

Reconstructing sociolinguistic variation

Jessica Kantarovich & Lenore A. Grenoble*

Abstract. In this paper we illustrate a methodology for reconstructing language in interaction from literary texts, demonstrating how they can serve as documentation of speech when primary linguistic material is unavailable. A careful incorporation of facts from literary dialect not only informs grammatical reconstruction in situations with little to no documentation, but also allows for the reconstruction of the sociolinguistic use of a language, an oft-overlooked aspect of linguistic reconstruction. Literary dialogue is often one of the only attestations of regional varieties of a language with a very salient standard dialect, where no primary sources are available. Odessan Russian (OdR), a moribund dialect of Russian, serves as a case study. OdR grew out of intensive language contact and differs from most other varieties of Russian, with substrate influences from Yiddish, Ukrainian, and Polish, and lexical borrowing from other languages. The only records of “spoken” OdR are found in fictional narrative. An analysis of works from several prominent Odessan writers, including Isaak Babel and Ze’ev Jabotinsky, reveals considerable variation among speakers of OdR; careful tracking of this variation shows how it was distributed among different social groups, and suggests how it may have been deployed to index and acknowledge different social roles.

Keywords. historical linguistics; language variation; sociolinguistics; historical reconstruction; language contact; Russian

1. Introduction: the problem. The goal of linguistic reconstruction has traditionally been to derive a complete and stable linguistic system, at a certain point in time. Whether we are concerned with the reconstruction of a hypothetical language, such as a proto-language, or with the reconstruction of a known language of which there has been little documentation, our methods are predicated on the idea of uniformity: throughout the system, and throughout the process of change. There are many advantages in reconstructing an ideal linguistic system that has consistent phonology and morphosyntax. It allows us to undertake comparative work that informs studies of typology as well as of the relatedness of different languages. In some cases, such reconstructions can be used in revitalization efforts, providing a standardized language that can be taught within the community.

However, we must accept that the results of traditional methods of historical reconstruction are necessarily idealized, theoretical entities that do not capture how people spoke. In outlining a uniform linguistic system, the linguist must ignore issues of variation, which in turn excludes a very important dimension of language use: the social and pragmatic factors that condition the selection of different variants. Thus, while reconstructed systems are of immense value to communities seeking to revitalize a moribund language, it is unlikely that new speakers are acquiring the same language that was lost, or even that the new system is a full-fledged language.

*We thank the Visiting Committee of the Humanities Division of the University of Chicago for their generous funding which has supported this research. We are especially grateful to the many Odessans who have helped us with our work; any errors are our sole responsibility. Authors: Jessica Kantarovich, University of Chicago (jkantarovich@uchicago.edu) & Lenore A. Grenoble (grenoble@uchicago.edu)

Additionally, by imposing an entirely uniform system on the bygone speakers of a language, we are almost certainly ignoring socially-conditioned variation in the distribution of linguistic features. Just as speakers today often differ systematically by region, socioeconomic class, gender, and education, we should expect these types of categories to project onto previous generations. It is important to understand whose linguistic variety is being reconstructed, especially when we consider the motivation for particular changes. This is a fallacy that is often found in language contact-dependent explanations for change: there is often both a failure to establish who within the community would have had access to the contact language, as well as what their own variety of the recipient language would have been like. Thus, the reconstruction of a language, and the origins of different features in that language, is not complete without the reconstruction of the community that used the language.

While we can anticipate the sociological categories that might condition variation, reconstructing the interactions between variation and multilingualism is likely to be more difficult. The literature on modern sociolinguistic variation has often focused on how monolingual populations differ in their use of a single language. Do we expect similar patterns in the treatment of variants from other languages? It is true that we can often easily point to the result of contact in a recipient language, but the origin and diffusion of these features is considerably less clear, in part because we are often lacking information about which subsets of the speaker population may have been bilingual, as well as their degree of access to the contact language.

In order to address these issues, we argue for a socially-anchored approach to historical linguistics and language reconstruction. The goal of this approach is to reconstruct not only the set of linguistic features that existed in a language—or dialect—when it was robustly spoken, but ideally also which subsets of the population used the language, whether different features were variably utilized by different groups, and the role of variable linguistic use in social interaction.

Of course, most linguists working with the reconstruction of ancient or endangered languages have come up against the difficulty of obtaining socio- and metalinguistic information about their speech communities. Even records that document the demographic composition of a community—such as census data—are often categorical and rigid. Failing to account for differences across an individual speaker's own speech patterns and shifting identities, they are unlikely to give a nuanced representation of variation in multilingual settings. For understudied languages, there are often even fewer records that describe language use, except perhaps where it may have been obvious to an outside anthropological observer, as in the case of ritual language. Even in these cases, however, the actual linguistic differences are unlikely to be clearly conveyed.

How, then, can we reconstruct the rich variation we expect of multilingual settings? The solution lies in reconfiguring strict notions of what counts as linguistic documentation, to include not only primary sources, recordings, and elicited data, but also literature. Although such data is invented, it is usually embedded within social contexts that can be inferred from a text, and which the author has deliberately chosen to represent in a particular way. In order to illustrate how literature can be incorporated as linguistic representation in a socially-anchored reconstruction, we conduct a case study of Odessan Russian, where literature actually serves as the best available source for the morphosyntax of the language. Odessan Russian is a moribund contact variety of Russian that emerged in Odessa (in what is now Ukraine) in the late 19th century. There is ample evidence that the dialect was considered low prestige when it was spoken, and this association has survived into the present. As a result, there is very little documentation of the dialect, but very strong and enduring notions about the types of people that spoke it.

In the rest of the paper, we outline our assumptions about language change as the result of language use (section 2), describe our case study and the historical situation that gave rise to the contact variety (section 3), and reconstruct the sociolinguistic variation of OdR (sections 4 and 5). Section 6 analyzes the current status of Odessan Russian within a well-documented, recurring cline of sociolinguistic change and gives a conclusion.

2. Usage-based theory in language change. We assume a *usage-based* approach to modeling language contact and change: the primary mechanism of diffusion of linguistic features is through person-to-person interaction. Thus, to an extent, we are not necessarily concerned with the reconstruction of a grammar, but rather the reconstruction of patterns of speech and usage, which in turn involves reconstructing the types of interactions that gave rise to different speech patterns. See Grenoble & Kantarovich (forthcoming) for more details.

2.1. ETHNOLINGUISTIC REPERTOIRE. In reconstructing sociolinguistically-conditioned variation for multilingual communities, we assume the existence of an *ethnolinguistic repertoire*, which is “defined as a fluid set of linguistic resources that members of an ethnic group use variably as they index their ethnic identities” (Benor 2010: 160). We propose a model for reconstructing sociolinguistic variation by drawing upon a range of materials to recreate social networks and ethnolinguistic repertoires. This approach emphasizes individuals’ selective use of linguistic features, i.e., that speakers can and do use different features strategically, depending on their interlocutors, domain, and register, to name just a few variables. Our data show variation at all linguistic levels, and variation across different writers and within an individual writer’s work. Moreover, when we track a single character’s speech patterns, we see that his or her usage of Odessan features varies according to general context of usage. The repertoire approach allows us to capture this flexibility of dialect use. Furthermore, it has several advantages over a static view of *ethnolects* inasmuch as it accounts for the fluid social indexing that we find in multilingual communities and can thus be used to account for both intra-speaker and intra-group variation.

3. Case study: Odessan Russian (OdR).

3.1. HISTORICAL BACKGROUND. The case study that informs our approach to sociolinguistic reconstruction is Odessan Russian (OdR), a dialect of Russian that was spoken in the city of Odessa from about the 1850s to the 1970s. OdR is entirely mutually intelligible with Russian, but it has clearly been influenced by the sociohistorical context in which it emerged. The most profound structural effects of this contact are on the morphosyntax of OdR, which shows signs of the existence of Ukrainian and Yiddish substrates. Polish, which was a *lingua franca* in Western Ukraine during the Middle Ages, also appears to have impacted the structure of OdR, if to a lesser extent (Stepanov 2004: 24).

During the period when the dialect was robustly spoken, Odessa was a major port and economic center within the Russian Empire, drawing a diverse group of merchants as well as laborers who worked aboard ships and at the docks. In particular, Greek (and later, Jewish) merchants dominated trade in the region. The port setting in Odessa had an overt impact on the variety of Russian that was spoken there, with numerous borrowings from Greek, Italian, English, and Turkish. The economic opportunities of the city were also attractive to Russians and Ukrainians, particularly men, from all across the Russian Empire—the presence of working class, uneducated Russian speakers likely also had an impact on the speech variety in Odessa, as these speakers were almost certainly not speaking the literary standard.



Figure 1. Map of Odessa¹

In addition to the importance of the port, Odessa was also considered an intellectual and cultural center, and as a result drew more white-collar migrants from western Europe, in particular German- and French-speaking women who worked as governesses in wealthier households. There is also evidence of contact effects from these two languages, most notably French forms of address and other socially-indexical terms and phrases. The presence of these features must be due, in part, to the prestigious status of French throughout all of Russia during this period; however, it is clear that the regular use of French phrases in discourse (particularly among the lower classes who likely did not learn or acquire the language) is an Odessa-specific speech pattern.

Language	Males	Females	Total	Percent Total Population
Russian	104,173	89,081	193,254	50.78
Yiddish	61,156	62,530	123,686	32.50
Ukrainian	13,224	8,302	21,526	5.66
Polish	11,174	5,864	17,038	4.48
German	5,253	4,680	9,933	2.61
Greek	3,166	1,847	5,013	1.32
Tatar	970	459	1,429	0.38
Armenian	929	470	1,399	0.37
French	423	701	1,124	0.30
Belorussian	799	296	1,095	0.29

Table 1. The Ten Largest Groups by Native Language in the City of Odessa, 1897
(*Perepis*, 1899-1905)

The ethnic group that is of particular interest when we consider the emergence of Odessan Russian is the Jewish population. Odessa is considered by Russians to be a “Jewish city.” Although the number of Jews has never exceeded the number of Slavs in Odessa, it is certainly the case that during the 19th and (early) 20th centuries the Jewish population in Odessa soared relative to that of other ethnicities.

Yiddish- and Ukrainian-speaking Jews were among the earliest settlers in the territory that

¹<http://media.web.britannica.com/eb-media/72/64472-004-333F1C28.gif>

would become Odessa, having been forcibly relocated to the region by Catherine the Great. However, the number of Jews in the city rose dramatically towards the end of the 19th century due to increased migration, and also the nature of Jewish migration relative to other groups': Jews tended to resettle in Odessa with their entire families, while other (predominantly male) economic migrants usually came alone (Herlihy 1986). Odessian Jews occupied the full socioeconomic spectrum—many of them were poor laborers, but they also held occupations that required education, such as doctors, journalists, and teachers, and there was a class of wealthier Jews (many of them merchants).

Year	Total Population	Jewish Population	Percent
1794	2,345	244	10.41
1827	32,995	4,226	12.81
1829	51,998	7,900	15.20
1841	73,888	10,775	14.58
1843	77,778	12,000	15.43
1854	90,319	17,080	18.91
1873	193,513	51,378	26.55
1880	219,300	55,300	25.22
1892	340,526	112,235	32.96
1897	403,815	138,935	34.41
1904	511,000	160,000	31.31
1912	620,143	200,000	32.25

Table 2. Percentage of Jews in Odessa's Total Population, 1794-1912²

3.2. WHY ODESSAN RUSSIAN? OdR is an ideal choice for demonstrating a sociolinguistically-oriented approach to reconstruction for a number of reasons. First, it is recognizable and highly salient for most modern Russian speakers, due to its enduring presence in Russian films, TV shows, and other media. However, the authenticity of these representations is debatable: since the dialect has not been spoken in Odessa since the 1970s, it is very unlikely that the speech patterns that are invoked in the media are anything more than stereotypes or caricatures.

The reason these stereotypical features are so recognizable and continue to occupy a prominent place in Russian culture is because speakers have strong extralinguistic associations with OdR. It is commonly thought of as an ethnolect that was spoken exclusively by poor Odessian Jews. It is also closely associated with stereotypes about the personality of the typical Odessian: jocular, good-natured, and devil-may-care. Russian speakers talk about both the Odessian Russian language and the culture of the city with a sense of nostalgia about a bygone era, which modern portrayals of the dialect evidently tap into. The ability to speak Odessian Russian even to a limited extent is seen as the marker of a "real Odessian." Thus, although the dialect is associated with lower socioeconomic classes and a lack of education, and would not likely have been seen as prestigious in its day, now that it has vanished it has taken on a kind of "covert prestige."

Overall, there are many preconceived notions about the sociolinguistic context and usage of Odessian Russian, and we have reason to doubt that these sociolinguistic associations and even the

²Sources: GAOO, *f. 274, op. 2, d. 4, ll. 26, 47, 68*; A.A. Skal'kovskii, *TOSK*, 1865, I, 134-35; *Rezul'taty*, 1894; *Perepis*, 1899-1905; NACRO, 25 July 1905; and Fedor, 1975, p. 202.)

linguistic depictions are entirely accurate. It is advantageous to reconstruct the actual social distribution and indexicality of a pragmatically rich dialect like OdR during the time period when it was robustly spoken, as it can inform our understanding of how metalinguistic awareness can evolve over time.

3.3. DATA SOURCES. Yet another reason why OdR is a particularly appropriate case study is that, like other extinct or endangered languages, it lacks both linguistic and sociolinguistic documentation. At the time, Russian authorities did not view dialectal differences (which were, in fact, widespread) as worthy of documentation. The Russian literary standard was highly dominant and any nonstandard speech was heavily stigmatized in formal or educational settings. As we will see below, wherever Odessan Russian speech was explicitly documented, it was treated as an oddity or a catalogue of mistakes.

We have not been able to find any primary sources in OdR, such as letters or diaries. We strongly suspect that such sources would be extremely rare and possibly never existed. At the time, speakers who were educated enough to be literate in Russian would have written in the standard language. Other OdR speakers would have likely had another more native language that they used to write—for Odessan Jews, who we know were certainly among the speakers of the dialect, it was probably Yiddish. Our archival research has turned up primary documents from Odessa during this era, but none of them are written in the dialect. In particular, the local government collected firsthand accounts of pogroms, most of which are written in Yiddish. Some of them are written in Russian, but do not have any obvious dialectal features. The authorship of these accounts is also uncertain—it is not clear whether the Russian reports are written by pogrom witnesses or by officials recording their testimony. There were also several Jewish newspapers in print during the 19th century in Odessa. While we might expect these newspapers to have more vernacular Odessan features to appeal to local readers, the language is actually quite standard. (This is not very surprising, as the newspapers were being produced by the Odessan Jewish intelligentsia, who would have been educated in standard Russian.)

Nevertheless, there are several decent sources of data containing actual OdR usage. The dialect seems to still be reliably preserved in some songs (Rothstein 2001), most notably those that were written in Odessa in the first half of the 20th century. Many of these songs belonged to a uniquely Odessan genre of “criminal songs” (*blatnyje pesni*) which glorified the underground that existed in the city. These songs would often not be published in anthologies and survived via an oral tradition—one song in particular, *Ja vam ne skažu za vsju Odessu* ‘I will not tell you about all of Odessa’, is still well known. (In fact, the title of this song contains an unusual construction that we can reliably reconstruct for OdR: the use of the preposition *za* ‘behind’ to mean ‘about’ instead of the standard *o* or *pro*.)

There do exist several reference materials that were written when Odessan Russian was still widely spoken, as it drew widespread attention from speakers of the standard language:

- K. Zelenetskii. 1855. *O russkom iazyke v novorossiiskom krae*. [On the Russian language in the Novorossiia Territory]. Odessa: Odesskii ochebnyi okrug.
- V. Dolopchev 1909. *Opyt slovaria nepravil'nostei v russkoi razgovornoj rechi*. [An attempt at a dictionary of incorrectness in Russian conversational speech]. (2nd ed). Warsaw: Tip. K. Kovalevskago.

The second of these, Zelenetskii 1855, is a 34-page description of the speech of the city that was written in response to a request from authorities in the Odessan Pedagogical District, who observed the use of “non-Russian words and entire expressions and phrases that do not correspond with the rules and spirit of the language of the fatherland (Zelenetskii 1855: 3). The academic Dolopchev (1909) expanded on Zelenetskii’s account and compiled an extensive lexical corpus of Odessan Russian, including both borrowings as well as lexical items that seem to have undergone phonological or morphological change. The title of the dictionary, which roughly translated means “an attempt at a dictionary of mistakes in Russian conversational speech”, confirms that Odessan speech was considered by authorities not to be within the range of acceptable variation of the Russian language.

Although Dolopchev’s dictionary seems very thorough, it is lacking important linguistic information to situate the lexical items within the overall grammar of OdR. Dolopchev did not systematically collect data through elicitation for particular features, but recorded words he overheard that sounded nonstandard to him. Even the morphologically or phonologically peculiar examples are given as dictionary entries, with the corresponding standard Russian form, so we are required to do a bit of forensic work to determine what makes the Odessan entry nonstandard. Occasionally, Dolopchev gives an example sentence containing the form, but it is unclear whether these are invented examples or examples that he actually overheard. The degree of uniformity of the morphological changes also cannot be conveyed through a single lexical item: for example, there are several instances where a noun appears to have shifted its declension class, but it is impossible to tell if this is a change across all cases in the paradigm because Dolopchev only provides one form. Similarly, when an OdR dictionary entry could be in citation form, e.g., an infinitive verb with different derivational morphology compared to the standard Russian equivalent, it is unclear if this is an attested form or if Dolopchev has reconstructed the infinitive from a finite verb form.

In addition to these sources, modern Russian speakers are very quick to direct us to more recent “references,” including an online dictionary, joke books, and a humorous pseudo-textbook. We are deliberately excluding these types of materials from our reconstruction because their sources are dubious and it is not clear that they document authentic OdR speech rather than stereotypes and popular beliefs about the dialect.

Our last source for direct Odessan Russian speech is fieldwork with the remaining speakers of the dialect, who now mainly live in the neighborhood of Brighton Beach in Brooklyn, NY. According to our consultants, the Jewish diaspora from Odessa was the major contributing factor to the loss of the dialect and the Odessan culture that is so strongly linked to it. Many of these speakers were displaced by World War II or else were motivated to leave Odessa due to antisemitic Soviet practices. In the 1980s and 1990s, some of the language and culture was evidently still preserved on Brighton Beach, but at the time of our fieldwork (beginning in 2010) Odessans were no longer the dominant group in the neighborhood, which has consistently been home to recent Russian-speaking immigrants from numerous former Soviet states. There is no real Odessan Russian community where the dialect can be overheard, and all of the speakers we were able to contact are mainly rememberers/overhearers of the dialect. (This is not entirely surprising, as anyone who would have been speaking the dialect as an adult in Odessa would be quite elderly now.) One of our consultants, who was in his 40s, had clearly begun acquiring the dialect in the home, but was punished and laughed at for his nonstandard speech once he got to school and avoids using these features today. Ultimately, we were able to record some

spontaneous Odessan Russian speech from this consultant, but most of our fieldwork has consisted of verifying that the features that have been attested in other sources were actually used, as well as interviewing them about how widespread these features were among different groups of Odessans.

3.4. LITERATURE AS LINGUISTIC REPRESENTATION. Given these more typical sources for language reconstruction, we are lacking information to reconstruct several dimensions of Odessan Russian: (1) facts about morphosyntax and discourse, (2) the extent of variation in the usage of the dialect or different dialect features, and (3) sociolinguistic dimensions of the use of the dialect. In order to address this limited availability of conversational and sociolinguistic data, it is necessary to expand the criteria for what linguists typically consider to be valid sources.

Ideal language documentation includes recordings of naturalistic speech events at the moment of speech, which enables future researchers to consult the data without having to be concerned with any distortion in a transcription. Of course, the pressure to record linguistic data (and make it publicly available) is a relatively recent development in work with lesser-studied languages, and such data is nonexistent for ancient languages. However, reconstruction should not be considered hopeless in these cases and should instead make use of any information is available about the language, including invented data such as literature.

The use of literature as linguistic documentation has already been applied to the study of English dialectology, particularly phonological variation (Culpeper & Kytö 2010, Schneider 2013). The depiction of nonstandard language in literature is called *literary dialect*, or “an author’s attempt to represent in writing a speech that is restricted regionally, socially, or both” (Ives 1971: 146). Literary works, unlike reference materials such as dictionaries, often depict direct quotation, dialogues, and other speech patterns. Although conversation between characters in a work of fiction is constructed by the author, it often realistically mirrors language the author might use or that the author has observed other people using. It is also often highly systematic: while the orthographic representations of nonstandard English phonology vary from author to author, they tend to be very consistent within a single author’s work. This systematicity extends to other aspects of the depiction of dialect in literature, including lexicon and morphosyntax.

Literary works also necessarily provide sociolinguistic information by virtue of being embedded in a narrative: the choice by the author to depict a character as speaking a certain way is deliberate, and is meant to convey information about that character’s history and how he or she fits into the social setting of the novel. The extent to which any sociolinguistic associations with language variation are highlighted by the author can vary. Some works make use of *selective reproduction*, in which the marked features are presented frequently enough to signal the use of a different dialect or language without interfering with intelligibility for the reader, who is not assumed to be proficient in both lects (Sternberg 1981). Other works employ *explicit attribution*, in which the narrator or another character directly points out that a speech act is nonstandard.

Not all invented data can be treated as an accurate representation of the language, and we want to avoid stereotypical or pejorative representations of certain socioeconomic or ethnic groups. Thus, the authorship of the literary work must be considered: is the author a member of the group being represented, or else would he or she have had sufficient contact with the group at the time the dialect or language was spoken?

For our reconstruction of OdR, we have limited the authors whose works we analyze to those who were likely themselves native speakers of OdR: those who were born and/or raised in

Odessa during the dialect's heyday. Two particular works serve as the basis of our reconstruction:

- Isaak Babel's *Odesskie Rasskazy* [*Odessan Stories*], which makes use of selective reproduction of OdR features
- Ze'ev Jabotinsky's *Pjatero* [*The Five*], which has both selective reproduction as well as explicit attribution in the form of direct commentary about OdR from the narrator

4. Reconstructing the linguistic system. The first component in a reconstruction of variation is deriving the actual linguistic system: making note of variants of features, rather than trying to explain them away or normalize the system.

4.1. USING MULTIPLE REFERENCES. When using non-ideal sources of data, a best practice is to keep a thorough catalog of where different features are attested. However, reconstructing variation should not be taken to mean that any feature, no matter how esoteric, should be included: ideally, the features that are reconstructed should be referenced in at least two sources to confirm that they are not idiolectal variants or mistakes by the author. (This is especially true of the literary sources, since identifying a feature for the relevant dialect is subject to misinterpretation by the researcher.)

We can also admit some variability in the examples we consider to be representative of a single feature. For example, a well-attested pattern for OdR in our sources is the overuse of the genitive case (extending the genitive to environments where another case is expected).

- (1) *Pust' vas ne volnu-et èt-ix glupost-ej*
may 2PL.ACC NEG worry-3SG these-GEN.PL silliness-GEN.PL
'Don't let this nonsense bother you' (Babel, *Odessan Stories*)

In (1), the genitive occurs where nominative case would be expected (since 'these sillinesses' is actually the subject of the verb 'bother'). This example also contains an instance of nonstandard agreement (a 3SG form where we actually would expect the 3PL).

Cross-referencing our other sources confirms that the pattern of using the genitive in nonstandard environments is well-attested; however the actual environments are fairly diverse:

- (2) *Nado pisat' stix-ov Nado myt' ruk-∅ i čistit'*
must write.INF poetry-GEN.PL must wash.INF hand-GEN.PL and brush.INF
zub-∅!
teeth-GEN.PL
'You have to write poetry. You have to wash your hands and brush your teeth!' (Shishkova 1973, 69, quoting a letter from Bagritskii)
- (3) *Čto znač-it det-ej?*
what mean-3SG child-GEN.PL
'What does children mean?' (Jushkevich, *Spasibo serdce*)

In (2), the genitive is being used in place of an expected accusative (the object of the verbs 'write', 'wash', and 'brush'), while in (3), the genitive occurs in place of an expected nominative, as in (1).

4.2. LINGUISTIC VARIATION IN ODESSAN RUSSIAN. The preceding examples, although slightly varied in terms of context, represent a single feature in OdR and do not constitute reconstruction of “variation.” In this section, we provide examples of the actual linguistic variation we believe existed when OdR was still spoken.

There is widespread evidence for the existence of significant variation in OdR. Based both on the direct descriptions of the dialect and our other sources, it is clear that even among speakers of OdR, not everyone spoke the same way. Zelenitski (1855: 11) described this variation explicitly, writing that “it is not possible to enumerate all the errors and inaccuracies because they occur, to a greater or lesser degree, in different places or among different people and, moreover, in the most diverse way.” Our consultants confirm this, describing intense microvariation in which each different *dvor* (courtyards outside apartment complexes, where Odessans would regularly gather) had distinctive speech patterns.

There does appear to be some subtle variation in the expression of the same OdR feature across our sources. For example, in expressing the equivalent of English ‘a day’, (4) shows use of the preposition *na* with the accusative case and (5) with the same preposition with the dative. In Standard Russian, one would find the preposition *v* with the accusative, and the preposition *na* does not occur with the dative:

(4) *Dvadcat’ raz na den’* (expected: *v den’* ‘in day-ACC’)
 twenty times on day.ACC
 ‘Twenty times a day’ (Babel, *Odessan Stories*; Dolopchev 1909)

(5) *Xod-ila k nemu po tri raz-a na dnju*
 go-PST.FEM to him three times-GEN on day.DAT
 ‘(She) went to him three times a day’ (Jabotinsky, *The Five*)

Dolopchev’s (1909) dictionary lends further support to the existence of variation, as it frequently provides variants for examples of nonstandard gender, stress patterns, lexical items, and morphology. For example, for the standard Russian word *tormošít’* ‘to shake awake, to pester’ the dictionary gives both *tormóšít’* (nonstandard stress) and *tormosit’* (*/š/ > /s/*) as attested OdR forms (Dolopchev 1909: 278). Typically Dolopchev does not contextualize these examples with information as to who may have been using which variants in which circumstances (although he does claim that, overall, the “mistakes” are more common among women and children). Given Zelenitski’s description of the impossibility of accounting for all the variation and the lack of extralinguistic details in Dolopchev’s dictionary, it is certainly possible that many of the forms were in free variation, or it was not obvious that they could be linked to a particular group. Ultimately, what is clear is that there was no such thing as a unified OdR grammar.

OdR variation is also expressed in the representation of these features among different authors’ works. More tellingly, the use of OdR varies in a single author’s work, with different characters who are shown to speak the dialect not necessarily sharing the same set of features. Consider the following examples from one of Babel’s *Odessan Stories*, “Father”:

- (6) *smejat'sja s* + GEN for *nad* + INST ‘to laugh at someone’
- a. ...*zajd-ite* *k nam na dvor, est' s čego*
 come.in-1PL.IMP to 1PL.DAT in courtyard be.INF from what.GEN
posmejat'sja
 laugh.INF
 ‘Come into our courtyard, there is something to laugh at’
- b. ...*neuželi ty smee-š'sja nado mnoj?*
 really you laugh-2SG.PRS over 1SG.INST
 ‘Is it possible that you are laughing at me?’

The character in (6a) uses the well-attested OdR expression for ‘laugh at someone’, using the preposition *s* and the genitive case. Later in that same conversation, a different character uses the standard form in (6b). Both characters belong to the same social class—they are poor Jews, who are the prototypical OdR speakers. It is unlikely that this is sociolinguistic variation, but it given these examples’ proximity in the text, it is also unlikely that the use of these variants is random or accidental. Babel seems to be signaling here that at least some OdR speakers have command of the standard language as well, which may indicate that OdR/standard Russian were part of a diglossic system.

5. Reconstructing sociolinguistic variation. From the preceding facts it is reasonable to conclude that there was variation in the use of OdR. The second component of our reconstruction is to determine the sociolinguistic dimension of this variation. Our goal here is to verify whether the stereotypical associations with OdR are grounded in historical fact, specifically: (1) which social groups used the dialect (i.e., was it restricted to Odessan Jews or did it spread to other groups in Odessa), and (2) is the modern pragmatic function of OdR (to signal an easygoing, jocular attitude) a recent innovation linked to the development of the stereotype?

5.1. LANGUAGE CONTACT IN ODESSA. To answer the first question, we rely on demographic and sociohistorical information about Odessa along with our literary corpus.

OdR was unquestionably spoken by Jews in Odessa. This is confirmed by historical accounts as well as our interviews with consultants in Brighton Beach, who all maintain that OdR was most prevalent among Jews, especially those living in the poorer neighborhoods Moldavanka and Peresyp. Why, then, do we suspect that OdR was not an ethnolect used exclusively by Odessan Jews? Part of the answer lies in the social organization of the city: contact among different social groups was extraordinarily high in Odessa (Herlihy 1986, Zipperstein 1986). Although the Jews did form a distinct community, many were upwardly mobile and participated in social gatherings with non-Jews of the middle and upper classes. Meanwhile, most of the city’s poor—Jews and non-Jews alike—lived together in cramped quarters in neighborhoods such as Moldavanka and Peresyp (Richardson 2008). In Jabotinsky’s *The Five*, a non-Jewish character who grew up in Peresyp (Valentinochka) is explicitly shown to use OdR, and she even makes note of her own use of “Jewish” expressions.

It is also important to recall that many of the contact effects in OdR cannot be traced to Yiddish, but are associated with the unique cosmopolitan nature of Odessa—these effects include borrowings from the port languages, the prevalence of French, and above all the influence of Ukrainian. Speakers of all ethnicities would have encountered this multilingualism, whether in the streets, the port, the market, or social gatherings. Thus, we expect that OdR was spoken to

some extent by everyone living in Odessa, though not necessarily in the same way. It is likely that different groups' dialectal features varied based on the unique nature of their contact with speakers of other languages.

5.2. THE DISTRIBUTION OF ODESSAN RUSSIAN FEATURES BY GROUP. Our literary corpus can be used to evaluate these hypotheses about OdR variation among different social groups in Odessa and to reconstruct potential sociolinguistic variation. This reconstruction requires careful tracking of the following information: (1) Instances of non-standard Russian features used by characters in the texts; (2) The demographics of both the speaker and the interlocutor, to the extent that they are explicitly laid out by the author, or can be deduced (e.g., education, socioeconomic status, ethnic group); and (3) The social setting in which the interaction takes place. Sorting by demographic criteria reveals that the different OdR features attested in a text are systematically distributed according to class, level of education, and ethnicity (Jewish/non-Jewish). This sociolinguistically-conditioned variation is given in Table 3.

FEATURE TYPE	REPRESENTATIVE EXAMPLE		SPEAKER GROUPS					
			upper middle class		middle class		lower class	
			educated	uneducated	educated		uneducated	
STANDARD	ODESSAN	Jews	Jews	Jews	non-Jews	Jews	non-Jews	
Odessan innovations	o + PREP	za + ACC					X	X
prepositions {general}	<Russian (e.g. k + DAT)	<Ukrainian (e.g. do + DAT)					X	X
collocations/ expressions	<i>slušajte</i> 'listen'	<i>slušajte sjuda</i> 'listen here'	X	X	X	X	X	X
Inflectional morphology	irregular NOM.PL (<i>synov'ja</i> 'sons')	"regularized" (<i>syny</i>)		X	X		X	X
Exuberant inflection	'whither, hither, thither' (uninflected adverbs)	Feminine nominal inflection		X			X	X
Lexicon	< Russian	< contact languages	X	X	X	X	X	X
French address	NA	<i>mamzel'</i> , <i>mos'e</i> {phon. varies with class}	X	X	X	X	X	X
Phonology (incl. stress)	[ft] (as in <i>što</i> 'what') / [g]	[ft] > [f:] / [g] > [ɣ]		X	X		X	X

Table 3. The plausible sociolinguistic distribution of OdR features from Jabotinsky's *The Five*

The reconstructed variation Table 3 confirms our hypothesis that OdR was spoken to some extent by a variety of groups in Odessa and not, as has often been claimed, exclusively by poor Jews. There are also other striking patterns, such as the overlap between lower class Jews and non-Jews (which is as expected, given their close contact), and the link between education and increased use of the standard language (educated Jews speak more like educated non-Jews). All of the groups make use of lexical borrowings and other Odessa-specific phraseology, including the use of French forms of address. However, only the uneducated groups seem to show significant nonstandard features at the level of grammatical structure.

5.3. PRAGMATIC FUNCTIONS OF VARIATION. Let us now turn to the second goal of our

sociolinguistic reconstruction: determining whether the use of OdR instead of standard Russian has any pragmatic significance. There is one example of a metalinguistic association with OdR that the narrator of *The Five* discusses explicitly, which is the use of nonstandard speech to signal deference toward the subject of discourse. In the following example, a maid employed by the narrator uses nonstandard verbal morphology to signify that she is speaking about someone of higher social standing (using the past active participle instead of the expected simple past tense):

- (7) *Fom-y Gavriľč-a nema vony uš-edši.*
 Foma-gen Gavriľč-GEN NEG 3.SG leave-PST.PRF.CVB
 ‘Foma Gavriľ is not here: he left.’ (Expected: *uš-el*, leave-PST.MASC)
- (8) *Pračka uš-la, a barynja – uš-edši.*
 laundress.NOM leave-PST.FEM while mistress.NOM leave-PST.PRF.CVB
 ‘The laundress left, while the mistress is “having left.”’ (Narrator’s explanation)

The second example of a pragmatic function of OdR emerges from the reconstruction in Table 3. There is a character in *The Five* (Serezha Milgrom) who does not seem to conform to the expected features for his class: although he is an upper-middle class Jew, his speech is highly nonstandard. In addition to the expected contact language borrowings and collocations, he also displays changes in inflectional morphology (e.g., nonstandard gender on nouns), exuberant inflection, and nonstandard preposition use. The narrator explicitly notes that Serezha’s use of these features is an affectation, which he only uses in relaxed social contexts. (Notably, when Serezha is in trouble with the law later in the novel, his speech is quite standard.) Thus, Serezha’s OdR use (while more authentic) seems to serve the same pragmatic function as in modern comedy routines and TV shows: he uses the dialect to sound more laidback, and to index an association with the Odessan culture of the lower class.

For the lower classes, OdR seems to index their lower social standing (which is perhaps why it is used to signal deference in (7-8)). The subject of the maids sentence in (7)—the narrators groundskeeper, Xoma—is actively attempting to abandon his OdR features, knowing that they index his lower social class. The maid seemingly notices this, and accommodates him by using the language she reserves for those of higher status. The dialect was also already linked to Odessan Jews: middle class non-Jewish characters in *The Five* are portrayed as mocking what they believe to be “Odessan” speech by specifically singling out Jewish lexical items.

Thus, as early as 1936, the language already apparently carried the conflicting social signaling it continues to have in the present day: members of the upper classes wanted to claim association with the language and its intrinsic link to the culture of Odessa, while members of the lower classes (who were viewed as uneducated when they used OdR) were attempting to standardize their speech. We can therefore reconstruct an ethnolinguistic repertoire for OdR when it was robustly spoken: speakers appear to have been able to switch among features of OdR (or to avoid the use of OdR entirely) in order to index extralinguistic traits.

6. Deriving the Odessan Russian stereotype. The social indexation of OdR likely formed the basis for what would become the modern Odessan Russian stereotype. Today, OdR invokes a unique association with not only the city, but its culture and its people—that is, use of the linguistic forms that define OdR indexes a specifically Odessan and Jewish identity, which modern speakers refer to as *kolorit* ‘color’. *Kolorit* includes the prototypical traits associated with people from Odessa: a sense of humor, southern (laid-back) temperament, and an entrepreneurial

spirit (Richardson 2008). As we have already noted, OdR features are widely used for this purpose and the use of a single salient feature is sufficient to index this Odessan identity.

Socially-salient dialectal features often follow a clear cline of linguistic change:

regional marker > identity marker > stereotype (Labov 1972, Irvine & Gal 2000)

This is the same pattern that we see in the development of OdR. Although it is commonly held to be a specifically Jewish ethnolect, many of the features are from Southern Russian dialects and Ukrainian—regional features. These were reinterpreted as markers of a Jewish Odessan identity, and have in turn become indexes of that stereotype. The variety of Odessan Russian described here is, as far as we can tell, no longer actually spoken in Odessa: since the break up of the Soviet Union in 1992, there has been steady Ukrainian immigration into Odessa, changing the city's demographics. The speakers we met with in Brighton Beach who had had occasion to visit Odessa in the last decade all spoke of the loss of “true” Odessan speech.

The result of these changes is that in popular Russian imagination, all that remains is the stereotype, which may be simply invoked by use of a handful of features and/or the use Jewish names, as described in Section 6.1.

6.1. INVOKING THE STEREOTYPE. There are several strategies that modern Russian speakers employ in the stereotypical use of “Odessan Russian.” The first is the over-regularization of attested features (which we have shown would have been subject to considerable variation even on an individual level). These include sound changes and certain lexical and grammatical items, as in (9)-(11):

- (9) [f] for SR *ft*, for example, *fo* instead of SR *fto* ‘what’ and *fto(by)* instead of SR *fto(by)* ‘in order to’
- (10) [ɣ] for SR *g*
- (11) Preposition *za* to mean ‘about’ (expected: *o*)

Another strategy is the use of features whose origins are unclear, as they are not attested and cannot be reconstructed from our sources at all:

- (12) (Over)use of *taki* ‘still’, ‘however’, ‘nonetheless’
- (13) Misuse of interrogative pronouns (e.g., *gde* ‘where’ for *čto*)
- (14) Pronunciation errors (e.g., *xodju* for *xožu* ‘I walk’)

Finally, it is often enough simply to include an explicit reference to the Jewish ethnicity:

- (15) Use one or more Jewish names, such as Sara, Abram, Moishe, Rabinovič

A common venue for the invocation of these stereotypes is in jokes (*anekdoty*). Odessans are known for their humor and numerous websites, videos (on YouTube), and publications catalogue Odessan jokes, often explicitly labeling them as Jewish jokes. Example (16) gives a fairly typical example of an Odessan joke. (Note the use of *taki*, an adverb that is frequently referenced in stereotypical Odessan speech but is not used in any of our sources, and the excessive use of a feature that is attested, *šo*.)

- (16) **Patient:** Doktor, ja sil'no terjaju volosy! **šo** mne sdelat', **šob** ix soxranit'?'
'Doctor, my hair is really falling out! What should I do, to save it?'
Doctor: Oj, èto **taki** prosto! *Kladite ix v korobočku.*
'Oj, it's so simple! Put them in a box!'

6.2. CONCLUSION. The case study of OdR justifies the need for a sociolinguistically-anchored approach to historical reconstruction. As we have seen, moribund languages and dialects undergo not only linguistic loss, but are also subject to changes in the social dimensions of their use. A dialect that was once used to perform a specific function in some contexts can become exclusively relegated to that function. OdR, once used throughout Odessa, has come to be used solely to index a particular attitude and only a subset of the original speech community. This change is recoverable only through reconstructing the original sociolinguistic variation of OdR.

References

- Benor, Sarah Bunin. 2010. Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics* 14.2. 159-183.
<https://doi.org/10.1111/j.1467-9841.2010.00440.x>
- Culpeper, Jonathan and Merja Kytö. 2010. *Early Modern English dialogues: Spoken interaction as writing*. Cambridge: Cambridge University Press.
- Herlihy, Patricia. 1986. *Odessa: A history, 1794-1914*. Cambridge, MA: Harvard University Press.
- Grenoble, Lenore A. and Jessica Kantarovich. *forthcoming*. *Methodological issues in documentation & reconstruction*. Amsterdam: John Benjamins.
- Irvine, Judith T. and Susan Gal. 2000. Language ideology and linguistic differentiation. In P.V. Kroskrity, (ed.), *Regimes of language: ideologies, politics, and identities*. 35-84. Santa Fe: School of American Research Press.
- Ives, Sumner. 1971. A theory of literary dialect. In Juanita V. Williamson & Virginia M. Burke (eds.), *A various language: Perspectives on American dialects*. 145-177. New York: Holt, Rinehart and Winston.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Richardson, Tanya. 2008. *Kaleidoscopic Odessa: History and place in contemporary Ukraine*. Toronto: University of Toronto Press.
- Rothstein, Robert A. 2001. How it was sung in Odessa: At the intersection of Russian and Yiddish folk culture. *Slavic Review* 60.4. 781-801. <https://doi.org/10.2307/2697495>
- Schneider, Edgar W. 2013. Investigating historical variation and change in written documents. In J.K. Chamber & Natalie Schilling-Estes (eds.), *Handbook of language Variation and change*. 5781. Malden, MA/Oxford: Wiley-Blackwell.
- Shishkova, Zinaida. 1973. *Eduard Bagritskij: Vospominanija sovremennikov*. Moscow: Sovetskij pisatel'.
- Stepanov, Ie. M. 2004. *Rosiiske movlennje Odesi: Mongraia [The Russian language in Odessa: A monograph]*. Odessa: Odessa National University.
- Sternberg, Meir. 1981. Polylingualism as reality and translation as mimesis. *Poetics Today* 2.4. 221-239. <https://doi.org/10.2307/1772500>
- Zipperstein, Steven J. 1986. *The Jews of Odessa: A cultural history, 1794-1881*. Stanford: Stanford University Press.