

Socially meaningful syntactic variation in sign-based grammar*

Emily M. Bender

1 Introduction

This study begins from two observations. The first is that linguistic variation is socially meaningful (see Labov 1963, Ochs 1992, California Style Collective 1993, Eckert 2000, Irvine 2001, Podesva et al. 2001, Bucholtz and Hall 2005, Campbell-Kibler 2006, *inter alia*). The second is that speakers use their knowledge of language (i.e., linguistic competence) to engage in various linguistic activities (i.e., linguistic performance) and therefore in modeling linguistic competence we are well-served to strive for models which are ‘performance-plausible’ (Bresnan 1978, Pollard and Sag 1994). Among other things, this means that, if sociolinguistic meaning-making is bound up with syntactic constraints, then our model of (syntactic) linguistic competence must interface smoothly with models of knowledge (and use) of social meaning, and that systems which look optimal under an artificial modularization might fare less well when viewed against this broader scope of data.

As we extend models of syntactic competence to encompass sociolinguistic variation, we must account for the following types of data:

1. The fact of variation itself, namely, that there may be multiple ways of saying the same thing, or multiple forms with the same (basic) meaning.
2. Linguistic constraints on variation, i.e., robust frequentistic patterns where the choice among variants is influenced by their grammatical context.
3. Social meaning, more precisely, that both *patterns of use* and *particular uses* of sociolinguistic variables can carry social meaning.

*This paper is a revision of Chapters 5 and 6 of Bender 2001. I am grateful to my dissertation committee (Penny Eckert, Tom Wasow, John Rickford, Ivan Sag and Arnold Zwicky) for their input and guidance. This paper has also benefited from thoughtful comments on an earlier draft by Tom Wasow, Fritz Newmeyer and two anonymous reviewers. The original work was supported by a Dissertation Fellowship from the Stanford Humanities and Sciences Graduate Alumni Association and a National Science Foundation Graduate Research Fellowship.

Formal models of variation (e.g., Anttila 1997, Adger 2006) typically attempt to derive relative frequency of occurrence of different variants from properties of the formal system itself or model it through including gradient information directly in the model (e.g., Labov 1969, Cedergren and Sankoff 1974, Boersma 1998, Bresnan et al. 2001). In contrast, I argue that a more complete understanding of the nature of robust non-categorical constraints on sociolinguistic variation will come from modeling them in the context of modeling the social meaning of linguistic variation. In Bender 2005, I provide preliminary evidence that speakers have knowledge of (non-categorical) linguistic constraints on variation, and that they use this information in interpreting the social meaning associated with uses of sociolinguistic variables (in particular copula presence/absence in African American Vernacular English [AAVE]).

In this paper, I present a brief overview of the literature on social meaning and argue that social meaning can attach to both particular instances as well as patterns of use (§2). After considering what grammars are models of (§3), I explore what would need to be included in a grammar that models socially meaningful variation (§§4–5), using AAVE variable copula absence as an example, and examine some heuristics for defining the boundaries of competence grammar (§6).

2 Social meaning¹

2.1 The social construction of identity and social categories

Eckert (2005) presents an overview of the evolution of sociolinguistic theory through three ‘waves’. The first wave is typified by large-scale community studies (e.g., Labov 1966, Wolfram 1969) documenting a systematic relationship between linguistic variation and macro social categories such as socioeconomic status and gender. The second wave (e.g., Labov 1963, Eckert 1989), is characterized by ethnographic studies looking at smaller communities and documenting the connection between sociolinguistic variation and local social categories. The local social categories often are related to the global or macro categories. For example, Eckert (1989) found that high school students’ use of particular variables (e.g., vowels involved in the Northern Cities Chain Shift as well as negative concord) correlated with their identification as ‘Jocks’ or ‘Burnouts’, two opposing social groups in the high school culture, associated with middle-class and working-class trajectories respectively.

The third wave (e.g., Ochs 1992, California Style Collective 1993, Eckert and McConnell-Ginet 1992, Irvine and Gal 2000) is concerned with how language use (and other symbolic practices) are actually constitutive of the categories in question. That is, rather than viewing an individual speaker’s style as a consequence of their ‘social address’ (in terms of global or local categories), this perspective investigates how speakers’ linguistic actions stake claims to particular stances and social categories, and in the process, helps shape those categories (both in their relationship to other social categories and in the properties which are associated

¹Some would object to this term on the grounds that all meanings are social. Although I can hardly disagree with this point, for lack of a better term, I will continue to use ‘social meaning’ and contrast it with ‘truth-conditional meaning’, for ease of exposition.

with them).

Ochs (1992) gives the example of two sentence-final particles in Japanese. The particle *ze* expresses coarse intensity and the particle *wa* expresses delicate intensity. The properties coarse and delicate are partially constitutive of the categories masculine and feminine, respectively. One thing about women, according to Japanese ideology, is that they are delicate, even when expressing intensity. One thing about men, according to the same ideology, is that they are (or can be) coarse about things such as expressing intensity. To use the particle *ze* is to be coarse, or in Ochs's terms to directly index coarse intensity. The use of the particle thus indirectly indexes gender: used by a man, it indexes masculinity. Used by a woman (against expectations) it probably indexes something more like anti-femininity. By the act of speaking one of these particles, the speaker has performed an instance of coarse or delicate intensity, which is part of being masculine, feminine, anti-masculine or anti-feminine, depending on what was said, by whom.

Synthesizing third wave sociolinguistic research as well as related work in allied fields, Bucholtz and Hall (2005) argue against a conceptualization of identity as an internal psychological fact which leads people to speak and act in certain ways. Rather, they present identity as something which results from linguistic and other behavior in interactions between individuals. Identity, for them, includes both the situation of self with respect to local and global social categories as well as stances and roles that participants take on in the course of interaction.

My purpose here is to examine the implications that sociocultural work on language, identity and social categories has for syntactic theory. As long as linguistic style was a consequence of (pre-determined) 'social address', formal linguists could view socially-conditioned linguistic variation as the result of speakers at different 'addresses' being exposed to different dialects. Intraspeaker variation could be seen as dialect-switching on the part of individuals moving between social categories (or desiring to do so). In a world where social categories are not predetermined but rather constructed, reinforced, contested and changed as they are performed (Butler 1990), social meanings attach to linguistic elements smaller than linguistic varieties, and we must ask what that means for our models of such elements and their combinatorics. That is, rather than linguistic varieties which belong to particular social categories (or situations), we have linguistic features associated with particular social values, which can be combined in a variety of ways to create styles (and social categories and situation types). Some of these styles endure and are called upon over and over, changing only slowly over time. Others are more fleeting. If we are to model this in a general theory of linguistic competence, what are the implications for the model of syntax?

Before I continue, I would like to point out that the claim that social categories are constructed does not entail that speakers have complete freedom to make them up as they please. This is tightly connected with speakers' experience of such categories as relatively natural and fixed.² Part of the reason they can appear to be fixed is what Irvine and Gal

²Though surely speakers actively engaged in contesting particular systems of categories have a different view. For example, Hall (1997) describes *hijras*, a transgender category in India who actively contest the dominant organization of (social) gender categories in their culture. One resource they use for this is the

identify as the process of *iconization*:

Iconization involves a transformation of the sign relationship between linguistic features (or varieties) and the social images with which they are linked. Linguistic features that index social groups or activities appear to be iconic representations of them, as if a linguistic feature somehow depicted or displayed a social group's inherent nature or essence. This process entails the attribution of cause and immediate necessity to a connection (between linguistic features and social groups) that may be only historical, contingent, or conventional. (2000:37)

Thus our understanding of socially significant linguistic forms is embedded in (and intricately involved with the construction of) our ideologies of social space. From this point of view, the connection between the particles *ze* and *wa* to coarse/gentle intensity and to masculinity/femininity is non-arbitrary to speakers of Japanese. *Ze* IS coarse and *wa* IS gentle, so of course men (who are coarse) say *ze* and women (who are gentle) say *wa*.

Ideologies are typically tacit and presupposed, which makes them difficult to challenge and constrains the ways in which speakers can use symbolic resources to make new styles and categories. Furthermore, it is not only the speaker's own ideologies that are at issue but also (and more importantly) the addressee's. Speakers rely on a certain amount of cooperation on the listener's part in their meaning-making, cooperation which may be more or less forthcoming in different situations. Eckert and McConnell-Ginet note that not everyone shares the privilege to "assume [one's] own positions to be norms toward which everyone else orients" and to "judge other positions while supposing [one's] own to be invulnerable to less privileged assessment" (1992:483). Campbell-Kibler (2006) presents data from matched-guise experiments as well as focus groups and argues that listeners actively participate in meaning-making, rather than merely faithfully decoding speakers' intentions. Looking in particular at the variable (ING) (e.g., *thinking* v. *thinkin*) she finds that listeners evaluate (ING) (in each of its guises) differently depending on several factors, including other features of the speaker's speech, the listener's current disposition, and their ideologies about social categories (e.g., about 'Southerners').

Finally, Bucholtz and Hall (2005) point out that just because social categories and identities are built by virtue of things that speakers do, doesn't mean that all of those component social actions are intentional (though some may be). Habitual actions are still interpreted by other people and still contribute to social reality, even if they are not each deliberate on the part of the speaker. Furthermore, we can be intentional about taking certain social stances without being intentional about the linguistic choices we use to do so. That is, just like we usually talk about cats without consciously choosing the word *cat*, it is enough for a Japanese speaker to decide to express coarse intensity—using the particle *ze* to do so doesn't require further deliberate planning.

Thus speakers do not have complete freedom as to construct styles, identities or social categories willy-nilly out of the existing stylistic resources. They are constrained in this endeavor by prevailing ideologies and by the need to gain the cooperation of listeners in

system of grammatical gender in Hindi.

meaning-making. Furthermore, such stylistic *bricolage* (see §6.3 below) is not (always) intentional or highly conscious. Nonetheless, these considerations do not detract from the points that social categories are social constructs and that symbolic behavior (including prominently linguistic behavior) is partially constitutive of these categories. This turn leads us to consider what the implications are for syntactic theory. To do so, we first need to consider where social meaning can reside. This topic is taken up in the next subsection.

2.2 Meaningful patterns, meaningful instances

Labov (1966) and Weinreich et al. (1968) find social meaning (such as it was for them) in the differing frequencies of each variant across different social groups. The discussion of social meaning in the construction of social categories and the social landscape above focused on individual forms. Where does social meaning reside? Given a sociolinguistic variable, there is both a pattern of variation and individual instances of each variant. The pattern forms the background for each new instance.³ This is reminiscent of Halliday's (1991) description of the interrelation between grammar and text with reference to the relationship between climate and weather:

To the 'instance' observer, the *system* is the potential, with its set of probabilities attached; each instance by itself is unpredictable, but the system appears constant through time. To the 'system' observer, each *instance* redefines the system, however infinitesimally, maintaining its present state or shifting its probabilities in one direction or the other (as each moment's weather at every point on the globe redefines the global climate). (1991:34)

This view of system and instance can be applied to the relationship between social actions and social structures as well as the relationship between linguistic actions and linguistic structures. This is schematized in Figure 1. By saying something, a speaker (S) produces an instance of linguistic form and an instance of verbal behavior (a speech act). As an instance of linguistic form, the utterance contributes to the system of how that speaker talks in situations like the current one. Moving up to broader contexts, it also contributes to the system of how that speaker talks in general, how people like the speaker talk and to the language of the speaker's whole community. As an instance of verbal behavior, the utterance contributes to the definition of the speaker's behavior in situations like the current one. As a situated action in the world, the utterance is also part of the practice that is the speaker's identity and the fabric of the social category or categories that the speaker belongs to. Finally, the social landscape of the community is populated with such categories.⁴

Thus each linguistic act is part of a whole series of larger systems, both in terms of linguistic systems (the patterns of language schematized on the left-hand side of the figure) and

³Cf. Bakhtin 1986.

⁴The separation of the linguistic and social hierarchies in this discussion is somewhat artificial. For an utterance to count as part of the system of how people like the speaker talk, it has to be interpreted by that speaker and others against the existing ideologized social background and the speaker's position in it.

language of S's community	↔	social structure of S's community
language of S's social category	↔	social content of S's social category
language of S in general	↔	S's identity
language of S in this kind of situation	↔	behavior of S in this kind of situation
linguistic form	↔	speech act

Figure 1: Parallel hierarchies of systems

in terms of social systems or categories (stances, identities, social categories and communities on the right-hand side). I don't claim to have identified all of the relevant levels on each side of the figure, nor that they are in strict correspondence. Nonetheless, the existence of multiple levels of patterning on both sides suggests that social meaning in linguistic variation may reside in both particular instances of use as well as in the larger patterns. Particular instances will be interpreted against the larger patterns, but also become part of the pattern against which future instances are interpreted.

For example, suppose that a middle manager named Kim gets upset and swears in a business meeting. In doing so, Kim is taking a particular stance. The impact that swearing will have on Sandy, another employee present, will depend on Sandy's knowledge of how Kim usually behaves in business meetings. If Sandy has often heard Kim swear in that situation, then Sandy might know that Kim is only somewhat upset. If Kim almost never swears in that situation, to Sandy's knowledge, then Sandy might deduce that Kim is extremely upset. Further, business meetings are not the only situation in which Kim speaks to Sandy. If Sandy knows that Kim swears a lot, but not at business meetings, then the import of that instance of swearing (for Sandy) is different than if Sandy almost never hears Kim swear.

Moving up to the broader categories, Kim is busy constituting the social category middle manager for Sandy, Kim, and the others present. Thus Sandy will also interpret Kim's swearing in the context of the behavior of other middle managers and other employees of the company. Finally, the taboo nature of the swear words is constituted in part by the combined practice of all members of the speech community, and Kim, in swearing, is making use of (and perhaps loosening) this taboo. Note that it's not just the act of swearing that has all of these interpretations. Suppose Kim has a particular favorite swear word (and Sandy knows this), but chose a different one on this occasion. The 'bite' or shock-effect that Sandy perceives in Kim's swearing will depend on that choice, but also on the general level of taboo associated with the particular word at all of the successively larger parts of the speech community.⁵

The preceding paragraph talked about patterns of behavior, including patterns of behavior of entire communities. These patterns can only be meaningful to the extent that they are known by speakers (and therefore available to be interpreted and to use in the interpretation

⁵For ease of exposition, this example has taken the notion of a 'business meeting' as a social situation to be invariant background, which of course it isn't. To the extent that Kim is a prominent or powerful individual, Kim's behavior also contributes to what a business meeting is in that company, and in general, for the people present.

of new utterances). Of course, no speaker will have perfect knowledge of the speech patterns of the rest of his/her community, or indeed, even of his/her own speech patterns. What counts are the patterns that the speaker believes to exist. Such beliefs are based on the actual interactions that the speaker has had, but certainly extend beyond that as the speaker extrapolates from his/her experience. As mentioned above, there is considerable potential for misunderstanding between speakers with different experience, who may or may not assume their position to be shared. Further, it is usually the prerogative of those in power to assume that their understanding of the patterns is indeed shared. With these caveats, I believe the model of parallel hierarchies of social structure and linguistic structure provides a basis for describing the relationship between instances of variants, patterns of variation, and social value.

Most variationist work focuses on the social value of the patterns, and not of the instance. In fact, Lavandera (1978:174) argues that because no speakers categorically use one variant or the other of any particular variable, the meaning must be in the frequency of use of a particular variant rather than in each individual instance. She contrasts this with cases of syntactic or lexical variation where each variant is marked as e.g., ‘informal’ or ‘formal’, and argues that the pattern of variants in formal or informal styles does not carry any meaning, but is rather an epiphenomenon derived from the fact that the forms are meaningful. On the other hand, Coupland (2001) argues that in the study of style it’s not aggregate patterns but individual situated acts that are most revealing. He draws an analogy to music, noting that a melody is poorly described by its average pitch.

I would like to argue that both the pattern and the instance can be meaningful at the same time. My intent with the hypothetical business meeting example is to show how patterns built up out of meaningful acts be called upon as the background against which new acts are interpreted. In her experimental work, Campbell-Kibler (2006) finds that listeners evaluate particular tokens of the (ING) variable (*-in* or *-ing*) differently depending on which variant they judge to be ‘natural’ for the speaker, on the basis of the rest of the speaker’s style. Thus meaningful instances can combine to produce meaningful patterns. To extend Coupland’s musical analogy, it is not possible to tell from only one note what key a piece of music is in.

2.3 Social meaning and linguistic structure

I have argued above that variation is (potentially) socially meaningful both at the level of the social import attached to individual variants, and the level of social import of patterns of those variants. In this section, I consider the relationship between linguistic structure and social meaning, highlighting three points: (i) Social meaning can attach to units of any level of linguistic structure. (ii) In some cases, social meaning is the only meaning associated with a particular contrast in linguistic form. (iii) Non-categorical grammatical constraints on the distribution of linguistic variables modulate social meaning.

Social meaning can attach to any level of linguistic structure Sociolinguistic studies have documented socially meaningful deployment of variable linguistic form across all

levels of linguistic structure: Phonological variables are the best-documented of course (see Campbell-Kibler 2006, Benor 2001, Eckert 1989 for some examples). Podesva (2006) has argued that gradient phonetic properties can also be deployed in socially meaningful ways (e.g., strong release bursts for released /t/ parodying its association with ‘precise’ speech to produce an effect of ‘prissyness’). Variationist work has also considered morphosyntactic variables such as variable agreement on auxiliaries in Basque (Haddican 2003) and verb-form honorifics in Japanese (Okushi 1997). Less work has been done on purely syntactic patterns, but negative concord (Eckert 1989, Eckert 2000) is one example. Social meaning can also attach to lexical choices (e.g., the choice of *say*, *be all*, *be like*, or *go* (Bucholtz and Hall 2005)). Likewise, social meaning can attach to intonation patterns (e.g., Lowry 2002) and discourse-level effects such as turn-taking strategies (e.g., Gumperz 1982).

Social meaning can be the only meaning associated with a particular contrast in form The list of linguistic units which can bear social meaning given in the preceding paragraph is clearly diverse. One way in which the various kinds of socially-meaning units differ is in whether or not they carry any other meaning. Clearly the well-studied phonological variables (as well as gradient phonetic features) do not: these variables all involve non-phonemic contrasts. On the other hand, many of the ‘larger’ units are also involved in the computation of the truth-conditional meaning of a sentence. Consider the sentences in (1):

- (1) a. I ain’t seen nobody.
- b. I didn’t see anybody.

For some speakers, these probably count as two ways of saying the same thing, as far as the truth conditions are concerned. In that sense, the syntactic variable is behaving like a phonological variable. At the same time, *nobody* and *anybody* are also meaningful words, something which can’t be said of the various possible productions of the variable (AY) in the greater Detroit area.

This is not to say that social meaning can only attach to forms which stand in contrast to alternate ways of saying the same thing. Consider the examples in (2):

- (2) a. She has multiple fractures in her fibula.
- b. She broke her leg bone all to bits.

While these two examples might be used to communicate about the same state of affairs, it would be a stretch to consider them truth-conditionally equivalent. This doesn’t preclude social meanings attaching to the various word choices, for example, an air of expertise or authority with the medical terminology and an sense of informality with *leg bone* and *all to bits*. Similarly taboo words don’t usually contrast with non-taboo ways of saying (exactly) the same thing, though their original or core senses can usually be paraphrased with more technical terminology.

At the same time, it is not the case that linguistic variables consisting of morphosyntactic or lexical units necessarily have meaning beyond their social meanings. *Be* and *do* (as used in (1)) are often analyzed as semantically empty units (though the particular forms of *be* and *do* in (1)) are also marked for tense and negation, which do have semantic reflexes). Verb-form honorifics in Japanese appear to only encode social stances of the speaker vis-à-vis the addressee (and related social meanings), as shown in the following minimal pair:

- (3) a. *Tabemashita ka*
Eat-HON-PAST Q
'Have you eaten?'
(Addressee is unfamiliar to speaker or socially superior to speaker or situation is formal or speaker wishes to present him/herself as typically polite and refined.)
- b. *Tabeta ka*
Eat-PAST Q
'Have you eaten?'
(Addressee is a peer to speaker or social subordinate and situation is informal and speaker is not presenting him/herself as typically polite and refined)

Similarly, as pointed out by Hudson (1996:257), words such as *please* (as used in requests), *goodbye* and *uh-huh* carry only social meaning.

Social meaning can be modulated by linguistic context A well-established property of sociolinguistic variation is that the distribution of variants is sensitive to both 'external' (social) and 'internal' (grammatical) constraints. For example, in the case of AAVE variable copula absence, numerous production studies (see Rickford 1998 for a summary and overview) have found robust effects of the following syntactic context, among other factors. In particular, NP predicates are most conducive to overt forms of the copula, followed by adjective and prepositional phrase predicates followed by verbal predicates (i.e. *-ing* form verbs).

On the view of socially meaningful variation presented here, the 'external' constraints are the result of speakers deploying socially meaningful forms in order to take on social stances and roles, construct personae and reinforce (or renegotiate) social space. How are the 'internal' constraints involved in this process, if at all? In Bender 2005, I report the results of an experiment designed to test the hypothesis that such 'internal' constraints can modulate the social meaning of a variable. The study involved a matched-guise experiment in which listeners (including AAVE speakers and non-AAVE speakers) judged AAVE speakers on seven personality trait scales on the basis of brief utterances, shown in (4).

- (4) a. Yeah I know her. She's teachin me piano at Music World.
b. Yeah I know her. She's my piano teacher at Music World.
c. Yeah I know her. She teachin me piano at Music World.

d. Yeah I know her. She my piano teacher at Music World.

In my experiment, African-American listeners judged AAVE speakers using copula presence in an unmarked environment (before NP) to be somewhat reliable/likeable/well educated, but not as much so as speakers using copula presence in the marked environment (before *V+ing*).⁶ Non-African-American listeners (even those who reported having experience with AAVE) did not show any effect of the grammatical environment, though they did judge copula absence to sound less well-educated etc. than copula presence.

To use the terminology introduced above, using copula presence appears to count as evidence that a person is reliable/likeable/well educated for these speakers, or index these properties. But it counts as better or stronger evidence (or indexes the properties more emphatically) in precisely those contexts where it is unusual. The study was small and the results need to be replicated with a larger number of speakers. Nonetheless, they seem to indicate that non-categorical grammatical constraints, so far ignored in this discussion of social meaning, are not a separate issue from social meaning but rather intertwined with it.

2.4 Further remarks on social meaning

So far in this section, I have argued that the social import of linguistic variation is best described in terms of speakers deploying socially meaningful linguistic elements as they perform stances and identities. Social meaning can reside both in patterns of use of linguistic variables in and particular uses of variables, and socially meaningful linguistic variables can appear at any level of linguistic structure. Finally, I have suggested that grammatical constraints on linguistic variation are intertwined with social meanings as well. I conclude this section with some remarks on the similarities and differences between social meanings and truth-conditional meanings.

One characteristic of social meaning is that the actual effect of a use of a socially meaningful form is dependent not only on the form's core meaning, but also on many other factors including the speaking situation, the speaker's identity in general and role in the situation, the listener's identity, attitudes and beliefs and even affect (Campbell-Kibler 2006). This makes it difficult to discover the core meaning of linguistic variables. The matched guise techniques (Lambert et al. 1975) used in Bender 2005 to detect effects of the syntactic environment on the interpretation of copula presence and absence are not able to isolate the core meaning of the variable. In her study of (ING), Campbell-Kibler uses similar matched-guise experiments but combines them with interviews of listeners to get a richer set of data from which to explore the meaning(s) of the variable. Even then, it is not possible to find a definitive answer to the question of its core meaning. While it is much easier (though still non-trivial) to describe the core truth-conditional meaning of lexical items (or syntactic

⁶This was true for both groups of African-American listeners, those who identified as AAVE speakers themselves and those who did not. In addition, those who did not self-identify as AAVE-speakers also showed a similar effect of the 'internal' constraint on their evaluation of copula ABSENCE, judging copula absence in the marked environment (before NP) to sound less well-educated etc. than copula absence in the unmarked environment.

constructions), these, too, interact with context such that the meanings of ordinary words get vivified in everyday use (Sag 1992). This vivification complicates considerably the task of discovering what exactly about the social meaning of an item speakers store in long term memory, but does not detract from the point that they do store something.

Another characteristic of social meaning is its possible specificity to particular groups of interlocutors. That is, it is plausible that copula absence/presence has some social meaning for AAVE speakers that it does not have for other groups, and that AAVE speakers are aware of this. Here again, there is a similarity with truth-conditional meaning. Take any ordinary word that also has a use as a technical term, for example, *daughter* (of a node in a tree), *mouse* (attached to a computer), *government* (of a dependent by a head), etc. Speakers who know the technical meanings of these words probably also know which interlocutors are likely to share that meaning. This is consistent with Harder's (1996, 2003) analysis of linguistic meaning as the potential of a form to cause a hearer to invoke some concept. If meaning involves a representation of the effect that using a word can have on an addressee, then such a representation could also specify the type of addressee one should expect to be so affected.

One would expect social meanings, however, to be subject to more disagreement between speakers and more active renegotiation. That is, while most English speakers probably have fairly similar meanings associated with the word *horse* (or in Harder's terms, similar beliefs about what concept its use evokes), certain kinds of social meaning appear to be much more fluid, more easily molded for local stylistic purposes, and more likely to be misunderstood. This is not to say that truth-conditional meanings are not subject to (re)negotiation, for they patently are. Words that have undergone semantic change provide evidence of past renegotiation: e.g., *will* (once meaning 'wish') or *mistress* (once parallel to *master*). Thus what we have here is not a difference in kind, but rather a difference in degree.⁷

One might think of this difference in terms of degree of conventionality. But what is convention? Lewis (1969) (see also Clark 1996) sees conventions as rooted in mutual knowledge, or a recursive sequence of "you know that I know that you know...". Searle (1995), on the other hand, argues that it is a matter of irreducibly collective intentions. This doesn't require any kind of community mind. Rather, individual speakers hold beliefs that start with "We intend that..." This notion of collective intentions is important because, as Harder (2003) points out, this is what accounts for the fact that, to speakers, the meaning of a word is not a matter of their choice, but rather a fact about the world.

Harder (2003) argues that there is still more to it, however. Conventions of language rely not only on collective intentions but also on communicative practice. The usefulness or function of any part of that system is dependent on the speakers' experience with communicative practice up to that point. Likewise, any given use of language provides input to the next stage of the interlocutors' systems. That *horse* can be used to evoke the concept HORSE among English speakers is a fact about the world. Humpty Dumpty deciding otherwise won't

⁷Changes in truth-conditional meaning are often highly tied up with aspects of social meaning, and changes in both kinds of meaning only take place in socially situated interaction. This is particularly clear in the case of words like *mistress* (McConnell-Ginet 1984, 1989).

change that. However, Humpty Dumpty could try to disrupt or change the feedback cycle by consistently using *horse* to refer to something else. This doesn't guarantee a change, but could lead to one if Humpty Dumpty has the right social status, etc.

So how does this apply to the difference in malleability between social and truth-conditional meaning? One thing to note is that it is possible for speakers from different English-speaking communities to have a conversation in which the truth-conditional semantics functions as intended without understanding all of the social meaning that may have also been expressed. For example, we expect the broader English speaking community to share a meaning for e.g., *turn right*, but not the social meaning of variable vowel pronunciations. Where might this expectation come from? I think in this regard one cannot ignore the influence of language standardization—schooling, dictionaries, widely available written sources—in shaping those expectations. But there may well be more to it, in particular, the kind of concepts that the different kinds of meaning refer to. Why should Belten High students (Eckert 1989, 2000) expect an outsider to understand the social meaning of (ay) raising, if that social meaning for them is based in a shared social landscape populated by individuals like Judy (the burned-out burnout, the social icon) whom the outsider doesn't know? On the other hand, I can expect any adult to be familiar with the notion of turning right.

The in-between cases also fall into line here. Politeness is a kind of social meaning that I expect most adult speakers to know about, and indeed Brown and Levinson (1987) are able to isolate universal aspects of politeness systems. I would suspect, then, that associations between form and politeness-based social meaning would remain constant across larger groups of speakers than other kinds of social meaning. Similarly, I don't expect all adult speakers to be familiar with the concept evoked by *c-command*, and technical terms have currency only in smaller communities of use.

2.5 Summary

This section has briefly reviewed the arguments that the relationship between social structure and linguistic variation is best treated in terms of linguistic variables whose meanings are co-constructed with local and global social categories by speakers and listeners. I then argued that social meanings can attach to both individual instances of linguistic variables as well as to patterns of use of collections of variables, and further suggested that social meaning can be sensitive to non-categorical (soft, frequentistic) linguistic constraints on the use of variables. I then suggested that social meanings, like truth-conditional meanings, boil down to expectations about the reaction of others to certain linguistic forms, and these expectations are grounded in actual practice and individual speakers' experience with that practice. Social meanings stored in speakers' long-term memory are somewhat abstract or underspecified and can get vivified differently depending on the situation of use. The social meaning recorded for a given item can be relativized to different types of addressee. In these two respects, social meanings differ from truth-conditional meanings only in degree: social meanings may be more fluid and more subject to vivification than truth-conditional meanings. However, it might well be that there is no clear distinction between social and truth-conditional meanings along these dimensions but that rather they merge into each

other.

With this background, we now turn to the question of what a linguist's grammar is supposed to model.

3 What grammar is a model of

At a sufficient level of abstraction, competence theories of syntax as well as sociolinguistic theory concern themselves with the same object of study: human language. In practice, however, theoretical work in these two subfields of linguistics is not usually integrated. I believe, however, that such an integrated theory would provide superior models of language, able to account for a broader range of data. My goal in this paper is to sketch the constraints on a theory of grammar (competence) that could form a part of a social theory of language, such as that described by Eckert:

It is impossible for a social theory of language to view *langue* as a pre-existing convention, for a social theory of language must be about the process of conventionalization. By the same token, it is impossible for a social theory of language to view the individual speaker's competence as a simple internalization of convention. Convention and individual competence are mutually produced and reproduced in practice, thus linguistic practice is not simply the consensual use of a common system. Convention is not a thing but a process, and the possibility of convention resides in speakers' ability to hypothesize about others' behavior and to take interpretable action, along with a commitment to doing so within a particular social unit. Our speaker, or speaking subject, can not be a clone but must be an agent in a process of convention-making. (Eckert 2000:45)

In §4 below, I suggest that such a model would need to include information usually considered to be either extragrammatical or superfluous by syntactic theory: social meaning, prefabricated 'chunks' of linguistic structure and some reflection of probabilistic (or frequentistic) aspects of speakers' experience with language. In §5-6, I reflect on how we define linguistic competence as an object of study, and argue that the traditional boundaries are arbitrarily drawn and thus should be reexamined. In this section, I ground that discussion with some general reflections on what a linguist's grammar is a model of. There are two dimensions of possibility: First, a grammar may model the language of a community or the language of an individual. Second, within models of the language of the individual, one may distinguish 'intensional' models that are meant to reflect the actual internalized knowledge of the individual and 'extensional' models that attempt only to generate the set of form/meaning pairs that the individual would accept. The next two subsections treat each of these dimensions in turn.

3.1 Community grammar v. individual grammar

Weinreich, Labov and Herzog (1968) argue that the structure in variation is only apparent when the language of the whole community is studied. What looks random and meaningless at the level of the individual speaker, according to them, becomes part of a pattern when that speaker is considered with the other speakers in his/her community. Thus the variable rules that they present are cast as rules of community grammar—partial descriptions of the language of a community.

Chomsky (1986) takes a radically opposed position, arguing that only grammars as theories of the internalized language of individuals (I-language) can have sound scientific basis. A grammar that generates the set of sentences in a language is a theory of that language. But how is a linguist to find out what the relevant set of sentences for a whole community (E-language) is? Chomsky concludes that linguists should be concerned with the language of individuals, since these systems at least are well-defined:

Theories of E-language, if sensible at all, have some different and more obscure status [than theories of I-language] because there is no corresponding thing in the world. (1986:27)

The thing-in-the-world that corresponds to a theory of I-language is an individual's internalized knowledge of the language. That is, the underlying system behind the patterns of grammaticality is the individual's own cognitive state. There is no cognitive state that contains an underlying system for the language of the whole community, no community mind for the community grammar to be an aspect of.

Part of the issue here is that Chomsky and Weinreich et al. differ in more than where they take the structure of language to reside. Chomsky defines language in terms of grammaticality, and studies the patterns of grammaticality across different forms. Weinreich et al. are studying differences in form that do not affect grammaticality, and are investigating how the patterns of use of the variant forms articulate with the structure of society. Thus what looks ill-defined to Chomsky from the point of view of patterns of grammaticality is the only way to study what Weinreich et al. are interested in.

Those of us concerned with sociolinguistic variation could retreat to the position that it is possible to study the language of a community, and to posit a grammar of that language, as long as that grammar is not construed as a model of linguistic competence. But this is not satisfactory. I argue, instead, that it is possible to develop a theory of language (and linguistic competence) that is both social and cognitivist: The thing in the world that corresponds to community grammar is individuals' (imperfect) understanding of how other people speak. In this case, it doesn't really matter that speech communities are vague or ill-defined. Every speaker has some well-defined (although ever changing) set of linguistic and social experiences. From these experiences, speakers build their imperfect models of the speech of other people in their social landscape. Through the process of iconization (Irvine and Gal 2000) forms or patterns of usage become (for the speaker) iconic of types of people, personality characteristics, stances, etc., analyzing or abstracting over both the linguistic and the social aspects of their experience.

3.2 Linguistic competence and psychological reality

In the preceding discussion, I have contrasted Chomsky's and Weinreich et al.'s positions along the community/individual dimension. Chomsky also takes a strongly cognitivist view. He recognizes that one and the same set of sentences can be described by multiple distinct grammars, and posits that a grammar generating the right set of sentences would nonetheless be wrong if it is not the same grammar as the one internalized by the speaker under study (1986:23). I will refer to models whose goal is to describe the set of sentences as 'extensional' and to those that additionally make claims of psychological reality as 'intensional'.

Wasow (1978) and Soames (1984) argue that the methodology of generative grammar can only provide evidence for extensional models. If linguists draw their evidence solely from the set of sentences accepted (grammaticality judgments), then grammars written on the basis of that evidence can only be equivalent in extension to actual internalized grammars of real speakers. While the extensional properties of the language being described constrain the space of possible grammars which model it, it is clear from the myriad competing syntactic frameworks available today that this type of data alone does not allow us to definitely settle on one particular model. Any similarity between linguists' grammars built on this kind of evidence and actual internalized grammars is a matter of chance.

However, cognitivist models are by definition intensional, and any theory of grammar that seeks to draw explanatory force from other aspects of human cognition (language acquisition, interfaces to PF and LF, etc.) is committed to modeling actual human knowledge of language rather than an arbitrary system which is equivalent at some level. In this connection, Bresnan (1978), Pollard and Sag (1994), and Sag et al. (2003) (among others) argue for 'performance-plausible' models of competence grammar, that are meant to be directly embeddable in models of linguistic performance. In such models, psycholinguistic results can inform the design of both the overall architecture and individual linguistic analyses.

I would like to suggest that the cognitivist stance requires questioning the traditional boundaries of competence grammar, which appear to have become set through historical accident: Chomsky (1955) applied the notion, developed by himself and others in Formal Language Theory, of language as a set of sentences and laid out a research program of building grammars that generate the appropriate strings.⁸ In this early work, the cognitivist position is not evident. Unconcerned with issues of linguistic knowledge, Chomsky is free to define the object of study in whatever way is convenient. Theories of linguistic competence, including the performance-plausible competence models, inherited this definition of the object modeled by a grammar. However, theories concerned with actual speakers' knowledge of language do not have the freedom to arbitrarily define the object of study.

3.3 Summary

Syntacticians have typically been concerned with modeling grammaticality judgments. This leads naturally to a view of grammar as only including that which affects grammaticality,

⁸See Pullum and Scholtz 2005 for a brief history of how Chomsky developed his notion of formal grammar on the basis of Post's (1944) production systems for logical proofs.

a view which is typically supported on grounds of economy or simplicity. But as Langacker (1987:41) points out, comparing the economy of two competing models presupposes that they account for the same range of facts. Sociolinguistic work over the past four decades has shown that native speakers' competence (writ broad) underlies not only the production and comprehension of novel utterances but also the manipulation of linguistic variables in ways which are both sensitive to social and linguistic constraints on the distribution of those variables and creative, performing moment by moment stances, styles, and personae. I see this as a challenge to competence models: can we extend our models to account for this broader range of data? There seem to be two general ways of doing so. The first is to add a second 'module' (Culy's (1996) "user's manual", see also Zwicky 1999) which interprets structures generated by an ordinary competence grammar in terms of social meaning. The second is to integrate the information into a single component. Such models could be compared to each other on grounds of overall economy, and also (potentially) with suitable psycholinguistic experiments. In the following section, I describe the ways in which a particular type of competence model could be extended to create an integrated model. §5 presents additional arguments supporting each of the extensions in §4.

4 Overview of the model

The experimental results reported in Bender 2005 and summarized in §2.3 above are but one more type of data showing that human language processing involves tightly integrated access to both standard grammatical information and other linguistic knowledge which does not bear directly on either distributional generalizations (grammaticality) or the construction of truth-conditional meaning from lexical items and their arrangement (semantic compositionality). In this section, I consider what could be added to a competence model to account for these results. In the remaining two sections, I review other arguments for including the various additional information within the purview of 'grammar' and then explore some heuristics for defining the boundaries of competence grammar.

4.1 Tentative formalization

Modeling social meaning associated with a linguistic form is straightforward in a sign-based framework such as HPSG (Head-Driven Phrase Structure Grammar, Pollard and Sag 1994). In HPSG (and similar frameworks), the grammar is a system of signs, which pair form (phonological, morphological, and syntactic information) with meaning. These signs include lexical items (pairing stems with lexical meanings), lexical rules (producing inflected forms with appropriate meanings) and phrase structure rules or constructions (which combine the meanings of their constituents into meanings for the whole phrase). In such a framework, there would seem to be no reason not to include social meaning along with 'ordinary' semantics and pragmatics. Furthermore, such social meaning could attach (in principle) to any kind of sign: lexical item, lexical rule, or phrase structure rule. As noted above, it is non-trivial to say the least to determine the (core) social meaning of any given form, but were we

to do so, sign-based frameworks provide appropriate loci in which to store the information.

To model the effect of the syntactic environment on the social meaning, however, we need to make two further extensions. The first involves including prefabricated chunks or over-specified types, that is, pieces of linguistic structure that can be calculated on the basis of other information in the grammar. The second is frequentistic or probabilistic information.

As shown in Bender 2001 (Chapter 3), the syntactic distribution of the empty copula is best handled in terms of a phonologically null lexical item. In addition, if the distributional facts were all we needed to account for, the most parsimonious analysis of the copula (silent or pronounced) in AAVE, as in other varieties of English, posits one lexical entry that can take complements of a variety of syntactic categories. AAVE differs from other varieties in having an additional possible inflected form of the copula with a null pronunciation.

However, the social meaning facts discussed here suggest that speakers store information about the null copula separately from the overt forms, and furthermore about the co-occurrence of copula forms with different syntactic categories of complement. This suggests that their working grammatical knowledge includes not just the supertype general lexical entry in Figure 2 and inflectional rules such as the one in Figure 3, but also additional pre-compiled instances of the inflectional rule such as those in Figure 4. Each of these is described in turn below.

The lexical entry in Figure 2 shows the constraints which describe copular *be*. As encoded in its argument structure feature (ARG-ST), copular *be* requires two arguments. The second argument is a predicative phrase which is still seeking a subject (as encoded in its non-empty SUBJ list). The first is a sign (word or phrase) which is identified with the missing subject of the predicative phrase (notated with the tag \square). This has the effect of making *be* ‘adopt’ for its subject the requirements placed on the subject of the predicative phrase. This type inherits various information from supertypes (e.g., *be-lxm* and above that *aux-verb-lxm*), including the part-of-speech *verb*, the auxiliary properties, the mapping of its ARG-ST elements to the SUBJ and COMPS lists, and phonological information.

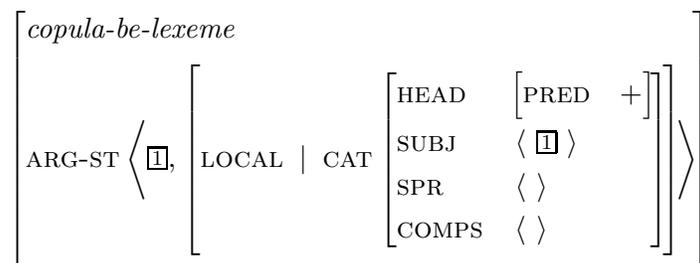


Figure 2: General lexical entry for the AAVE copula (subtype of *be-lxm*)

Organizing the lexeme types into a hierarchy such that *be-lxm* subsumes *copula-be-lxm* as well as *existential-be-lxm*, *cleft-be-lxm* and others, also allows us to write a lexical rule for the silent form which captures that fact that it is allowable across all of those different uses of *be*. This lexical rule is shown in Figure 3. This lexical rule maps its input (the DAUGHTER value) to an output which differs from the input as specified: (i) The PHON value is the

empty list, indicating the ‘silent’ pronunciation. (ii) The form is constrained to be finite and non-inverted (accounting for various constraints on the distribution of silent *be*). (iii) The subject may not be first person singular. (iv) The complement may only be elided if something (i.e., *not*) is left behind.⁹ (v) Finally and most importantly here, the rule carries some representation of the social value in its context (CTXT) feature. In Figure 3 this is glossed as ‘*not educated*’, but this is merely meant to stand in for whatever the actual social meaning of silent *be* might turn out to be. Note that the lexical rules producing overt forms of the copula will also carry (presumably contrasting, if not directly so) social meaning as well.

$$\left[\begin{array}{l} \textit{silent-be} \\ \text{PHON} \quad \langle \rangle \\ \\ \text{SYNSEM} \quad \left[\begin{array}{l} \text{CAT} \quad \left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l} \text{FORM} \quad \textit{fin} \\ \text{INV} \quad - \end{array} \right] \\ \text{COMPS} \quad \textit{nelist}(\textit{canonical}) \\ \text{ARG-ST} \quad \langle [\textit{non-1sg}], \dots \rangle \end{array} \right] \\ \text{CTXT} \quad \left[\text{SOCIAL} \quad \textit{'not educated'} \right] \end{array} \right] \\ \text{DAUGHTER} \quad \textit{be-lxm} \end{array} \right]$$

Figure 3: Silent *be* inflectional type (Note that ‘*not educated*’ is meant to stand in for whatever representation of the actual social meaning turns out to be appropriate.)

The lexical entry in Figure 2, the lexical rule in Figure 3, and similar lexical rules for producing the overt forms of the copula are enough to capture both the basic (non-statistical, categorical) distributional facts about AAVE copula absence. In order to capture the effect of non-categorical constraints modulating the social meaning of each variant, however, more is needed. One way to handle this is to posit ‘overspecified’ types in the grammar, which combine information from the lexical entry and the lexical rule, but then add additional constraints. These ‘overspecified’ types won’t change the generative capacity of the grammar (the strings it accepts as grammatical). Rather, they provide a locus to which to attach the information about the dependency of social value on grammatical context/the frequency of the silent copula with NP complements. An example is given in Figure 4 where the DAUGHTER of the lexical rule is instantiated as *copula-be-lxm* and the predicative complement is further constrained to be an NP.¹⁰

⁹For an explication of the details of this analysis, see Bender 2001, Chapter 3.

¹⁰In principle, the dependency between syntactic context and social meaning could be handled by making the frequency value and/or some aspect of the social value a function of the part of speech of the complement. The precompiled instance approach appears preferable in that it is consistent with other work examining the possibility of linguistic prefabs, as discussed below.

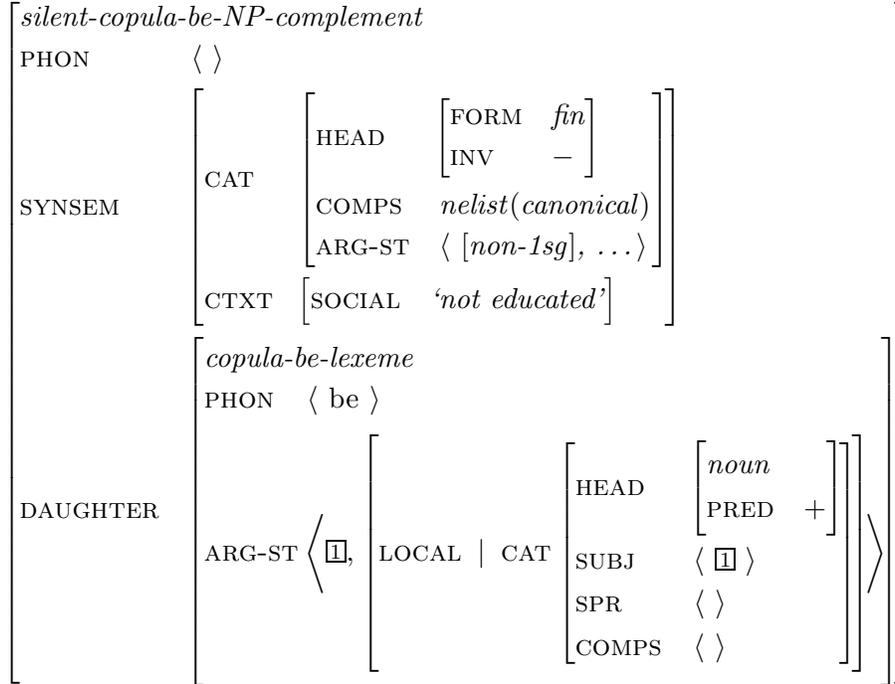


Figure 4: Example precompiled instance

4.2 Probabilities in grammar

The preceding discussion has sketched how the grammar could include the appropriate grammatical units for both social meaning and probabilities to attach to. The next step is to describe how the probabilities can be involved in modeling the modulation of social meaning according to context. There are at least three ways that probabilities have been included in linguistic models. The first is models (e.g., Kemmer and Israel 1994) which claim that people themselves are behaving probabilistically. That is, on these models, human linguistic performance involves choosing among competing alternants randomly, but according to the probability of occurrence of each alternant. The second type of model does not ascribe probabilistic behavior to human speakers, but includes probabilities in the model to stand in for various factors which are not yet well-enough understood to be included directly in a (deterministic) model of language production. This type of model does not ascribe probabilistic knowledge to speakers. Finally, the third type of model (e.g., Jurafsky 1996) has speakers using probabilistic (or frequentistic) knowledge to choose the most probable among competing alternatives (such as alternative parses/interpretations of a string or alternative realizations of a communicative message). In this case, speakers do have probabilistic knowledge, but they use it in a deterministic fashion.

Models of the first type would seem to be incompatible with evidence that individual instances of a linguistic variable are meaningful. If all a speaker can set is the probability with which s/he will produce a given variant, then listeners can only ascribe meaning to

the overall pattern of use. Models of the second type would seem to be incompatible with psycholinguistic evidence (discussed below) that speakers do store information about the relative frequencies of different linguistic objects, and use this information in sentence processing. I will therefore adopt the third type of model here, and adapt the type-one model of Kemmer and Israel 1994, to present a sketch of an account of how speakers and hearers can assign social meaning to copula presence/absence differentially depending on its syntactic context.

Kemmer and Israel (1994) present an analysis of grammatical and social constraints on t/d deletion within a usage-based model (Langacker 1987, 1990, 2000). The essential properties of their analysis (for present purposes) are as follows: All common words subject to variable t/d deletion have two representations in speakers' grammars. The existence of many such pairs leads to the abstraction of a general schema of t/d deletion. However, here we will concentrate on the individual words. In any given pair, say /læs/ and /læst/ for *last* the more frequent alternant will be more entrenched, that is, have higher resting activation. Both forms compete for selection when the speaker goes to pronounce the word, but the more entrenched form tends to win out, in proportion to how entrenched it is. However, both forms may be associated with information about the social situation, etc. If the characteristics of the social situation 'match' the specification on the less frequent item,¹¹ then the less frequent item will be promoted with respect to the more frequent item, and vice versa.

I propose to keep the notion that the probability of a form in context is a function not only of its overall frequency but also of its social value and the speaker's intention to express or not express that social value. Rather than having the forms 'compete for selection', speakers choose the most probable form to express their intended meaning, including social meaning. Since the probability of a form depends on its linguistic environment, it will take more to promote a form in a disfavoring environment than otherwise. Listeners know this, so that when they hear a form in a disfavoring environment, they attribute more emphasis to its social value. In this way, stored information about the relative frequency of occurrence of a variant in different environments can give rise to the differential interpretation of the social value of that variant across different environments.

4.3 Summary

In this section I've briefly outlined a model which I believe could account for listeners evaluating the social meaning of instances of linguistic variables differently depending on their syntactic context. There is clearly more to be done to get from this general picture to the details of a working model. Nonetheless, the above remarks suggest that an account of this phenomenon will require three extensions to the type of information usually included in competence grammars: social meaning, overspecified types, and probabilities. All three of these extensions have been suggested before. In the following section, I review arguments

¹¹Or, more accurately in my opinion, if the speaker wishes to do his/her part in constructing the social situation to match that specification.

for their inclusion in the grammar.

5 Arguments for inclusion in the grammar

I suggested above that there are two general ways to extend a competence model in order to capture socially meaningful variation: one is based on separate modules whereas the other is integrated. In this section, I review arguments that each of the three kinds of information proposed above (social meaning, overspecified types and frequency information) are better handled in an integrated model. When I speak of information being ‘in the grammar’, then, I mean integrated with the rest of the model of linguistic competence.

5.1 Social meaning

The first kind of additional information to consider is social meaning itself. The first argument for including social meaning in the grammar was already mentioned above: There are words that appear to only have social meaning (e.g., *please*, *good-bye*, *uh-huh*) (Hudson 1996:257). Signs are pairings of form and meaning, and the only meaning available here is the social meaning. Exclude it, and the status of these words becomes obscure indeed.

Likewise, certain contrasts in form correspond only to contrasts in social meaning. A clear case of this is verb-form honorifics in Korean and Japanese. On Pollard and Sag’s (1994) analysis of Korean referent honorifics, the contrast in form corresponds to a contrast in the presuppositions. That is, a verb with a honorific ending introduces an *owe-honor* relation into the BACKGROUND. This last is a feature within CONTEXT that, among other things, encodes the presuppositions associated with using a given sign. The manner in which such presuppositions introduced by lexemes combine to give the presuppositions of the sentence remains to be worked out, but as properties of signs, they are a matter of grammar.

I believe that the results of the experiment reported in Bender 2005 suggest that the social meaning of copula absence/presence in AAVE is an entity similar to the social meaning of honorifics in Korean or Japanese. That is, the contrast in form corresponds to a contrast in social meaning. The obvious difference is that, like much truth-conditional meaning, the proper use of honorifics is a matter of formal study in these cultures. In this respect, the meaning of honorifics can be expected to behave more like a truth-conditional meaning in its currency among large groups of people and in its inertia with respect to local renegotiation. However, I conjecture that a corpus study of honorifics in either language would turn up internal constraints that, in turn, interact with the social meaning much like what I found in AAVE.

A second argument comes from the idea that the conventions of language are rooted in communicative practice. Given that social meanings are part of our experience of language, why should we expect them to be stripped away (cf. Langacker 1987:63)? This argument presupposes a view of language acquisition that is at odds with that of Principles and Parameters theory. Rather than seeing language acquisition as a matter of discrete parameter setting in response to triggers in the language learner’s experience, the view adopted here

emphasizes gradual accumulation of knowledge of language through experience and generalization over that experience. However, regardless of the theory of acquisition of syntax, acquisition of lexical knowledge—detailed and language-specific as it is—must follow more of a gradual accumulation model. Thus this argument should go through at least for social meaning attached to individual lexical items, even if syntax is acquired more discretely.

5.2 Overspecified types

The second extension proposed in §4 was linguistic prefabs or overspecified types. These appear necessary if frequentistic information and/or modulations in social meaning attach to structures larger than the minimal pieces needed in a grammar strictly concerned with grammaticality. The example in the AAVE case study involves precomposed bits of linguistic structure combining a lexical entry, a lexical rule, and some additional constraints (on the part of speech of the resulting word’s complement). Here I review and present arguments that these, too, should be integrated with the rest of the competence model.

Most theories of syntax exhibit rich deductive structure, where relatively spare stored units combine to derive relatively elaborate syntactic structures. However, as Henderson (1989), Jurafsky (1996) and others have noted, psychological models tend to view storage as relatively cheap and processing as relatively costly.¹² This suggests an alternative model where speakers store larger building blocks from which to create utterances, as well as generalizations over those building blocks which assist in the creation and interpretation of more novel utterances. Just such a model has been proposed by Langacker (1987, 1990, 2000) as the usage-based model (see also Barlow and Kemmer 2000).

In HPSG terms, the general picture is as illustrated in Figure 5. In this schematic picture, the top part represents the minimal grammar—those types that are required for modeling the bare distributional facts (grammaticality judgments) usually targeted by syntactic analyses. However, the more specified types can also fit into the same hierarchy as further specializations of the existing types. Examples are frequently used phrases like *I don’t know* (discussed more below), the canonical forms of idioms and collocations (see Riehemann 2001), and the complement-specified subtypes of the copula proposed here.

Now, it should be possible in principle to “prune” the hierarchy shown in Figure 5 to produce the minimal grammar. However, since all of the information in this hierarchy is represented in a uniform format, it is difficult to see why speakers would partition their knowledge of language along that pruning line. Thus if speakers can be shown to have knowledge of more specific types, as is predicted by the usage-based model, then there is reason to believe that that knowledge is a matter of grammar. We now turn to two kinds of evidence that speakers have this kind of knowledge: the existence of collocations and distributional facts about the domain of certain linguistic rules.

It would seem that the very existence of collocations is evidence of storage of redundant types in the grammar. Examples include the following:

- (5) avoid *x* like the plague

¹²See also Bresnan 1978:14.

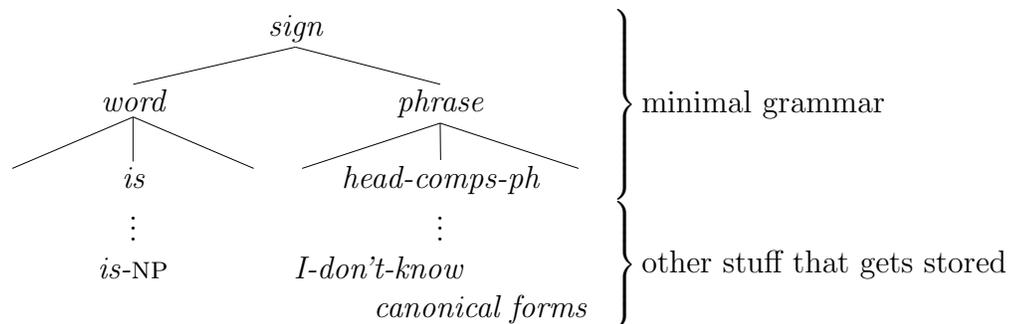


Figure 5: Extended grammar

- (6) everything you always wanted to know about x but were afraid to ask
- (7) get up on the wrong side of the bed

These phrases all sound familiar, and indeed are listed in collocation dictionaries (e.g., Kirkpatrick 1996). Further, (8), while clearly meaning the same as its more idiomatic counterpart, sounds odd, or at least distinctly less typical.

- (8) avoid x like a tax audit

Likewise, (7) is subject to dialect variation, so that for some speakers (9) sounds far more natural (and for others vice versa).

- (9) get out of bed on the wrong side

Anyone making such a judgment must have a representation of the collocations as phrasal chunks. But such a representation would be redundant in a minimal grammar because the phrases can certainly be built up out of smaller pieces.

An example of distributional evidence for extra types comes from Bybee and Scheibman's (1999) production study of the reduction of *don't* to /rō/ or /ə/. They found that this reduction occurred by far the most frequently in common, fixed phrases containing *don't*, such as *I don't know* or *why don't you*. From a strictly distributional point of view, this indicates that these phrases must be stored as linguistic units in order to be the domain of application of the *don't* reduction rule. If the common phrases are stored as units, the fact that *don't* reduction occurs here and not elsewhere (or, eventually, here more than elsewhere) acquires a natural explanation: If these units constitute single production routines that are recognizable as single chunks by addressees, there is less pressure to articulate each piece clearly. Further, frequent use of these units on the part of speakers could result in phonetic change that doesn't spread to other instances of the word *don't*. Further evidence for the existence of such over-specified types comes from the fact that social meaning can attach to them. For example, Penelope Eckert (p.c.) suggests that /aõno/ is a socially meaningful way of pronouncing *I don't know*.

5.3 Probabilities

The third and final proposed extension is the inclusion of probabilistic or frequentistic information in the grammar. This has been proposed many times before. In my analysis, I invoke knowledge of probabilistic information in order to model the modulating effect of grammatical context on social meaning. In this section, I briefly discuss the long-standing debate about this point. It is abundantly clear at this point that speakers have and use knowledge of frequencies over various linguistic properties. The open question is the relationship between this knowledge and competence grammar. Again, is it integrated or separate? As the use of the information appears to be tightly integrated, if we are to build models of competence meant to be embedded into models of language use, it would seem that the representations should be tightly integrated as well.

Psycholinguistic studies of both perception (e.g., Trueswell et al. 1993, MacDonald 1994, Jennings et al. 1997) and production (e.g., Gahl and Garnsey 2004) have provided evidence that speakers' linguistic performance correlates with the corpus frequency of various linguistic items, in a variety of ways, including subcategorization preferences for particular verbs influencing the ease with which readers/listeners handle local ambiguity (i.e., potential garden path sentences) and the pronunciation of the verbs themselves in different subcategorization frames. Computational linguists (e.g., Jurafsky 1996, Levy 2005) have implemented systems which model human difficulty in processing certain sorts of sentences by incorporating probabilistic information attached to lexical items and syntactic structures.¹³

Newmeyer (2003) takes up the argument from the point of view of formal syntax, and argues that probabilities are not a matter of linguistic competence. This paper kicked off a lively discussion in the pages of *Language*, including Gahl and Garnsey 2004, Clark 2005, Meyer and Tao 2005, Laury and Ono 2005, Newmeyer 2005 and Newmeyer 2006. There is not room to summarize the whole discussion here, but I would like to address Newmeyer's (2006) objections to Gahl and Garnsey's (2004) argument that their production evidence establishes that speakers have knowledge of probabilities.

Gahl and Garnsey (2004), in a carefully constructed sequence of experiments, show that when speakers read the same sequence of words in different syntactic contexts, their pronunciation (specifically, *t/d* deletion and various duration measures) varies according to the probability of the verbs in question appearing in the syntactic context represented. More specifically, verbs which are biased towards taking direct objects showed fewer reduction effects in subcategorization frames where they took sentential complements instead, and vice versa. The stimuli in this experiment (which the subjects read aloud after reading silently), were carefully designed to control for plausibility effects: the results could not be chalked up to some sentences being more sensible than others.

¹³Jurafsky (2003) notes that there is only preliminary evidence for probabilities attaching to delexicalized syntactic structures (i.e., bits of phrase structure, as opposed to subcategorization frames for particular verbs). The particular probabilistic information at issue in modeling AAVE copula absence/presence likewise involves particular subcategorization frames for given lexical items (forms of the copula). It is an interesting open question for future research whether sociolinguistic variation can be used to construct an argument for probabilities attached to delexicalized phrase structures.

Newmeyer (2006) argues that just because the particular sentences tested were matched for plausibility does not mean that the effects in question aren't due to the underlying meanings. He writes:

I am quite sure that there are many more contexts in real life where the situation would dictate that I use the verb *understand* with a bare NP object than with a sentential complement and many more contexts where the situation would dictate that I use the verb *claim* with a sentential complement than with a bare NP object. In other words, the frequency effects observed in these sentences might well be a simple consequence of the real-world fact that we are more likely to understand ideas and people (encoded by NP) than propositions (encoded by S) and more likely to claim propositions than physical objects. (2006:400–401)

Newmeyer notes that “[i]t is a truism that the more often we do something, the faster we are able to do it”, but disputes that the evidence presented constitutes evidence of knowledge of anything, grammar or otherwise. He argues on the basis of an analogy to cooking (just because one might bake a blueberry pie faster the second time doesn't mean there's any reason to posit probabilities in the recipe), that “[a] more parsimonious hypothesis is that these effects are a simple consequence of our doing more familiar things (e.g., talking about claiming propositions) faster than less familiar things (e.g., talking about claiming objects).” (p.401).

This rhetorical move begs the question: what are the cognitive consequences of something (in particular, of certain kinds of talking) being familiar? If humans did not record in their cognitive states some trace of past experience, we could do something a thousand times without its becoming familiar. Clearly, if something is familiar, it is because doing it before (once or many times) has altered our internal representation of it. The only open question, to my mind, is whether the internal representations being altered by experience are the same as those that linguists intend to model with (cognitivist, competence) grammar, or if there is some principled way to separate the two kinds of knowledge. Newmeyer asserts that the recipe does not change as the baker gets more adept and efficient at creating the pie. But recipes are a poor analogy for grammatical knowledge: The former are explicit instructions where the latter involves tacit knowledge. Furthermore, language does change over time, and sociolinguistic work dating back to Weinreich et al. 1968 suggests that this change is not quantized through new generations of language learners but taking place throughout all age groups all the time.¹⁴

5.4 Summary

This section has reviewed arguments that all three of the proposed extensions involve information which should be integrated with the usual knowledge captured in syntactic competence models. A final argument which applies to all three is that this additional information needs to refer to units of grammatical information already described in the grammar. If

¹⁴For an opposing view, see Lightfoot 1991 and Kroch 2001.

social meaning, overspecified types and probabilistic knowledge are segregated into separate modules, then those modules will need to replicate much of the information in the grammar ‘proper’ (cf. Hudson 1996).

This brings us back to Newmeyer’s (2006)’s use of the term ‘parsimonious’ (in the quote above) and Langacker’s (1987)’s point about economy/parsimony only being a meaningful metric if the data to be accounted for are held constant. In a general model of language, responsible for not only grammaticality judgments but also language production and perception as well as the ways in which speakers manipulate linguistic variation to negotiate social meaning and construct social categories and the resulting linguistic change over time, we may find that the model as a whole is more parsimonious if its component representation of linguistic knowledge includes probabilistic information, as well as social meaning and linguistic prefabs.

6 The boundaries of grammar: Some heuristics and examples

In the preceding section, I have suggested that speakers have knowledge of three kinds of linguistic information not usually included in competence grammars—social meaning, overspecified types, and probabilities of a certain kind—and I explored possible evidence for including this knowledge in grammar. However, by opening up the boundaries of competence grammar to include these kinds of information, I do not mean to imply that grammar is essentially boundless. Clearly, there are kinds of knowledge that are not grammatical. I would like to propose that grammar includes just that knowledge of language that is used in linguistic processing. Definitive answers to whether a particular kind of knowledge is used in linguistic processing (and if so, whether it is used in an integrated fashion) require psycholinguistic tests. Here, however, I would like to discuss some possible heuristics for identifying the sort of linguistic knowledge likely to be used in processing. The discussion should serve to illuminate the commonalities between ‘standard’ grammatical knowledge and the additional information included in the extended competence model I have been arguing for here. §6.1 discusses the notion of arbitrariness or language-specificity, §6.2 considers language as a system of contrasts, §6.3 considers creativity of use, and §6.4 applies the notion of Heideggerian *thrownness* of use and acquisition.

6.1 Arbitrariness/language specificity

One fairly standard criterion for including something in a competence grammar is that of arbitrariness or language specificity (in the sense of specific to a given language). That is, linguistic facts that are conventional in Lewis’s (1969) sense—they could have been otherwise—must therefore be learned and in the grammar. This criterion is meant to distinguish grammatical facts from natural consequences of functional pressures. It leaks, however, in several respects.

The first problem is it really only works as a sufficient (not necessary) condition for inclusion in the grammar. That is, linguistic facts such as that the word for HORSE in French is *cheval* can be shown by this criterion to be in the grammar, but the non-arbitrariness of other facts does not preclude their grammaticization. That is, just because a pattern could follow from independent principles doesn't mean that speakers don't learn it as a pattern. Indeed, this process is probably what is behind many cases of grammaticization.

The second problem, related to the first, is that language is not as arbitrary as linguists often make it out to be (Langacker 1987). Examples of non-arbitrariness range from onomatopoeia and phonesthemes to all morphologically complex words. That is, given the meaning of a stem and some affixes, the meaning of a word composed of that stem and those affixes is generally not arbitrary.

A third problem arises from the process of iconization. If Irvine and Gal (2000) are right, then the connection between form and social meaning is non-arbitrary FOR THE SPEAKERS CONCERNED, even though it could well have been otherwise. For example, for many English speakers, copula absence and other features of AAVE are iconic of lack of education (or worse, low intelligence) in that they are believed to show it directly. Indeed, this kind of social value probably arises as an extension of the stigmatization (by outsiders) of AAVE speakers to the stigmatization AAVE features (again, in the minds of those outsiders). Thus, this kind of social meaning is non-arbitrary in its inception, although it is still arbitrary in the sense that it could have been otherwise. (For example, if AAVE had a different collection of constructions, or if AAVE was not stigmatized but rather exalted.)

To summarize, arbitrariness is at best a one-way criterion for inclusion in the grammar. Many grammatical facts are either not completely arbitrary or not experienced as arbitrary by speakers.

6.2 Structure: A system of contrasts

The second heuristic for identifying the boundaries of competence grammar I would like to consider is the extent to which linguistic knowledge forms a coherent system. Some authors (e.g., Langacker 1987, Langacker 1990, Hudson 1996) ask, Why should grammar be a separate module of knowledge? They suggest instead a view in which linguistic knowledge is embedded in other knowledge. I think that the motivation for distinguishing (some part of) knowledge of language (the grammar) comes from the fact that the system of contrasts gives internal coherence to grammar, even if it is embedded in a larger knowledge system. I am not asserting that all other kinds of knowledge lack the property of being systems of contrasts, only that the fact that linguistic forms enter into such contrasts with each other gives linguistic knowledge a kind of internal coherence.

This criterion can certainly be interpreted as placing social meaning in the grammar. Social meaning, like truth-conditional meaning, exploits contrasts in form. Further, Irvine (2001) argues that style is a matter of distinctiveness, i.e., contrasts between different groups of people in the social landscape. As building blocks of style, socially meaningful linguistic elements thus pair contrasts in form with contrasts in social space.

One might also be tempted to talk about certain probabilities as a matter of contrasts.

That is, copula presence and copula absence are necessarily complementary in their frequency of occurrence because the variable is defined in such a way that choosing one precludes the other. A similar thing could be said for frequencies of valence patterns and the like. However, in this case, it seems that the contrast in frequency merely follows from the contrast in form or in valence. As far as I know, nothing else builds on the contrast in the way that contrasts in social space build on or attach to contrasts in form.

6.3 Creativity of use

Another possible criterion is generativity in the sense of creativity of use. Knowledge of syntax and semantics allows speakers to create and understand novel utterances. If other knowledge about language functions in a similar way, by this criterion it should be in the grammar.

Individuals create their own styles out of existing symbolic resources (linguistic and otherwise) (Hebdige 1979, California Style Collective 1993, Bucholtz and Hall 2005). Hebdige dubs this process *bricolage*, a French word which evokes a picture of building something out of whatever resources and raw materials that one can find. Socially meaningful forms can thus combine to create novel styles in a manner that is analogous at a certain level of abstraction to the way that lexemes and constructions combine to create novel utterances (though the rules of compositionality for social meaning clearly differ from those for truth-conditional meaning). This is equally true of linguistic as of non-linguistic symbolic resources. The difference is that linguistic stylistic resources are deployed in the confines of the linguistic signal, and, as argued above, find their symbolic value sensitive to surrounding contrasts in linguistic form. Thus linguistic stylistic resources display creativity of use that is parallel to the creativity of use of linguistic resources in general, and, unlike non-linguistic stylistic resources, constrained by or intertwined with other linguistic facts.

It is somewhat less clear how to apply this criterion to non-categorical constraints and overspecified types, but it might be seen to suggest that both belong in the grammar. As for non-categorical constraints, inasmuch as their effect on social meaning allows them to be a tool in the construction of style, they should be considered grammatical for the same reason that social meaning is. With over-specified types, possible evidence for including them by this criterion could come from evidence of people making reference to collocations in the creation of new utterances. Merely using collocations doesn't provide any evidence with respect to this criterion as, by definition, collocations are semantically compositional and can be built up with smaller pieces. However, making reference to collocations means involving them in some sense in the production of novel utterances. An example would be Mastercard's 'Priceless' marketing campaign, consisting of more than 160 ads over 9 years (Stevenson 2006), all using the template (*NP: _ dollars.*) + *NP: priceless.* After 160 commercials, this formula is well-known and thus eminently available for parody. Another example is Kraft's marketing of a macaroni and cheese product as *cheese and macaroni*. The unusual placement of *cheese* as the first conjunct is meant to suggest especially cheesy macaroni and cheese. This effect of this suggestion comes partially from the contrast with the usual way of saying it, that is it relies on the knowledge that one usually says *macaroni*

and cheese.

6.4 Heideggerian *thrownness* of use/acquisition.

A fourth property which potentially distinguishes linguistic competence from other kinds of knowledge about language, is the fact that linguistic competence is acquired and used in a state of Heideggerian *thrownness*. Winograd and Flores gloss Heideggerian *thrownness* as “the condition of understanding in which our actions find some resonance or effectiveness in the world” (1986:32) and remark:

[T]he essential feature of language activity (the processes of saying and listening) is the thrownness of a person within language. When we are engaged in successful language activity, the conversation is not present-at-hand, as something observed. We are immersed in its unfolding. Its structure becomes visible only when there is some kind of breakdown. (1986:68)

Thrownness contrasts with detached contemplation. Isolated and categorized representations of objects are creatures of detached contemplation. Most language use, however, is entirely transparent. In conversation people pay attention to the messages they are exchanging and not the details of how they are said. Indeed, it takes some training for linguists to learn to pick out examples of phenomena of interest to us from the flow of language use that we encounter in our everyday lives. A “breakdown”, in this context, is some kind of problem in the functioning of a tool, such as language. Experiencing a misunderstanding can cause even non-linguists to focus on the words and structures used, to produce and categorize (conscious) representations of the problematic linguistic elements.

By this criterion, the three types of knowledge described in Section 5 above are similar to (other) grammatical knowledge. As Ochs (1992) notes, we learn, comprehend and deploy socially meaningful contrasts in language without paying much attention. That is, it is enough to decide to be informal, or to be polite, or be distant. As competent speakers, we need not pay attention to the individual linguistic choices that implement such decisions. Likewise, we acquired the relevant knowledge with the same seeming effortlessness as the rest of our grammars. Similarly, whatever frequentistic information we’ve stored is clearly not the result of formal study or conscious tabulation (cf. Haser and Chromiak 1977). Indeed, part of the reason some authors have doubted that people store frequentistic information is the sheer difficulty of learning it by these other means. Finally, although sayings and proverbs are studied in school in some cultures (notably in Chinese-speaking cultures), this is not the case for the extra types representing such collocations as *avoid like the plague*.

What this criterion does rule out is knowledge of language that is not used in language processing. For example, take the fact that *honcho* in *head honcho* is a borrowing from Japanese. Speakers of English who know this fact about English know it because they were told so explicitly, or because they also speak Japanese and noticed the connection. Both cases involve focusing attention on the word in question, and thus do not constitute *thrown* acquisition. Likewise, it does not seem likely that this knowledge could naturally figure

in language processing, and certainly not without treating it with the kind of detached contemplation that is inconsistent with thrownness. I picked an example from Japanese and not, say, French, because something like the knowledge that *je ne sais quoi* or *raison d'être* are borrowings from French does seem to figure more subtly into their use. French has a certain *cachet*, i.e., air of sophistication, among many English speakers, and a speaker going for this sociolinguistic effect might quite naturally reach for a French phrase. Whether this process is as *thrown* as the rest of language use is open to investigation.

6.5 Summary

Previous thought on what belongs ‘in the grammar’ seems to take as its primary criterion or heuristic the notion of parsimony (alternatively, Ockham’s Razor). As noted in §3.2, the primary data to be accounted for have been taken to be grammaticality judgments. Therefore, any linguistic information that does not impact distributional facts (i.e., grammaticality) is automatically out: it complicates the theory with no ‘empirical’ gain. I have argued, however, that for cognitivist theories of language, the focus on grammaticality judgments is artificially narrow. In attempting to study the actual linguistic knowledge of speakers, we cannot define the the range of knowledge we will study in an *ad hoc* fashion and then declare that that knowledge forms an isolated module.

A more sensible way forward would be to first explore whether speakers have knowledge of various kinds of linguistic facts (including, for example, non-categorical constraints on sociolinguistic variation) and then ask if that knowledge forms an integrated whole with other linguistic knowledge, or if it is best modeled as separate. The kind of model in question would necessarily be a psycholinguistic model of language acquisition, production and comprehension. Until we have a comprehensive model of this type, this debate remains a matter of philosophy. I hope the discussion of heuristics above represents an interesting contribution to that debate.

7 Conclusion

In this paper, I have started from the point of view of performance-plausible sign-based grammar, and explored the implications of sociolinguistic variation for theories of competence grammar. The particular data I’ve been focusing on come from a case study AAVE copula absence, and in particular, preliminary experimental results indicating that speakers with sufficient experience with AAVE evaluate the social value of copula presence (and sometimes absence) differentially depending on whether the syntactic environment favors the variant in question. Given this interaction of syntactic constraints and social meaning, I have argued that a full account of sociolinguistic variation will only be possible if we take social meaning seriously. I have attempted to take some steps down that path by synthesizing research on social meaning, exploring how socially meaningful syntactic variation could be handled in sign-based grammar, and then reflecting on the consequences these moves bring for the boundaries we draw around competence grammar. I conclude that it is possible for a theory

of language to be both social and cognitivist, but that such a theory must not define the object of study *a priori*, as the interactions between language and social reality are only clear in the broader picture.

References

- Adger, David. 2006. Combinatorial variation. *Journal of Linguistics*.
- Anttila, Arto. 1997. Deriving variation from grammar. In F. Hinskens, R. van Hout, and W. L. Wetzels (Eds.), *Variation, Change and Phonological Theory*, 35–68. Amsterdam: John Benjamins.
- Bakhtin, Mikhail Mikhaïlovich. 1986. *Speech Genres and Other Late Essays*. Austin: University of Texas Press. Translated by Vern W. McGee. Edited by Caryl Emerson and Michael Holquist.
- Barlow, Michael, and Suzanne Kemmer. 2000. *Usage Based Models of Language*. Stanford, CA: CSLI.
- Bender, Emily M. 2001. *Syntactic Variation and Linguistic Competence: The Case of AAVE Copula Absence*. PhD thesis, Stanford University.
- Bender, Emily M. 2005. On the boundaries of linguistic competence: Matched-guise experiments as evidence of knowledge of grammar. *Lingua, Special Issue on Data in Syntax, Semantics, and Pragmatics* 1579–1598.
- Benor, Sarah B. 2001. The learned /t/: Phonological variation in Orthodox Jewish English. In T. Sanchez and D. Johnson (Eds.), *Penn Working Papers in Linguistics: Selected papers from NWAV 29*, 1–16.
- Boersma, Paul. 1998. *Functional Phonology. Formalizing the interaction between articulatory and perceptual drives*. PhD thesis, University of Amsterdam.
- Bresnan, Joan. 1978. A realistic transformational grammar. In M. Halle, J. Bresnan, and G. Miller (Eds.), *Linguistic Theory and Psychological Reality*, 1–59. Cambridge, MA: MIT Press.
- Bresnan, Joan, Shipra Dingare, and Christopher D. Manning. 2001. Soft constraints mirror hard constraints: Voice and person in english and lummi. In M. Butt and T. H. King (Eds.), *Proceedings of the LFG01 Conference*, Stanford, CA. CSLI Publications.
- Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge University Press: Cambridge.
- Bucholtz, Mary, and Kira Hall. 2005. Identity and interaction: a sociocultural linguistic approach. *Discourse Studies* 7(4–5).

- Butler, Judith. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.
- Bybee, Joan, and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37(4):575–596.
- California Style Collective. 1993. Variation and personal/group style. Unpublished ms., Paper presented at NWAVE-22.
- Campbell-Kibler, Kathryn. 2006. *Listener Perceptions of Sociolinguistic Variables: The Case of (ING)*. PhD thesis, Stanford University.
- Cedergren, Henrietta J., and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50(2):333–355.
- Chomsky, Noam. 1955. The logical structure of linguistic theory. Unpublished ms., Published 1975 by New York: Plenum.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Clark, Brady. 2005. On stochastic grammar. *Language* 81:207–217.
- Clark, Herbert H. 1996. *Using Language*. Cambridge: Cambridge University Press.
- Coupland, Nikolas. 2001. Language, situation and the relational self: Theorizing dialect-style in sociolinguistics. In P. Eckert and J. R. Rickford (Eds.), *Style and Sociolinguistic Variation*, 185–201. Cambridge: Cambridge University Press.
- Culy, Christopher. 1996. Null objects in English recipes. *Language Variation and Change* 8:91–124.
- Eckert, Penelope. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1:245–267.
- Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Oxford: Blackwell.
- Eckert, Penelope. 2005. Variation, convention, and social meaning. Unpublished ms., Paper Presented at the Annual Meeting of the Linguistic Society of America. Oakland, CA.
- Eckert, Penelope, and Sally McConnell-Ginet. 1992. THINK PRACTICALLY AND LOOK LOCALLY: Language and gender as community-based practice. *Annual Review of Anthropology*.
- Gahl, Suzanne, and Susan M. Garnsey. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80(4):748–775.
- Gumperz, John J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.

- Haddican, Bill. 2003. Dialect contact in a Southern Basque town. *Language Variation and Change* 15:1–35.
- Hall, Kira. 1997. "go suck your husband's sugarcane!": Hijras and the use of sexual insult. In A. Livia and K. Hall (Eds.), *Queerly Phrased: Language, Gender, and Sexuality*, 430–460. New York: Oxford University Press.
- Halliday, M. A. K. 1991. Corpus studies and probabilistic grammar. In K. Aijmer and B. Altenberg (Eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, 30–43. Longman.
- Harder, Peter. 1996. *Functional Semantics: A Theory of Meaning, Structure and Tense in English*. Berlin: Mouton de Gruyter.
- Harder, Peter. 2003. The status of linguistic facts: Rethinking the relation between cognition, social institution, and utterance from a functional point of view. *Mind and Language* 18(1):52–76.
- Haser, I., and W. Chromiak. 1977. The processing of frequency: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior* 16:173–184.
- Hebdige, Dick. 1979. *Subculture: The Meaning of Style*. London: Routledge.
- Henderson, L. 1989. On mental representation of morphology and its diagnosis by measures of visual access speed. In W. Marslen-Wilson (Ed.), *Lexical Representations and Processes*, 357–391. Cambridge, MA: MIT Press.
- Hudson, Richard A. 1996. *Sociolinguistics*. Cambridge: Cambridge University Press. Second edition.
- Irvine, Judith T. 2001. "Style" as distinctiveness: The culture and ideology of linguistic differentiation. In P. Eckert and J. R. Rickford (Eds.), *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Irvine, Judith T., and Susan Gal. 2000. Language ideology and linguistic differentiation. In P. V. Kroskrity (Ed.), *Regimes of Language: Ideologies, Politics, and Identities*, 35–83. Santa Fe, NM: SAR Press.
- Jennings, F., B. Randall, and L.K. Tyler. 1997. Graded effects of verb subcategory preferences on parsing: Support for constraint-satisfaction models. *Language and Cognitive Processes* 12:485–504.
- Jurafsky, Dan. 2003. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay, and S. Jannedy (Eds.), *Probabilistic Linguistics*, 39–96. Cambridge, MA: MIT Press.

- Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20:137–194.
- Kemmer, Suzanne, and Michael Israel. 1994. Variation and the usage-based model. *CLS* 30:165–179.
- Kirkpatrick, Betty. 1996. *Clichés*. New York: St. Martin's Griffin.
- Kroch, Anthony. 2001. Syntactic change. In M. Baltin and C. Collins (Eds.), *The Handbook of Contemporary Syntactic Theory*, 699–729. Oxford: Blackwell Publishers.
- Labov, William. 1963. The social motivation of a sound change. *Word* 19:273–309.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4).
- Lambert, Wallace E., Howard Giles, and Omer Picard. 1975. Language attitudes in a French-American community. *Linguistics* 158:127–152.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, Vol I: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, Ronald W. 1990. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. New York: Mouton de Gruyter.
- Langacker, Ronald W. 2000. A dynamic usage-based model. In M. Barlow and S. Kemmer (Eds.), *Usage-Based Models of Language*, 1–63. Stanford: CSLI.
- Laury, Ritva, and Tsuyoshi Ono. 2005. Data is data and model is model: You don't discard the data that doesn't fit your model! *Language* 81:218–225.
- Lavandera, Beatriz R. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7:171–182.
- Levy, Roger. 2005. *Probabilistic models of word order and syntactic discontinuity*. PhD thesis, Stanford University.
- Lewis, David K. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Lightfoot, David. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: The MIT Press.
- Lowry, Orla. 2002. The stylistic variation of nuclear patterns in Belfast English. *Journal of the International Phonetic Association* 32:33–42.

- MacDonald, Maryellen C. 1994. Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes* 9:157–201.
- McConnell-Ginet, Sally. 1984. The origins of sexist language in discourse. In S. J. White and V. Teller (Eds.), *Discourses in Reading and Linguistics*, 123–135. New York: New York Academy of Sciences.
- McConnell-Ginet, Sally. 1989. The sexual (re)production of meaning: A discourse-based theory. In F. W. Frank and P. A. Treichler (Eds.), *Language, Gender, and Professional Writing*, 35–50. MLA.
- Meyer, Charles F., and Hongyin Tao. 2005. Response to Newmeyer’s ‘Grammar is grammar and usage is usage’. *Language* 81:226–228.
- Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79:682–707.
- Newmeyer, Frederick J. 2005. A reply to the critiques of ‘Grammar is grammar and usage is usage’. *Language* 81:229–236.
- Newmeyer, Frederick J. 2006. On Gahl and Garnsey on grammar and usage. *Language* 82:399–404.
- Ochs, Elinor. 1992. Indexing gender. In A. Duranti and C. Goodwin (Eds.), *Rethinking Context: Language as an Interactive Phenomenon*, 335–358. Cambridge: Cambridge University Press.
- Okushi, Yoshiko. 1997. *Patterns of Honorific Use in the Everyday Speech of Four Japanese Women*. PhD thesis, University of Pennsylvania.
- Podesva, Robert J., Sarah J. Roberts, and Kathryn Campbell-Kibler. 2001. Sharing resources and indexing meanings in the production of gay styles. In K. Campbell-Kibler, R. J. Podesva, S. J. Roberts, and A. Wong (Eds.), *Language and Sexuality: Contesting Meaning in Theory and Practice*, 175–189. Stanford, CA: CSLI Publications.
- Podesva, Robert Jay. 2006. *Phonetic Detail in Sociolinguistic Variation: Its Linguistic Significance and Role in the Construction of Social Meaning*. PhD thesis, Stanford University.
- Pollard, Carl, and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: Chicago University Press.
- Post, Emil. 1944. Recursively enumerable sets of positive integers and their decision problems. *Bulletin of the American Mathematical Society* 50:284–316.

- Pullum, Geoffrey K., and Barbara C. Scholtz. 2005. Contrasting applications of logic in natural language syntactic description. In P. Hájek, L. Valdés-Villanueva, and D. Westerståhl (Eds.), *Logic, Methodology and Philosophy and Science: Proceedings of the Twelfth International Congress*, 481–503. London: King's College Publications.
- Rickford, John R. 1998. The creole origins of African-American vernacular English: Evidence from copula absence. In S. S. Mufwene, J. R. Rickford, G. Bailey, and J. Baugh (Eds.), *African-American English*, 154–200. New York: Routledge.
- Riehemann, Susanne. 2001. *A Constructional Approach to Idioms and Word Formation*. PhD thesis, Stanford University.
- Sag, Ivan A. 1992. Taking performance seriously. In C. Martin-Vide (Ed.), *VII Congreso de Lenguajes Naturales y Lenguajes Formales*.
- Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI. Second edition.
- Searle, John R. 1995. *The Construction of Social Reality*. New York: The Free Press.
- Soames, Scott. 1984. Linguistics and psychology. *Linguistics and Philosophy* 7:155–179.
- Stevenson, Seth. 2006. The end of a played-out ad campaign? priceless. *Slate.com*. March 13, 2006; <http://www.slate.com/id/2137884/>.
- Trueswell, John C., Michael K. Tanenhaus, and Christopher Kello. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition* 19:528–553.
- Wasow, Thomas. 1978. On constraining the class of transformational languages. *Synthese* 39.
- Weinreich, Uriel, William Labov, and Martin I. Herzog. 1968. Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (Eds.), *Directions for Historical Linguistics: A Symposium*, 95–188. Austin: University of Texas Press.
- Winograd, Terry, and Fernando Flores. 1986. *Understanding Computers and Cognition: A New Foundation for Design*. New York: Addison-Wesley.
- Wolfram, Walter A. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Center for Applied Linguistics.
- Zwicky, Arnold M. 1999. The grammar and the user's manual. Unpublished ms., 1999 Linguistic Institute Forum Lecture.