski,

2020), the polarity in the use of *going to* and *will* for future temporal reference, such that *will* is favored for negative constructions, as in "He *won't* be back until next Sunday" (Roussel & Tagliamonte, 2022a), and, more recently, the use of *you know* as an utterance final tag (Roussel & Tagliamonte, 2022b). Both phenomena indicate that Kapuskasing is a type 6 community in Trudgill's (2011) model and has a distinctive place in the panoply of English varieties in Ontario. The blend of English and French features suggests dual cultural comfort among interlocutors from both heritage language groups for communication with a French language flavor. The cross-system inventory of linguistic phenomena engaged in social alignment means that SD is only one of a suite of features that Kapuskasing Anglophones employ for this purpose and reinforces the idea that it is a vernacular norm.

The comparative sociolinguistic analysis between Kapuskasing and Parry Sound offers an insightful new perspective on SD. First, this comparison confirms that it is a low-frequency vernacular feature of English. Second, SD in these communities exhibits a prominent internal linguistic constraint—type of subject—reported for other languages as well as in Canadian French. Moreover, the patterning of this constraint is largely parallel between the reports in the literature and the English varieties reported here, which embody an animacy hierarchy from proper name to human subject to animate subject to inanimate subject, a common typological continuum for animacy across the world's languages (e.g., Li & Thompson, 1976). Third, the contrasts *within* the type of subject continuum differ between the Anglophones and Francophones, mirroring similar differences in the typology of this universal tendency across languages. Fourth, Kapuskasing and Parry Sound demonstrate similar usage of SD according to the syntactic structure where it occurs. In both communities, SD arises most frequently in contexts of intervening linguistic material, whether adverbs and clauses or nonlexicalized spoken language phenomena (hesitation or pausing). Both cases produce a break in the syntax and a separation between the subject and the co-referential pronoun, showing that SD is produced by the same underlying mechanism in each situation. These results are notably in synch with those of Torres Cacoullos and Travis (2019) where traits that distinguish typologically distinct languages, that is, null subject language (Spanish) and nonnull subject language (English), share the same structured variability. Fifth, the antithetic social patterning of SD in Kapuskasing and Parry Sound demonstrates that its cultural embedding and local meaning differ depending on community type and the local ecology, specifically, in this case, because of increasing bilingualism among both Anglophone and Francophone heritage individuals in Kapuskasing.

Taking all these results into consideration, we can affirm the previous explanation of SD in Kapuskasing as a noteworthy indicator of ongoing alignment between French and English in northern Ontario. This finding helps elucidate the language situation in northern Ontario and demonstrates how language contact in small communities with tight social networks and high language contact may be catalysts for patterns of social alignment. We have also established the underlying sociolinguistic patterns of SD in a small English community with tight social networks but low language contact, Parry Sound, which lays the foundation for further comparison of this feature in other types of English-speaking communities where SD is also part of the local milieu. In this endeavor, we, like Torres Cacoullos and Travis (2019:684) advocate for a "flesh

and bones" approach—analyses that include not only attention to the presence and frequency of linguistic features, but also nuances of the variable context, the internal linguistic constraints, and importantly, the local social situation in order to determine if the patterns of language use are universal, typological, or social.

To conclude, the use of a dislocated subject for topics in conversational discourse is a feature common to many languages. In vernacular parlance in Ontario, when speakers pause, embellish or restart when they tell a tale, SD facilitates introducing the protagonist(s) and/or focal object(s) of the story, as in (22).

22. That one year, when *the ministry* said there was no wolves, uh, *the Elton boys and us and some of the other trappers,* **we** got sixty-two wolves in that area. And *the old deer* just come ahead and *the beaver* and everything. *Wolf* eats a lot of beaver.
(male, 75, b. 1943, Parry Sound, Anglophone)

We suggest this phenomenon deserves more scrutiny in other communities and languages as a touchstone attribute to probe more deeply how sociolinguistic typology can help elucidate language variation and change, but also how grammar, usage, and cognition are related.

Competing interests.   The authors declare none.

Notes

1. According to Pérez Guerra and Tizón-Couto (2004:fn 1), the label "left dislocation" belongs to the subtype of "hanging topic left dislocation" (Ross, 1967; Van Riemsdijk, 1997; Vat, 1997).

2. Tagliamonte and Jankowski (2019) used the term "subject doubling." Here we will use the term "dislocation" as a more general phenomenon, consistent with the literature on Canadian French.

3. This is a popular contrast in the community where members distinguish each other by heritage group: "Anglophones," "the English," "les Anglais," or "Francophones," "the French, "les Français."

4. Prosodic cues and pragmatic function/discourse context also distinguish doubling from dislocation in studies of French (Culbertson, 2010). Nagy et al. give several reasons why they "did not make a distinction between subject doubling and left dislocation" in their study of Montréal French. They referred to the feature as "subject doubling" based at least in part on the ability of quantified subjects in their data to be doubled (Nagy et al., 2003:78).

5. De Cat's data (2007:217-246, appendices A and B) comprise fifteen adults and six children, recorded from 1997-2001, representing three varieties of French: Belgian (Brussels, Liège), Canadian (Montréal), and France (Paris, Pyrénées). It is unclear how many adult informants were ultimately used in the 5,613 token adult speaker analysis, which "corresponds to three recording sessions per country, selected randomly but from across the whole time span of the recording period" (De Cat, 2007:219).

6. Not to be confused with the closely related Gallo-Romance language Picard.

7. An anonymous reviewer pointed out that the 1972 children's book *Timothy the Terror*, by Ruth Calvin, makes considerable use of SD in the representation of AAVE speech.

8. See Don Harron discuss Charlie Farquharson in a video from the Canadian Broadcasting Corporation at: https://www.cbc.ca/player/play/2209488288.

9. http://ontariodialects.chass.utoronto.ca/ Accessed 2-17-21.

10. These legacy materials were collected by local historian John Macfie.

11. Bilingualism in Kapuskasing is the community norm. However, the level of fluency in the two languages varies by generation, with more widespread bilingualism among younger generations. We used self-identification as the measure of Anglophone versus Francophone heritage.

12. In Parry Sound, 6/49 individuals, three men and three women, have no SD tokens. In Kapuskasing, 2/46 individuals never use SD despite the fact that these individuals used many lexical subjects, and their interviews were as informal as any other.

13. T. Lovely, an animated speaker with exceptional storytelling skills, used more SD than anyone, suggesting an element of style may also be involved in subject doubling use.

14. Modeling speaker date of birth as continuous or by decade does not capture the linear decline in SD in Parry Sound nor the age-graded U-shaped curve for Kapuskasing cohorts. No models including these factors returned significant results. The factor of "bymo.age" used in Figure 9 and for the remaining analyses bins speaker age at time of interview into three categories: *young* = 17-29, *mid* = 30-59, and *older* = 60+ or as in the GLMER model in Table 3: 1936-1956, 1957-1983, and 1984-2001.

15. model_data <-ctree(dep_var ~ Occ + Edu + Community + Animacy + Intervening, data=sub-doub, control=ctree_control(minbucket=100)).

16. glmer_5_13_22 <- glmer(dep_var ~ Age_Group + Community + Animacy + Between + Community + Occupation + Community * Animacy + Community * Age_Group + (1|shortform.name), data = subd2, family = "binomial," control = glmerControl(optCtrl = list(maxfun = 2e4), optimizer = "bobyqa"))

References

Auger, Julie. (1998). Le redoublement des sujets en français informel Québécois: Une approche variationiste. *Canadian Journal of Linguistics* 43:37–63.

Auger, Julie. (2003). Les pronoms clitiques sujets en picard: Une analyse au confluent de la phonologie, de la morphologie et de la syntaxe. *Journal of French Language Studies* 13(1):1–22.

Auger, Julie, & Villeneuve, Anne-José. (2010). La double expression des sujets en français saguenéen. In W. Remyse, & D. Vincent (eds.), *Hétérogénéité et homogénéité dans les pratiques langagiéres: Mélanges offerts à Denise Deshaies*. Québec: Presse d'l'Université Laval. 67–86.

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

Bates, Douglas M., Maechler, Martin, & Bolker, Ben. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375–42.

Blanquart, Eleonore. (2012). *Subject doubling in spoken French*. Master of Arts thesis, Arizona State University.

Carvalho, Ana Maria. (2016). The analysis of languages in contact: A case study through a variationist lens. *Cadernos de Estudos Linguísticos* 58(3):401–424.

Chandler, Daniel. (2007). *Semiotics: The basics*. London: Routledge.

Cheshire, Jenny. (1998). English negation from an interactional perspective. In P. Trudgill & J. Cheshire (eds.), *The sociolinguistics reader, vol. 1: Multilingualism and variation*. London: Edward Arnold. 127–144.

Coveney, Aiden. (2005). Subject doubling in spoken French: A sociolinguistic approach. *The French Review* 79(1):95–111.

Culbertson, Jennifer. (2010). Convergent evidence for categorial change in French: From subject clitic to agreement marker. *Language* 86(1):85–132.

De Cat, Cécile. (2007). *French dislocation–Interpretation, syntax, acquisition*. Oxford: Oxford University Press.

Dunlap, David W. (2017). A New York Times company town deep in the Ontario wilderness. *New York Times*. https://www.nytimes.com/2017/06/30/world/canada/a-new-york-times-company-town-deep-in-the-ontario-wilderness.html.

Fonseca-Greber, Bonnie B. (2000). *The change from pronoun to clitic and the rise of null subjects in spoken Swiss French*. Doctoral dissertation, University of Arizona.

Godfrey, John J., & Holliman, Edward. (1993). Switchboard–1 Release 2 LDC97S62. Web Download. Philadelphia: Linguistic Data Consortium.

Goffman, Erving. (1967). *Interaction ritual*. New York: Anchor Books.

Gries, Stefan Th. (2018). On classification trees and random forests in corpus linguistics: Some words of caution and suggestions for improvement. *Corpus Linguistics and Linguistic Theory* 16(3):617–647.

Jankowski, Bridget L., Needle, Jeremy M., & Tagliamonte, Sali A. (2022). Is it a camp or a cottage?: The cultural evolution of a lexical item in Ontario. Paper presented at the *American Dialect Society Annual Meeting*. Washington, D.C., January 6-9.

Labov, William. (1972). The study of language in its social context. In W. Labov (ed.), *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press. 183–259.

Labov, William. (1982). Building on empirical foundations. In W. P. Lehmann & Y. Malkiel (eds.), *Perspectives on historical linguistics*. Amsterdam: John Benjamins. 17–92.

Liberman, Mark. (2008). Left dislocation. *Language Log: Syntax*. https://languagelog.ldc.upenn.edu/nll/?p=629.

Li, Charles N., & Thompson, Sandra A. (1976). Subject and topic: A new typology of language. In C. N. Li (ed.), *Subject and topic*. New York: Academic Press. 457–489.

Mougeon, Raymond, Beniak, Edouard, & Valois, Daniel. (1985). A sociolinguistic study of language contact, shift and change. *Linguistics* 23(3):455–487.

Nadasdi, Terry. (1995a). Subject NP doubling, matching, and minority French. *Language Variation and Change* 7(1):1–14.

Nadasdi, Terry. (1995b). *Variation morphosyntaxique et langue minoritaire: Le cas du français ontarien*. Doctoral dissertation, University of Toronto.

Nagy, Naomi, Blondeau, Hélène, & Auger, Julie. (2003). Second language acquisition and "real" French: An investigation of subject doubling in the French of Montreal Anglophones. *Language Variation and Change* 15(1):73–103.

Pérez Guerra, Javier, & Tizón-Couto, David. (2004). "These hands, they are apt enough to dislocate and tear they flesh": On left dislocation in the recent history of the English language. In B. Shaer, W. Frey, & C. Maienborn (eds.), *Proceedings of the Dislocated Elements Workshop*. Berlin: ZAS Berlin. 449–464.

Poplack, Shana, & Tagliamonte, Sali A. (2001). *African American English in the diaspora: Tense and aspect*. Malden: Blackwell Publishers.

R Core Team (2022). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rizzi, Luigi. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17(3):501–557.

Roberge, Yves. (1990). *The syntactic recoverability of null arguments*. Kingston: McGill-Queen's University Press.

Ross, John Robert. (1967). *Constraints on variables in syntax*. Bloomington: Indiana University Linguistics Club.

Roussel, Basile, & Tagliamonte, Sali A. (2022a). Frenching it up: The English future temporal reference in a French dominant town. Paper presented at the *Linguistic Symposium on Romance Languages (LSRL 52)*. Madison, Wisconsin. April 21–23.

Roussel, Basile, & Tagliamonte, Sali A. (2022b). "More French than anything, you know?" A variationist account of you know in Northeastern Ontario. Paper presented at the *Atlantic Provinces Linguistics Association*. Halifax, Nova Scotia, Canada. November 4-5.

Sankoff, David. (1978). Probability and linguistic variation. *Synthèse* 37:217–238.

Sankoff, Gillian. (1982). Usage linguistique et grammaticalisation: Les clitiques sujets en français. In N. Dittmar (ed.), *La sociolinguistique dans les pays de langue romane*. Tübingen: Gunter Narr. 81–85.

Southard, Bruce, & Muller, Al. (1998). Blame it on Twain: Reading American dialects in The Adventures of Huckleberry Finn. In D. Oaks (ed.), *Linguistics at work: A reader of applications*. San Diego: Harcourt Brace. 565–573.

Statistics Canada. (2017). Parry Sound, Census Profile. Statistics Canada Catalogue no. 98–316-X2016001. Ottawa. Released November 29, 2017. In Secondary Statistics Canada (ed.), Secondary Parry Sound, Census Profile. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa. Released November 29, 2017.

Strobl, Carolin, Malley, James D., & Tutz, Gerhard. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*,14(4):323–348.

Tagliamonte, Sali A. (2002). Comparative sociolinguistics. In J. K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *Handbook of language variation and change*. Oxford: Blackwell Publishers. 729–763.

Tagliamonte, Sali A. (2006). Conversations from the speech community: Exploring variation and change in synchronic dialect corpora. Paper presented at the *International Computer Archive of Modern and Medieval English (ICAME)*. Helsinki, Finland, May 24-28.

Tagliamonte, Sali A. (2012). Comparative sociolinguistics. In J. K. Chambers & N. Schilling-Estes (eds.), *Handbook of language variation and change* (2nd ed.). Oxford: Blackwell Publishers.

Tagliamonte, Sali A. (2013-2018). Social determinants of linguistic systems. Insight Grant: Social Sciences and Humanities Research Council of Canada #435130065.

Tagliamonte, Sali A., & Baayen, R. Harald. (2012). Models, forests and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(2):135–178.

Tagliamonte, Sali A., & Jankowski, Bridget L. (2019). Grammatical convergence or microvariation? Subject doubling in English in a French dominant town. *Proceedings of the Linguistic Society of America (PSLA)* 17(4):1–15.

Tagliamonte, Sali A., & Jankowski, Bridget L. (2020). Up north *there*: Discourse-pragmatic deixis in Northern Ontario. *Journal of Pragmatics* 170:216–230.

Thibault, Pierrette. (1983). *Équivalence et grammaticalisation*. Doctoral dissertation, Université de Montréal.

Torres Cacoullos, Rena, & Travis, Catherine E. (2019). Variationist typology: Shared probabilistic constraints across (non-)null subject languages. *Linguistics* 57(3):653–692.

Town of Kapuskasing. (2010). "Kapuskasing Community Profile." http://www.kapuskasing.ca/en/Living/resources/Documents/Community_Profile_Kapuskasing.pdf.

Traugott, Elizabeth Closs. (2007). Old English left-dislocations: Their structure and information status. *Folia Linguistica* 41(3–4):405–441.

Trudgill, Peter J. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Van Riemsdijk, Henk. (1997). Left dislocation. In E. Anagnostopoulou, H. Van Riemsdijk, & F. Zwarts (eds.), *Materials on left dislocation*. Amsterdam: John Benjamins. 1–10.

Vat, Jan. (1997). Left dislocation, connectedness and reconstruction. In E. Anagnostopoulou, H. Van Riemsdijk & F. Zwarts (eds.), *Materials on left dislocation*. Amsterdam: John Benjamins. 67–92.

Wolfram, Walt & Christian, Donna. (1976). *Appalachian speech*. Arlington: Center for Applied Linguistics.

Zahler, Sara. (2014). Variable subject doubling in spoken Parisian French. *University of Pennsylvania Working Papers in Linguistics* 20(1):361–371.