■ *Asif Agha*
UNIVERSITY OF PENNSYLVANIA

Voice, Footing, Enregisterment

The article argues that a fuller consideration of voicing phenomena clarifies the nature of processes whereby registers of language expand, change, or remain constant in the socialized competence of language users. The first half of the article describes the semiotic processes whereby voicing effects are recognized or identified by interactants. The second half discusses the role of discursive voices and figures in large-scale sociohistorical practices involving registers. The overall goal is to show that the social existence of registers depends on the semiotic activities of language users, particularly those characterized in this article as matters of "alignment." [voice, registers, footing, stereotypes, discourse analysis].

Introduction

In recent work in linguistic anthropology the concepts of register and voice are seen increasingly as linked. Formulations such as "the registers described here represent . . . voices a speaker takes on in different social situations" (Irvine 1990:153) or "every individual has a repertoire of cultural registers or voices" (Mannheim 1997:218) highlight the fact that contrastive patterns of register use index distinct speaking personae in events of performance.

My goal here is to develop some consequences of this fact for larger-scale processes of enregisterment, processes whereby distinct forms of speech come to be socially recognized (or enregistered) as indexical of speaker attributes by a population of language users.[1] I have discussed these processes in some detail in recent work, arguing that registers are not static facts about a language but *reflexive models of language use* that are disseminated along identifiable trajectories in social space through communicative processes (Agha 2002, 2003). My goal here is not to review that discussion but to elaborate on a single aspect of these social-reflexive processes: In the course of any process of social dissemination, register models undergo various forms of revalorization, retypification, and change. What role do voicing phenomena play in encounters with registers and in the maintenance and transformation of register values over time?

My general argument is that we cannot understand macro-level changes in registers without attending to micro-level processes of register use in interaction. More specifically, I argue that encounters with registers are not merely encounters with voices (or characterological figures and personae) but encounters in which individuals establish forms of footing and alignment with voices indexed by speech and thus with social types of persons, real or imagined, whose voices they take them to be.

Voicing Contrasts versus Typifiable Voices

The concepts of register and voice overlap—both involve speaker-focal indexicality—but are not equivalent. When Bakhtin (1981, 1984) speaks of voices, he is concerned with the ways in which utterances index typifiable speaking personae; similarly many registers index social attributes of speaker such as gender, class, caste, and profession. Yet the Bakhtinian conception is looser and more inclusive. Bakhtin uses the term *voice* for speech forms that index widely recognized register distinctions (which he terms *social speech types* or *social voices*; his examples include the speech of particular classes and professions, slangs, trade jargons) but also for speech that indexes event-specific, potentially unique images of personhood (which he calls *individual voices*). The formulation leads to certain obvious questions: Why do some voices appear individual, unique, event-specific? What, on the other hand, are social voices?

The typifiability of *voices* (whether as "individual" or "social") presupposes the perceivability of *voicing contrasts,* or the differentiability of one voice from another. Thus a more basic question concerns the variety of ways—including the variety of semiotic channels and media—in which voicing contrasts can be expressed and recognized. Bakhtin's terminology of voice and dialogism is based on metaphors of oral speech and conversation that are rather misleading if taken literally (I do not suggest that Bakhtin himself is misled by them, however). The term *voice* is based on a corporeal metaphor of phonation—the friction of air over vocal chords—even though the phenomenon it names is not restricted to, and hence has no necessary connection to, oral speech; Bakhtin's own examples involve written texts in any case. Similarly, the term *dialogic* is more general than the special case of dyadic conversation to which it metaphorically alludes. The term describes any structure of entextualization that juxtaposes images of speaker-actor as contrasting with or appearing to react against each other. Dialogic relations are manifest in oral conversation but also in a variety of other discursive and semiotic genres, including novels, other literary works, even "images belonging to different art forms" as long as they are "expressed in some *semiotic* material" (Bakhtin 1984:185). Voicing contrasts involve figures of personhood that are juxtaposed within structures of entextualization composed of many types of signs, including linguistic signs (whether written or spoken) and nonlinguistic ones. In the case of nonlinguistic semiosis, the metaphor of "voice," already stretched for written discourse, becomes too thin to be usable. I shall use the term *figures of personhood* to speak of indexical images of speaker-actor in general terms, using the term *discursive figure* or *figure performed through speech* as an equivalent of *voice* in the special case of linguistic semiosis.

In the following sections I argue that voicing contrasts, once perceived, are construed as typifiable voices (i.e., as positively characterizable "types") on the basis of reflexive cues contained within the text segments that formulate them. Yet the Bakhtinian dichotomy of "individual" versus "social" types (or rather, the Durkheimean dichotomy that Bakhtin presses into service here) provides an incomplete framework for reasoning about this process. In the next section I show that Bakhtin's individual voices are not a unified phenomenon but a class of cases, involving different degrees and varieties of "individuality," cued by co-occurring signs. The "social" pole of the dichotomy requires reanalysis as well. In general terms, Bakhtin's social voices are discursive figures that permit characterization through a metadiscourse of social types of person or persona attributes. Within this set lies a particularly interesting subclass, the class of social voices linked to registers, which I call *enregistered voices*. These have a "social" character in two quite different senses. First, a register's forms are social indexicals in that they index stereotypic social personae (viz., that speaker is male, lower-class, a doctor, a lawyer, an aristocrat, etc.), which can also be troped upon to yield hybrid personae of various kinds; thus every register has a *social range,* a range of figures performable through its use. Second, registers are social formations in the sense that some language users but not

others are socialized in their use and construal; thus every register also has a *social domain*, a group of persons acquainted with—minimally, capable of recognizing—the figures performable through use.[2]

The fact that registers are used by social persons and index social personae introduces an inherent reflexivity in the social life of registers. Encounters with registers are not merely encounters with characterological figures indexed by speech but events in which interlocutors establish some footing or alignment with figures performed through speech, and hence with each other. I refer to these as matters of *role alignment.* The class of processes is quite large and I discuss only a few illustrative types in this article. My goal is to suggest that registers are living social formations, susceptible to society-internal variation and change through the activities of persons attuned to alignments with figures performed in use, and that macrosocial regularities of enregisterment—facts of demographic growth or decline (changes in social domain) or of value maintenance or counter-valorization (changes in social range)—are large-scale effects of alignments that unfold one communicative event at a time.

Entextualized Voicing Contrasts

No figure of personhood is typifiable as a discrete voice (of whatever type) unless it is differentiable from its surround. The typifiability of voices presupposes the perceivability of voicing contrasts. Voicing contrasts are made perceivable or palpable by the metrical iconism of co-occurring text segments—the likeness or unlikeness of co-occurring chunks of text—which motivate evaluations of sameness or difference of speaker. Such entextualized contrasts are wholly emergent and nondetachable: They are figure-ground contrasts that are individuable only in relation to an unfolding text structure (hence emergent) and are not preserved under decontextualization (hence nondetachable). Moreover, such effects may or may not be typifiable as the voices of particular, nameable persons.

Unnamed Voices

The following example from Dickens' *Little Dorrit* is one of Bakhtin's illustrations of voicing contrasts that do not map clearly onto any named persons in the text. Bakhtin analyzes the excerpt into a number of differentiable voices or speaking personae that do not correspond to the speech of the novel's named characters (i.e., its official dramatis personae). Here, then, are voices that are individuable but not nameable. How do such effects occur?

> It was a dinner to provoke an appetite, though he had not had one. The rarest dishes, sumptuously cooked and sumptuously served; the choicest fruits, the most exquisite wines; marvels of workmanship in gold and silver, china and glass; innumerable things delicious to the senses of taste, smell, and sight, were insinuated into its composition. *O what a wonderful man this Merdle, what a great man, what a master man, how blessedly and enviably endowed*—in one word, what a rich man! [*Little Dorrit,* bk. 2, ch. 12; Bakhtin 1981:304]

Since the effects in question depend on a framework of contrasting text segments, I have rewritten the example in Table 1 in order to highlight the main contrasts.

The first part of the passage (segment$_1$, roman type) exhibits what Bakhtin calls "a parodic stylization of high epic style." Here we have a series of epithets (underlined) that characterize the items comprising the banquet (the food, the china) in intensely hyperbolic terms.

The hyperbolic epithets continue in segment$_2$, but with shifts in topical referents and represented speakers. The topic of segment$_2$ is no longer the banquet display but the banquet giver, Merdle. And the impersonal, third-person scene descriptions of segment$_1$ give way to a series of emotive speaker exclamations in segment$_2$ (e.g., three tokens of the phrase *(O) what a __ man!*) suggesting a plurality of exclaiming voices. But whose voices? The pattern of metrical continuity and substitution be-

Table 1
Voicing effects based on metrical contrasts.

Segment$_1$	It was a dinner to provoke an appetite, though he had not had one. The <u>rarest</u> dishes, <u>sumptuously</u> cooked and <u>sumptuously</u> served; the <u>choicest</u> fruits, the <u>most</u> <u>exquisite</u> wines; <u>marvels</u> <u>of</u> <u>workman-ship</u> in gold and silver, china and glass; <u>innumerable</u> things <u>deli-cious</u> <u>to</u> <u>the</u> <u>senses</u> of taste, smell, and sight, were insinuated into its composition.
Segment$_2$	*O what a <u>wonderful</u> man this Merdle, what a <u>great</u> man, what a <u>master</u> man, how <u>blessedly</u> and <u>enviably</u> endowed—*
Segment$_3$	**in one word what a <u>rich</u> man!**

tween the two segments formulates segment$_2$ as a series of appreciative cries by on-lookers to the scene described in segment$_1$. The overall text pattern (segment$_1$ + seg-ment$_2$) models a form of interactional uptake on the part of onlookers, a sudden grasp of what the banquet items imply about the banquet giver to those who behold them.

The exclamative form is preserved in segment$_3$ (bold italics) in the expression *what a rich man!*. Metrical parallelism here maintains the effect that someone is exclaiming upon Merdle, the banquet giver. But who is exclaiming now? The represented speech frame *in one word* formulates the exclamation as the narrator's gloss on the preced-ing exclamations. The metricalized substitution formulates a specific footing be-tween the voiced speaker(s) of segment$_2$ and segment$_3$. The onlookers' use of hyperbolic epithets to glorify Merdle (*wonderful, great, master, blessedly-and-enviably endowed*) is now summed up by the narrator's use of a single epithet (*rich*). The con-trast formulates the onlookers as somewhat crass and vulgar, as persons who would mistake wealth for refinement. Bakhtin sums this up by saying that the phrase *what a rich man!* implements an "ironic" voice, a voice in which the narrator effects the "unmasking" of another's speech.

It is very important to see, however, that the voice in question is not implemented by the phrase *what a rich man!* alone but by an entextualized structure (segment$_1$ + segment$_2$ + segment$_3$) of which it is a part. This point is critical. The phrase *what a rich man!*—taken now as a phrase of English—does not convey "irony" or "unmasking" in any intrinsic sense (e.g., by virtue of its sense, denotation, or illocutionary force); these effects do not occur regularly for every token of the phrase in ordinary English usage. Yet the effect is fairly well motivated for the current token of the phrase by its framing text structure. If we isolate the phrase from the larger structure of which it is a part, the effect vanishes; the remainder, the isolable expression *what a rich man!*, conveys no stable (detachable) voice of "irony" or "unmasking" by itself. These ef-fects are, rather, emergent projections from a metricalized text structure of which the exclamatory phrase is a fragment.[3]

The Voices of Named Individuals

In the foregoing examples the phenomenon of metrical contrasts in text consti-tutes a metapragmatic framework for delineating voicing contrasts that is highly ro-bust but implicit; it permits the *contrastive individuation* of voices but not their biographic identification. Under these conditions Bakhtinian voices occur as the in-dexical effects of textual scopes or segments, metrical contrasts among which moti-vate construals of stance, footing, and alignment among entextualized figures. But such voices are virtual speaking personae; they are not clearly identifiable as the speech of named biographic persons.

Any sort of *biographic identification* requires that we use a system of person deixis to name the textual zones that convey the voicing contrast. In such cases, person deixis functions as a second metapragmatic framework linking facts of textually zoned voicing contrasts to facts of named personhood. But the effect is not equally transparent for all types of represented speech (see Table 2).

In the case of direct reports, the reporting and reported voices are distinguished very clearly because several cues co-occur together to reinforce the voicing contrast. First, such constructions involve the contiguity of two textual zones, the framing and framed material; if we decontextualize the framed material from its frame, the voicing contrast vanishes (i.e., the utterance is no longer understood as another's speech). Second, since the two textual zones are grammatically linked clauses, the voicing contrast is demarcated by a clause boundary. Third, direct reports exhibit an independence of deictic and other indexical patterning across clause boundaries. For example, in a case like the following,

John$_i$ promised Alice$_j$ "I$_i$'ll go to the bank for you$_j$"

we have two distinct zones of deictic patterning, configured into two grammatical clauses. The matrix clause contains proper names and past tense; the subordinate clause has participant pronouns and future tense. The person-referring forms are co-referential in the way shown by subscripts. The independence of deictic choices across the clause boundary—both for person deictics (John$_i$ vs. I$_i$, Alice$_j$ vs. you$_j$) and for tense deictics (*-ed* "past" vs. *-'ll* "future")—entails that the clauses do not share the same zero point or origo of deictic reckoning; this implies that two distinct speech centers, or occasions of speaking, are at issue. Finally, noun phrases in the matrix clause identify the participants of the reported event, in this case by name, as persons distinct from the participants of the reporting speech event.

In other words, direct reported speech is a highly transparent voicing structure because all four conditions in (a) through (d), Table 2, are met: Metrical contrasts individuate distinct textual zones, configured into discrete clauses-propositions, with independent indexical origos or speech centers, with framing clause units identifying voiced participants.

As we move away from the highly transparent case of direct reported speech, the contrast between reporting and reported voices becomes problematic due to the absence of specification of one or more of the cues listed in Table 2, yielding various types of elision of frame boundaries and blurring of dependent effects at the level of notional voicing structure. This yields cases such as indirect reports, free direct speech, and free indirect speech (columns II-IV, Table 2).

Thus if we compare the preceding direct report to the corresponding indirect report,

John$_i$ promised$_j$ Alice that he$_i$'d go to the bank for her$_j$

we find that subordinate clause deictics are no longer independent but are anchored to the matrix clause by cross-clausal anaphora (cf. *he$_i$/her$_j$* rather than *I$_i$/you$_j$*) and sequence of tense rules (yielding subjunctive *-'d* rather than future *-'ll*). The notional independence of voices is compromised as well: The subordinate clause is no longer understood as another's wording.

In the case of free direct speech (column III), no distinct framing clause occurs at all, and framed/framing relationships are distinguished only through metrical contrast (independence) of deictic origos across textual zones. We have seen an example of free direct speech in Table 1, namely segment$_2$, which occurs in metrical apposition to segment$_1$ but is neither syntactically linked to, nor described in, segment$_1$. Yet the two segments are quite distinct in deictic patterning: Segment$_2$ consists of tenseless speaker-indexing exclamatives and segment$_1$ of past-tense, third-person statements. Hence, although a voicing contrast is indeed discernible, the absence of a clausal frame entails that the voices of segment$_2$ cannot be identified in any locally

Table 2
Voicing contrasts in represented speech.

	Representing voice		Represented voice	
	I. direct report	II. indirect report	III. free direct speech	IV. free indirect speech/thought
(a) metrically contrastive text segments	+	+	+	+
(b) segments linked/demarcated by clause boundary	+	+	–	–
(c) segments differ in deictic/ indexical origos	+	–	+	–
(d) matrix clause NPs denote voiced participants	+	+	–	–
Transparency of voicing contrast	HIGH ←		→	LOW

explicit framework of named biographic identities; such identity may well be inferable from other co-textual cues but not necessarily uniquely (see n. 3).

In the case of free indirect speech (column IV), shifts in deictic origo do not map neatly onto clause boundaries (see Voloshinov 1973, Banfield 1982, and Lee 1997 for examples); hence a single clause may have multiple origos in this written, literary style, a feature generally absent in oral speech. Here, the internal fractionation of the speech center entails that a distinct centering frame must be supplied for a retelling to occur. In the tradition of metacommentary on literature called "literary criticism," the descriptive framework usually supplied is one of mental states, not speech events (viz., a merger of "subjectivities," a stream of "consciousness," an "omniscient" narrator, etc.). In such cases, discursive effects within the novel are enregistered as psychological-mental states by literary critics and other readers. But this is just a particular genre of metasemiotic construal for a form of entextualization that lacks a differentiating framework of person deixis capable of imposing an unambiguous structure of biographic identities onto facts of textually zoned, metrically individuated voicing contrasts.

Individuation, Naming, and Characterization of Voices

The preceding considerations show that voices are not attributes of persons but entextualized figures of personhood whose recognition depends on distinct metasemiotic processes (Table 3).

The first of these is the *contrastive individuation* of one voice against another, or the delineation of a voicing contrast, by a text-metricalized formulation of juxtaposed figures. This effect can be diagrammed by any metrical or poetic organization of text that delineates contrastive textual zones as unlike each other and where such likeness/unlikeness of text segments motivates the construal of likeness/unlikeness of the default variable of co(n)text, namely speaker. In the previous examples the contrastive individuation of voices depends on patterning of referential indexicals (deictics); however, such individuation can be achieved by contrasts of social indexicals (registers) as well, as I show in the next section.

The second process is the *biographic identification* (e.g., naming) of voices individuated by the first process. This strength and clarity of identification depends on the number and types of metapragmatic cues that map contrastive textual scopes onto biographical identities. I have illustrated this point through examples involving per-

son deixis (e.g., proper names, pronouns, reported speech frames). Thus when proper names can be used to associate textual scopes with biographic identities, we have a framework for construing entextualized voices as the speech of biographical persons (real or fictional).

In some cases, as with the preceding unnamed voices, textual individuation is robust but identity ascription is not. Bakhtin observes that, in the novel, contrasts among a vast range of text-forming devices—parentheticals, tense, person, mood, report frames of varying degrees of fragmentariness—draw implicit text-internal boundaries that cannot always be mapped onto biographical identities in a clear way but are nonetheless critical to the novel's dramaturgical work. This is not just a feature of novels, however, but of any narrative. Jane Hill (1995) has shown that a similar range of voicing contrasts is detectable in everyday oral narratives, where the text-internal organization of a single person's speech contains many individuable voices, each linked to describable textual scopes but not always to named biographical identities.

In other cases, both textual individuation and identity ascription are possible, but their results do not converge. Bakhtin uses the term *character zones* (Bakhtin 1981:316) for stretches of text in which a character's speech is particularized (i.e., differentiated from its co-textual surround) through specific locutions, styles, idioms, and the like. The character zones of a novel are often wider (i.e., involve larger text segments) than those assigned to a named character through a system of proper names and reported speech deixis. Indeed, a character's textually implicit voice may overwhelm surrounding framing material and thus interpenetrate the named voices of other characters, or of the author, even though the identities of named persons now at issue and the characteristics of their speech may be clear and unambiguous elsewhere in the text. In such cases the two frameworks for delineating voicing in text-one wholly implicit, the other sometimes explicit-fail to converge.

We can think of the forms of voicing construal discussed previously—and summarized in Table 3, (a) through (c)—as a series of construals that may be carried out in isolation, or conjointly, as a series of steps. Thus in the case of unnamed voices, it is possible to carry out step (a) but not (b). In the case of character zones, both steps (a) and (b) are possible but diverge in textual scope (do not involve precisely the same text segments). In the case of free indirect speech, more than one biographic identity is assigned to a textual scope, and the resulting figure construed as a "double-voiced thought" or "merged subjectivity."

Although the foregoing examples require only a two-way distinction, namely (a) versus (b), a third type, (c), is critical to the discussion that follows. We have already observed its relevance in an implicit way. We saw, for example, that even when step (b) is not possible, step (c) frequently is. Thus the "unnamed voices" in Table 1 do not permit clear biographic identification; this does not prevent us, however, from using some other description, such as "a parody" for $segment_1$, the voice of "onlookers" or

Table 3
Segmentation and typification of voices.

(a) Contrastive individuation:	Recognizing a voicing contrast, e.g., recognizing that metrical contrasts among text segments imply a *difference* of speaker
(b) Biographic identification:	Typifying an individuable voice as the speech of a biographic person, e.g., using a system of person deixis to link text segments to biographic identities
(c) Social characterization:	Assigning an individuable voice a social character, e.g., using a metalanguage of social types to describe text segments

of "crass person(s)" for segment$_2$, or the voice of "irony" or "unmasking" for segment$_3$. Such descriptions employ a metalanguage of social types—whether types of interactant (onlooker), of persona/stance/attitude (parody, irony), or of social kind of person (crass, vulgar)—in typifying voices individuated by text-metrical contrasts.

We saw previously that Bakhtin's "individual voices" are textually individuated discursive figures that are typified through a system of person deixis as biographic individuals of some kind. His "social voices" are textually individuated figures that are recognized or typified through social-characterological descriptions. But recognized by whom? In some cases such entextualized effects may be recognized by interactants as "social voices" unique to that occasion; here the social domain of recognition is simply current participants.

But the cases to which I now turn are more restrictive. These are cases where a repertoire of speech forms is widely recognized or enregistered as indexing the same "social voice" by many language users. In such cases we have a social regularity of typification—a system of metapragmatic stereotypes—whereby a given form, or repertoire of forms, is regularly treated as indexical of a social type by a given social domain of persons. In some among these cases the process of *social characterization* (Table 3, c) operates more specifically in terms of categories of social-demographic classification. These are the cases traditionally called *registers.*

Enregistered Voices

Encounters with registers are encounters with characterological figures stereotypically linked to speech repertoires (and associated signs) by a population of users. Language users typify such figures in social-characterological terms when they say that a particular form of speech marks the speaker as masculine or feminine, as high- or low-caste, as a lawyer, doctor, priest, shaman, and so on.

Some examples are given in Tables 4 through 7. In all four cases distinct speech repertoires are treated as indexing different social-characterological types of speaker (small caps). Tables 4 and 5 illustrate registers of speaker gender in two Native American languages. In both cases, column A forms are understood as stereotypically FEMALE; those in column B, MALE. Tables 6 and 7 illustrate two occupational registers of English, a register of MILITARY discourse in 6 and the case of SPORTS ANNOUNCER register in 7. The actual form alternations are quite different in the four cases;[4] similarly the characterological figures associated with speech are also quite distinct (viz., speaker's gender in Tables 4 and 5, speaker's profession in Tables 6 and 7).

Table 4 Registers of speaker gender (Koasati).

(a) Repertoire contrasts

Gloss	*A*	*B*
'he will lift it'	lakawwą	lakawwáˑs
'I am lifting it'	lakawwíl	lakawwís
'Don't lift it'	lakawčin	lakawčiˑs
'he is lifting it'	lakáw	lakáws
	↓	↓
(b) metapragmatic stereotypes:	FEMALE	MALE

(Source: Haas 1964)

Table 5 Registers of speaker gender (Lakhota).

(a) Repertoire contrasts

Illocutionary force	*A*	*B*
Formal questions	huᵃwe	huᵃwo
Familiar command	nitʰo	yetʰo
Opinion/emphasis	yele, ye	yelo
Emphatic statement	kßto	kßt
	↓	↓
(b) metapragmatic stereotypes:	FEMALE	MALE

(Source: Trechter 1995)

Table 6
Pentagon military register.

Pentagon lexicon ('Militarese')	Standard English
aerodynamic personnel decelerator	parachute
frame-supported tension structure	tent
personal preservation flotation device	life jacket
interlocking slide fastener	zipper
vertically deployed anti-personnel device	bomb
portable handheld communications inscriber	pencil
manually powered fastener-driving impact device	hammer

(Source: Lutz 1990)

Table 7
Register of sports announcer talk in English.

a) Omission of sentence-initial deictics (e.g., anaphors, determiners) and present-tense copula: e.g., *[It's a] pitch to uh Winfield. [It's] a strike. [It's] one and one*

b) Preposed location & motion predicates: e.g., <u>Over</u> <u>at</u> <u>third</u> is Murphy. <u>Coming</u> <u>left</u> <u>again</u> is Diamond

c) Preponderance of result expressions: e.g., *He throws <u>for</u> the out.*

d) Epithets and heavy modifiers: e.g., <u>left-handed</u> <u>throwing</u> <u>Steve</u> <u>Howe</u>

e) Use of the simple present to describe contemporaneous activities: e.g., *Burt ready, <u>comes</u> to Winfield and it'<u>s</u> lined to left but Baker'<u>s</u> there and <u>backhands</u> a sinker then <u>throws</u> it to Lopez*

(Source: Ferguson 1983)

Such registers are reflexive models of the effects of speaking. They are differentiable as discursive formations within a language only as a function of the fact that they are so differentiated by language users. Their identifiability by linguists relies on the metapragmatic ability of language users to discriminate forms across register boundaries and assign pragmatic values to variant forms. The data used to identify a register's repertoires are, at the same time, data providing some indication of stereotypic figures associated with use.

The unit data point on which register identification depends is an act of metapragmatic typification by a language user, whether the act be descriptively explicit or implicit, naturally occurring or elicited, articulated discursively or through other semiotic media.[5] But any such register is a social regularity: A single individual's metapragmatic activity does not suffice to establish the social existence of the register unless confirmed in some way by the evaluative activities of others. The data of socially recurrent typifications amount to an order of *metapragmatic stereotypes*—folk models of indexical value—associated with a repertoire of forms.

Any such register is a model of language use that links a semiotic repertoire of some describable characteristics (Table 8, A) to a range of stereotypic social-indexical effects, its social range (Table 8, B). Such a model is inevitably a model *for someone*; that is, it involves a social domain of persons who recognize it as a model enactable through speech (Table 8, C).

Registers have a social existence only insofar as—and as long as—the metapragmatic stereotypes associated with their repertoires continue to be recognized by a criterial population of users, that is, continue to have a social domain. But any social collectivity—anything that we might call "a society" or "a subgroup (within a society)"—is continuously changing in demographic composition due to many

Table 8
Three aspects of register organization: Repertoires, Social Range and Social Domain.

A. characteristics of **repertoires**:
  - *Repertoire size*: number of forms;
  - *Grammatical range*: number of form-classes in which forms occur;
  - *Semiotic range*: types of linguistic & non-linguistic signs that appropriately co-occur in use

B. stereotypes of indexical effectiveness, typically exhibiting a **social range**
  - *Stereotypes* of speaker/actor kind; of enactable relationship (e.g., deference, intimacy); of appropriateness to specific social occasions and scenarios of use

C. **Social domain**(s) of user
  - categories of persons that can recognize (at least some of) the register's forms/indexical effects
  - categories of persons fully competent in the use of the register

processes, such as births, deaths, and migrations. Thus, registers exist continuously in time only as a function of communicative processes that disseminate awareness of and competence in such registers to *changing* populations. Institutional processes of various kinds frequently seek to stabilize features of registers—their repertoires, indexical stereotypes, social domain of users—by codifying their normative values or restricting access to them; yet registers frequently change in their defining features through communicative activities that mediate their social existence (Agha 2003).

I focus, in the remainder of this article, on a claim I made earlier: We cannot understand macro-level changes in registers without attending to micro-level processes of register use in interaction. One of the key features of everyday use is that effects of register token use are not always consistent with the stereotypic values associated with the register's form types. This flies in the face of a common folk theory about registers, a kind of folk assumption of contextual invariance, typically subscribed to by language users and often adopted uncritically by linguists as well. Taken very strictly, this view implies that the construable context, or co-text, of any particular token use is always irrelevant to the overall construal of that use. Let us consider this issue in more detail.

Congruence of Voicing Effects: Tropic and Appropriate Use

The assumption of contextual invariance is false for the simple reason that enregistered voices are encountered in social life only as fragments of entextualized voicing effects, and the two voicing effects may or may not be congruent. In actual events of language use, enregistered voices may exhibit various types of congruence/noncongruence with more implicit nonce images of personhood (entextualized voices) that are less easily reportable out of context. Let's look at some examples.

In the Lakhota example in Figure 1, a form token of the female register is uttered by a man who unexpectedly sees his two-year-old nephew at his house one evening. The linguistic utterance is reproduced in the box. Notice that the utterance ends with the female speech form *wele* (in boldface), but its co-textual frame contains a token of male speech, the form *wąlewą* (italics). Moreover the nonlinguistic context of utterance, its visible semiotic surround, also makes clear that the one speaking is a man (and the one spoken to, a child). These linguistic and nonlinguistic signs comprise a *multichannel co-text* for the female speech token, *wele*; and, in the case at hand, this multichannel co-text specifies that the speaker is male. Thus the boldface text segment is indexically noncongruent from its co-text.
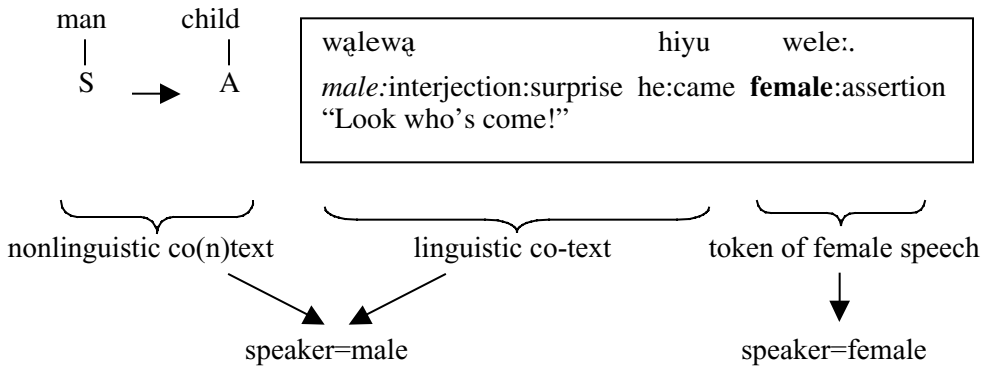
man   child
| |
S   →   A

ważewą   hiyu   wele:.
*male:*interjection:surprise he:came **female**:assertion
"Look who's come!"

nonlinguistic co(n)text   linguistic co-text   token of female speech

speaker=male   speaker=female

Figure 1
Non-congruence of voicing effects within a speaking turn: Gender tropes in Lakhota.

The man's use of female speech is tantamount to an interactional trope, the performance of an affective, caring persona often associated with women speaking to young children. But the trope of maternal concern is recoverable only to someone who attends to the multichannel sign configuration of which the token of female speech is a fragment; the entextualized construal (that speaker is a maternal, affective male) vanishes if the female token is decontextualized from the co-textual frame that motivates the construal. Hence, in tropic uses of this kind, there are sharp differences of reportability between enregistered and entextualized voices. The enregistered voice associated with the form *wele*—namely that *speaker is female*—is highly detachable from context and reportable by any native speaker acquainted with female register; it is a commonplace, easily reportable stereotype about the form. However, the entextualized voicing effect—that *male speaker is maternal, affective, et cetera*—is contrastively recoverable only by someone who has access to the larger entextualized structure (viz., that the one speaking is a man, that he is speaking to a two-year-old, that the child has turned up unexpectedly, etc.), which motivates the "maternal, affective" construal.

Now, the general point that *enregistered voices are always and only experienced in the course of entextualized voicing effects* is no less important in the case of appropriate use. I have just illustrated this point for the case of tropes of voicing—namely cases where entextualized voicing is manifestly noncongruent with enregistered voices—by citing a case where the one using "female speech" is co-textually identifiable as a man. Yet the issue is equally important (though less foregrounded) in the case of appropriate use. The term *appropriate use* never describes a token-level phenomenon; it is a name for a token-to-text relationship. For in the absence of an evaluation that links a register token to surrounding or entextualized semiotic effects—for example, without evaluating whether the one using the male forms is everywhere, in every co-textual semiotic respect, a "man"—we can never evaluate the usage as appropriate in any meaningful sense of the term. We call a register's usage "appropriate to context" when co-occurring signs are congruent with, or satisfy, the model of context indexed by the register token. We perceive a usage as tropic when co-occurring signs have noncongruent indexical effects.[6]

The two cases are illustrated diagrammatically in Figure 2. In (a) we have the case where the effect indexed by a register token is congruent with effects projected from co-text; the resulting composite sketch is thus indistinguishable from its component elements. In the tropic case in (b) the component effects are distinct from each other and the composite sketch is different from both, yielding a kind of superimposed figure. We saw an example of type (b) in the case discussed in Figure 1, where the effect of the register token (speaker is female) is noncongruent to the effect of its
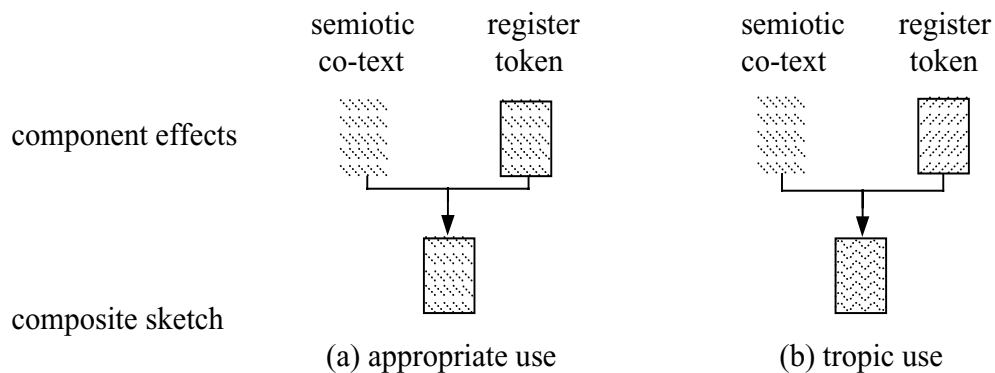
Figure 2
Congruence vs. non-congruence of co-occurring indexical effects.

semiotic surround (speaker is male), yielding a composite figure different from both (male speaker is female-like, i.e., maternal, affective, etc.).

Now, anyone acquainted with a register can employ it in acts of strategic manipulation of roles and identities and achieve effects that, although *dependent on stereotypic values* of text segments, are *significantly at odds with such values* at the level of text configurations. In all such cases, the stereotypic values of a register's forms may function canonically in discourse—that is, make criterial personae recognizable through speech—but the co-occurrence of framing devices formulates co-textually superposed personae that differ from those indexed by local text fragments. In such cases, entextualized voices differ from, yet dominate, enregistered ones.

So far I have considered the issue of the congruence and noncongruence of voicing effects within single speaking turns. But the same logic is of critical importance in stretches of discourse that involve many interactional turns.

Congruence across Interactional Turns: Role Alignment

Any semiotic activity that implements a voicing effect is always subject to uptake and potential ratification in a subsequent semiotic act that may itself index features of speaker persona and, to this extent, may itself implement a voicing effect. Any two voicing effects linked together in a *speech chain* of this kind (Agha 2003) can themselves be compared by criteria of congruence or lack thereof. I refer to patterns of congruence/noncongruence of voicing effects across interactional turns as *patterns of role alignment.*

The data in Table 9 illustrate an interaction between two young boys who perform a series of persona displays in a turn-by-turn engagement, each turn segment of which implements a trope of voicing through the use of sports announcer talk. Neither child is a sports announcer by profession, of course. Yet both have the passing acquaintance with the register possessed by anyone exposed to radio and television broadcasts. In the course of a game of ping-pong the two boys, Ben and Josh, switch to the register of sports announcer talk in a spontaneous manner, each using last names rather than pronouns to formulate each other as sports figures and several among the features noted in Table 7 to inhabit the persona and mantle of a sports announcer.

When problems arise within the ping-pong game itself (e.g., scorekeeping disputes, arguments about the rules, external events that interfere with the game), the boys switch back to everyday speech, thus abandoning the sportscaster persona in favor of these more pressing concerns (see Hoyle 1993 for further details). Hence the switching back and forth between sportscasting and everyday registers corresponds to a switching between imaginary and real identities keyed to specific interpersonal ends within this complex bout of play. But in stretches of talk where the sports an-

Table 9
Roles alignment across speaking turns: Tropes of sports announcer speech in English.

Context: The two participants are young boys (an eight- and a nine-year-old), who describe their own game-playing activities in sports announcer speech, each seeking in a turn-by-turn engagement to reframe "what (just) happened" in a voice more authoritative than his own.

Josh:	So eleven eight, Hoyle's lead.
	Hoyle serves it!
	Ben Green cannot get it...over the net
	and it's twelve eight Hoyle's lead now.
Ben:	Hoyle takes the lead by four.
Josh:	[fast] Green serving.
	[fast] Hoyle returns it.
	THEY'RE HITTING IT BACK AND FORTH!
Ben:	Ach-boo:m!
Josh:	And Ben Green hits it over the table!
	And it i:s thirteen eight.
	Hoyle's lead.

(Source: Hoyle 1993)

nouncer register is used, as in the preceding excerpt, there is clearly a second-order game going on—quite distinct from the ping-pong itself!—a game played entirely through talk, whose object is to control representations of the first-order game in personae more authoritative than the boys' own.

In the data shown in Table 9, each turn segment exhibits an interactional trope—evident from turn-internal noncongruence of performed and presupposed identities, akin to the case in Figure 1—to which the interlocutor responds with a comparable trope in a subsequent turn, the two turn segments thus exhibiting a *symmetric* form of role alignment.

Notice also that through a particular artifactualized speech chain, which began through the recording, transcription, and analysis of this interaction by Hoyle (1993) and continues in my own discussion of her work here, this particular semiotic engagement between Ben and Josh reaches you, gentle reader, allowing you to take a particular stance on this register, on its actual uses by Ben and Josh, and perhaps on other, similar uses of this kind that you may have witnessed—or even performed!—in the past and may yet witness again.

More generally, it should be evident that when occasions of register use are represented through further metadiscursive activity, such representations map one participant framework onto another, linking them through a speech chain. I observed at the beginning of this article that the phenomena of "voicing" and "dialogicality" are not confined to events of dyadic conversation. It should now be clear that they are not confined to isolable "events" at all; such relations potentially hold across speech chains as well. Indeed, any mechanism of speech representation that is capable of representing speech events within speech events is, at the same time, capable of bringing images of narrated participants into dialogic interaction with images of current participants (including their self-images). All of this suggests that processes of role alignment play a part in macro-level processes of discourse circulation as well.

Patterns of Role Alignment in Public Sphere Discourses

I have argued elsewhere that register competence is disseminated across social populations by institutionalized practices of speech typification. Such practices bring into circulation images of social personhood linked to speech through the circulation of discursive artifacts—I discuss examples involving oral narratives, printed cartoons, newspapers, magazines, novels, et cetera in Agha 2003—that connect persons

together in communicative interactions. The practices themselves are quite diverse (see Agha 2002, 2004), but a few common examples will suffice to illustrate the larger point.

It is well known that prescriptive socialization within the family plays a critical role in the dissemination and early acquisition of many registers. By communicating register distinctions to children, such metapragmatic activity expands the social domain of register competence, one child at a time, from one generation to the next within the family unit. But processes of register socialization continue through adult life as well. One cannot become a doctor or a lawyer, for example, without acquiring the forms of speech appropriate to the practices of medicine or law or without an understanding of the values—both cognitive and interactional ones—linked to their use. In these cases, the process of language socialization typically involves a cohort whose members acquire competence in the use of profession-specific registers of a language through extended affiliation with an educational institution, such as a law school or medical school. In societies with written scripts and mass literacy, a variety of normative public institutions—such as educational institutions, traditions of lexicography and grammatology, school boards, and national academies—serve as loci of public sphere legitimation and replication of register stereotypes over segments of the population. The effect is particularly marked for prestige registers such as the Standard Language. When effective, such methods may result in the growth or rise of a register formation by extending a more or less uniform competence in its use over relatively large segments of the population.

Despite their evident diversity, such large-scale practices share certain common features. They unfold one semiotic event at a time, each event involving persons in interactional roles of sender and receiver of messages; the messages conveyed or exchanged in these events typify speech forms and their pragmatic values, though with varying degrees of explicitness. For receivers of such messages, the voicing structure of the message constitutes a set of directions for locating one's own speech in relation to those of others.

In the special case of prescriptivist discourses, such metapragmatic directions are often formulated as explicit "instructions" in the narrow, most literal sense of the term. For instance, the following discussion of uses of the pronoun *tu* (from a French etiquette guide of 1937) consists of a series of metapragmatic descriptions of speech events, some formulated as commonplace, others as possible, hypothetical, to-be-avoided, or to-be-desired.

> In general . . . the use of *tu* between two people indicates their extreme familiarity. . . . In friendship . . . to use *tu* and let oneself be called *tu* is to open the private doors to one's heart and personality. Now, in a man or woman who is scrupulous [*délicat*] or well-mannered, these doors should be very hard to open and thus should open very rarely, only for a lucky few . . .
>
> So many people who have never seen or taken the time to know each other call each other by their first names after a first "car ride" together! After the second, they call each other *tu*! . . . In this way, they deprive others-who until then enjoyed a *tutoiement* [use of *tu*] that they thought was justified by a very long, sincere relationship—of the attraction of this distinction. Indeed, for a sensitive person, what value can there be in the handsome proof of deep friendship which *tutoiement* is, if one can obtain it so easily?
>
> If you want to be seen as well-mannered, then be "very hard to address as *tu*." [Saunier 1937; cited by Morford 1997:9-10]

Over the course of this account, the text reanalyzes the pronoun's first-order indexical values, that it indexes "familiarity" with *interlocutor,* into second-order values of *speaker* indexicality by describing two characterological figures, two types of *tu* users: These are formulated through descriptions of when, how often, and with whom a speaker exhibits "familiarity." The first of these is the person who recognizes that *tu* usage entails "open[ing] the private doors to one's heart and personality" and permits such intimacy only in the rarest circumstances; this is the traditional figure of "scrupulous or well-mannered" reserve. The other is a manifestly modern figure, who switches to first name after the first "car ride" (itself a modern image in 1937!) and to *tu* after the second; this is

52

a contrastive figure, given to a lifestyle that is carefree, easygoing, without restraint. In the last line, the reader is given explicit instructions on how to become the first type (how "to be seen as well-mannered"), namely, to exhibit reserve preemptively, so as to prevent others from initiating *tu* and thus to avoid having to reciprocate.

The reader of this text is thus given a contrastive paradigm of two personae, each linked to a distinct pattern of pronominal usage and associated behaviors. What is distinctive about prescriptivist discourses generally is that they encourage or enjoin readers/hearers to undertake a specific course of conduct. But the formulation of characterological figures linked to speech is a much more general kind of metasemiotic work, commonplace even in genres of public sphere discourse that lack any prescriptivist intent.

A particularly important source of such folk stereotypes in modern societies is the circulation of representations of speech and speakers in the mass media. Here, the depiction of language use by socially identifiable persons is itself a form of metasemiotic representation, which members of mass audiences construe for its characterological content and to which they respond in a variety of ways.

A well-known example of this is the role of British mainstream newspapers and BBC broadcasters in disseminating images of Standard British Received Pronunciation. Both Standard and non-Standard accents are objects of everyday commentary in public sphere media in Britain, as the following quote indicates (see Agha 2003 for more details):

> In our serious newspapers political columnists and other journalists regularly pass comment on the accents of public figures, while television critics discuss the accents of actors, programme presenters, and other television personalities. The correspondence columns of both national and local newspapers frequently carry letters from readers commenting on various forms of accent—favourably, or, more often, unfavourably—and when the BBC uses people with marked regional accents to present radio programmes or to read the news, waves of protest are expressed in letters of complaint to the BBC and sacks of hate-mail to the presenters themselves. . . . Writers of contemporary novels and memoirs use observations about accent as a crucial part of the description of character. . . . Most of the characters in Anthony Burgess's recent memoirs are introduced with reference to their accent. [Honey 1989:10]

Public sphere discourses thus comprise a vast domain of characterological work, though different genres of such discourse differ in their degree of explicitness. In the newspaper and television examples discussed previously, various social-characterological figures are explicitly linked to accents through *descriptions* of the speech and personality of public figures. In the case of novels, accent is used to portray fictional characters; here social types are formulated through *contrastive depictions* of accent in reported speech frames, often with little or no independent description of personality characteristics.

Audiences respond to these representations by performing various types of alignment to specific character types. In some cases, implicit typifications are rendered more explicit through uptake and response in subsequent speech events: In the case of the BBC announcers who speak with regional accents, particular social personae are only implicitly palpable in the announcers' performances; however, in their subsequent letters of complaint and hate mail, the audiences of these broadcasts describe such enacted personae in highly explicit, sometimes vituperative terms. In all of these cases, the responders themselves exhibit a form of role alignment—whether symmetric or asymmetric, whether expressing praise or contempt—vis-à-vis the figures to which they respond.

In some cases, media audiences report patterns of *self-reported usage* that are modeled on widely circulating *representations of others' usage.* The case of pronominal registers in Egyptian Arabic (Alrabaa 1985) is particularly interesting from this point of view (Table 10). Alrabaa's study is a questionnaire-based investigation of metapragmatic stereotypes of use associated with informal-solidary and polite-formal pronouns, not a study of actual use.

Alrabaa observes that upper-class and working-class youths appear to self-report patterns of usage that are mirror images of each other. Upper-class youths claim to

Table 10
Inverse icons in Egyptian Arabic: Reciprocal, mirror-image alignments between two groups, each claiming to use pronouns they regard as the other's usage.

	Group₁: Upper-class youths	**Group₂**: Lower-class youths
Stereotype of self-report	Claim to use solidary *inta/inti* forms	Claim greater use of the formal *ḥaḍritak/ḥaḍritik* pronouns
Stereotype of others' usage	Say that lower-class speakers use the *inta/inti* forms	Say that upper/middle-class speakers use the *ḥaḍritak/ḥaḍritik* forms
Ideological positioning	Egalitarian (self-lowering)	Stratificational (self-raising)

(Source: Alrabaa 1985)

use the solidary-informal forms *inta/inti* (you [m./f.]), which they believe lower-class speakers to use, and lower-class speakers lay claim to the more polite-formal lexemes *ḥaḍritak/ḥaḍritik* (you [m./f.]; polite), which they perceive as upper/middle-class usage.

Upper-class youth appear to get their images of lower-class speech through the mass media: "In off-the-record comments during our interviews, both older and younger upper-class informants did often express a conviction that lower-class informants would be 'looser,' less formal, etc. This upper-class belief is also reflected in many movies and television comedies, which frequently present a stereotype of the bawdy, raucous lower-class character who addresses all listeners as *inta/inti* = [German] *Du*, [French] *tu*" (Alrabaa 1985:648).

In interviews, upper-class youth describe themselves as adopting what they perceive to be "the system of 'the people' (*al-sha'b*)," thus professing an egalitarian impulse, whereas lower-class youth are drawn toward "what they presume to be the middle-class values" (Alrabaa 1985:649), thus exhibiting a more stratificational ideology. Each group ideologically professes an alignment to the stereotypic voice of the other, the two together exhibiting an ironic form of role reversal. Alrabaa judges the egalitarian pattern of upper-class youth to be the institutionally more dominant pattern and thus the likelier pattern of overall change. I return to the role of institutions in simplifying patterns of enregisterment in my concluding discussion.

Aspects of Role Alignment

In earlier discussion, I defined role alignments as patterns of congruence/non-congruence across interactional turns among semiotic behaviors expressing voicing effects. To speak of "alignments" here is to speak of patterns of relative behavior; to speak of "role" alignments is to focus on the expression of voices and figures in the behaviors in question.

Like voices, the phenomena of role alignment are effects formulated through patterns of discursive and other semiotic behaviors; both are attributes of biographic persons only in a derivative sense. The special cases where voices and alignment are attributable to individuals—the case where entextualized voices are uniquely associable with biographic persons, or the case where some role alignment is identifiable as the causal result of an individual's conscious, strategic choices—do, of course, occur in everyday life, and, indeed, they are commonplace. But the variety of patterns of voicing and role alignment discernible (or "individuable," Table 3) in the discursive activities of persons far exceeds the variety describable ("identifiable" or "characterizable," Table 3) through simple labels in our everyday metapragmatic terminology for roles and relationships (see Agha in press for further discussion of this issue).

Role Alignment as Generalized Footing

Goffman uses the term *footing* for a special type of role alignment, the case where alignments emerge between persons copresent in participant roles linked to spoken utterances: "A change of footing implies a change in the alignment we take up to ourselves and the others present as expressed in the way we manage the production or reception of an utterance" (Goffman 1981:128). Yet Goffman is aware that relations of footing are neither limited to spoken utterances nor to relations among copresent interlocutors but extend to any sign-mediated interaction, or "coordinated task activity," some among which contain "no speech event at all" (1981:144).

Once we see that interpersonal alignments can be negotiated through behavioral responses to any semiotic display, and once we note that technologies for fashioning semiotic artifacts and bringing them into circulation can connect persons to each other across greater spatiotemporal removes, it is readily seen that the purview of role alignments extends well beyond semiotic interactions of the face-to-face, conversational kind. I have suggested that even large-scale social practices can be understood in terms of patterns of role alignment, where the things compared in order to discern the pattern are, again, performed arrays of signs, frequently multichannel texts.

I have suggested also that role alignments mediate more than participation frameworks. When Goffman talks of footing, he is mainly interested in a particular kind of role, namely a "participant role"; he shows that our folk metalanguage of speaker, hearer, and so forth, is inadequate for characterizing such roles, since a much larger variety of role projections is individuated through processes of entextualization in everyday talk. Yet Goffman variously suggests that a much larger range of phenomena—such as performed interpersonal stances, attitudes, forms of irony and parody, relations of respect and formality, and other, more vividly "social" effects—can also be analyzed in the same terms. This, then, is a much larger class. But larger in what respect?

Recognizability of Role Distinctions

In all cases, the *individuation* of facts of role alignment requires the metrical comparability of indexical cues in stretches of semiotic text (i.e., voicing zones, in the sense discussed previously) across interactional turns, and their *identification* and *characterization* require a denotational metalanguage for talking about the unit kinds thus individuated. Thus on Goffman's account the role speaker is decomposable into a number of role fractions; these are individuable only by attention to entextualized indexical cues that contrastively differentiate, or fail to differentiate, the one physically producing the message from the one responsible for its content, from the one responsible for its wording, from the one depicted in the message, and so on, and his labels—animator, principal, author, figure—comprise a metalanguage for characterizing the participant roles thus individuated.[7] Now the same point applies to displays and descriptions of any other unit kind of speaker-actor role, that is, a social persona, character, a stance toward another, a social-demographic identity, and so on.

Yet our folk intuitions tell us that participant roles are completely different in kind from social roles. Why does this seem so? Part of the reason is that when we identify and characterize discursive figures we rely on everyday terminologies (and associated folk ontologies) that make participant roles and social roles appear significantly different in culture-internal terms. But the process of the individuation of difference is the same for all; here, they do not differ. What I wish to suggest, in other words, is that across this larger range of cases we are engaged in similar semiotic evaluations at *the tier of entextualized individuation* (Table 3, a) but are socialized to different habits at *the tier of descriptive identification and characterization* (Table 3, b and c). At the first tier, we are concerned with metrical differences among performed entextualized figures, or voicing contrasts, indexically articulated through arrays of signs in discursive interaction. At the second tier, however, we are concerned with describing such effects, with fixing or stabilizing them through a classification; we thus impose some further structure on the phenomenon at hand.

When we encounter others in interaction we are concerned with both tiers, not just the latter one; yet the latter is more transparent to subsequent reportability in folk consciousness. In everyday talk, we do not normally describe the first tier at all (i.e., we do not ordinarily ask, "How does the voicing structure of this utterance-fraction compare with that one's?"). Our everyday habits of talk about these experiences are limited to a commonplace descriptive lexicon, sometimes reorganized into higher-order taxa by scientific or other institutional codifications. Thus we may observe that one text fraction semiotically conveys agreement/disagreement with or sympathy/antagonism to another; or switch to quasi-technical hypernyms and speak of interpersonal "stance" or "affect." Or we may observe that the second voice/figure appears more refined, elegant, prudent, or wise than the first and group these differences under psychosocial categories, such as "character" or "personality." Or observe that the second figure is younger, or lower-class, or male rather than female and group these matters under social-demographic rubrics, such as "social status" or "social identity." But our ability to offer *any* such social characterizations (whether through everyday or quasi-technical terminologies) presupposes that *differences* among voices or figures are contrastively perceivable in the first place. Thus metrical processes of entextualized individuation both *underlie* the identifiability and characterizability of roles and are *less transparent* than them in subsequent report and discussion.

Self-Descriptions

Like the notion of footing, the generalized notion of role alignment does not seek to explain self-descriptions. Take the case of legal register. I argued earlier that the law school classroom is an institutionalized site of socialization to legal register. This suggests that students who acquire the register are performing a kind of role alignment with the characterological figures linked to the legal register; that when, over the course of some period of socialization, a law school student acquires some proficiency in legal register, the student has learned to align his or her self-image with the characterological figures of legal register.

Such an account is, of course, wildly at odds with any self-description that a law school student might volunteer as an account of conscious, strategic choices. Thus a person may consciously intend to go to law school to acquire wealth and power, to serve civil rights causes, or for some other reason; he or she may never attend focally to questions of register acquisition. Yet the capacity of a lawyer to acquire wealth and power (or to serve civil rights causes, or to pursue whatever ends he or she has in learning the law) nonetheless depends on the acquisition of the register. It depends on entitlements acquired through acquiring the register. The register is itself a form of semiotic capital that advances certain rights and privileges. And to be able to speak the register is to be able to perform an image of social personhood as one's own image and to perform it in a register-dependent way. Thus the notion of role alignment here describes the acquisition of register competence in empirically consequential ways, even though the process may not be transparent at all times to all participants caught up in the process itself.

In other cases, common self-descriptions may be partly or even entirely correct. We have already seen a case of partial correctness in the aforementioned Egyptian example. Alrabaa (1985) found that in interviews, the statistically most common self-description by upper-class informants is the claim that the speaker is aligning with lower-class usage. But socioeconomic class is not the only factor relevant here; for, among upper-class speakers, both older and younger informants are aware of stereotypes of lower-class speech, but only younger upper-class informants align their own self-images with lower-class stereotypes in the mid-1980s milieu on which Alrabaa is reporting. For them, the ideological stance is one of egalitarianism vis-à-vis those perceived as lower-class; yet the stance of these *younger* upper-class speakers is also one of generational differentiation within the ranks of the upper-class itself, that is,

younger versus older. The two stances are mutually consistent here—they are empirically inhabitable through a single strategy—though the egalitarian impulse is the more widely reported by upper-class youth in their interview responses. Such multiplicity of role alignments simply reflects the multiplicity of engagements with real or imagined others that is characteristic of social life. The more common ideological stance may simplify or distort what it describes, in one sense; yet, in general, its greatest social importance may well lie not in its degree of correctness but in its *efficacy*, its capacity to bring more and more of the group's future discursive history into conformity with itself.

Conclusions

I have argued that registers have a dynamic social life (i.e., can change in social domain, range, or repertoires) mediated by metadiscursive practices of speech typification, reception, and response. The unit events over which such practices unfold are speech events (more generally, semiotic events) in which particular voices and figures are metadiscursively linked to performable signs, such as utterance types. Insofar as such practices disseminate discursive figures and personae, they are capable of expanding the social domain of their recognition or enregisterment.

For receivers of such messages, the voicing structure of the message constitutes a set of directions for locating one's own speech in relation to those of others. Of particular interest is the way in which receivers of such messages recognize the forms and values of the register (i.e., treat them as ones already encountered in prior socialization) or seek to incorporate them in their own discursive habits, whether by bringing their personae into conformity with them or by playing upon them in various tropes of parody, irony, recognizable hybridity, and the like.

Any such encounter is mediated by institutional processes that influence its social domain. Yet institutions do not simply "speak down" to individuals. They live through them. Macrosocial processes of register expansion always operate through microsociological encounters, or interactions, whether of the face-to-face type or ones mediated by artifacts that connect senders and receivers of messages at greater spatiotemporal removes from one another; even messages that are highly institutionalized (thus widely disseminated or even highly codified) are subject to further negotiation—through processes of ratification, counter-valorization, and other forms of role alignment—in moments subsequent to those where they are first encountered.

A register grows in social domain when more and more people align their self-images with the social personae represented in such messages. The stereotypic social range of the register may change during the process of its demographic expansion when those exposed to it seek to formulate additional, partly independent, or even counter-valued images of what its usage entails. The repertoires of a register can similarly change as well, whether through analogical extension, "borrowing," changes in "reference standards" (such as changes in exemplary speaker), changes in practices of codification (cf. dictionaries), or even the substitution of the speech of one group by the speech of another under the same metapragmatic label.

Although such changes are almost continuously in progress in the social life of most registers, not all such alterations are equally consequential from the point of view of widespread patterns of social life. For ultimately not every form of alteration or change is taken up by those metasemiotic practices that are most highly institutionalized in society. Only some among the changes that do occur can, through the mediation of institutions, become widely circulated images of speech and thus can become sources of potential response through the logic of role alignment-and thus, of uptake, fractionation, change, revalorization—by significant parts of the population. For many registers, competing models are common in social life; however, only some among them—or even just one—may come to count as the "official" model for

a given group at a given time and thus become the model to which more and more of the subsequent social history of the group is an intertextual response.

Notes

1. In my technical usage, the term enregisterment is derived from the verb *to register* ("recognize; record"); the noun form *a register* refers to a product of this process, namely a social regularity of recognition whereby linguistic (and accompanying nonlinguistic) signs come to be recognized as indexing pragmatic features of interpersonal role (persona) and relationship. My technical terms are cognate with, but differ from, their everyday homonyms. Thus the verb *to register* corresponds, in ordinary English, to at least two verbal lexemes: (1) a verb of cognition and recognition that takes a dative experiencer (viz., "the point didn't register on him at all") and (2) a *verbum dicendi* meaning "to (institutionally) record, inscribe, write down" (viz., "he hasn't registered to vote," "she is a registered user," etc.). The everyday lexeme *a register* (cf. "a book containing [official] records") brings together the cognitive, discursive, and institutional senses to some extent; thus we have registers of births, deaths, and marriages and, in contexts of class differentiation, a "social register" recording recognized distinctions of rank.

2. It is sometimes assumed that if a register exists, it has a universal social domain (i.e., is known by all members of a language community), but this is false. All members of a language community do not have identical competence over all of its registers. For any given register, the social domain of the register (the set of persons acquainted with it) changes over time in ways mediated by mechanisms of language socialization; for some registers, the social domain is very tightly delimited by institutions that confine register competence to specific demographic locales within a population, thus maintaining sharp asymmetries of register competence within a language community. Indeed, the *competence to recognize* a register's forms/effects may have a much wider social domain than the *competence to speak* the register fluently (cf. Table 8, C); in the case of prestige registers, this type of asymmetry is often a principle of value maintenance that preserves the register as a desirable commodity in which fluency is desired by those who lack it and may be purchased for a price (see Agha 2003).

3. Once the structure of antiphony or contrast is clear, a rereading of the excerpt permits more than one such emergent projection, thus exhibiting "double-voicedness" or "hybridity" in Bakhtin's sense. Thus, once segment$_2$ is seen as a dialogic response to segment$_1$, it can also be viewed as an interior monologue by Merdle himself, his ruminations, as onlooker to his own banquet, on what others will say of him. Similarly once segment$_3$ is seen as marking authorial irony, this stance can be read backward so that the hyperbolic epithets that pervade the first two segments appear inflected with the author's ironic stance throughout. And so on.

4. Tables 4 through 7 exhibit a wide range of formal contrasts. In Table 4 we see a contrast of verb stems in the indicative and imperative mood; in Table 5, a contrast of particles marking illocutionary force (IF); in Table 6, complex nominalizations in place of simple everyday lexemes; in Table 7, a contrastive patterning of co-occurrence styles involving a range of devices (anaphors, determiners, prepositions, constituent order, tense, etc.) that marks the sportscaster register as deviating from everyday English.

5. For example, several kinds of *explicit metapragmatic activity* occur naturally in all language communities. These include verbal reports and glosses of language use, practices of naming registers, accounts of typical or exemplary speakers, proscriptions on usage, standards of appropriate use, and positive or negative assessments of the social worth of the register. Other types of more *implicit metapragmatic behavior* also serve as data points. These include utterances that implicitly evaluate the indexical effects of co-occurring forms (as "next turn" responses to them, for example) without describing what they evaluate; such behavior may include *nonlinguistic semiotic activity* as well, such as gestures, or the extended patterning of kinesic and bodily movements characteristic of ritual responses to the use of many registers. Metapragmatic data can also be elicited through the use of queries, interviews, questionnaires, and the like. A detailed discussion of these issues may found in Agha 2002:24–32.

6. Such contrary-to-stereotype effects are not felt to be tropes when they are entextualized in a denotationally explicit voicing frame such as a direct reported speech construction. Such constructions denotationally distinguish the utterer from the character reported, thus allowing men to utter women's speech, and vice versa, without taking on the characterological attributes of the other gender. Thus for the case of Koasati gender indexicals (Table 4), Mary Haas observes, "If a man is telling a tale he will use women's forms when quoting a female character; similarly, if a woman is telling a tale she will use men's forms when quoting a male

character" (Haas 1964:229–230). When the metapragmatic frame is more implicit, however, the non-transparency of the frame suggests that the contrary-to-stereotype effect is an effect of the register token *all by itself.* But this is an illusion. When such tropes occur, it is the construal of a text configuration, a co-textual array of signs, that globally superposes an effect contrary to the stereotypic effect of the register token; the semiotic basis of the construal is not the register token alone but a text configuration of which the register token is a fragment. This is precisely what is illustrated in the Lakhota example in Figure 1.

7 Moreover, recent work (Hanks 1996; Irvine 1996) shows that there is no upper bound on the complexity or delicacy of role distinctions performable in context and that no final, decontextualized inventory of role labels can be given—or is analytically necessary—because such effects are individuated by entextualized semiotic cues and are recoverable only by persons having access to such cues in the event itself; they are therefore highly nondetachable for purposes of construal, like entextualized voices in general.

References Cited

Agha, Asif
  1998 Stereotypes and Registers of Honorific Language. Language in Society 27(2):151–193.
  2002 Honorific Registers. *In* Culture, Interaction and Language. Kuniyoshi Kataoka and Sachiko Ide, eds. Pp. 21–63. Tokyo: Hituzisyobo.
  2003 The Social Life of Cultural Value. Language and Communication 23(3/4):231–273.
  2004 Registers of Language. *In* A Companion to Linguistic Anthropology. Alessandro Duranti, ed. Pp. 23–45. Oxford: Blackwell.
  In press. Language and Social Relations. Cambridge: Cambridge University Press.
Alrabaa, Sami
  1985 The Use of Address Pronouns by Egyptian Adults. Journal of Pragmatics 9(5):645–657.
Bakhtin, Mikhail M.
  1981 Discourse in the Novel. *In* The Dialogic Imagination: Four Essays. Michael Holquist, ed. Caryl Emerson and Michael Holquist, trans. Pp. 259–422. Austin: University of Texas Press.
  1984 Problems of Dostoevsky's Poetics. Caryl Emerson, ed. and trans. Minneapolis: University of Minneapolis Press.
Banfield, Ann
  1982 Unspeakable Sentences. London: Routledge and Kegan Paul.
Ferguson, Charles A.
  1983 Sports Announcer Talk: Syntactic Aspects of Register Variation. Language in Society 12:153–172.
Goffman, Erving
  1974 Frame Analysis. Cambridge, MA: Harvard University Press.
  1979 Footing. Semiotica 25:1–29.
Haas, Mary
  1964 Men's and Women's Speech in Koasati. *In* Language in Culture and Society. Dell Hymes, ed. Pp. 228–233. New York: Harper and Row.
Hanks, William F.
  1996 Exorcism and the Description of Participant Roles. *In* Natural Histories of Discourse. Michael Silverstein and Greg Urban, eds. Pp. 160–200. Chicago: University of Chicago Press.
Hill, Jane
  1995 The Voices of Don Gabriel: Responsibility and Self in a Modern Mexicano Narrative. *In* The Dialogic Emergence of Culture. Bruce Mannheim and Dennis Tedlock, eds. Pp. 97–147. Urbana: University of Illinois Press.
Honey, John
  1989 Does Accent Matter? The Pygmalion Factor. London: Faber and Faber.
Hoyle, Susan M.
  1993 Participation Frameworks in Sportscasting Play: Imaginary and Literal Footings. *In* Framing in Discourse. Deborah Tannen, ed. Pp. 114–145. New York: Oxford.
Irvine, Judith T.
  1990 Registering Affect: Heteroglossia in the Linguistic Expression of Emotion. *In* Language and the Politics of Emotion. Lila Abu-Lughod and Catherine A. Lutz, eds. Pp. 126–161. Cambridge: Cambridge University Press.

1996 Shadow Conversations: The Indeterminacy of Participant Roles. *In* Natural Histories of Discourse. Michael Silverstein and Greg Urban, eds. Pp. 131–159. Chicago: University of Chicago Press.

Lee, Benjamin
  1997 Talking Heads: Language, Metalanguage and the Semiotics of Subjectivity. Durham: Duke University Press.

Lutz, William
  1990 Doublespeak. New York: Harper Collins.

Mannheim, Bruce
  1997 Cross-talk. Journal of Linguistic Anthropology 7(2):216–220.

Morford, Janet
  1997 Social Indexicality and French Pronominal Address. Journal of Linguistic Anthropology 7(1):3–37.

Saunier, Baudry de
  1937 Principes et Usages de Bonne Éducation Moderne. Paris: Éditions Flammarion.

Silverstein, Michael
  1996 Indexical Order and the Dialectics of Sociolinguistic Life. *In* SALSA, vol. 3. Risako Ide, Rebecca Parker, and Yukako Sunaoshi, eds. Pp. 266–295. Austin, TX: Department of Linguistics, University of Texas.

Trechter, Sara
  1995 Categorical Gender Myths in Native America: Gender Deictics in Lakhota. Theme issue, "Sociolinguistics and Language Minorities," Issues in Applied Linguistics 6(1):5–22.

Voloshinov, Valentin N.
  1973 Marxism and the Philosophy of Language. Ladislav Matejka and I. R. Titunik, trans. Cambridge, MA: Harvard University Press.