Both (of) the Variants Show a Couple (of) Different Patterns: Social Conditioning of *of*-Variation across Multiple Linguistic Environments

Molly Cutler, Auromita (Disha) Mitra, Marc Barnard, Jessica Göbel, and Laurel MacKenzie*

1 Introduction

A longstanding question in sociolinguistics is whether the social patterning of a variant is consistent across linguistic environments. It is traditionally assumed that language-external (i.e., social) factors do not interact with language-internal factors (i.e., linguistic environment) in the conditioning of variation (Labov 1993, Labov 2001:28, Labov 2010:265), but this is largely untested with modern statistical methods and large data sets.

In this paper, we report on a study of a single variable alternation in English—between *of* and ∅, henceforth "*of*-variation"—that is instantiated in several distinct linguistic environments. We find that its social patterning differs by environment, suggesting a counterexample to the proposed independence of external and internal conditioning factors.

2 Background

2.1 Internal/External Interaction in the Conditioning of Sociolinguistic Variation

Decades of work in quantitative variationist sociolinguistics has revealed that variation in language is not random, but is systematically shaped by a number of factors (Weinreich et al. 1968, Bayley 2013). These "conditioning factors" include social characteristics of the speaker and of the setting ("language-external factors") and linguistic characteristics of the utterance containing the varying item ("language-internal factors").

Dating back to at least the late 1970s, a question of interest in the variationist literature has been the extent to which internal and external conditioning factors can interact. Weiner and Labov (1983), in a paper that was first circulated in 1977, examined the alternation between passive and active sentences in English, and found that the linguistic constraints on this variation were consistent across sex, class, age, and ethnic groups. That is, all social groups showed the same hierarchy of internal constraints on the variation, suggesting that internal and external factors are independent from one another. Although subsequent researchers have cast doubt on whether the passive/active alternation in English should indeed be modeled as a sociolinguistic variable (Lavandera 1978, Romaine 1981), the idea of internal/external independence was influential. It led Labov to assert that the sociolinguistic monitor, the cognitive mechanism hypothesized to process frequency of variable use, "does not, perhaps can not, access information on linguistic structure" (Labov 1993:24), and that "for the most part, linguistic structure and social structure are isolated domains, which do not bear upon each other" (Labov 2001:28). A counterexample to this hypothesis in the domain of language production could take the form of a variable alternation showing social conditioning (e.g., women using one form more than men) in one linguistic environment but not another, or showing opposite social patterns in two environments. In language perception, a counterexample could take the form of a variant being socially evaluated—say, sounding uneducated—in one linguistic environment but not another.

To determine whether internal and external factors interact for a given linguistic variable, a researcher needs to test the significance of a statistical interaction between one or more internal and one or more external predictors. This was not generally not possible in early sociolinguistic work,

*This work began as a class project in LING-GA 2530 Linguistic Variation at New York University, Spring 2024. We gratefully acknowledge Valentina Cojocaru for her role in contributing to this initial stage. A version of this work was presented at New Ways of Analyzing Variation 52. We thank the audience of that conference and members of the NYU Sociolinguistics Lab for their feedback.

because statistical software at the time precluded it. Researchers instead added or subtracted social factors from a statistical model and determined whether the effect of linguistic factors changed as a result (Cedergren and Sankoff 1974, Weiner and Labov 1983). There is thus an opportunity for researchers working with modern statistical tools to revisit the question of hypothesized internal/external independence.

One study that has explored this question in the domain of perception is Maddeaux and Dinkin (2017). They performed a matched-guise study to elicit social evaluations of the form *like* in several different linguistic contexts (e.g., when functioning as a discourse particle, as an approximative adverb, as a quotative…). Their results were somewhat inconclusive: "[T]here is no evidence for the hypothesis that the different functions of *like* have similar social meanings, though also not very strong evidence against it" (Maddeaux and Dinkin 2017:4).

In the domain of production, we have more convincing evidence that social and linguistic factors can interact. Göbel (2020) examined /t/-glottaling in Edinburgh English and found an interaction between social class and the following phonological environment. Specifically, Working Class and Middle Class speakers had a different ordering hierarchy of the following phonological context (vowel, consonant, or pause) than speakers that had been upwardly socially mobile throughout life, also known as the "New Middle Class" (Dickson and Hall-Lew 2017). Göbel suggested that the same variant (glottal /t/) could have entered Scottish English from two separate sources: Standard Scottish English (SSE) and Scots. Plausibly, different levels of dialect contact across the classes has resulted in the different constraint orderings. Given that the question of internal/external independence remains an open one, we bring new data from *of*-variation to bear on it.

2.2 *Of*-variation in English

This paper's focus is on environments in some varieties of English where *of* can variably be absent. In (1–7), we exemplify the environments we selected for study based on this criterion. For each, we provide one example with *of* and one without (i.e. with ∅). All examples are taken from the Philadelphia Neighborhood Corpus (PNC, Labov and Rosenfelder 2011), our data source (described further in section 3.1), and speaker identification numbers are given in parentheses.

(1) Inverted-degree construction
   a. It wasn't **that big of a thing.** (PH00-1-4)
   b. It was **that close ∅ a relationship.** (PH79-3-6)

(2) *couple*
   a. She was fine for a **couple of months**. (PH90-2-5)
   b. He was working there for a **couple ∅ months**. (PH00-1-3)

(3) *all*
   a. I mean we were always respectful, respect **all of our neighbors** and stuff. (PH10-1-2)
   b. They want to be able to know **all ∅ their neighbors**. (PH82-1-10)

(4) *half*
   a. **Half of the time**, he wouldn't be there. (PH81-0-3)
   b. **Half ∅ the time** I'll just say they can just sleep overnight. (PH12-2-1)

(5) *both*
   a. Well, **both of our parents** were in the air force. (PH80-2-4)
   b. But **both ∅ our parents** were born here. (PH10-2-4)

(6) *off*
   a. He's been knocked **off of his bike** and stuff. (PH84-1-2)
   b. Like if he fell **off ∅ his bike** he'd say, "You see him wreck out on his bike?" (PH74-0-8)

(7) *out(side)*
   a. Today, you can't even put your head **out of your door** at night without fearing that someone's going to come in and hurt you. (PH12-2-10)
   b. You look **out ∅ your door** and if you need any help, you can holler. (PH84-1-4)

The apparent surface similarity of the alternation between *of* and ∅ across the environments in (1–7) masks different diachronic trajectories, regional preferences, and levels of prescriptivism. In some environments, *of* is diachronically innovative, meaning that the variability reflects a process of *of*-insertion; in others, ∅ is the innovative variant, suggesting *of*-deletion. In still other environments, both variants have been present in the language for centuries. The rest of this section summarizes these diachronic and social patterns for the seven environments under study, grouping those with similar profiles.

The inverted degree construction (henceforth IDC) in (1) is a clear example of diachronic *of*-insertion. This construction consists of a degree word (e.g., *so*, *as*, *too*, or *how*) modifying an adjective, followed by a singular noun introduced by an indefinite determiner. (See Blass 2020 for a recent review of work on this construction, where it is called the Big Mess Construction.) Though this construction has been used productively since Early Modern English, it did not regularly occur with *of* in that era (Sommerer 2022). The earliest documented instances of IDCs with *of* date to the 1940s; it is a specifically American variant (Merriam-Webster 2002, s.v. *of a*) that is favored in spoken over written language (Allen 1989, Rapp 1991), is increasing in acceptability over time (Nylund and Seals 2010), and is disfavored by prescriptivists (e.g., Garner 2022, s.v. *of*).

Showing a nearly identical diachronic and social profile, but for the ∅ variant, is *couple*, as in (2). *Couple* has been attested with *of* since Middle English, but with the ∅ variant only since the 19th century (Oxford English Dictionary, s.v. *couple*). The ∅ variant after *couple* is more common in American than British English (Merriam-Webster 2002, s.v. *couple*, adj.; Wood 2019), favored in vernacular language (ibid.), and disfavored by prescriptivists (e.g., Bernstein 1977:54; Garner 2022, s.v. *couple*).

The other environments in our study have more complicated profiles. *All*, as in (3), is attested with *of* as far back as Middle English, and with ∅ since Early Modern English (Oxford English Dictionary, s.v. *all*), though *of* is nonetheless perceived as an innovation (Zwicky 2009). *Of* after *all* is also associated with American English (Garner 2022, s.v. *all*), and is indeed used more in that variety than in British English, though still at a very low rate (Estling 2004). It is disfavored by prescriptivists (e.g., Flesch 1964:19; Garner 2022, s.v. *all*). *Half*, as in (4), is attested with both *of* and ∅ as far back as Middle English (Oxford English Dictionary, s.v. *half*, adj.; *half*, n.). As with *all*, the *of* variant is used more in American than British English (Estling 2000, 2004) and is disfavored by prescriptivists (e.g., Garner 2022, s.v. *half*). Likewise, *both* is attested with *of* and ∅ as far back as Middle English (Oxford English Dictionary, s.v. *both*), though the Oxford English Dictionary's earliest attestations of *both of* precede a pronoun, an environment where *both* invariably takes *of* in Mainstream American English today (Merriam-Webster 2002, s.v. *both*), perhaps suggesting that the variation with *both* before non-pronominal determiner phrases, as in (5), is more recent. As with *all of* and *half of*, *both of* is used more in American than British English (Estling 2000, 2004), though the same prescriptivists who disfavor *all of* and *half of* do not prescribe against *both of* (Flesch 1964, Garner 2022).

Off, as in (6), shows a profile that is reminiscent in some ways of those of *all, both,* and *half*. After *off*, the ∅ variant goes back to Old English, but the *of* variant also has a long history, with attestations dating back at least to the 16th century (Vartiainen and Höglund 2020; Oxford English Dictionary, s.v. *off*). Still, the *of* variant is perceived as an Americanism—despite its widespread use across the UK (Vartiainen and Höglund 2020)—and is disparaged by prescriptivists (e.g., Flesch 1964:211; Bernstein 1977:153; Garner 2022, s.v. *off of*). Its use in American English appears to have been increasing since the 1970s (Vartiainen and Höglund 2020).

Finally, we group *out* and *outside* together, as in (7). In the case of *outside*, the *of* variant is again disparaged by prescriptivists (e.g., Garner 2022, s.v. *outside*, and see additional examples in Merriam-Webster 2002, s.v. *outside of*, and Zwicky 2007). Attestations of *outside of* and *outside* ∅ first appear in the 1700s (Oxford English Dictionary, s.v. *outside*), but no diachronic or synchronic corpus work on this specific environment seems to exist. Despite the prescriptivist awareness of *outside of*, the analogous variation between *out of* and *out* ∅ goes unmentioned by prescriptivists (e.g., Flesch 1964, Bernstein 1977, Garner 2022, none of whom have entries for *out of*). This may be because, in Mainstream American English, the variation is restricted to cases where the following noun is what Estling (1999) calls an "aperture," namely nouns like *door* and *window*

(Cappelle 2001:321, Zwicky 2007). When *out* is combined with determiner phrases containing *door* or *window*, the ∅ variant is used more in American than British English (Estling 1999, Cappelle 2001:fn. 5). In British English, its use is mostly confined to spoken language (Estling 1999, Cappelle 2001:fn. 5), though it may be increasing in real time in written British English (Estling 1999). We also note that the ∅ variant of *out* can be used with non-aperture nouns (as in *out the house, out the car*) in many other varieties of English (Sánchez 1986, Cappelle 2001, Merriam-Webster 2002, s.v. *out of*, Vasko 2007), and appears to have been possible in early English as well, with attestations going back to the 1300s (Oxford English Dictionary, s.v. *out*). *Out of* is similarly longstanding, with attestations back to Old English (Oxford English Dictionary, s.v. *out of*).

In sum, in most of the environments selected for study here, the *of* variant is prescriptively disfavored and/or associated with American English, either in actual or perceived use. This is the case for IDCs, *all, half, both*, and *off*. In some of these cases (IDCs and *off*), *of* is also diachronically innovative and/or increasing in real time. In the case of *couple*, ∅ is the newer, American, prescriptively disfavored variant. And, finally, *out* appears to be a case where both *of* and ∅ have been in use for centuries, with ∅ before non-apertures no longer possible in Mainstream American English, but ∅ before apertures characteristic of American English and potentially increasing in real time. These different diachronic trajectories and levels of prescriptivism already suggest that *of*-variation may provide a counterexample to Labov's hypothesized segregation of internal and external constraints. Sociodemographic patterns of production—which we analyze in section 4—will provide further support.

We close this section by noting that *of*-variation has been observed in more contexts than just those in (1–7). Kjellmer (2007) finds evidence of the ∅ variant in British, American, and Australian Englishes after partitives including *lots* (e.g., *lots ∅ factors involved*), *pair* (e.g., *a pair ∅ clogs*), *variety* (e.g., *a variety ∅ things*), and *number* (e.g., *a number ∅ points needed to be clarified*), though generally at very low rates. And Sánchez (1986) provides evidence of the ∅ variant in African American Vernacular English after partitives including *one* (e.g., *one ∅ my partners*) and *bunch* (e.g., *a bunch ∅ s.b.'s*) and prepositions including *in front* (e.g., *in front ∅ a woman*) and *upside* (e.g., *upside ∅ your head*). We restrict our study to the environments in (1–7) because their prescriptivism and diachrony are well documented, but future research could explore whether additional environments show the same variation in our data.

3 Methods

3.1 Materials

We examine *of*-variation in data from the PNC (Labov and Rosenfelder 2011), which contains sociolinguistic interviews from speakers of Philadelphia English from a variety of economic, educational, and ethnic backgrounds, collected over a period of several decades between 1973 and 2012.

We restricted our data extraction to environments within the envelopes of variation for *of*-variation in the various contexts under study. For instance, quantifiers are said to invariably surface with *of* before object pronouns (e.g., *all of it/*all ∅ it, both of them/*both ∅ them*) (Merriam-Webster 2002, s.v. *all of*; *both*), so our search did not target these environments. These same quantifiers also invariably surface with ∅ before bare plural nouns (e.g., *all ∅ people/*all of people, both ∅ kids/*both of kids*). Accordingly, for *all, both,* and *half*, we searched the corpus only for these forms when preceding determiner phrases starting with *the, a(n),* a possessive pronoun, or a demonstrative pronoun. *Of* or *o'* was allowed to optionally intervene between quantifier and determiner phrase.

We similarly restricted our corpus searches for *out(side)* and *off* to only tokens preceding determiner phrases starting with *the, a(n),* a possessive pronoun, or a demonstrative pronoun. Again, *of* or *o'* was allowed to optionally intervene between preposition and determiner phrase. This was an attempt to weed out irrelevant hits, though it did have the downside of excluding from data extraction instances where *out(side)* and *off* precede a bare noun phrase, where variation can in fact

occur (e.g., *out (of) windows, off (of) Facebook*). Still, we deemed that preferable to wading through thousands of spurious hits. We did not attempt to restrict the nouns following *out (of)* to apertures as Estling (1999) did (and, indeed, we did find instances of *out* ∅ with non-aperture nouns).[1]

For tokens of IDCs, we searched the corpus for instances of *how, as, that, so, too, very, quite, rather,* or *more* followed by a single word (to capture the following adjective)—or a word ending in *-er* (to capture a comparative adjective)—followed by *a(n)*. Again, *of* or *o'* was allowed to optionally intervene between adjective and determiner phrase. Finally, tokens of *couple* were found simply by searching for *couple*, since variation in this environment occurs only with bare non-pronominal nouns (e.g., *couple (of) friends* but *couple of the kids*/**couple* ∅ *the kids*, *couple of them*/**couple* ∅ *them*), and there was no way to target such nouns directly.

In addition to the searches described above, we also searched for the fused forms *botha*, *halfa*, *alla*, *coupla*, and *outta*/*outa*. Searches were executed via Python scripting using regular expressions. We excluded any tokens for which sociodemographic information about the speaker was not available. The speakers that this applied to were generally interviewers, for whom sociodemographic metadata was not recorded by the corpus. We then read through all of the resulting hits and manually omitted instances of the following:

- lexical/idiomatic uses, e.g., *uh three hundred dollar houses and all that jazz* (PH86-3-1) (compare ??*all of that jazz*)
- false string matches, e.g., *It's not all her fault* (PH88-1-2) (compare ??*It's not all of her fault*)
- utterances that were mistranscribed, interrupted, or lacked sufficient context to reliably judge

Table 1 summarizes the number of usable tokens for each environment following these linguistic exclusions.

	all	*couple*	*out(side)*	*off*	*half*	*both*	IDC
Hits with sociodemographic information	2382	726	955	264	160	78	623
Linguistic exclusions	890 (37%)	180 (25%)	554 (58%)	86 (33%)	29 (18%)	40 (51%)	614 (99%)
Remaining (usable) tokens	1492	546	401	178	131	38	9

Table 1: Token counts across environments.

Given the low numbers of usable hits for *off, half, both,* and IDCs, we exclude them from further study and report only results for *all, couple,* and *out(side)*.[2] The final data set of 2,439 tokens comes from 352 unique speakers, of whom 56% are female[3]. 26% of speakers in the sample have greater than a high school education; the mean years of education is 12, with a minimum of 0 years (for a three-year-old child) and a maximum of 21 years. The birth years of the speakers in our sample range from 1888 to 1998, and the breakdown of ethnicities in the sample, as reported in corpus metadata, is 9% African American, 4% Hispanic, 2% unknown, and 85% European American, Jewish, or "white."

The dependent variable was coded as a binary variable (*of*-presence vs. *of*-absence, i.e. ∅). Instances transcribed as *o'* or a fused form (e.g. *coupla, outta,* etc.) were coded as *of* presence. We generally made our judgments of *of* presence or absence based on the corpus transcripts, though we did consult audio files when a transcription was unclear or improbable. A direction for future work is to code all data with reference to the audio, potentially treating reduced variants of *of* as their own category separate from *of* and ∅.

[1] We encountered three speakers, recorded separately in 2002, all of whom were asked to repeat the sentence *It will be nice to get out of the house for a while*, and all of whom repeated it without *of*, i.e., *It will be nice to get out* ∅ *the house for a while*.

[2] Of the 401 *out(side)* tokens, 53 are *outside*; the rest are *out*.

[3] No genders other than male or female are reported in the PNC metadata.

3.2 Research questions and hypotheses

As outlined in section 1, the main aim of this study is to assess whether social conditioning differs across linguistic environments. Towards this end, we ask the following research questions. First, to understand if linguistic structure plays a role in the variation, we ask Q1: Does the overall rate of ∅ vary by linguistic environment (*all* vs. *couple* vs. *out*)? Second, to assess the possible role of social factors, we ask Q2: Does the rate of ∅ in each linguistic environment vary by any social demographics tracked in the corpus? We specifically consider gender, years of formal education, and year of birth. Given the prescriptivism summarized in section 2.2, we expect that years of formal education in particular may be significant. Thus, we ask Q3: Does the effect of demographic factors on rates of ∅ differ by linguistic environment? This last question addresses our main aim, and we hypothesize that if social and linguistic factors do not interact in conditioning variation (the null hypothesis), we should find a negative answer to Q3 (i.e., no significant difference in effect), but if social and linguistic factors do interact in conditioning variation, we should find a positive answer to Q3 (i.e., a significant difference in effect).

3.3 Statistical Methods

We fitted binomial mixed-effects logistic regression models using the `lme4` package (Bates et al. 2015) in R (R Core Team 2024). The dependent variable was the probability (log odds) of a ∅ realization. Modeling the probability of ∅ (as opposed to *of*) is an arbitrary choice, and we follow this throughout the analysis, discussion, and in all figures for the sake of consistency. For Q1, the relevant predictor was linguistic environment (categorical variable: *all* vs. *couple* vs. *out*). For Q2, the relevant predictors were gender, years of schooling (numeric, centered at the mean), and year of birth (numeric, scaled to decade and centered at the mean).[4] To assess Q3, the question of social conditioning across environments, we tested the significance of interactions between each social factor and linguistic environment.

We also included two control predictors in modeling. One was speaking rate, calculated as words per second based on the transcript line in which a given token was uttered.[5] The second was the length of the constituent following the variation site, in number of words (e.g., *all of the women*: 2; *a couple of pretty bad fights*: 3; *out of that house across the street*: 5).

4 Results

Figure 1 shows the distribution of variants across environments. Rates of ∅ use are different in the three environments, with ∅ being used 97% of the time after *all*, 70% of the time after *couple*, and 28% of the time after *out*. The regression model confirms that these differences across linguistic environments are significant: *couple* and *out* both differ from *all*, the reference level (and best-represented environment), at $p < 0.001$. *Couple* and *out* also significantly differ from one another, based on a post-hoc test with Tukey correction. The very high rate of ∅ with *all* mirrors findings from previous corpus work (Estling 2000, 2004).

Turning to Q2, we find no significant effect of year of birth or years of education on the overall rates of ∅ use (when averaged over all the linguistic environments). Instead, bearing on Q3, these two social factors condition ∅ use in a more complex way. Specifically, year of birth conditions ∅ use only after *couple* (p=0.037); the positive model coefficient (β=0.184) indicates that speakers

[4] A concern could be raised that years of schooling and year of birth may be collinear, if speakers born later in the twentieth century were more likely to attend higher education than those born earlier. Indeed, there is a low but statistically significant correlation between the two predictors in our data (Spearman's ρ=0.16, $p < 0.001$). Accordingly, we used the `car` package to calculate the standard- error inflation factor (SIF) for each predictor, following Sonderegger (2023:274). No predictor exceeds the 3.2 SIF threshold for "potentially problematic collinearity" given by Sonderegger (2023:275); the relevant SIF values are all under 2.

[5] In cases where the dependent variable was realized as *of*, we subtracted 1 from the number of words in the transcript line, to correct for the increased length relative to ∅.

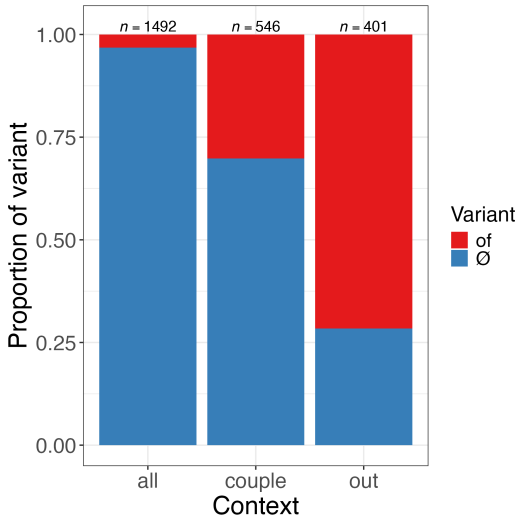

Figure 1: Distribution of variants across environments.

born more recently use more ∅, suggesting change in progress toward ∅. *All* shows no effect of year of birth (p=0.827), and *out* does not significantly differ from *all* (p=0.285), i.e., also shows no effect of year of birth. On the other hand, years of education conditions ∅ use in all three environments, but, importantly for Q3, these effects differ in both strength and direction. *All* shows decreased ∅ use among more educated speakers (β=−0.128, p=0.022), *out* does not significantly differ from this (i.e., also shows decreased ∅ use among more educated speakers; p=0.361), and *couple* shows the opposite pattern (increased ∅ use among more educated speakers; β=0.152, p=0.022). Figures 2 and 3 show these relationships.

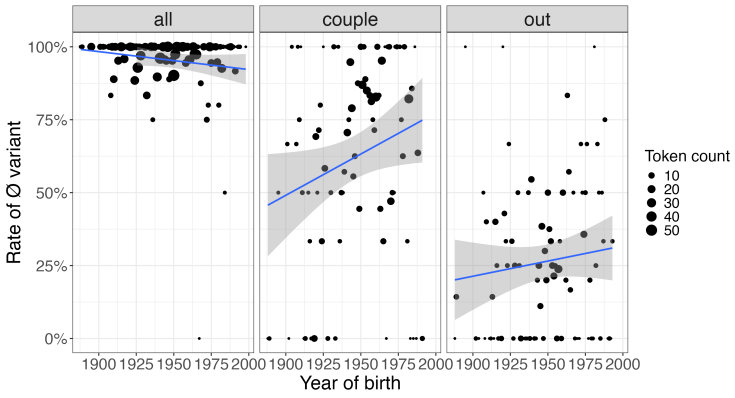

Figure 2: Rates of ∅ by year of birth.

We find no effect of gender on the variation (either averaging over environments or in interaction with environment), and no effect of following constituent length. There is a significant effect of speaking rate for *all*, such that a higher speaking rate predicts more ∅. This effect is significantly weakened (though not erased entirely) for *couple* and *out*.

5 Discussion

Alternation between *of* and ∅ is possible in a variety of linguistic constructions. The statistical measures reported in section 4 show differences in the behavior of the alternation after the quantifiers

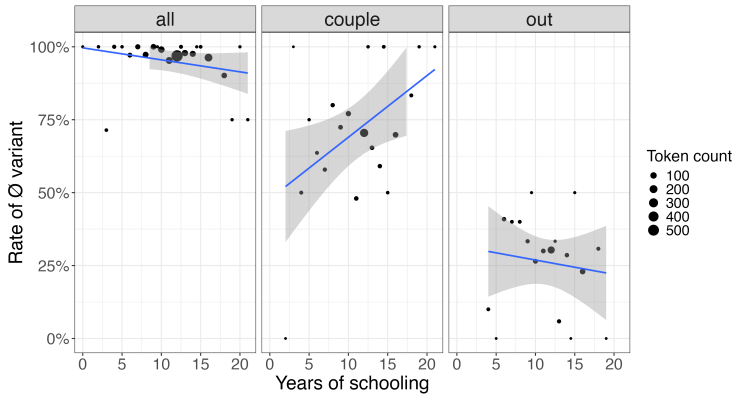

Figure 3: Rates of ∅ by years of education.

all and *couple* and the preposition *out* in the PNC: the ∅ variant occurs most frequently after *all*, followed by *couple*, and occurs least frequently after *out*. Since this variation has not been studied across these environments at this level of detail before, our results provide evidence that the phenomenon is conditioned by linguistic environment.

In light of existing literature suggesting that the choice between *of* and ∅ in some of these constructions is socially meaningful, with one variant often sounding more "correct," "formal," or "literate," and/or possibly changing over time (see section 2.2), it is perhaps not surprising that years of education and age also affect the realization of the variable. That said, education and age do not predict the choice of variant across the board. Instead, the strength and magnitude of any effect of social factors becomes apparent only within each environment. Year of birth conditions rate of use of the ∅ variant after *couple*: speakers who are born later are more likely to use the ∅ variant. However, with *all* and *out*, there is no significant effect of year of birth on rate of ∅ variant use: speakers of all ages are equally likely to use the ∅ variant. These findings mirror the diachronic trajectories of these variants: in the case of *couple*, ∅ is a relatively recent innovation; in the cases of *all* and *out*, both *of* and ∅ variants have been in the language for a long time. It is plausible, then, that *of*-variation after *couple* reflects a slow change in progress toward ∅, while after *all* and *out*, it reflects stable variation.

Years of education conditions rate of ∅ variant use in all three linguistic environments, but has different effects in each of them. Speakers with more years of education are less likely to use the ∅ variant with *all* and with *out*, but **more** likely to use it with *couple*, compared to speakers with fewer years of education. These findings are surprising in light of the prescriptivism summarized in section 2.2: with *all*, the *of* variant is prescriptively disfavored, but this is the variant that more educated speakers use more. Similarly, with *couple*, the ∅ variant is prescriptively disfavored, and again this is the variant that more educated speakers use more. This suggests that prescriptions in usage guides have little effect on vernacular spoken language. A direction for follow-up work is to examine whether listeners have any overt or covert social evaluations of *of*-variation in these contexts that mirror prescriptivist attitudes.

These patterns suggest that although social factors affect the variation, this effect is mediated by linguistic environment, providing a positive answer to Q3 (Does the effect of demographic factors on rates of ∅ differ by linguistic environment?). This means that in each instance of use, the choice between *of* vs. ∅ depends on both who the speaker is (social factors) and what the construction is (linguistic factors). Thus, if we consider occurrences of *of*/∅ in all three environments to be instances of the same variable, we find evidence to say that social and linguistic factors interact in conditioning it, which would run counter to previous assumptions (as summarized in section 2.1).

However, another way to interpret these results would be to preserve the assumption that social and linguistic factors do not interact. Under this assumption, our results can be taken as evidence that *all of* ∼ *all* ∅, *couple of* ∼ *couple* ∅, and *out of* ∼ *out* ∅ are three separate variables, despite their surface similarities. This is supported by their different diachronic and prescriptive profiles

(summarized in section 2.2). This provides an interesting counterpoint to studies like that of Dinkin (2016), who finds change in progress apparently targeting a single variant (*like*) across multiple contexts, suggesting consistent social evaluation of the variant regardless of variable structure. The *of*-variation case appears to be the opposite.

Ultimately, our findings raise the following question: what counts as different instances of "the same" sociolinguistic variable? Must different instances permit the same set of variants? Must they be perceived by language users as part of the same alternation? Must they show identical or parallel effects of social factors? Identifying more instances of variations which (appear to) span different linguistic environments will provide insight into this question.

6 Conclusion

In this work, we examined a pattern of variation between *of* and ∅ which can occur in multiple linguistic environments, in order to reexamine existing assumptions about whether linguistic and social conditioning interact. We found evidence that the choice between variants is influenced by both linguistic and social factors, and that the strength and direction of the effect of social factors significantly varies across linguistic environments. Our findings allow for two interpretations: either all the instances of this alternation across environments are instances of the same variable, and therefore linguistic and social conditioning *do* interact, or they are instances of three different variables that share surface similarities, and therefore the differences in social conditioning across linguistic environments may be expected.

Based on this data alone, both interpretations seem reasonable, and the answer depends on the larger theoretical and methodological issue of how one defines a sociolinguistic variable. Future work is likely to provide more empirical data to support one or the other interpretation, potentially by examining more of the linguistic environments for this alternation, examining other populations, and examining other alternations which span multiple environments. More data may also come from perceptual work, which can experimentally investigate how language users perceive and evaluate such cross-environment alternations, and whether those evaluations also differ by environment. Existing literature on the formal properties of these constructions is another potential source of evidence, as these can provide clues about the degree to which these different linguistic constructions are likely to register as a unified phenomenon to language users.

References

Allen, Harold B. 1989. The *big of* syndrome. *American Speech* 64(1):94–96.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Bayley, Robert. 2013. The quantitative paradigm. In *The Handbook of Language Variation and Change*, ed. J. K. Chambers and N. Schilling, 85–107. Malden, MA: Blackwell.

Bernstein, Theodore M. 1977. *Dos, Don'ts & Maybes of English Usage*. New York, NY: Barnes & Noble Books.

Blass, Anne-Katrin. 2020. The Big Mess Construction: Forms and Functions in Present-Day English. Doctoral dissertation, Universität Trier.

Cappelle, Bert. 2001. Is *out of* always a preposition? *Journal of English Linguistics* 29(4):315–328.

Cedergren, Henrietta, and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50(2):333–355.

Dickson, Victoria, and Lauren Hall-Lew. 2017. Class, gender and rhoticity: The social stratification of non-prevocalic /r/ in Edinburgh speech. *Journal of English Linguistics* 45(3):229–259.

Dinkin, Aaron J. 2016. Variant-centered variation and the *like* conspiracy. *Linguistic Variation* 16(2):221–246.

Estling, Maria. 1999. Going out (of) the window? *English Today* 15(3):22–27.

Estling, Maria. 2000. Competition in the wastebasket: A study of constructions with *all*, *both* and *half*. In *Corpus Linguistics and Linguistic Theory*, ed. C. Mair and M. Hundt, 103–116. Brill.

Estling, Maria. 2004. Syntactic Variation in English Quantified Noun Phrases with *All, Whole, Both* and *Half*. Doctoral dissertation, Växjö University.

Flesch, Rudolf. 1964. *The ABC of Style: A Guide to Plain English*. New York, NY: Harper & Row.
Garner, Bryan A. 2022. *Garner's Modern English Usage*. Oxford: Oxford University Press.
Göbel, Jessica. 2020. Don't glottal stop me now: A comparative study of /t/-glottaling in Edinburgh English in the 1970s. *Lifespans and Styles* 6(2):32–43.
Kjellmer, Göran. 2007. *A couple hours* – on non-occurrence of *of* after partitive English nouns. *Zeitschrift für Anglistik und Amerikanistik* 55(3):293–312.
Labov, William. 1993. The unobservability of structure and its linguistic consequences. Paper presented at NWAV22, University of Ottawa.
Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Malden, MA: Blackwell.
Labov, William. 2010. *Principles of Linguistic Change: Cognitive and Cultural Factors*. Malden, MA: Blackwell.
Labov, William, and Ingrid Rosenfelder. 2011. The Philadelphia Neighborhood Corpus of LING 560 Studies, 1972–2010. With support of NSF contract 921643.
Lavandera, Beatriz. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7(2):171–182.
Maddeaux, Ruth, and Aaron Dinkin. 2017. Is *like* like *like*? Evaluating the same variant across multiple variables. *Linguistics Vanguard* 3(1).
Merriam-Webster. 2002. *Merriam-Webster's Concise Dictionary of English Usage*. Springfield, MA: Merriam-Webster.
Nylund, Anastasia, and Corinne Seals. 2010. "It's not that big (of) a deal": The sociolinguistic conditioning of inverted degree phrases in Washington, DC. In *University of Pennsylvania Working Papers in Linguistics 16.2: Selected Papers from NWAV 38*, ed. M. Lerner, 133–140.
Oxford English Dictionary. 2024. Retrieved February 28, 2025. Oxford University Press.
R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Rapp, Linda L. 1991. The *big of* syndrome: An update. *American Speech* 66:213–220.
Romaine, Suzanne. 1981. On the problem of syntactic variation: A reply to Beatriz Lavandera and William Labov. In *Working Papers in Sociolinguistics 82*. Austin: Southwest Educational Development Laboratory.
Sánchez, Francisca. 1986. *Of*-reduction in Black English: A quantitative study. In *Diversity and Diachrony: Current Issues in Linguistic Theory 53*, ed. D. Sankoff, 59–72. Philadelphia: John Benjamins.
Sommerer, Lotte. 2022. 'So great a desire': Investigating the BIG MESS construction in Early Modern English. *Journal of Historical Syntax* 6(2):1–34.
Sonderegger, Morgan. 2023. *Regression Modeling for Linguistic Data*. Cambridge, MA: MIT Press.
Vartiainen, Turo, and Mikko Höglund. 2020. How to make new use of existing resources: Tracing the history and geographical variation of *off of*. *American Speech* 95(4):408–440.
Vasko, Anna-Liisa. 2007. Up Cambridge: Prepositional Locative Expressions in Dialect Speech: A Corpus-based Study of the Cambridgeshire Dialect. Doctoral dissertation, University of Helsinki.
Weiner, E. Judith, and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19(1):29–58.
Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. *Empirical Foundations for a Theory of Language Change*. Austin, TX: University of Texas Press.
Wood, Johanna L. 2019. A couple (of) changes in the 'Brown family': British and US English compared. In *The Sign of the V – Papers in Honour of Sten Vikner*, ed. K. R. Christensen, H. Jørgensen, and J. L. Wood, 759–787. Aarhus: Aarhus University.
Zwicky, Arnold. 2009. WOO: The War On Of. Accessed 5 June 2024. URL `https://arnoldzwicky.org/2009/10/20/woo-the-war-on-of/`.
Zwicky, Arnold M. 2007. P+of. Class handout, LSA.376, 15 July 2007.

10 Washington Place
New York, NY 10012
mic6457@nyu.edu
auromita.mitra@nyu.edu
jsg6312@nyu.edu
laurel.mackenzie@nyu.edu

ArtsOne, Mile End Road
London E1 4PA
United Kingdom
m.e.barnard@qmul.ac.uk