

Rethinking the Observer's Paradox and Data "Purity"*

SUZANNE WERTHEIM

University of California, Berkeley

In this paper, I will examine some of the difficulties faced by the linguistic fieldworker who is attempting to observe and record "natural" conversations, and I will reconsider the long-held sociolinguistic notion of the observer's paradox by recasting it within Bell's (1984) framework of audience design theory. Using data gathered during my own fieldwork, I will once again call into question the idea of a single, unmarked, unperformed vernacular, the access to which is supposedly blocked by the observer's paradox. Finally, I will demonstrate that "performed" or "self-conscious" speech produced for the fieldworker can be useful in systematic linguistic analysis, and in gaining insights into local language ideologies and linguistic norms.

The usual goal of sociolinguistic investigation is to gather recordings of "natural" speech, which is to say, somehow accessing and recording what is generally referred to as the vernacular, as "untainted" by interactions with the fieldworker as possible. Although specific definitions vary, among sociolinguists it is generally agreed that the vernacular is "the relatively homogenous, spontaneous speech reserved for intimate or casual situations...taken to reflect the most systematic form of the language acquired by the speaker, prior to any subsequent efforts at (hyper-) correction or style shifting..." (Poplack 1993:252). Historically, sociolinguistic investigators have also used the term "vernacular" to refer to a low-prestige variety in contradistinction to a standard and high-prestige variant, associating the vernacular both with social groups (e.g., African Americans) or with localities (e.g., Belfast English) (Milroy 1987b:58).

My fieldwork situation was bilingual rather than bidialectal, and without a ready binary distinction of standard vs. vernacular varieties. The language I am studying, Tatar, is a is a Turkic language currently spoken by one quarter of the

* The fieldwork upon which this paper is based was made possible by generous grants from the International Research and Exchanges Board and from the Academy for Educational Development.

four million residents of Tatarstan,¹ a semi-autonomous republic of the Russian Federation, and it has been under significant stress from Russian for hundreds of years. However, since the fall of the Soviet Union and the declaration of Tatarstan's autonomy, the republic's language policies have been promoting the use of Tatar and the expansion of its functional domains. Russian remains the dominant language and is high prestige, but in the milieu of the Tatar Social Club that was the locus of my fieldwork, as well as among the Tatar intelligentsia, the Tatar language, more precisely "pure" Tatar (*saf tatar tele*) and "literary" Tatar (*ädäbi tatar tele*), is also awarded high prestige. Additionally, as will be seen later, there is a continuum of linguistic performance for urban Tatar bilinguals, ranging from completely Russian to completely Tatar, further complicating the designation of a single "unmarked" vernacular.

1. Observation and audience design

While some linguistic fieldworkers (e.g., Baugh 1993, Besnier 1994) claim that informants can forget about recording and produce naturalistic speech, I found that in my own fieldwork, the presence of the language investigator, with or without recording materials, would often be enough in and of itself to precipitate "performed" speech rather than "unperformed." Performance speech, as defined by Schilling-Estes (1998), is "associated with speakers' attempting to display for others a certain language or language variety, whether their own or that of another speech community" (53). Only after my return home from the field, and extensive review of my recordings and fieldnotes, did I realize that the linguistic performance and style shifting of the Tatar speakers I had observed and recorded could be best interpreted by taking into account two major factors: (1) the speaker's assessment of my social role, particularly as in-group or out-group member, and (2) my participant role in the speech event in question.

The first sociolinguistic investigations conceptualized style shifting on a single continuum ranging from careful to casual speech. Labovian sociolinguistic interviews were designed to elicit more- and less-careful styles, and topics introduced by the interviewer were meant to create contexts for casual speech.

Bell, following Brown and Levinson (1979), dismissed Labov's attention-to-speech continuum as an "impoverished" view. He proposed an alternate explanation, that of *audience design*, which holds as a basic tenet that "at all levels of language variability, people are responding primarily to other people. Speakers are designing their style for their audience" (1984:197). Variables such as topic and setting are seen to have less effect upon stylistic variation than audience, which is the "responsive, critical forum before whom the utterances are performed" (161): for empirical testing of this claim cf. Rickford and McNair-Knox (1994) and Lewis (2002), inter alia. The audience design framework is generally held to be superior to the attention-to-speech continuum, which has fallen into disuse.

¹ And elsewhere in the territories of the former Soviet Union.

Rethinking the Observer’s Paradox and Data Purity

Refining upon Goffman’s (1981) “participant framework,” Bell offers five participant roles for any given speech situation, which are as follows:

(1) Bell’s Participant Roles (Bell 1984)

1. Speaker – uses the 1st person,
2. Addressee – addressed in 2nd person,
3. Auditor – referred to in 3rd person,
4. Overhearer – unratified to participate in the speech event, not addressed or referred to,
5. Eavesdropper – both unratified and unknown.

Each step down on the list increases the distance from the speaker, sometimes physically. Figure 2 below summarizes the attributes of audience roles in a speech event. A fieldworker can participate in any of these audience roles.

(2) Hierarchy of attributes and audience roles (adapted from Bell (1984))

	<i>Known</i>	<i>Ratified</i>	<i>Addressed</i>	<i>Person</i>
Addressee	+	+	+	2 nd
Auditor	+	+	-	3 rd
Overhearer	+	-	-	n.a. (unratified)
Eavesdropper	-	-	-	n.a. (unratified)

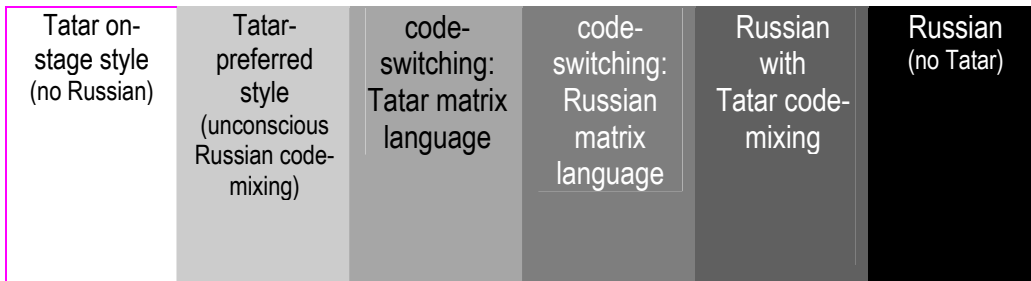
Audience roles are assigned by the speaker, and will have different levels of salience for the speaker’s style design – for example, auditor effect is usually lower than addressee effect. Style shifting can result in convergence to or divergence from the audience – convergence is seen as accommodation, and an expression of what Brown and Gilman (1960) call “solidarity,” while divergence is interpreted as “referee design,” which marks the speaker as a member of a social group not present in the speech situation – but referred to by his or her divergent style – and marks the audience (most commonly an addressee) as a non-member of the referred-to group.

Therefore, one must take into account both the fieldworker’s participant role in a speech event and her or his position as a socially located being. As a member of neither the majority nor minority group, yet somehow aligned with minority group interests and culture, my own social position was unique. My attempts to speak Tatar would elicit commentary on my performance, on the linguistic performance of other Tatars, and on the Tatar language itself. It quickly became clear that Tatar speakers felt responsible for presenting me with the best possible Tatar, both so I could have appropriate models for learning, and so I could represent the language well in my research. I received invitation after invitation to be taken “back to the village,” where I could hear “real” Tatar spoken, Tatar that was “purer” than the urban dialects, which had been “tainted” by Russian. As my Tatar competence improved, I found that this high level of language awareness, with its stated ideals of “pure” and “literary” Tatar, was not merely provoked by

the presence of the language learner and investigator. This “discourse of purity” and constant awareness of the level of Russification of Tatar, combined with the idea of “saving” the Tatar language from both misuse and disuse, was not only presented to me, but also found in newspaper articles and opinion pieces, on television, on internet bulletin boards, and on the radio.

Over time, as speakers’ familiarity with me increased and my positioning within the community became increasingly in-group, I eventually gained access to a wide range of styles. What I found was that urban bilingual Tatars had a continuum of linguistic performance that was related to language mixing, as shown in Figure (3), below.

(3) Cline of language mixing for urban Tatar bilinguals ²



Tatar on-stage style is often found in the public sphere in formal registers, particularly when aiming for a high literary standard, but can also found in informal register. In this style, speakers will de-Russify their Tatar to the best of their ability. Performances in Tatar on-stage style will generally take place at Tatar cultural events, in radio and television interviews, in political speeches, and in public presentations or comments at the Tatar Social Club. Private use of Tatar on-stage style seems to only occur in conversations with or for investigators of Tatar language and culture.

Tatar-preferred style is always in informal register, and is found in private conversations where Tatar has been in some way established as the preferred language of communication: for example, this is the main style of intergenerational family communication in Tatar-speaking homes, both urban and rural. For intragenerational family interactions, generally Tatar-preferred style will be used by parents, while siblings will interact in one of the code-switching

²I would like to acknowledge up front that these styles are arranged according to a single stylistic variable, the level of use of Russian, even though there are other markers that differentiate, for example, Tatar on-stage style from Tatar-preferred style. Additionally, linguistic performance within each style can vary, e.g., Tatar on-stage style can be both in formal register and in informal register. And within formal register of Tatar on-stage style there are further nuances, such that speeches given at a literary tribute evening are different from the on-stage performance of the master of ceremonies at a concert. I in no way wish to suggest that this single variable, the level of Russification, is the sole differentiating marker of Tatar style and performance.

styles, or even in Russian with Tatar code-mixing. In Tatar-preferred style, the level of language awareness and verbal hygiene (Cameron 1995) is lower than in Tatar on-stage style, and although speakers believe themselves to be speaking “pure” Tatar, they actually are unconsciously using what I call Russian “metalinguaging” words (following Maschler 1994). These code-mixed words – not standard borrowings, and mostly function words, adverbs of manner, and interjections – all have pragmatic functions and both structure and comment upon discourse (for a more complete examination of Tatar-preferred style and code-mixing, see Wertheim 2003).

2. The role of the language investigator in style shifting

Upon reviewing my data, I found that I was able to discern patterns in style shifting that showed that a speaker's choice of style was dependent in part upon the audience role assigned to the fieldworker participating in a speech event. This can be demonstrated by the style shifting of one Tatar bilingual, who I will call Galimä,³ a 46-year old Tatar philologist with whom I had both a professional and personal relationship, and who I observed in a variety of situations, locations, and speech events. I believe that for Galimä my social role remained reasonably constant – I was non-peer (15 years younger), outsider, language learner and language investigator. Although our relations were quite warm, and we discussed personal matters frequently, she would only address me using the formal version of ‘you’ (*sez*) and never once used the informal ‘you’ (*sin*).

(1) In dyadic conversations (where I was both a speaker and an addressee) in either a professional or a personal context, Galimä was always in Tatar on-stage style. As with many other Tatars who felt responsible for helping me in my Tatar language learning and investigations, she would choose ideology over communication, and never used Russian with me, opting for a simplified Tatar – or even drawings – to explain words or phrases I hadn't understood, when a single Russian word would have sufficed.

(2) Galimä would use Tatar on-stage style with colleagues, friends, and acquaintances when ratifying them as participants in conversations in which I would be speaker, addressee, or auditor. She would do this even with speakers whom she knew to have limited Tatar competence. This was clearly a self-conscious “performance” of Tatar, and arrangement of performance by others, one that was staged for me. Very often it was only Galimä's participation in the conversation that kept it in Tatar on-stage style – if she would leave, speakers would frequently ask in Tatar if I knew Russian, and when I answered in the affirmative, would either switch to Russian with no Tatar in it, or would code-switch with Russian as the matrix language.

(3) If I was an auditor of a conversation with family and friends, Galimä would speak in Tatar on-stage style, with no Russian whatsoever. For example, if we were drinking tea in her kitchen and talking, and her son came in and asked a

³ A name I have chosen for its meaning of “scholarly, knowledgeable.”

question, she would use only Tatar with him. Additionally, if he used any Russian, she would upbraid him and tell him to speak in Tatar only. However, if I was not an auditor but rather an overhearer, and thus unratified as a conversation participant, Galimä would code-switch with Tatar as the matrix language. For example, one time I arrived early for a meeting with Galimä and entered the room while she was still consulting with her younger cousin, advising her on how to best teach Tatar periphrastic verbs. Galimä saw me enter the room – I sat at the opposite end of a long table, read a newspaper, and surreptitiously listened as the two spoke mostly in Tatar, but with code-switching into Russian of higher-level grammatical constituents and code-mixing of isolated words. As soon as their meeting was over, I became a ratified conversation participant, and when Galimä introduced me to her cousin, she switched into Tatar on-stage style, with no Russian, and her cousin followed suit.

(4) Galimä would speak Russian with those bilingual Tatars with whom Russian is the usual language of interaction, but if I was an auditor of one of these Russian-language interactions, my auditing would cause linguistic meta-commentary on Galimä's part. For example, one time we stopped off at the post office, where the transaction was conducted in Russian. As we turned away to leave, resuming our Tatar-only conversation, Galimä said to me, "I don't know why I speak Russian with that woman. I've been coming here for years. She knows that I speak Tatar, and I know that she speaks Tatar. So why do we speak in Russian?" I don't want to say that my presence brought this fact to Galimä's attention for the first time, but I believe that my auditing did trigger the meta-commentary. Perhaps this is because of the conflict that my auditing caused in Galimä's style-shifting. Recall that Galimä's usual behavior when I was a ratified conversation participant was to use Tatar whenever possible with anyone whom she knew to be a Tatar speaker, regardless of their competence. However, local linguistic norms required her to shift to Russian when transacting post office business, and this requirement seems to have superseded the style-shifting patterns that were based on my presence. Perhaps this conflict in linguistic presentation of identity – culturally competent citizen on the one hand, and speaker of pure Tatar on the other – is what caused her explicitly stated dissatisfaction with her linguistic performance.

(5) When conversing with Russian monolinguals, Galimä would speak Russian only, regardless of my participation role. This would sometimes lead to an interesting phenomenon – a sort of Russian hangover, where Galimä would be "out of phase" in her style shifting, such that after the Russian-language conversation had ended, she would return to our conversation and address me in Russian. However, this Russian performance would only last for one conversational turn, because regardless of the language of my response, Galimä would become immediately aware of her "inappropriate" style. I could respond in Russian, or I could respond in Tatar, but her response, always in Tatar, would be the same: "Why am I speaking Russian with you? We don't speak Russian

Rethinking the Observer’s Paradox and Data Purity

together. Let’s speak Tatar.” And we would continue our conversation with her in Tatar on-stage style.

In summary, we can see that my various participation roles seem to have influenced both Galimä’s style shifting and her awareness of style shifting. Perhaps this can be seen more clearly in Figure (4), below.

(4) Effect of changing participant roles of fieldworker on Galimä’s speech

<i>Fieldworker role</i>	<i>Speech event participants</i>	<i>Situation</i>	<i>Galimä’s speech style</i>
Speaker/ addressee	friends; family; colleagues; acquaintances; no other participants	private conversation	Tatar on-stage style
Auditor	friends; family; colleagues; acquaintances	private conversation	Tatar on-stage style
Overhearer	friends; family; colleagues; acquaintances	private conversation	Code-switching: Tatar as matrix language
Auditor	service personnel	business transaction	Russian with post-transaction metacommentary
Auditor	Russian monolinguals	all	Russian with periodic “Russian hangover” followed by metacommentary

3. The role of the language investigator in recording speech

The various performance styles I have just described were produced in the presence of a fieldworker who was not explicitly in her role as fieldworker at the time of the speech events in question, and who was without any obvious note-taking or recording equipment. Now I would like to turn to the recorded speech event, and briefly examine this too from within the framework of audience design.

A recorded private-domain speech event, regardless of the physical presence or absence of the fieldworker, is an atypical, even extraordinary situation that is not classifiable within Bell’s ordinary hierarchy of audience roles. And here is why: the recording equipment, previously analyzed by some sociolinguists as a participant itself in the speech event, actually represents an end-listener or listeners whose identity is not known at the time of the speech event. This means that the speech event participant represented by the recording equipment is simultaneously *ratified* (providing that permission to record has been requested and granted) and *unknown* – a participant role that is not analogous to any found in Bell’s framework because it is unique to the experience of being investigated by a fieldworker of some sort, linguistic or otherwise. In Figure 5 below I have added the end-listener of recorded speech to the audience role hierarchy so it can be easily compared with the other standard audience roles. The “strange” or “unnatural” behavior of recorded speech event participants can thus, in part, be interpreted as speakers trying to grapple with a participant role they have never dealt with before, that of the unknown eavesdropper who is nonetheless ratified. Perhaps what is unnatural is not so much the behavior and performance of

speakers as the uniquely conflicting audience role attributes of the recording's end listener, for whom speech must be designed, but whose social role and identity remains cloaked in mystery (or at least unclear).

(5) Adjusted hierarchy of attributes and audience roles for a recorded speech event

	<i>Known</i>	<i>Ratified</i>	<i>Addressed</i>	<i>Person</i>
Addressee	+	+	+	2 nd
Auditor	+	+	-	3 rd
Overhearer	+	-	-	n.a. (unratified)
Eavesdropper	-	-	-	n.a. (unratified)
End-listener of recorded speech	-	+	-/+	2 nd /3 rd /n.a.

Recall that according to Bell, speakers design their speech with audience members in mind, and those participants whose audience roles are higher up on the hierarchy (as seen in Figure 1) will usually have a greater effect on the linguistic performance of the speaker. I submit that the unnatural audience attributes of the end-listener, highlighted by the act of recording and the presence of recording equipment, can cause this audience role to be of primary salience and effect. The fieldworker may be entirely absent and yet still the most salient participant. By understanding the identity and traditional audience role assigned to this end-listener by the speakers being recorded, it is possible to more systematically account for her or his effect on recorded speech. The end-listener can be seen as the fieldworker, and the fieldworker alone; can be seen as the fieldworker in combination with other language investigators; and can be seen as a person or persons completely unknown to the speaker. For example, some of the people I recorded pictured me alone as the end listener, such that the end-listener took on the attributes of addressee. In three separate recorded conversations, people performed for me in the most standard sense of the word, singing Tatar songs directly into the recording equipment. Several of them later volunteered the information that they had sung as a memento, so that later on when I was back home in America, I would have these nice songs to listen to, and think of them fondly. We can extrapolate from this performance and subsequent meta-commentary that these speech events were performed with consciousness of my eventual review of the recorded material and subsequent linguistic analysis. These people were speaking “good Tatar” both for me and for posterity, and their level of verbal hygiene remained quite high.

4. The value of performance speech

Bauman (1975, 1977) argued that in addition to “vernacular” speech, “performed” speech is available and meaningful for analysis, yet it remains understudied by sociolinguists and linguistic anthropologists. Ochs (1988) used performance speech to illuminate the linguistic counterpart to the Samoan system of spatial

contrasts, and her observations prompted her to study the relationship between children's acquisition of language and their acquisition of culture. Schilling-Estes (1998) found performance speech in Ocracoke English to be both regularly patterned (contra Labov), and useful in gaining insight as to what elements of the dialect were most salient to its speakers.

Tatar on-stage style, particularly when combined with metacommentary and speaker evaluations by members of the speech community, was absolutely indispensable in my understanding of a variety of local norms and practices. Performance speech pointed the way to the "discourse of purity" that was the most significant of the language ideologies that I found, and exemplified speaker's attempts to produce "pure" Tatar. It demonstrated verbal hygiene in practice, and showed me who were the community's standards keepers. When speakers remained in Tatar on-stage style rather than accommodating their uncomprehending audience⁴ by switching into Russian, it alerted me to the existence of Tatar speakers who choose ideology over communication, for whom out-group referee design supersedes accommodation and convergence. Tatar on-stage style, when produced by less-competent Tatar speakers who feel compelled to refrain from using any Russian at all, can provide excellent examples of morphosyntactic interference. For fully competent speakers, a comparison of their linguistic performance in Tatar on-stage style and in Tatar-preferred style gives evidence for what appears to be two different underlying competences and grammars: the one accessed by Tatar on-stage style has limited Russian influence, and the one accessed by Tatar-preferred style appears to be what is referred to as a "composite matrix language," a language with lexical, morphological, and syntactic elements of both languages in contact that is posited by some linguists to be an intermediary step in multi-generational language shift. Using only one of these styles for a grammatical description would lead to an incomplete and erroneous analysis. Finally, placing "performance" speech within a continuum of language-mixing styles gives insight into the identities and roles that are associated with certain kinds of styles, and how choosing styles when constructing identity may play a role in the stylistic and domain shrinkage of a language, and thus play a role in language shift.

5. Conclusions

In her early critique of the Labovian sociolinguistic interview as a method for collecting spontaneous speech, Wolfson called into question the idea of a "single, absolute entity answering to the notion of natural/casual speech" (1976:202). And indeed, not one of the styles that I observed among urban Tatar bilinguals can be regarded as more natural or unmarked than any of the others: they are all designed to be appropriate for the audience, topic, and setting of the speech event in which they are produced.

⁴ Sometimes, but not always, me.

Abandoning the search for a natural and unmarked vernacular has important ramifications for linguistic fieldwork and analysis. The first is that no speech should be dismissed as unnatural and thus unusable, but that instead speech events should be analyzed according to audience and participant roles in the context of local norms and ideologies. Second, performance speech can be utilized for linguistic analysis in a variety of ways, for example, to elucidate verbal hygiene patterns, or to demonstrate which elements of a language or dialect are most salient to its speakers. Third, the fieldworker should try to gain access to as many styles and registers as possible by using her or his unique social status. Varying one's recording methodology – recording when the fieldworker is both present and absent, recording in different settings, and having community members record for you – can increase access to various styles, particularly those that are in-group only, and reduce the salience of the “unnatural” audience role of the unknown yet ratified end-listener.

To conclude, I have briefly addressed just one of the difficulties of sociolinguistic fieldwork, the gathering and contextualizing of spontaneous speech. I have attempted to demonstrate that by framing the observer's paradox within audience design theory, it is possible to more systematically account for and analyze the effect of the fieldworker upon speech production. I have once again called into question the idea of a single, unmarked, unperformed vernacular, the access to which is supposedly blocked by the observer's paradox. Finally, I have shown that a “performance” or “on-stage” style can be just as useful for sociolinguistic analysis as other styles. By maintaining what Hyman (2001) calls the “fieldworker mentality,” which keeps one open to unanticipated phenomena, the language investigator may find that performance speech will lead to both unexpected and unexpectedly fruitful avenues of linguistic inquiry.

References

- Baugh, John. 1993. Adapting dialectology: The conduct of community language studies. In D. R. Preston (ed.) *American dialect research*. Amsterdam: John Benjamins.
- Bauman, Richard. 1977. *Verbal art as performance*. Rowley, MA: Newbury House.
- _____. 1975. Verbal art as performance. *American Anthropologist* 77:290-311.
- Bell, Alan. 1984. Language style as audience design. *Language in Society* 13, 145-204.
- Besnier, Nico. 1994. The truth and other irrelevant aspects of Nukulaelae gossip. *Pacific Studies* 17:3.
- Brown, Roger and Albert Gilman. 1960. The pronouns of power and solidarity. In T. A. Sebeok (ed.) *Style in language*. Cambridge, MA: MIT Press. 253-276.

Rethinking the Observer's Paradox and Data Purity

- Brown, Penelope and Stephen Levinson. 1979. Social structure, groups, and interaction. In K. R. Schere and H. Giles (eds.) *Social markers in speech*. Cambridge: Cambridge University Press.
- Cameron, Deborah. 1995. *Verbal hygiene*. London/New York: Routledge.
- Goffman, Erving. 1981. *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Geertz, Clifford. 1983. *Local knowledge: Further essays in interpretive anthropology*. New York: Basic Books.
- Hyman, Larry. 2001. Fieldwork as a state of mind. In Paul Newman and Martha Ratliff (eds.) *Linguistic fieldwork*. Cambridge/New York: Cambridge University Press.
- Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1, 97-120.
- Lewis, Julie. 2002. *Social influences on female speakers' pitch*. Unpublished University of California, Berkeley dissertation.
- Maschler, Yael. 1994. Metalanguaging in bilingual conversation. *Language in Society* 23, 325-366.
- Milroy, Lesley. 1987a. *Language and social networks*, 2nd edition. Oxford/New York: B. Blackwell.
- _____. 1987b. *Observing and analysing natural language: A critical account of sociolinguistic method*. Oxford/New York: Blackwell.
- Ochs, Elinor. 1988. *Culture and language development: Language acquisition and language socialization in a Samoan village*. Cambridge: Cambridge University Press.
- Poplack, Shana. 1993. Variation theory and language contact. In D. R. Preston (ed.) *American dialect research*. Amsterdam: John Benjamins.
- Rickford, John and Faye McNair-Knox. 1994. Addressee- and topic-influences style shift: A quantitative sociolinguistic study. In D. Biber and E. Finegan (eds.) *Sociolinguistic perspectives on register*. New York/Oxford: Oxford University Press.
- Schilling-Estes, Natalie. 1998. Investigating 'self-conscious' speech: The performance register in Ocracoke English. *Language in Society* 27, 58-83.
- Wertheim, Suzanne. 2003. *Linguistic purism, language shift, and contact-induced change in Tatar*. Unpublished University of California, Berkeley dissertation (in progress).
- Wolfson, Nessa. 1976. Speech events and natural speech: Some implications for sociolinguistic methodology. *Language in Society* 5(2), 189-210.

Suzanne Wertheim
1203 Dwinelle Hall
University of California, Berkeley
Berkeley, CA 94720-2650

wertheim@socrates.berkeley.edu