English Language and Linguistics 11.2: 301–346.
doi:10.1017/S1360674307002274 Printed in the United Kingdom

Typology in variation: a probabilistic approach to *be* and *n't* in the *Survey of English Dialects*[1]

JOAN BRESNAN
Stanford University

ASHWINI DEO
Yale University

and

DEVYANI SHARMA
King's College London

(Received 30 August 2006; revised 6 March 2007)

Variation within grammars is a reflection of variation between grammars.[2]

Subject agreement and synthetic negation for the verb *be* show extraordinary local variation in the *Survey of English Dialects* (Orton et al., 1962–71). Extracting partial grammars of individuals, we confirm leveling patterns across person, number, and negation (Ihalainen, 1991; Cheshire, Edwards & Whittle, 1993; Cheshire, 1996). We find that individual variation bears striking structural resemblances to invariant dialect paradigms, and also reflects typologically observed markedness properties (Aissen, 1999). In the framework of Stochastic Optimality Theory (Boersma & Hayes, 2001), variable outputs of individual speakers are expected to be constrained by the same typological and markedness generalizations found crosslinguistically. The stochastic evaluation of candidate outputs in individual grammars reranks individual constraints by perturbing their ranking values, with the potential for stable variation between two near-identical rankings. The stochastic learning mechanism is sensitive to variable frequencies encountered in the linguistic environment, whether in geographical or social space. In addition to relating individual and group dialectal variation to typological variation (Kortmann, 1999; Anderwald, 2003), the findings suggest that an individual grammar is sensitively tuned to frequencies in the linguistic environment, leading to isolated loci of variability in the grammar rather than complete alternations of paradigms.

A characteristic of linguistic variation that has emerged in distinct fields of enquiry is that variation within a single grammar bears a close resemblance to variation across grammars. Sociolinguistic studies, for instance, have long observed that 'variation within the speech

[1] We wish to thank Lieselotte Anderwald, Brady Clark, Richard Coates, Andrew Garrett, Jane Grimshaw, Bruce Hayes, Wouter Kusters, Hanjung Lee, Roger Levy, Chris Manning, Elizabeth Traugott, two anonymous reviewers, and audiences at the OT Symposium of the English Linguistic Society of Japan (Kobe, November 2000), the Bay Area Typology Workshop (University of California at Berkeley, March 2001), the 2003 Nijmegen Lectures (Max-Planck Institute, Nijmegen, December 2003), the University of Sussex (February 2005) and ICLCE1 (University of Edinburgh, June 2005) for many comments which have improved earlier versions of this work. We remain solely responsible for all errors of fact and interpretation. This work is based in part on work supported by the National Science Foundation under Grant No. BCS-9818077.

[2] Emmon Bach's characterization of the theme of Bresnan's OT lectures at the Vilem Mathesius Lecture Series 13 in Prague, November 1998.

of a single speaker derives from the variation which exists between speakers' (Bell, 1984: 151). In the present study, individual patterns of variation in subject–verb agreement with affirmative and negative *be* extracted from the *Survey of English Dialects* (*SED*, Orton et al., 1962–71) show striking structural resemblances to patterns of interdialectal, or categorical, variation.

Recent developments in Optimality Theory (Boersma, 1997; Anttila, 1997; van Oostendorp, 1997; Nagy & Heap, 1998; and others) have problematized the assumption that the phenomena of variation – variable outputs for the same input – must be external to formal grammatical theory. In the framework of Stochastic Optimality Theory (Stochastic OT; Boersma, 1997, 1998, 1999a; Boersma & Hayes, 2001), for example, it is expected that variable outputs across dialects and within individual speakers should be constrained by the same kinds of typological generalizations that are found crosslinguistically. Typological variation across languages is explained in OT by means of language-particular rankings of universal constraints, and variation across dialects should thus derive from the same typological space. In Stochastic OT, the noisy evaluation of candidates reranks constraints by temporarily perturbing their ranking values along a continuous scale; this inherent variability in grammars may lead to either categorical or variable grammars depending on the environment a speaker is exposed to. In this framework, therefore, both dialectal variation and individual variation sample the typological space of possible grammars.

1 Background

1.1 Previous work

A number of studies have examined verb agreement patterns in nonstandard varieties of English (Ihalainen, 1991; Cheshire, 1991, 1996; Cheshire, Edwards & Whittle, 1993; Schilling-Estes & Wolfram, 1994; Anderwald, 2001, 2002, 2003). Many of these studies have observed a reduction of variation with plural (vs. singular) subjects and negative (vs affirmative) sentences.[3]

Leveling of distinctions in paradigms of *be* with plural subjects is widespread, and is also instantiated in Standard English, which assigns the form *are* to all plural subjects. Cheshire (1991: 55) observes that in many nonstandard dialects of English leveling across number and person results in either the present tense *-s* suffixed form of verbs or the suffixless form of verbs generalizing across verbal paradigms. Trudgill & Chambers (1991: 52), Cheshire et al. (1993: 73), and Trudgill (1999: 104) also observe that the negative counterparts of present tense *be* paradigms in many modern nonstandard dialects of British English have reduced distinctions and employ just one form, *ain't*, for the negative present tense of both auxiliary *be* and auxiliary *have*. In many of these varieties, this single form covers all

[3] Another type of leveling in *be* inventories involves a reduction of variation in past tense marking relative to present tense marking (Cheshire et al., 1993: 71–2; Schilling-Estes & Wolfram, 1994: 280; Trudgill, 1999: 106; Anderwald, 2003: 520). We restrict the present study to present tense inventories, but the pattern of leveling in past tense would be straightforwardly subsumed under the analysis here, as past tense morphology can also be seen as marked in ways similar to plural and negative morphology.

subject persons and numbers, despite the fact that the affirmative paradigms for these two auxiliary verbs retain person and number distinctions. Schilling-Estes & Wolfram (1994: 287) note that some nonstandard varieties of American English that have leveling of *be* distinctions in the past tense also restrict this leveling to negative sentences.

These patterns of dialect variation have recently been related to typological markedness by Kortmann (1999), Anderwald & Kortmann (2002), and Anderwald (2003). Studies in typology have shown that contrasts are often categorically neutralized across languages in marked contexts, and many of the grammatical contexts in which British dialects exhibit leveling correspond to marked grammatical categories: plural number, negation, and past tense.

Our goals in this study are twofold: first, we aim to verify whether variation in affirmative and negative leveling in English dialects does indeed reflect more general typological patterns, and if so, why; second, we offer a unified formal analysis of variable leveling in the grammars of dialects as well as of individuals, using a probabilistic model.

Following a description of the data extraction methodology used, we first present a summary of all categorical affirmative and negative *be* paradigms (interspeaker variation) and present an analysis of this space of variation. Next, we present a summary of all variable affirmative and negative *be* paradigms (intraspeaker variation) and offer a Stochastic OT analysis of individual variation. As the data do not include frequency distributions, they do not make full use of the Stochastic OT apparatus; however, we adopt Stochastic OT as a useful conceptual and theoretical model of localized, individual variation.

1.2 Data extraction from the Survey of English Dialects

Although *be* variation is attested in many varieties of English, the dialects of England may exhibit the widest variety of *be* inventories (Schilling-Estes & Wolfram, 1994: 277), and this was our motivation for selecting the *Survey of English Dialects* (Orton et al., 1962–71) as a data source.

We should note that the *SED* was compiled during the 1950s (first published in 1962 for the University of Leeds) and thus constitutes a relatively old data source. Some studies have attempted to relate *SED* findings to more recent survey work. For instance, Cheshire et al. (1993) compare the *SED* to the Survey of British Dialect Grammar (conducted 1986–9) and Anderwald (2003) briefly compares the *SED* to the British National Corpus (completed in 1994). The primary finding of both comparisons is that selected features which were originally regional have spread to many urban areas and now constitute a set of generalized nonstandard urban British dialect features, while other traditional regional features are being lost. As we are specifically concerned with the typological range of possible paradigms of *be*, a slightly earlier stage of regional variation is no less appropriate for study than a more contemporary one, and as the *SED* offers explicit and organized detail of over 300 individual grammatical systems along with their regional groupings, it lends itself particularly well to an examination of intra- and inter-group variation.

The questionnaire data in the *SED* are organized by county and survey question, but also include an index of individual respondents for each set of responses to a given question. To

Figure 1. Counties of England

extract partial grammars for each individual, we entered all of the responses to questions that elicited present-tense forms of the verb *be* into a database, collapsing the fine-grained phonetic variations in pronunciation recorded in the transcriptions into an orthographic representation of distinct morphosyntactic forms (see appendix A for a list of the relevant *SED* questions).

In the construction of this database, we coded for construction type (interrogative/tag/declarative, with/without ellipsis, affirmative/negative), predicate type, subject person, subject number, region, and site/speaker. Figure 1 shows the regional divisions used in the *SED* and appendix B gives a list of abbreviations used for these regions. Assuming a 'grammar' to be a set of construction types used by an individual, the total number of individual grammars present in the *SED* is 312.[4]

[4] Individual data points in the *SED*, e.g. Sr5, usually represent responses by one individual; however, in a few cases they represent the composite responses of two or three demographically similar individuals from a single locality. It

For the present study we used a subset of each grammar, restricting our attention to affirmative declarative constructions and their synthetic negation counterparts and excluding from the present analysis other forms of positional variation such as *wh-*, yes/no, or tag question formation. In order to isolate individual partial grammars for declarative clauses, we sorted the data by respondent and construction type.

Some speakers in the *SED* have fixed paradigms for *be* with pronominal subjects and these speakers comprise the set of invariant inventories. Other speakers give multiple answers for a single subject type, and these individuals form the group of variable inventories. We classified speakers with identical paradigms, whether invariant or variable, as sharing a single inventory. Each inventory discussed in the article thus represents the grammar of an individual speaker or a group of speakers from whom the same input/output pairs were elicited.

Because of systematic gaps in the *SED* survey questionnaires, the following subject types were the maximum possible data extractable for a given speaker:

Affirmative declarative:	singular:	1sg, 2sg, 3sg
	plural:	1pl, 3pl
Negative declarative:	singular:	1sg, 3sg
	plural:	3pl

Aside from these intrinsic constraints on the *SED* data, we were obliged to impose two additional criteria on the initial data set in order to ensure a reliable basis for comparison of dialect systems. Dialect inventories were only included for analysis if (a) the inventory had a complete set of affirmative and synthetic negative forms recorded and (b) each combined affirmative and synthetic negative paradigm was attested in an identical form for at least two speakers.

According to the first criterion, any speaker with an incomplete affirmative or negative paradigm was omitted. For the affirmative part of speakers' *be* paradigms, this simply applied to speakers for whom a form had not been recorded by the fieldworker in one or more of the cells. The criterion is slightly more specific in the case of speakers' negative paradigms. The *SED* includes either synthetic negation such as *isn't* or *ain't*, analytic negation such as *am not* or *'s not*, or both synthetic and analytic forms. The hypothesis in the present article regarding leveling only applies to synthetic forms, as the claim pertains to overloading of a single lexical form with multiple semantic features such as negation, person, and number. As analytic negation such as *am not* or *'m not* reserves separate morphemes for the marking of nominal features and negation, leveling is not predicted for such constructions. Complete synthetic negation paradigms are thus needed to test our prediction, so speakers for whom only analytic negation or incomplete synthetic negation had been recorded in the *SED* were excluded, as we could not verify which synthetic negation forms they would favor for different subject types. This first criterion reduced the total number of individuals included in the study to 216.

would be slightly more accurate to refer to these points as localities rather than individuals, but as we are discussing regions as well, we retain the term 'individual' in referring to distinct data points collected in a given region.

The second criterion was designed to isolate patterns in the *SED* data that are reliably systematic. In the present article we are primarily interested in systematic and stable dialect paradigms. Although Stochastic OT grammars can model a certain degree of noise and instability, evident during periods of massive constraint reranking, they can also model the stable systems that speakers may ultimately converge on and they make typological predictions about these. As we are interested in the typology of stable dialect paradigms, we sorted all the *SED* speakers into groups that shared affirmative *and* synthetic negative paradigms and omitted speakers that had unique or idiosyncratic paradigms, treating their data as less reliable. As a result, the subset of data analyzed includes all speakers who share their affirmative and negative declarative paradigms with at least one other speaker.

The only exception to the second criterion is the inclusion of two invariant inventories that are represented by only one speaker each in the *SED*: Kent (speaker K7) and Sussex (speaker Sx5). We include these two inventories as other research in these regions has shown evidence of these two paradigms having once been robust systems.[5]

The total number of speakers remaining after both selection criteria were applied was 119. No other individuals were excluded, so the group represents the entire *SED* data that conform to the two criteria. These speakers were separated into two groups: speakers with invariant affirmative paradigms (89 total) and speakers with variable affirmative paradigms (30 total).

Additional methodological considerations include the analysis of contracted forms and of null forms. Where contracted forms are provided by speakers in addition to full forms (e.g. *am, 'm* or *is, 's*), the contracted form is treated as an allomorph of the full form, rather than as a distinct dialect variant of *be*. Similarly, the paired set *'r, Ø* occurred in some paradigms, and here Ø is also treated as a reduction of *'r* rather than as a completely distinct null form of *be*.[6] Where a contracted form is clearly not an allomorph of another variant in its cell within a paradigm, e.g. non-1-sg *'m* (figure 22), it is included in the analysis as a distinct form.

2 Interspeaker variation in affirmative and negative declaratives

This section presents all *be* paradigms in the *SED* which are instantiated in more than one speaker, have complete data sets for affirmative and synthetic negation paradigms, and are

[5] Support for the existence of the all-*be* paradigm of Sx5 and the *I are* paradigm of K7 comes from dialect literature as well as the *SED*. A number of early texts support the view that invariant *be* existed in the Somerset area for all subject types (Elworthy, 1877: 55, Barnes, 1863: 24, Hewett, 1894: 3, Wilson, 1913: 30; all references cited in Ihalainen, 1991: 104). Richard Coates (p.c., 4 August 2004) similarly suggests that the regional dialect in Sussex and neighboring regions had an all-*be* paradigm that began to be replaced in the nineteenth century by more general vernacular forms and gradually came to be largely limited to stylized dialect writing. Evidence of the earlier robustness of the all-*be* paradigm also comes from the fact that several *SED* speakers other than Sx5 do in fact exhibit the all-*be* pattern but have additional variants and thus are either included as variable systems (Bk3, O3) or excluded due to their having unique negative paradigms (Sx1, Sx3, Brk1, Brk4, Ha7, O2, So1). The *I are* system of K7 is similarly cited as an attested, once robust system in Kent and Surrey (Gower, 1893: vi; Trudgill, 1999: 106). Additional evidence of its wider distribution comes from its presence in the paradigms of other *SED* speakers as well, who also either had to be classed as variable due to the presence of other variants (K3, Bd1, Bd2, Bd3, Sr2, Sr4) or excluded due to their having different negative paradigms (K1, K4).

[6] The null form is not treated as a distinct form because (i) it does not occur independent of reduced *'r* in the *SED* and (ii) it is not generally attested as an independent verbal form in British dialects (Wolfram, 2000: 54).

Derbyshire: Db1(thee),Db6(thee),Db7,St1,Y22(she)

(I) am	(we) are	(I) amnt	
(thou) art			
(her) is	(they) are	(her) isnt	(they) arent

Cornwall: Co5,Co7

(I) am	(we) are	(I) arent	
(thee) art			
(she) is	(they) are	(she) isnt	(they) arent

Figure 2. All person distinctions in singular

Devon: D2,D6,Do3(we),Co1,So13(we)

(I) be	(us) be	(I) baint	
(thee) art			
(her) is	(they) be	(her) isnt	(they) baint

Wiltshire: Gl4,W2,W4,W5(she),W6(isnt),W8(she,isnt)

(I) be	(we) be	(I) baint	
(thee) beest			
(her) is	(they) be	(her) aint	(they) baint

Figure 3. Leveling of first person

invariant. The paradigm tables in figures 2–8 present affirmative and synthetic negative paradigms, listing at the top of each table all individual *SED* respondents (e.g. Db6) who exhibit the given pattern. Slight differences in lexical form for a speaker are included in parentheses following the speaker index. The figure headings separate tables according to the type of leveling in the affirmative paradigm. When the affirmative paradigm is identical but the negative paradigm is distinct, two separate tables are listed, both being under the general heading that describes their affirmative pattern (e.g. Derbyshire and Cornwall).[7]

A striking aspect of the data is that the same abstract paradigm is sometimes instantiated with different morphs. For instance, Devon and Wiltshire share the same abstract paradigm, as do Kent and Somerset. Similarly, the complete loss of all agreement contrasts is leveled to the form *be* in the Sussex inventory, but parallel systems using *am*, *are*, and *is* have also been reported, although we did not find these in our data: *I/you/she/we/you/they am here*, *I/you/she/we/you/they are here*, *I/you/she/we/you/they is here* (Trudgill, 1999: 10b). Past tense in West and East Midlands shows a similar loss of all agreement contrasts, again with

[7] Regional names assigned to inventory tables are somewhat arbitrary and are based on their representation among *SED* respondents. For instance, Devon, Somerset, and Sussex have significant overlaps in their *be* patterns, and the all-*be* pattern we refer to as 'Sussex' has been described as characteristic of Somerset as well. These regional names should therefore be treated simply as tags for inventories rather than accurate geographical delineations.

Northumberland: Nb1,Y26(thou)

(I) am	(we) are	(I) amnt	
(you) are			
(she) is	(they) are	(she) isnt	(they) arent

Norfolk: Nf1-2,Nf5,Nf9-13,Sf2,Ess1,L6(isnt),
Nf3(isnt),Nf6(isnt),St4(ina)

(I) am	(we) are	(I) arent	
(you) are			
(she) is	(they) are	(she) aint	(they) arent

Suffolk: Sf1,Sf3-5,Nf4,MxL2,Lei1-2,Lei4-6,Lei8,
Ess2-3,Ess5,Ess8-9,Ess11-13,Hu1-2,K5,Ha4,
Sr1,Sr3,M6,C1-2,L14-15,R1-2,Hrt1-2,Nth2-4

(I) am	(we) are	(I) aint	
(you) are			
(she) is	(they) are	(she) aint	(they) aint

Figure 4. Leveling of second person

Kent: K7

(I) are	(we) are	(I) aint	
(you) are			
(she) is	(they) are	(her) aint	(they) aint

Somerset: So12

(I) be	(we) be	(I) baint	
(you) be			
(she/her) is	(they) be	(she) baint	(they) baint

Hampshire: D8,So6,Ha2,Ha5,Bk5(aint3sg)

(I) be	(us) be	(I) baint	
(you) be			
(her) is	(they) be	(her) isnt	(they) baint

Figure 5. Leveling of first and second person

Berkshire: Brk1,Brk2,W7

(I) be	(us) be	(I) baint	
(thee) beest			
(her) be	(they) be	(her) baint	(they) baint

Figure 6. Leveling of first and third person

Yorkshire: Y2,Y6,Y13,Y24,La1,Cu2

(I) is	(we) are	(I) isnt	
(thee/thou) is			
(she) is	(they) are	(she) isnt	(they) arent

Figure 7. Leveling of person but not number

Sussex: Sx5

(I) be	(we) be	(I) baint	
(you) be			
(she) be	(they) be	(she) baint	(they) baint

Figure 8. Leveling of person and number

a different morph performing the leveled function: *I were singing. So were John. Mary weren't singing* (Cheshire et al., 1993: 80). These abstract parallels in dialect systems are unlikely to be explicable in terms of simple sound changes ('accidental homonymy' in Carstairs-McCarthy's (1987: 91) and Kusters' (2003: 27) terminology). They are better understood in terms of changes at the paradigmatic level in the system for expressing semantic content. Therefore we distinguish between the inventory of specific forms and the inventory of abstract contrasts; it is the latter that this article is concerned with.

Nevertheless, it is worth noting in passing that the choice of lexical forms is affected by regular sociohistorical processes. Figures 2–8 show that certain forms, such as *be* and *ain't*, are quite widespread. While *be* is an archaic form and is being replaced in some regions by newer forms (Trudgill, 1999: 106), *ain't* is commonly cited as one of several supralocal nonstandard features currently spreading across parts of the British Isles, replacing more regional forms. The use of this latter type of nonstandard urban form tends to be determined more by social class than region (Hughes & Trudgill, 1987; Coupland, 1988; Cheshire et al., 1993), and the resulting leveling has often been associated with 'a reduction of marked, socially heavily stigmatised, highly localised, or minority forms in favour of unmarked, less stereotyped, supralocal, majority variants' (Britain, 2002: 35). A number of social and historical factors are thus instrumental in the processes of selection and adoption of particular forms.

We emphasize that these processes are not the focus of the present study; our focus rather is on the typological range of possible abstract contrasts revealed by paradigms of specific morphs. Three key observations can be drawn from the data in figures 2–8 regarding abstract systems of contrasts and leveling of distinctions:

Observation 1:

There are 0–3 person distinctions made in the singular;
There are 0 person distinctions made in the plural; therefore
⇒ Person distinctions are leveled in the plural.

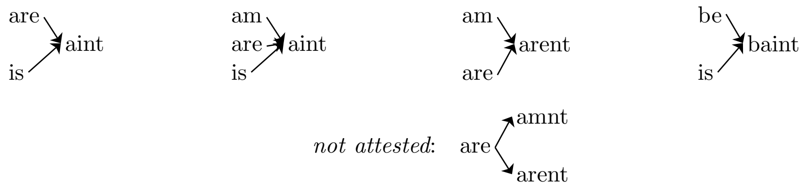

Figure 9. Leveling in negation

*	(I) are	(we) are	(I) amnt	
	(you) are			
	(she) is	(they) are	(her) isnt	(they) arent

Figure 10. Paradigm unattested in the *SED*

Observation 2:
Regardless of whether verb forms are leveled, pronominal subjects do not undergo leveling.

Observation 3:
The negative paradigms never express more information about person or number than their corresponding affirmative paradigms, and they frequently express *less*, as illustrated in figure 9. The type of paradigm shown in figure 10 – with leveling of *be* forms in the first person in the affirmative but with no leveling in the first person in negation – is not attested.

3 Optimality Theory analysis of leveling

We now turn to the framework we use for formally analyzing the surveyed inventories and the three observations noted above. In the present section we restrict the analysis to conventional OT, and in the later discussion of individual variation we introduce the stochastic component.

3.1 Optimality Theory

An OT grammar can be viewed as a function from INPUTS to OUTPUTS. We take the morphosyntactic INPUT to be language-independent content drawn from the space of possible lexical and grammatical contrasts and the OUTPUT to consist of language-specific forms with varying expressions of that content. INPUTS are fully specified for person and number features. Candidate expressions for each INPUT are generated by GEN and evaluated according to an EVAL function. Given a set of violable constraints hypothesized to be present in all grammars, and a language-particular ranking of these constraints, the EVAL function defines the OUTPUT to be the candidate which best satisfies the highest ranked constraint on which it differs from its competitors (Grimshaw, 1997a; Prince & Smolensky, 2004).[8]

[8] Note that Stochastic OT as a framework does not require that constraints be universal and/or innate, and in fact Boersma's theory of functional phonology (1998) is a well-articulated alternative.

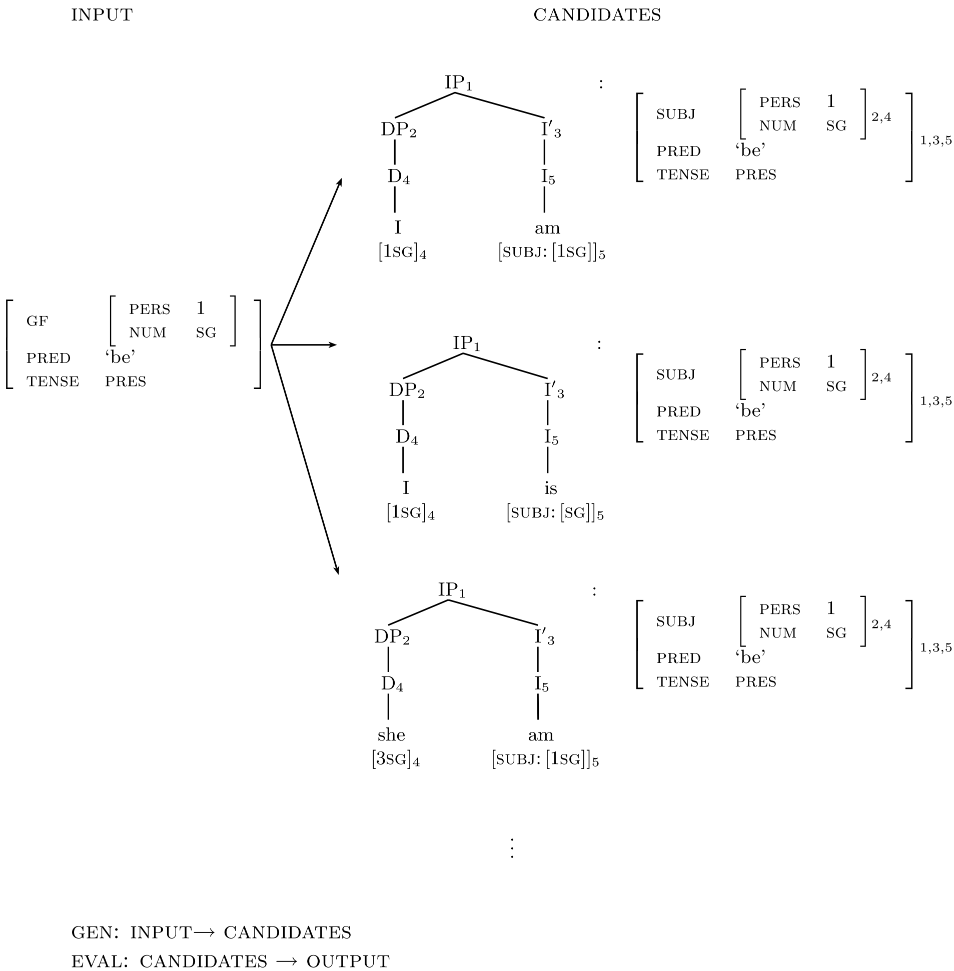

Figure 11. OT grammar of English subject–verb agreement

The overall structure we assume for syntactic expressions in OT is shown in figure 11. The INPUT is represented here as an abstract specification of semantic features, while the candidate set comprising the OUTPUT is represented by pairings of c(categorial)-structures and f(feature)-structures in correspondence. This conception of INPUT and OUTPUT draws on a mathematically and empirically well-understood representational basis, OT-LFG (see Bresnan, 2000, 2001a, b, c, 2002; Kuhn, 2000, 2001, 2002, 2003; Clark, 2004).[9]

[9] In figure 11 the customary attribute-value notation is used in which $+feature$ is rendered [$feature$ +] (Johnson, 1988). The verb forms paired with each f-structure actually consist of an abstract characterization of word-class properties, such as V^0 or I^0, and a language-particular pronunciation, such as *is*. The choice of phonological representations is outside the scope of our study.

We assume that the INPUT is an underspecified f-structure which semantically subsumes the candidate f-structures, an assumption justified by considerations of decidability and learnability (Kuhn, 2002, 2003). The INPUT feature structure contains only semantically relevant features; thus GF (for 'grammatical function') denotes any argument of the predicator BE and does not specify syntactic role.[10] GEN provides additional purely grammatical features as well as particular argument realizations (SUBJ, for example) to the candidate analyses, which thus contain the INPUT. The terminal string of the c-structure consists of fully inflected words which represent morpholexical choices to be optimized against the candidate f-structure. The lexical choices of the sentence are optimized in parallel, so that in figure 11 both the subject pronoun and the verb must be optimized against the given features [1SG] belonging to the SUBJ argument of the candidates. Lexical choices may be unfaithful to the INPUT to varying degrees.[11]

Since the candidate feature structures are all semantically subsumed by the input in this model, the lexical optimizations can be carried out against the candidate f-structure, which in general contains the input together with purely grammatical features provided by GEN. More precisely, then, the faithfulness constraints will relate the morpholexical f-structures of the c-structure terminals to the global feature structures of the candidates. Again, different lexical optimizations (for example, those for the subject pronoun and for the verb) may proceed in parallel and degrees of faithfulness to pronominal INPUT information and to verbal INPUT information may vary.

[10] As observed in Bresnan (2000), an underspecified f-structure is a formal representation of the idea that the OT INPUT for syntax is an argument structure with annotations of additional semantically relevant information (Legendre, Raymond & Smolensky, 1993; Grimshaw, 1997a). One advantage of this formalization is the availability of generation and parsing algorithms, recursive enumeration of the candidate set, a formal constraint language, and other useful computational and mathematical properties (Kuhn, 2002, 2003). Another advantage is the typological expressiveness of the theory of representations (Bresnan, 2001a).

[11] In a feature-logic basic theory of syntactic representation such as this, the formalism may be viewed as a feature-checking system which is output oriented ('declarative') rather than derivational ('procedural'). The basic workings of the system of feature-structure comparison are as follows. The numerical subscripts coindexing the tree nodes and feature structures show the correspondence relations between the two parallel structures, which follow from general principles of tree-to-feature-structure correspondence (Bresnan, 2001b; Kuhn, 1999). For example, the feature structures associated with the I nodes in these particular trees are indexed by 5, which is identified with the index of I′ (=3) and IP (=1) by a principle that identifies the f-structures of heads with those of their mothers. Similarly, the feature structures of the D nodes are indexed by 4, which is identified with the index of DP (=2) by the same head principle. The DP and IP f-structures are related by the specifier principle, which says here that f-structure 5's SUBJ has f-structure 2 as its value. (Other principles apply to the exocentric and nonconfigurational constructions found in many languages: see Bresnan, 2001a; Nordlinger, 1998.)

In faithfulness evaluations, the lexical feature structure of a terminal node is compared with the f-structure corresponding to (coindexed with) its preterminal node in the c-structure. By the syntactic correspondences in figure 11 just discussed, this comparison will hold for the f-structures of the phrasal projections of these terminals (IP in the case of *am*, *is*, and DP in the case of *I, she*). By the uniqueness principle, which states that every f-structure attribute must have a unique value, the verb's inner agreement feature structure [1SG] in $[\text{SUBJ}\ [\text{1SG}]]_5$ can be inferred to correspond to the subsidiary f-structure 4 (=2) in the sentential feature structure, which also corresponds to the lexical feature structure of the subject pronoun. For more details of the LFG representational basis adopted here, see Bresnan (2001a) and references.

3.2 Analysis for observation 1: leveling in plural

Observation 1 noted that all of the varieties of English surveyed here show loss of person distinctions in the plural. This leveling in the plural in British dialects reflects a more general, crosslinguistic markedness pattern (Greenberg, 1966: 28–9; Croft, 2003: 126), though there are exceptions (see n. 15). For the reasons given earlier (absence of explanation in terms of simple sound changes, presence of the same abstract leveling pattern in very different inventories of forms), we represent leveling by changes in the inventories of expressions of abstract semantic contrasts.

To model these contrasts, we assume that each form of *be* is represented by *the intersection of person and number values of all of the cells of the paradigm it occurs in.* The examples listed in (1) illustrate this mapping between semantic content and lexical form.

(1)

Yorkshire ⇒

	SG	PL
1	is	are
2	is	
3	is	are

Yorkshire feature values

	SG	PL
1	[SG]	[PL]
2	[SG]	
3	[SG]	[PL]

Derbyshire ⇒

	SG	PL
1	am	are
2	art	
3	is	are

Derbyshire feature values

	SG	PL
1	[1 SG]	[PL]
2	[2 SG]	
3	[3 SG]	[PL]

Wiltshire ⇒

	SG	PL
1	be	be
2	beest	
3	is	be

Wiltshire feature values

	SG	PL
1	[ ]	[ ]
2	[2 SG]	
3	[3 SG]	[ ]

Somerset ⇒

	SG	PL
1	be	be
2	be	
3	is	be

Somerset feature values

	SG	PL
1	[ ]	[ ]
2	[ ]	
3	[3 SG]	[ ]

A possible alternative would be to assume that perfect faithfulness between the input and the candidates' morphosyntactic features is maintained, as in (2).

(2)

Yorkshire ⇒

	SG	PL
1	is	are
2	is	
3	is	are

Yorkshire feature values

	SG	PL
1	[1SG]	[1PL]
2	[2SG]	
3	[3SG]	[3PL]

This approach would posit extensive, arbitrary homonymy, and would deprive us of a means for explaining the extension and retraction of forms by feature neutralization and feature generalization which recurs across the dialect varieties and is a common typological feature of languages (Greenberg, 1966: 28–9; Croft, 2003: 126). We assume that our paradigms are not based on arbitrary homonymy and instead we allow candidate feature structures to be unfaithful to the input.

Examples of morphosyntactic faithfulness violations (Grimshaw, 1997b: 193–4, 2001) are Romance clitic inventories where number and gender features 'float' onto adjacent clitics in certain circumstances (Bonet, 1995). When the divergence between the form and content of the candidate is contextually restricted, as in the Romance example, the output alternates between a faithful form and an unfaithful form that replaces it in limited circumstances. The contentful features of the input are thus only contextually neutralized, and are still transparent in most output forms.

In the case of the Yorkshire grammar, we could similarly posit unfaithful use of *is*[3SG] in non-3SG contexts in order to satisfy other higher ranked constraints. Morphosyntactic faithfulness violations can produce such divergences between form and content. However, the Yorkshire grammar in fact gives us absolute (context-free) neutralization of person features in the output, such that the candidate's person feature could be opaque in every context of its use.[12] In this situation 'remorphologization' or 'lexicon optimization' of the system may occur, i.e. although the set of candidates is technically unconstrained, the lack of evidence for the speaker/learner of person distinctions in the Yorkshire system can induce a 'rewriting' of input feature values in the output, replacing the candidate's unfaithful features with a more faithful, and therefore meaningful, analysis. This leads to generalization of the lexical form through remorphologization of its syntactic features as simply bearing a [SG] value.

We will see below how remorphologizing can arise through continuous constraint reranking in a Stochastic OT grammar. The point of interest here is that gradual changes on the continuous ranking scale can give rise to apparently categorical changes in content – without any derivational operations or procedures. This approach also allows inflectional changes to arise from morphosyntactic feature simplification independently of phonological erosion (Kusters, 2000).

In the same way, the analysis of *are* as a general form lacking PERS and NUM features may be the result of historical remorphologization of an earlier more specific plural form. In the Yorkshire and Derbyshire/Cornwall inventories, *are* is restricted to the plural. But elsewhere in our data *are* generalizes into the singular column of the paradigm, expressing the second person or both second and first persons.

The generalization or spread of a form in the *be* paradigm can proceed in the present theory by (the OT equivalents of) either feature deletion or, less commonly, feature change. The generalization of *are* across both number and persons in some dialects requires the deletion analysis, under which the form eventually lacks both PERS and NUM values. Although we do not have clear instances of feature change in the present data,

[12] Because our data set, like our constraint set, is small and incomplete, we cannot of course be certain that there are not relevant alternations elsewhere in the grammar. Indeed, the 'Northern Rule' affecting verb agreement when a subject pronoun is adjacent would be relevant in some Yorkshire inventories. All of our *SED* inventory verbs come from sentences with pronoun subjects.

a lexical form can also undergo this process, such that it becomes specialized to a new person and/or number value.

3.2.1 The constraint set

In OT there are two broad types of constraints: faithfulness constraints, which compare a candidate to the input, and markedness constraints, which assess the well-formedness of the candidate in terms of its featural complexity. Markedness constraints penalize complex or 'difficult' structures, and so tend to erode contrasts. Faithfulness constraints, by contrast, require that features of the input content be preserved in the output expression; they thus serve the communicative function of expressing contrasts in content, protecting content against the eroding effects of markedness constraints on forms. A particular language harmonizes these conflicting constraints by prioritizing (ranking) them.

Different faithfulness constraints may be instantiated for various morphosyntactically defined domains; this is called 'positional faithfulness' in phonology (Urbanczyk, 1995; Benua, 1995). English has three inflectional classes for present-tense verbs (*be*, modal verbs, and lexical verbs), for which there are three families of separately rankable faithfulness constraints (Bresnan, 2001b, 2002). We will be concerned here mainly with faithfulness in the domain of *be*. The faithfulness constraints that follow are thus implicitly indexed to this domain.

The faithfulness constraints in (3) ensure the expression in the output of person and number features present in the input.[13] This faithfulness may be achieved in different grammars by either fusional or nonfusional forms. Each of these constraints represents a family of more specific constraints. For instance, Express (persValue) includes Express (1), Express (2), and Express (3).

(3) Nonfusional faithfulness: Express (numValue), Express (persValue)
Fusional faithfulness: Express (persValue, numValue)

If we consider the sample input in (4), candidate 1 violates both the nonfusional constraints – Express (numValue) and Express (persValue) – and the fusional constraint Express (persValue, numValue). Candidate 2, by contrast, satisfies the nonfusional constraint Express (numValue), but violates the nonfusional constraint Express (persValue) as well as the fusional constraint Express (persValue, numValue).

(4) example input: $\begin{bmatrix} \text{NUM} & \text{SG} \\ \text{PERS} & 2 \end{bmatrix}$

candidate 1: *be*: $\begin{bmatrix} \text{NUM} & \\ \text{PERS} & \end{bmatrix}$

candidate 2: *is*: $\begin{bmatrix} \text{NUM} & \text{SG} \\ \text{PERS} & 3 \end{bmatrix}$

[13] These constraints differ somewhat from those in the preliminary study by Bresnan & Deo (2001) which were based in part on Grimshaw (1997b, 2001). The present constraints are conceptually preferable in postulating word-class differences in faithfulness to agreement values rather than arbitrary markedness differences among person values.

The two markedness constraints in (5), again indexed to the domain of the verb *be*, impose restrictions on the featural complexity of candidates regardless of their input features. We interpret these as constraints to avoid informational density. Thus, although candidate 2 in (4) satisfies faithfulness to number, in doing so it violates *NUM. By contrast, candidate 1 violates all faithfulness constraints, but satisfies both markedness constraints.[14]

(5) Avoid informational density: *PERS, *NUM

Increased leveling in plurals, as evidenced in the present data and in typological studies, can be captured by constraint subhierarchies, within which the relative rankings are fixed across languages, either extrinsically (Prince & Smolensky, 2004; Aissen, 1999; Kager, 1999) or by use of constraint semantics (de Lacy, 2002). The relevant subhierarchy for the present study is shown in (6).

(6) EXPRESS (PERSVALUE, SG) ≫ EXPRESS (PERSVALUE, PL)

The fixed ranking of constraints within this subhierarchy allows us to capture the crosslinguistic generalization that languages exhibit fewer distinctions among plural forms than singular forms in verbal agreement inventories.[15] The subhierarchy in (6) expresses the observation that, because plurality is a marked feature, it is universally dispreferred to mark plurality in addition to another feature, such as a person feature. In other words, there is a preference to highlight the marked status of plurality at the cost of other features.

A markedness constraint such as *PERS may intervene at any point in a constraint subhierarchy. As a result, the expression-constraint subhierarchy in (6) sets up implicational structures that permit leveling of plurals before singulars, but not the reverse. This effect is shown in (7).[16]

(7) *PERS ≫ EXPRESS (PERSVALUE,SG) ≫ EXPRESS (PERSVALUE,PL)
EXPRESS (PERSVALUE,SG) ≫ *PERS ≫ EXPRESS (PERSVALUE,PL)
EXPRESS (PERSVALUE,SG) ≫ EXPRESS (PERSVALUE,PL) ≫ *PERS

A secondary observation that can be made with regard to the present data is that there are ‘column generalizations’ leveling person distinctions within a single number category – the Yorkshire system has column generalizations for both SG and PL and Derbyshire has a column generalization for PL – but there are no ‘row generalizations’ leveling number distinctions within a single person category. This distinction is illustrated in (8).

[14] Of course, derivational operations of feature deletion and rewriting are not involved when candidates ‘omit’ input features; rather, these are epiphenomenal consequences of the parallel optimization of candidates that may diverge from the given input in various ways.

[15] This is sometimes said to be a general property of Germanic, but in Modern Icelandic, and in Old Icelandic as well to a lesser extent, in most paradigms there is only one person distinction in the singular – 1st against 2nd and 3rd, or 1st and 3rd against 2nd person – while 1st, 2nd, and 3rd person are distinguished in the plural (Wouter Kusters, p.c., 6 April, 2001). Thus, we can only provisionally interpret the constraint subhierarchy in (6) as universal, pending detailed study of the relevant grammars.

[16] See Kager (1999) for further exemplification of this type of factorial typology.

(8) Column generalizations

	SG	PL
1	a	b
2	a	b
3	a	b

Row generalizations (not attested)

	SG	PL
1	a	a
2	b	b
3	c	c

The faithfulness constraints EXPRESS(PERSVALUE) capture 'row forms'. In the analysis of our data, these constraints are always ranked below constraints favoring the expression of number. They are consequently inactive in grammars of all our varieties, and the candidates they select – with person/number values of [1], [2], [3] – are always suboptimal. For expository simplicity, we omit these inactive constraints and candidates, as well as those that would produce person contrasts in the plural. We do not, however, structure this secondary observation as a general typological property of language. There is plenty of evidence that these constraints can be active, leading to leveling of number distinctions within a single person category (as occurs in the future and the present progressive in Bengali, for instance).

3.2.2 Constraint rankings and dialect outputs

In this section we present a simplified OT account of constraint rankings, omitting details of stochastic evaluation which are assumed to be part of the grammar; we later elaborate on the mechanism of stochastic evaluation in relation to variable inventories. Here, we present detailed constraint rankings for three invariant dialect systems – Yorkshire, Derbyshire, and Suffolk – to illustrate the varied outcomes of constraint reranking. Aspects of each of these three analyses extend to all the other systems of contrast and neutralization in figures 2–8.

Yorkshire

The constraint ranking for Yorkshire *(is,is,is,are,are)* levels the expression of all person contrasts, both in the singular and in the plural. In figure 12,[17] we see that the high rank of *PERS disfavors the selection of any candidate bearing person features, regardless of whether the input is singular or plural. However, the relatively high rank of EXPRESS(SG) and EXPRESS(PL) favors the choice of lexical forms indexed for SG when a SG input is involved and PL when a PL input is involved, as opposed to the selection of a completely underspecified form such as *be* [ ].

Derbyshire

Figure 13 shows that the same constraints reranked for Derbyshire *(am,art,is,are,are)* preserve all singular person contrasts and level the expression of all plural contrasts. The relatively high rank of *PERS, EXPRESS(SG), and EXPRESS(PL) leads to a result for PL

[17] Note that in this and subsequent tableaux the candidate set forms 'is', 'art', etc. are merely convenient mnemonic tags for the feature structure which is the actual input.

	input: [1SG]	*PERS	EXP(SG)	EXP(PL)	EXP(PERSVALUE,SG)	*NUM	EXP(PERSVALUE,PL)
	‘am’: [1SG]	*!				*	
☞	‘is’: [SG]				*	*	
	‘are’: [PL]		*!		*	*	
	‘are’: [1PL]	*!	*		*	*	
	‘be’: [ ]		*!		*		
	‘am’: [1]	*!	*		*		

	input: [2SG]	*PERS	EXP(SG)	EXP(PL)	EXP(PERSVALUE,SG)	*NUM	EXP(PERSVALUE,PL)
	‘am’: [1SG]	*!			*	*	
☞	‘is’: [SG]				*	*	
	‘are’: [PL]		* !		*	*	
	‘are’: [1PL]	*!	*		*	*	
	‘be’: [ ]		* !		*		
	‘am’: [1]	*!	*		*		

	input: [1PL]	*PERS	EXP(SG)	EXP(PL)	EXP(PERSVALUE,SG)	*NUM	EXP(PERSVALUE,PL)
	‘am’: [1SG]	*!		*		*	*
	‘is’: [SG]			*!		*	*
☞	‘are’: [PL]					*	*
	‘are’: [1PL]	*!				*	
	‘be’: [ ]			*!			*
	‘am’: [1]	*!		*			*

Figure 12. Tableaux of a Yorkshire grammar

inputs that is identical to that of the Yorkshire grammar, namely a form specified for number but unspecified for person. However, the higher rank of the fusional constraint EXPRESS(PERSVALUE,SG) means that when a SG input is involved, the grammar will

	input: [1SG]	EXP(PERSVALUE,SG)	*PERS	EXP(SG)	EXP(PL)	*NUM	EXP(PERSVALUE,PL)
☞	'am': [1SG]		*			*	
	'is': [SG]	*!				*	
	'are': [PL]	*!		*		*	
	'are': [1PL]	*!	*	*		*	
	'are': [ ]	*!		*			
	'am': [1]	*!	*	*			
	'art': [2SG]	*!	*			*	

	input: [2SG]	EXP(PERSVALUE,SG)	*PERS	EXP(SG)	EXP(PL)	*NUM	EXP(PERSVALUE,PL)
	'am': [1SG]	*!	*			*	
	'is': [SG]	*!				*	
	'are': [PL]	*!		*		*	
	'are': [1PL]	*!	*	*		*	
	'are': [ ]	*!		*			
	'am': [1]	*!	*	*			
☞	'art': [2SG]		*			*	

	input: [1PL]	EXP(PERSVALUE,SG)	*PERS	EXP(SG)	EXP(PL)	*NUM	EXP(PERSVALUE,PL)
	'am': [1SG]		*!		*	*	*
	'is': [SG]				*!	*	*
☞	'are': [PL]					*	*
	'are': [1PL]		*!			*	
	'are': [ ]				*!		*
	'am': [1]		*!		*		*
	'art': [2SG]		*!		*	*	*

Figure 13. Tableaux of a Derbyshire grammar

always select a distinctive lexical form that uniquely marks both person and singular number.

Suffolk
Finally, the Suffolk system *(am,are,is,are,are)* is the Standard English system, which is similar to the Derbyshire system but avoids a distinct form for second person. In figure 14, the low rank of the fusional constraint EXPRESS(2,SG) and the higher rank of the markedness constraints *PERS and *NUM leads to the selection of a completely underspecified form *are* [ ]. This constraint is frequently low-ranked, reflecting the avoidance of too direct reference to the second person, a recurrent crosslinguistic phenomenon, with pragmatic and/or sociolinguistic motivations (Brown & Levinson, 1987) which may become formally crystallized in grammars.

3.3 Analysis for observation 2: no leveling in pronominal subjects

Observation 2 noted that within the context of clauses with pronominal subjects there appears to be no leveling of pronoun forms competing with leveling of *be* forms.[18] In other words, the expression of person is more faithful in the class of pronouns than in verbs. The present data show numerous instances of leveling of person distinctions in *be*; however, no dialect grammar levels pronominal forms along the lines proposed in the second column of (9).

(9)

Yorkshire:	Nonoccurring equivalents:
I is	she am
thee/thou is	she art
she is	she is

We propose that this asymmetry is a result of faithfulness constraints being relative to word classes. The architecture of Optimality Theory does not itself rule out pronominal unfaithfulness to person, as it permits both verbal and pronominal unfaithfulness, indicated earlier in figure 11. Different expressions in the lexical string may be variably faithful in terms of feature specifications; for instance, a first person subject pronoun may co-occur with a verb form specified for [1SG] in one dialect but [SG], or even [3SG], in another. In general, however, faithfulness to the referentially classificatory feature of person is much stricter for pronominal expressions than for verbal expressions.

This point is illustrated by the fact that, in figure 11 earlier, the first two candidates *I am* and *I is* are both possible expressions of the input with its first-person singular argument, while the third candidate *She am* is always suboptimal. (Note that *She am* is an optimal expression of a *third*-person subject in some English varieties; we suggest that it

[18] Verbal agreement may differ with pronominal and nonpronominal subjects in some varieties (Ihalainen, 1991: 107–8) by the so-called 'Northern Rule' (n. 12); see Börjars & Chapman (1998) for a formal syntactic analysis. The present study is limited to agreement in simple declarative affirmative and negative sentences with pronominal subjects.

	input: [1SG]	EXP(1,SG)	*PERS	*NUM	EXP(SG)	EXP(PL)	EXP(2,SG)	EXP(PERSVALUE,PL)
☞	'am': [1SG]		*	*				
	'is': [SG]	* !		*				
	'are': [PL]	* !		*	*			
	'are': [ ]	* !			*			
	'are': [1PL]	* !	*	*	*			
	'am': [1]	* !	*		*			
	'art': [2SG]	* !	*	*				

	input: [2SG]	EXP(1,SG)	*PERS	*NUM	EXP(SG)	EXP(PL)	EXP(2,SG)	EXP(PERSVALUE,PL)
	'am': [1SG]		* !	*			*	
	'is': [SG]			* !			*	
	'are': [PL]			* !	*		*	
☞	'are': [ ]				*		*	
	'are': [1PL]		* !	*	*		*	
	'am': [1]		* !		*		*	
	'art': [2SG]		* !	*				

	input: [1PL]	EXP(1,SG)	*PERS	*NUM	EXP(SG)	EXP(PL)	EXP(2,SG)	EXP(PERSVALUE,PL)
	'am': [1SG]		* !	*		*		*
	'is': [SG]			*!		*		*
	'are': [PL]			*!				*
☞	'are': [ ]					*		*
	'are': [1PL]		* !	*				
	'am': [1]		* !			*		*
	'art': [2SG]		* !	*		*		*

Figure 14. Tableaux of a Suffolk (Standard English) grammar

is suboptimal only as an expression of a *first*-person subject.) This generalization can be captured by the following subhierarchy:

(10) $\text{EXPRESS}_{pron}(\text{PERS}) \gg \text{EXPRESS}_{verb}(\text{PERS})$.

These two positional faithfulness constraints are indexed respectively to the morphosyntactic domains of pronominal and verbal expressions. The verbal and pronominal positional faithfulness constraints are separately rankable, but the subhierarchy ensures that the subject pronoun cannot be less faithful to the input person of the subject argument than the verb is.[19]

Further support for the claim that faithfulness constraints are generally indexed to word classes comes from within verbal word classes, namely the greater faithfulness to expression of person in some verb classes as against others. The table in (11) shows that agreement with subject person in Standard English is most differentiated with *be*, slightly less so with lexical verbs, and least so with modal verbs, resulting again in a class-based ranking of faithfulness:

$$\text{EXPRESS}_{be}(\text{PERS}) \gg \text{EXPRESS}_{verb}(\text{PERS}) \gg \text{EXPRESS}_{modal}(\text{PERS}).$$

(11)

	be:		(main) verbs:		modal verbs:	
	SG	PL	SG	PL	SG	PL
1	am	are	hit	hit	will	will
2	are	are	hit	hit	will	will
3	is	are	hits	hit	will	will

As this paper focuses on forms of *be*, observation 2 is less central to our analysis than observation 1, but this short discussion demonstrates the need for faithfulness constraints to be specifically indexed to particular word classes.

3.4 Analysis for observation 3: leveling in negation

Observation 3 noted that if leveling occurs, it occurs to an equal or greater degree in the negative paradigms of *be*. As with plural leveling, this parallels the typologically attested markedness of the negative (Greenberg, 1966: 50; Givón, 1978: 70; König, 1988: 161; Croft, 2003: 202).

Again, as with leveling in the affirmative, the leveling seen in negation cannot all be attributed to purely phonological simplification: for instance, $\{be, is\} \Rightarrow$ *baint*. We therefore treat variation in negation also as an instance of changes in the inventory of content.

In our analysis of this phenomenon, we draw a crucial distinction between synthetic and analytic negation. We treat synthetic negation as any single verb form that contains both the

[19] It is noteworthy that, unlike person, number and gender are categories in which pronominal expressions may be less faithful than verbal expressions. In Golin, a Papuan language of New Guinea, both bound and free pronouns are undifferentiated for number contrasts but there is a verbal suffix specialized for first-person-singular subjects (Foley, 1986: 70). In Jersey French, the pronoun for both singular and plural first-person subjects is *je* but the verb maintains distinct forms (Jones, 2001: 115). Similarly, in many Indo-Aryan languages, e.g. Hindi, third-person pronouns are undifferentiated for gender, but subject gender is marked on the verb.

verbal content of *be* and the negation feature value. This primarily involves forms bearing the contracted negative *-n't*. Payne (1985: 226) distinguishes between negative auxiliaries and negated auxiliaries, the former having inherent negative meaning and the latter simply involving an added inflectional marker to a non-negative morpheme. Kortmann (1999: 10) suggests that although English synthetic negation forms such as *isn't* clearly start out as negated auxiliaries, their patterns of leveling and phonological reduction make them comparable to negative auxiliaries. Zwicky & Pullum (1983) similarly argue that these forms have properties more typical of bound morphemes than of clitics, such as allomorphic variation (*will* vs. *won't*, *do* vs. *don't*).

This article is primarily concerned with synthetic negation forms rather than analytic negation constructions, as we argue here that person/number leveling is a process predicted to apply specifically in synthetic negative morphology (e.g. *ain't*) due to the increase in the 'load' of semantic values borne by a single morphological item. Naturally, if the semantic values of *be* and NEG are carried by different morphological forms, as in an analytic construction such as *am not*, this over-burdening does not occur.

Based on this reasoning, leveling of *be* in analytic negation, as in (12), is not predicted to occur.

(12)
I am not ↘
→ *[I ai not, She ai not]
She is not ↗

I am ↘
→ *[I are not, They are not]
They are ↗

If leveling of *be* does occur in negation, it will occur in the synthetic negative paradigm first. This leveling may occur alongside continued differentiation of forms in the paradigm of analytic negation, as in (13).

(13) I am → I am not, I ain't

Our hypothesis is supported by the fact that we found no instances of leveling in analytic (but not synthetic) negation in the *SED*, whereas dozens of cases of leveling in synthetic (but not analytic) negation were found. The more detailed grammar for speaker K5 given in (14), showing both synthetic and analytic negation, illustrates restricted leveling in the synthetic negation paradigm only.

(14)

(I) am	(we) are	(I) 'm not, aint	
(you) are			
(she) is	(they) are	(she) aint	(they) aint

In the discussion that follows, we restrict our focus to leveling in synthetic negation. Further constraints, not included in the analysis here, would regulate the choice of analytic or synthetic expressions of negation (Bresnan, 2002).

3.4.1 The constraint set

Two contextual markedness constraints, given in (15), formalize the intuition discussed above.[20] The high ranking of *[NEG+NUM] would lead to leveling of number distinctions in negative forms of the verb (e.g. *I ain't, we ain't*), while the high ranking of *[NEG+PERS] would lead to leveling of person distinctions in negative forms of the verb (e.g. *we ain't, you ain't, they ain't*).

(15) Avoid overloaded morphology: *[NEG+PERS], *[NEG+NUM]

These two constraints interact with the faithfulness constraints already discussed to yield the typological structure shown in (16).

(16) *PERS, *[NEG+PERS] ≫ EXPRESS (PERS. . .)
EXPRESS (PERS. . .) ≫ *PERS, *[NEG+PERS]
*[NEG+PERS] ≫ EXPRESS (PERS. . .) ≫ *PERS

The first ranking in (16) levels person contrasts, regardless of whether the clause is affirmative or negative. The second ranking expresses person contrasts, regardless of whether the clause is affirmative or negative. The final ranking, crucial to our discussion here, levels person contrasts only in the context of negative morphology. Equally crucial is the observation that no ranking of these constraints will level person contrasts only in affirmative contexts, as there is no markedness constraint to impose restrictions exclusively on the unmarked affirmative context.[21]

3.4.2 Constraint rankings and dialect outputs

The interaction of the negation constraints with the constraints already introduced generates a typological space that permits a range of possible contrasts and neutralizations in affirmative and negative paradigms. Below we extend the grammars described for the three sample cases earlier – Yorkshire, Derbyshire, and Suffolk – to include negation constraints. These expanded grammars instantiate the typological possibilities predicted by the rankings in (16). We also present a grammar for Cornwall, as it represents a subtler interaction of negation constraints with person and number constraints.

[20] These constraints can be conceived of as a type of constraint conjunction (Smolensky, 1995). The present data do not permit a closer exploration of whether a complete subhierarchy of conjoined constraints (e.g. sensitive to the person or number hierarchy) is attested, and we retain a simple formulation for the present discussion.

[21] We also never find leveling of the affirmative–negative distinction in order to retain person contrasts in synthetic negative verb forms. We might argue that in situations of morphological overload within a verbal domain, faithful expression of verbal features is universally preferable to the expression of nominal features; this asymmetry would resemble the preferred faithfulness to person features in the domain of pronouns as opposed to verbs, discussed earlier in (10). However, negation also has special properties that can be argued to require expression even where other verbal agreement features may not. Affirmative and negative propositions are fundamentally opposed semantically – they cannot be true in the same world – so an output without formal negation marking cannot be considered to be underspecified for affirmative or negative sense (unlike underspecification of person or number). From a functionalist perspective, the expression of negation is fundamental to the clause and may be considered inviolate.

	input: [1 SG NEG]	*PERS	EXP(SG)	EXP(PL)	EXP(PERSVALUE,SG)	*NUM	EXP(PERSVALUE,PL)	*[NEG+PERS]	*[NEG+NUM]
	'amn't': [1SG NEG]	*!				*		*	*
☞	'isn't': [SG NEG]				*	*			*
	'aren't': [PL NEG]		* !		*	*			*
	'aren't': [1PL NEG]	*!	*		*	*		*	*
	'ain't': [NEG]		* !		*				

Figure 15. Yorkshire grammar including negation constraints

	input: [1SG NEG]	EXP(PERSVALUE,SG)	*PERS	EXP(SG)	EXP(PL)	*NUM	EXP(PERSVALUE,PL)	*[NEG+PERS]	*[NEG+NUM]
☞	'amn't': [1SG NEG]		*			*		*	*
	'isn't': [SG NEG]	*!				*			*
	'aren't': [PL NEG]	*!		*		*			*
	'aren't': [1PL NEG]	*!	*			*		*	*
	'ain't': [NEG]	*!		*					

Figure 16. Derbyshire grammar including negation constraints

Yorkshire

As witnessed earlier, Yorkshire has leveling across person, retaining only the number distinction of singular and plural. This division is maintained in the negative paradigms of these speakers as well. As we saw in figure 12, the constraint ranking for Yorkshire (*is, is, is, are, are*) levels the expression of all person contrasts, both in the singular and in the plural; the same constraints determine the choice of candidate for negative inputs. The constraints on overloaded morphology in synthetic negation do not play a part in the evaluation and are low ranked (figure 15).

Derbyshire

In the Derbyshire type of paradigm, a number of contrasts are made in the affirmative paradigm. Although this affirmative paradigm is very different from that of Yorkshire, as there is no leveling in the singular, there is a similarity between Derbyshire and Yorkshire in the context of negation, as the amount of leveling in negation mirrors the amount of leveling in the affirmative in both dialects. In terms of constraint ranking for Derbyshire, this again translates into a low ranking for the two negation constraints (figure 16).

	input: [1SG NEG]	*[NEG+PERS]	*[NEG+NUM]	EXP(1,SG)	EXP(3,SG)	*PERS	*NUM	EXP(SG)	EXP(PL)	EXP(2,SG)	EXP(PERSVALUE,PL)
	‘amn’t’: [1SG NEG]	* !	*			*	*				
	‘isn’t’: [SG NEG]		* !	*			*				
	‘aren’t’: [PL NEG]		* !	*			*	*			
	‘aren’t’: [1PL NEG]	*!	*	*		*	*	*			
☞	‘ain’t’: [NEG]			*				*			

Figure 17. Suffolk grammar including negation constraints

Although Devon, Wiltshire, Northumberland, Hampshire, Berkshire, and Sussex all have different amounts of leveling in their affirmative paradigms, their negative systems are all accounted for in the same way: the synthetic negation constraints are low ranked and the amount of leveling in affirmative and negative paradigms emerges as identical in each of these systems.

Suffolk
Several distinctions are made in the affirmative *be* paradigm of Suffolk, but this group diverges from those previously discussed in exhibiting complete leveling in negation. The ranking of person and number constraints was seen earlier in figure 14; when a synthetic negative input is involved, the high rank of *[NEG+PERS] and *[NEG+NUM] becomes apparent, as a general form is always selected (figure 17).

Cornwall
Finally, the affirmative pattern of the Cornwall group is identical to that of Derbyshire, but it differs in its negation pattern. The Cornwall system exhibits more leveling in negation than in the affirmative, but this leveling is not absolute as in the case of *ain't* in Suffolk. This type of partial leveling in negation also occurs in the negative paradigm of Norfolk.

The one distinction that is maintained in the negative paradigm of Cornwall is the third singular form. In this case, it is necessary to posit that the Cornwall system prioritizes a single constraint out of the family of EXPRESS(PERSVALUE,SG) constraints, namely EXPRESS(3,SG), above the negation constraints. With the exception of this very high-ranked constraint, the constraints on morphological overloading in synthetic negation outrank other person and number faithfulness constraints, forcing the selection of a general form in all other cases (figure 18). This ensures that in the affirmative all singular person distinctions are maintained – due to the relatively high rank of EXPRESS(PERSVALUE,SG) – but in negation only a distinct form for 3SG inputs is maintained.

	input: [1sg neg]	Exp(3,sg)	*[neg+pers]	*[neg+num]	Exp(persValue,sg)	*pers	Exp(sg)	Exp(pl)	*num	Exp(persValue,pl)
	'amn't': [1sg neg]		* !	*		*			*	
	'isn't': [3sg neg]		* !	*	*	*			*	
	'aren't': [pl neg]			* !	*		*		*	
	'aren't': [1pl neg]		* !	*	*	*	*		*	
☞	'aren't': [neg]				*		*			

	input: [3sg neg]	Exp(3,sg)	*[neg+pers]	*[neg+num]	Exp(persValue,sg)	*pers	Exp(sg)	Exp(pl)	*num	Exp(persValue,pl)
	'amn't': [1sg neg]	* !	*	*	*	*			*	
☞	'isn't': [3sg neg]		*	*		*			*	
	'aren't': [pl neg]	* !		*	*		*		*	
	'aren't': [1pl neg]	* !	*	*	*	*	*		*	
	'aren't': [neg]	* !			*		*			

Figure 18. Cornwall grammar including negation constraints

This analysis of negation predicts that there may be variable systems in which the general form *ain't* is alternating with, and is in the process of replacing, a specific form such as *amn't*. We do indeed frequently find this type of variability in the *SED*. These systems are directly accounted for by the current analysis, but as these alternations were very idiosyncratic, with no single type of alternation occurring for more than one speaker, they did not satisfy our criterion for including only stable systems attested in more than one individual and so we do not list all of them here.

To summarize, the extraction of all invariant (categorical) paradigms for the verb *be* in the *SED* has yielded two significant patterns in the data which confirm previous studies of leveling in English: there is more leveling of person/number contrasts in the plural than in the singular and more leveling in synthetic negatives than in affirmatives.

We have constructed an OT model of person leveling and negation leveling in present-tense English *be* which allows for degrees of leveling in these domains, but which precludes the occurrence of more leveling in the singular (than in the plural), or more leveling in the affirmative (than in the negative). Even though it is far from complete, we have adopted the minimal constraint set needed to account for our present data and to exclude grammars that appear to be unattested. Furthermore, the architecture of OT captures relations between inter-speaker variation and crosslinguistic typological patterns (see also Kusters, 2003; Deo & Sharma, 2006).

4 Intra-speaker variation in affirmative and negative declaratives

All individual *be* paradigms in the *SED* which were found to contain internal variation, and which were also instantiated in more than one speaker and had complete data sets for affirmative and synthetic negation paradigms, are presented below. As before, the paradigm tables present affirmative and synthetic negative paradigms, with all individuals who exhibit the pattern listed at the top of the table, and slight differences in lexical form for a speaker given in parentheses following the speaker index.

We treat figure 22, the plural *am* paradigms, as distinct from the others. We cannot characterize the plural *am* varieties purely in terms of person/number information, as plural *am* is always a variant and never occurs as the sole plural form in any person in any of the grammars here. In over half of the paradigms with plural *am*, its distribution is precisely coextensive with another form (*be* or *are*), so person and number features are not sufficient to distinguish its distribution and some other factors must be involved. Ihalainen (1991: 107–8) observes that in the generalized *am* dialects in east Somerset, *am* is used as an unstressed allomorph of *be*, and so its occurrence appears to be dependent on phonological constraints. We therefore set aside the plural *am* systems in figure 22 from our analysis.

For the remaining variable paradigms, we can see that observation 1 (plural leveling), observation 2 (no pronoun leveling), and observation 3 (negative leveling) from the previous section still hold. In addition, we can make three further observations:

Observation 4:
Choices of variant forms of *be* and of pronominal forms are often at least partially independent.

We do not discuss this observation further save to note that it forms part of a more general finding here that grammatical variables in the present data do not appear to alternate as systematically as a competing grammars view (Kroch, 2000) would anticipate. Although instances of covariation do occur in the data, e.g. *thee art, you be* in the speech of Do5, a single pronoun frequently occurs with variant verb forms. Some examples are given in (17).

(17) *thee art, thee are* (Variable Yorkshire)
thee be, thee art (Variable Somerset)
I am, I be (variable Monmouthshire)
I am, I are (variable Bedfordshire)
she is, she be (variable Oxfordshire, variable Dorset)
her is, her be (variable Gloucestershire)

Mixing of variant pronominal forms with variant verbal forms has also been illustrated in extracts from taped Somerset speech in Ihalainen (1991), repeated in (18).[22]

(18) a. *You taught theeself, didn't ee?* (Ihalainen, 1991: 115)
b. *I'm not under no obligation about this, be I?* (Ihalainen, 1991: 109)
c. *They're not ready, be 'em?* (Ihalainen, 1991: 116).

[22] (18d) is used by Ihalainen to illustrate the fact that *thee* is used more frequently in stressed positions than in unstressed ones.

Variable Yorkshire: St3,Y21,Y29,La6

(I) am	(we) are	(I) amnt	
(thee) art/are			
(she) is	(they) are	(she) isnt	(they) arent

Variable Somerset: So7,Do5(thee art/you be)

(I) be	(we) be	(I) baint	
(thee) be/art			
(her) is	(they) be	(her) isnt	(they) baint

Figure 19. Variable second person singular

Variable Monmouthshire: M1,Gl7(she aint)

(I) am/be	(we) be	(I) baint	
(thee) beest			
(her) is	(they) be	(her) aint	(they) baint

Variable Bedfordshire: Bd1,Bd2,Bd3,K3(aint)

(I) am/are	(we) are	(I) aint/ent	
(you) are			
(she) is	(they) are	(she) aint/ent	(they) aint/ent

Figure 20. Variable first person singular

d. B.I. *What be you, Herb? Seventy-two?*
H.T. *Gone seventy-five.*
B.I. *Seventy-five! Thee!*
W.B. *Thee! Thee! I didn't know you were gone seventy-five.* (Ihalainen, 1991: 115)

Observation 5:

i. The variable patterns can be decomposed into combinations of the invariant patterns already seen.
ii. The general verb form is often in free variation with more specific forms.

Each variable inventory can be represented as a partial intersection of two invariant systems. This is not to say that these systems are direct sources of the variable system in geographical space or historical time, but rather that each alternant in the variable system gives rise to one of two grammars very close in terms of pure typological space. All of the variable systems listed in figures 19–22 can be described in this way. Two detailed examples of observation 5 are given in figure 23.

Variable Oxfordshire: O3,Bk3(her aint/ent, her is/she be)

(I) be	(us/we) be	(I) beaint	
(you) be			
(she) is/be	(they) be	(she) aint	(they) beaint

Variable Gloucestershire: Gl5,Gl6,Ha1

(I) be	(us/we) be	(I) baint	
(thee) beest			
(her) be/is	(they) be	(her) aint	(they) baint

Variable Dorset: Do2,Do4(her is/she be)

(I) be	(us/we) be	(I) baint	
(thee) art			
(she) is/be	(they) be	(she) isnt	(they) baint

Figure 21. Variable third person singular

Surrey: Sr2,Sr4

(I) are/am	(we) are/am	(I) aint	
(you) are			
(she) is	(they) are/am	(she) aint	(they) aint

Cornwall: Co3,Co4(she); So8('m only pl),Co2('m only pl)

(I) be/'m	(we) be/'m	(I) baint	
(thee) art			
(her) is	(they) be/'m	(her) isnt	(they) baint

Devon/Wiltshire: D1,D3(us),W9(bist)

(I) be	(we) be	(I) baint	
(thee) art			
(her) is	(they) be/'m	(her) isnt	(they) baint

Devon/Hampshire: D5,Ha3(she,we,isnt)

(I) be	(us) be/'m	(I) baint	
(thee) art			
(her) is	(they) be	(her) aint	(they) baint

Figure 22. Plural *am* varieties

Variable Bedfordshire = Suffolk + Kent

Variable Bedfordshire

	SG	PL
1	am, are	are
2	are	
3	is	are

Suffolk

	SG	PL
1	am	are
2	are	
3	is	are

Kent

	SG	PL
1	are	are
2	are	
3	is	are

Variable Yorkshire = Suffolk + Derbyshire

Variable Yorkshire

	SG	PL
1	am	are
2	art, are	
3	is	are

Suffolk

	SG	PL
1	am	are
2	are	
3	is	are

Derbyshire

	SG	PL
1	am	are
2	art	
3	is	are

Figure 23. Decomposition of variable systems

Summary of decomposition of all variable inventories:
Variable Bedfordshire = invariant Kent + invariant Suffolk
Variable Yorkshire = invariant Derbyshire + invariant Suffolk
Variable Somerset = invariant Hampshire + invariant Devon
Variable Monmouthshire = invariant Cornwall + invariant Wiltshire
(abstractly: *art* ≈ *beest*, *are* ≈ *be*)
Variable Oxfordshire = invariant Hampshire + invariant Sussex
Variable Gloucestershire = invariant Wiltshire + invariant Berkshire
Variable Dorset = invariant Devon + invariant Berkshire
(abstractly: *art* ≈ *beest*)

Observation 6:

i. Most of the variable inventories *cannot* be decomposed into two geographically adjacent dialects.
ii. Instead, every case of variability but one appears to involve variation between a vernacular form and a standard (Suffolk-type) form, the latter generally resulting in a system that resembles some other non-Suffolk dialect.

Reference to figure 1 confirms the generalization that the decomposition of variable inventories does *not* point to two geographically adjacent inventories. Rather, almost all cases of variability involve variation of a vernacular form with a standard (Suffolk-type) form. The one exception is the variable Somerset inventory, in which a variant from a neighboring dialect (Devon) infiltrates the system.

Thus social prestige of the standard variety and geographical continuity of vernacular varieties appear to be the two forces placing environmental (as opposed to typological) constraints on the types of inventories that arise. The former appears to be a far stronger factor in the *SED* data. A natural sociolinguistic explanation of this situation is that the learning data or environment is composed of the local vernacular system and the global standardized system, and that this sociohierarchical structure may have a greater effect than physical proximity and spatial diffusion.

However, as noted already, the data do not show global covariation of standard and vernacular paradigms, but rather very local alternations in parts of the *be* paradigms. The decomposition of variable inventories showed that the intrusion of an isolated standard form into an otherwise nonstandard inventory does not lead to a completely standard paradigm. Instead, the second system of contrasts that arises from the inclusion of a single standard form almost always resembles another nonstandard system. For instance, the intrusion of the standard form *am* into M1's Wiltshire-like grammar leads to the resulting paradigm resembling the inventory of Cornwall in terms of abstract contrasts, despite the lack of any significant contact with that variety.

The present analysis predicts that in theory any combination from the typological space of possible grammars may occur for a single variable speaker, and the two forces of social prestige and geographical proximity are simply external constraints restricting expression of the full typological range of possible inventories.

If this interpretation is correct, it suggests a model of variation in which the standard grammar is perturbing the vernacular grammar but not necessarily replacing it. The perturbed grammar appears to vary between the vernacular and a second grammatical system that is very close to it in the space of possible grammars, if not in geographical space. The second system usually does not have the overall structure of the standard grammar, but rather merely one additional resemblance to it.

Stochastic evaluation of constraints with stochastic learning as in the Gradual Learning Algorithm (Boersma, 1998; Boersma & Hayes, 2001; Jäger, in press; cf. Keller & Asudeh, 2002) provides a way of formally modeling this kind of variation. The section that follows offers an account linking observation 5 and observation 6 as consequences of the stochastic nature of individual grammars.

5 A Stochastic OT model of individual variation

5.1 The framework: generalizing from the categorical to the quantitative

In this final section, we present a formal model to account for localized individual variability in grammars as witnessed in the *SED* data. As mentioned at the outset, the full power of the Stochastic OT apparatus is not needed in the present analysis as we do not have frequency distributions for each variant in variable systems. However, we believe that this approach is useful conceptually and theoretically even in the absence of frequency data, as it allows us to formalize what is meant by individual variation and to offer an account of localized variation, as opposed to the systematic covariation predicted by competing grammars.

Optimality Theory with stochastic evaluation was originally developed by Paul Boersma as part of a theory of functional phonology that addresses the learning of categories, variation, optionality, and probability (Boersma 1997, 1998, 2000; Boersma & Hayes, 2001). It is one of a family of generalized OT frameworks that address variation

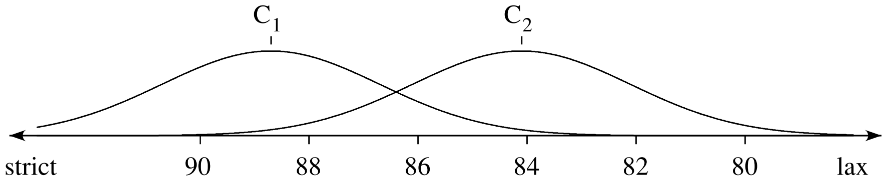

Figure 24. Constraint ranking on a continuous scale with stochastic evaluation

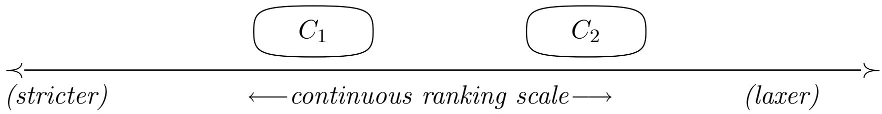

Figure 25. Categorical constraint ranking with ranges of variation

(see Anttila, 1997, Boersma, 1999b; Hibiya, 2000; and Boersma & Hayes, 2001 for reviews). Stochastic OT is distinguished by a particularly well-developed underlying theory, including an associated Gradual Learning Algorithm.

Stochastic OT differs from standard OT in two essential ways:

i. **ranking on a continuous scale:** Constraints are not simply ranked on a discrete ordinal scale; rather, they have a value on the continuous scale of real numbers. Thus constraints not only dominate other constraints, but are specific distances apart, and these distances are relevant to what the theory predicts.
ii. **stochastic evaluation**: At each evaluation the real value of each constraint is perturbed by temporarily adding to its ranking value a random value drawn from a normal distribution. For example, a constraint with the mean rank of 99 could be evaluated at 98.12 or 100.3. It is the constraint ranking that results from these new disharmonic values that is used in evaluation. The rank a constraint has in the grammar is the mean of a normal distribution or 'bell curve' of these variant values that it has when applied in evaluations; this is illustrated in figure 24.[23]

As explained by Boersma & Hayes (2001), an OT grammar with stochastic evaluation can generate both categorical and variable outputs. Categorical outputs arise when crucially ranked constraints are spread far apart on the continuous scale, so that the stochastic variation in ranking values has no discernible effect. In figure 25, for example, $C_1 \gg C_2$ and the two constraints are spread far enough apart that the bulk of their ranges of variation (illustrated in a simplified way by the ovals) do not overlap. As the distance between constraints increases, interactions become vanishingly rare, reaching a point where variant outputs lie beneath any given error threshold, or beyond the life expectancy of the speaker. (A distance of five standard deviations ensures an error rate of less than 0.02 percent; Boersma & Hayes, 2001: 50.)[24]

[23] The diagrams in figures 24–7 are adapted from Boersma & Hayes (2001).

[24] Units of measurement are arbitrary. With standard deviation = 2.0, a ranking distance of 10 units between constraints is taken to be effectively categorical.

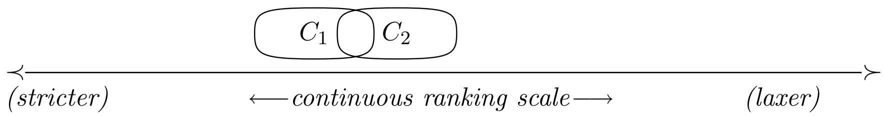

Figure 26. Free constraint ranking with ranges of variation

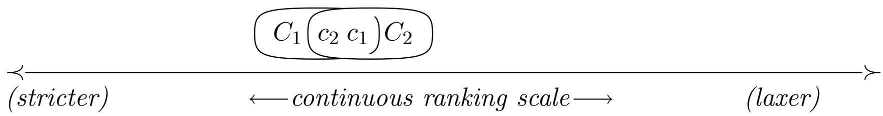

Figure 27. Reversal of constraint dominance

Variable outputs arise when crucially ranked constraints are close enough together for the variation in their ranking values to interact with some observable frequency. This possibility is illustrated in Figure 26, where the bulk of the ranges of variation of two constraints overlaps. Here again $C_1 \gg C_2$, but with some discernible frequency during stochastic evaluation C_1 will be ranked at a point in its lower range, call it c_1, while C_2 is simultaneously ranked at a point c_2 in its higher range. As shown in Figure 27, C_2 will then temporarily dominate C_1 in selecting the optimal output, possibly producing a different output.

The frequency of this reversal depends on the ranking distance between constraints and the standard deviation in ranking variance during evaluations (which is assumed to be the same across constraints). If we take the standard deviation to be zero, the constraints are always evaluated in the same strict domination sequence, and we have ordinal OT (Prince & Smolensky, 2004). Stochastic OT is thus a generalization of ordinal OT. Its associated learning algorithm can learn grammars robustly from variable data (Boersma, 1997, 1998, 2000; Boersma & Hayes, 2001), as illustrated in the next section.

5.2 Stochastic grammars and the Gradual Learning Algorithm

Boersma's stochastic grammars are based on the optimization function of ordinal Optimality Theory (Prince & Smolensky, 2004).[25] The effective ranking ('selectionPoint') of a constraint C_i is given by the equation (Boersma, 2000: 483):

$$selectionPoint_i = rankingValue_i + noise$$

The *noise* variable represents unknown factors that are independent of the linguistic theory embodied in the constraint set. We assume that there is in fact a deterministic function

[25] Other optimization functions have also been explored. See Goldwater & Johnson (2003), Jäger (in press), Jäger & Rosenbach (2003).

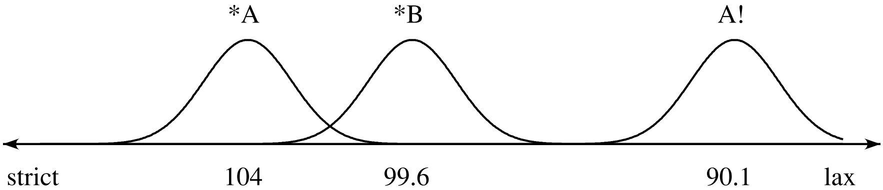

Figure 28. Sample stochastic grammar

from the total context plus the input to the output, but many aspects of the context are too complex to know in detail. The random noise variable simply models our ignorance of the total context which includes nonlinguistic factors that determine the probability of an output (for example by affecting the speaker's sensitivity to aspects of the current context).

The Gradual Learning Algorithm (GLA), implemented within the freely available cross-platform Praat computer program (Boersma, 1999a; Boersma & Weenink, 2000),[26] models stochastic grammars given particular constraints and exposure to learning data. Starting from an initial state grammar in which all constraints have the same ranking values (arbitrarily set to be 100.0), the GLA is presented with learning data; this may, for instance, consist of input–output pairs having the statistical distribution of (in the present case) a sample of spoken English.

For each learning datum (a given input–output pair), the GLA compares the output of its own grammar for the same input; if its own output differs from the given output, it adjusts its grammar by moving all the constraints that differentially disfavor its own output upward on the continuous ranking scale by a small increment, and moving all constraints that differentially disfavor the given output downward along the scale by a small decrement. The increment/decrement value is called the 'plasticity' and may be assumed to vary stochastically and to change with age (Boersma, 2000). In the case of constraint subhierarchies, the adjustment process applies recursively in order to preserve their local ordering relations.

Figure 28 and the tableaux in (19) and (20) illustrate this process. In Figure 28, the markedness constraints **A* and **B* are ranked fairly close together and the faithfulness constraint *A!* is ranked lower. If the 'selectionPoint' of **A* is higher than that of **B* in a given evaluation, then the representative tableau is (19). If the 'selectionPoint' of **A* is lower than that of **B* in a given evaluation, then the representative tableau is (20).

(19)

		*A	*B	A!
☞	$cand_1$		*	
	$cand_2$	*!		

(20)

		*B	*A	A!
	$cand_1$	*!		
☞	$cand_2$		*	

26 The GLA is also implemented in OTSoft, also freely available (Hayes, Tesar & Zuraw, 2000).

Given exposure to data in the environment, the grammar can compare its own output to the output of the learning data for the same input and gradually adjust its own ranking to match external evidence.

If cand$_1$ is always correct in the learning data, i.e. if the surrounding grammars all have the ranking in (19), then each time cand$_2$ is produced by the grammar, the countervailing evidence from the categorical learning data will progressively repel constraints A* and B* further apart, fixing their ranking in that order. If cand$_2$ is always correct in the learning data, then when cand$_1$ is produced by the grammar, the countervailing evidence from the categorical learning data will cause *A and B* to gradually rerank and then continue spreading apart, fixing this reverse order over time.

If both cand$_1$ and cand$_2$ are encountered in the learning data as correct outputs for the same input, i.e. if there is variation in the environment, then the variable data will cause the constraints *A and *B to attract and repel, as in (21), eventually attaining a holding pattern that matches the frequency of variation in the data to which the individual is exposed.

(21)

		*A⇒	⇐*B	A!
☞	cand$_1$		*	
	cand$_2$	*!		

		*B⇒	⇐*A	A!
	cand$_1$	*!		
☞	cand$_2$		*	

Crucially this means that the Stochastic OT model analyzes the acquisition of categorical and variable systems in exactly the same way, and variation is latent in every grammar.

5.3 Analysis for observations 5 and 6: localized variation

The present data were subjected to this learning process using idealized categorical and variable frequencies. The noise parameter is arbitrarily set at 2.0 which, as mentioned earlier, models our ignorance of the complete set of factors that may probabilistically influence selection of an output.

A total of 3,200,000 input–output pairs for each British dialect grammar was used to train the Gradual Learning Algorithm (Boersma, 1997, 1999a; Boersma & Hayes, 2001), starting from an initial state grammar in which all constraints have the same ranking values (arbitrarily set to be 100.0). The learning data for categorical dialect systems consisted of 3,200,000 input–output pairs with the same output for a given input 100 percent of the time. For instance, the categorical system of Standard English consisted of learning data in which 100 percent of the outputs for [1SG] were the fully faithful feature structure [1SG] abbreviated by the tag '*am*'; 100 percent of the outputs for [2SG] were the general feature structure [ ] abbreviated by the tag '*are*', and so on.

The output distributions of the earlier and later grammars for Standard English, shown in figure 29, were learned by the GLA in this way.[27] The earlier grammar was learned

[27] The output forms '*am*', '*are*', etc. are mnemonic tags for the abstract feature structure; see n. 17. Only a sample of candidate outputs is included for each input.

Output distributions (outputs > 1%)				
input	output	% in learning data	% (stochastic) Earlier	Later
[1SG]	am[1SG]	100	69.7	99.9
	are[ ]	0	30.2	
[2SG]	art[2SG]	0	21.8	
	is[SG]	0	10.0	
	are[ ]	100	68.1	99.9
[3SG]	is[3SG]	100	74.2	99.9
	are[ ]	0	25.7	
[1PL]	are[ ]	100	95.8	
	are[PL]	0	4.2	
[3PL]	are[ ]	100	95.7	99.9
	are[PL]	0	4.3	

Figure 29. Output distributions of earlier and later grammars for Standard English

from only 8,000 input–output pairs, while the later grammar was learned by additional exposure to 3,200,000 quantities of categorical data, given the earlier grammar as the initial grammar. The figure shows that the choice of outputs begins to converge towards categoricality.

For the same grammar, figure 30 shows that the ranking of constraints also becomes more strict with increased exposure to categorical data. The constraint ranking values are shown on the vertical axis; constraint names are horizontally spread out merely for readability. Greater vertical distance between constraints represents decreasing likelihood of ranking reversal. The earlier and later grammars have the same crucial ordinal constraint rankings, but these constraints are spread out differently on the scale. Greater exposure to categorical data incrementally shifts these rankings further apart.

By contrast, exposure to variable data would cause constraints to become closer, as long as there is still plasticity in the system.[28] In the case of variable paradigms, we lacked frequency information for the *SED* inventories and so we simply assumed that each variant form was used 50 percent of the time. In the case of Variable Monmouthshire, for example, we provided the GLA with data in which the output form *am* was selected 50 percent of the time with a [1SG] input and the output form *be* was selected for the other 50 percent of [1SG] inputs, as shown in (22).

[28] Boersma & Hayes (2001) demonstrate how the GLA approximates variable distributions in the environment for a number of test cases.

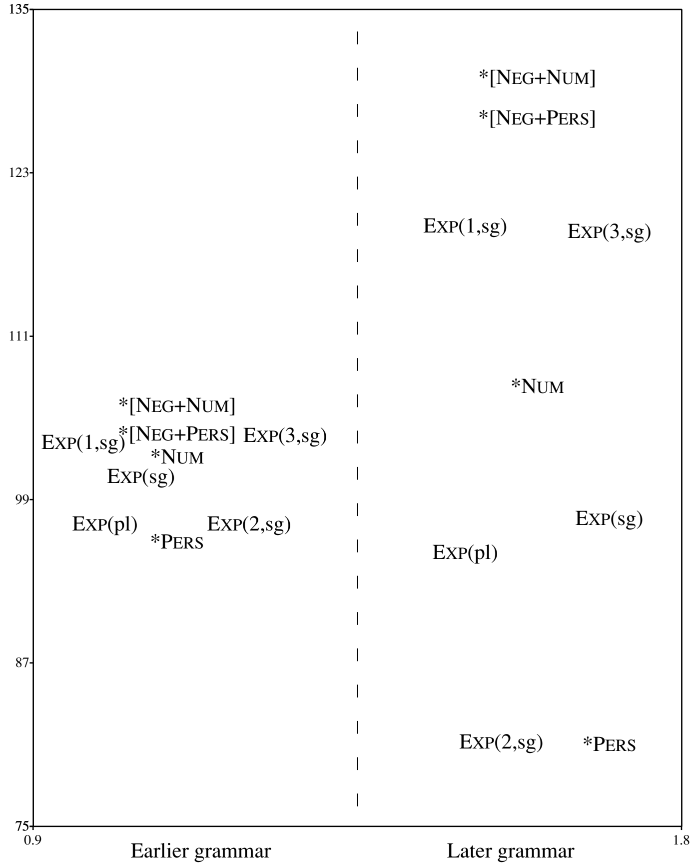

Figure 30. Reduction of variation under exposure to categorical data during 'first-language' stochastic learning by GLA

(22) /1SG/ → *be*[ ] 50
/1SG/ → *am* [1SG] 50
/2SG/ → *beest* [2SG] 100
. . .

Recall observation 5 that the variable grammars in the data can be decomposed into two invariant grammars, for instance: Variable Monmouthshire *(am/be, beest, is)* = Wiltshire *(be, beest, is)* + Cornwall *(am, art, is)*. Figure 31 represents the GLA acquisition of this variable grammar and the two component invariant grammars. Again the constraint ranking values for the three varieties of English are shown on the vertical axis, while the horizontal spread within each variety is simply for readability. The learned distribution of constraints exemplifies observation 5, as the reranking of two constraints results in two different categorical grammars – not necessarily geographically adjacent – and variation between the two rankings gives rise to an individually variable grammar. These three

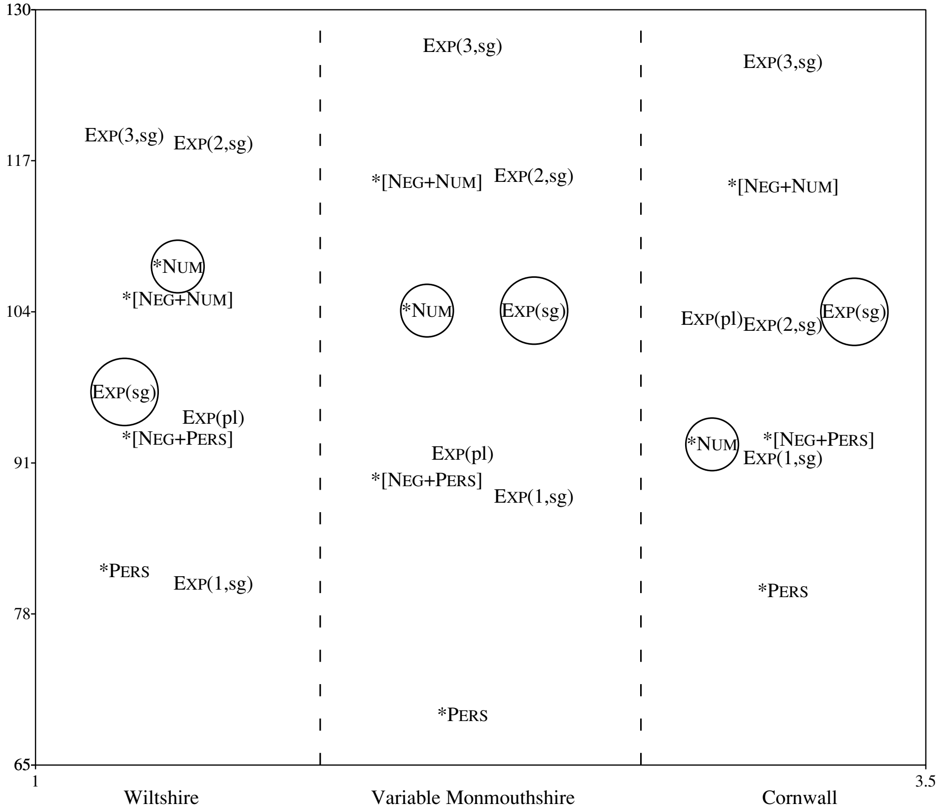

Figure 31. Decomposition of a variable grammar

grammars need not arise through direct contact: all three are simply typologically predicted systems whose attestation in the actual inventory of British dialects may be conditioned by social and historical factors.

The example in figure 31 also shows that the different rankings of constraints frequently select between candidates that are either more or less specified for certain input features, i.e. they may frequently choose between specific forms and general forms, which was the second aspect of observation 5. An important correlate of this observation is that reranking of constraints can lead to feature deletion and feature change in the lexical inventory, as a form can come to be partially or wholly underspecified if it comes to always be selected for a range of different inputs, as in Yorkshire. Both of these processes lead to remorphologization, as the lexical entries gain or lose featural specifications.

This highly variable range of systems is not naturally explained by a model using blocking of general forms by more specific forms, nor by an ordinal ranking of violable constraints (ordinal OT) or by a systematically covarying competing grammars scenario.

Finally, as we saw in observation 6, when a standard form of *be* is variably included in a vernacular grammar, the resulting grammar usually has neither the overall structure of the standard grammar nor that of geographically adjacent grammars. The account given here shows that the two component systems are simply close to one another in the typological space of possible rankings, and the intervention of a standard form leads to an alternation between these two similar rankings. In the case of Variable Monmouthshire in figure 31, the intrusion of standard *am* favors the Cornwall-like ranking, while the underspecified dialect variant *be* favors the Wiltshire-like system.

The point of interest here is that with stochastic evaluation of constraints, rankings and hence grammars are inherently variable. There is a region of variant grammars closely surrounding every grammar. The variant grammars belong to the factorial typology of OT constraints. Stochastic evaluation is, in effect, always sampling the typological space of grammars.

6 Conclusions

This analysis of inter- and intraspeaker paradigms has covered all systems present in the *SED*, excluding only those ruled out by our two initial criteria – the requirements that a full set of affirmative and synthetic negative *be* forms be recorded and that at least two speakers be attested per system. Our initial analysis of invariant systems in the *SED* found that interspeaker (dialectal) leveling in the *SED* occurs in the plural and in the negative, mirroring crosslinguistic typology. This parallel was accounted for in our analysis by the typological space generated by universal constraint subhierarchies in OT.

Intraspeaker (individual) variation in the *SED* was found to frequently involve alternation of individual forms rather than alternations of two complete dialect grammars. Covariation was not found to always happen systematically, and we did not always find a comprehensive switch of all nonstandard forms to all standard forms, but rather a piecemeal variation in isolated forms. The standard does not therefore appear to be replacing the vernacular in a robust competing grammars scenario; rather, variation is idiosyncratic and inherent in individual grammars. There are many possible sociolinguistic reasons for the adoption of individual forms, including salience of forms, frequency, access to the standard, and conscious selectivity on the part of the speaker (LePage & Tabouret-Keller, 1985; Trudgill, 1986).

The choice of using a particular isolated form (such as a pronoun or a verb) may thus be made for reasons entirely external to its particular linguistic content of agreement values. Those values permit the form to be fit into its appropriate place in the speaker's grammatical system, and the constraints that govern them must be ranked appropriately to allow this fit. If a form is frequently used by an individual, either due to frequent use in the environment or due to its particular social value, it will become a permanent fixture of the speaker's inventory, through gradual movement of the active constraints in the ranking space.

Stochastic OT, together with an appropriate output-oriented system for syntactic representation such as optimality-theoretic LFG (OT-LFG), is a model that allows for such partial intrusion/perturbation by the standard variety. Stochastic OT treats individual grammars as highly plastic cognitive systems sensitively tuned to frequencies in the linguistic environment. While typology determines the space of possible grammars, individual exposure determines which forms and grammar(s) are instantiated in a given individual. The structure and acquisition of categorical and variable grammars are formally identical under this analysis, simply differing in their degree of variability, which is treated as an inherent property of all grammars.

More systematic covariation can also be captured within the Stochastic OT framework. Such variation may reflect substantive constraint dependencies, seen in phenomena such as the 'constant rate effect' in historical syntactic change in English (Clark, 2004). Systematic covariation may also reflect style sensitivity parameters which boost or depress the ranking values of groups of constraints (Boersma & Hayes, 2001: 83–4) as in the morphosyntax of case ellipsis in Korean and Japanese (Lee, 2002, 2003, 2006). In an extreme case, such parameters could define quantal jumps in ranking that would create entirely distinct grammars, modeling diglossia.

The detailed paths of historical change producing the English systems studied here remain a topic for further research, as are the implications for the learnability of morphology. Important work in language development has adopted the central assumptions that there is only one correct form for each slot in a paradigm and that over-regularizations are corrected by exposure to the correct form (Pinker, 1984). Yet, as we have seen, the Gradual Learning Algorithm of the Stochastic OT model allows for robust learning from variable outputs of the same input.

We should note in closing that questionnaire responses, like other data collected through elicitation of linguistic intuitions, may inaccurately reflect the use of these forms in actual speech and should be treated with caution (Ihalainen, 1991: 110; Schilling-Estes & Wolfram, 1994: 297; Cornips, 2006). Our primary interest in these data has been to map the typological diversity in British dialects. The Stochastic OT model of individual dialectal variation that has been presented here should ultimately be tested against genuine frequencies of use as attested in robust sociolinguistic data.

Authors' addresses:
Department of Linguistics
Stanford University
Stanford CA 94305 USA
bresnan@stanford.edu

Department of Linguistics
Yale University
370 Temple Street
PO Box 208366
New Haven CT 06520
USA
ashwini.deo@yale.edu

English Department
King's College London
London WC2R 2LS
UK
devyani.sharma@kcl.ac.uk

Appendix A: Questions from the *SED* used to create database of forms of the verb *be*

VIII.2.8 HOW ARE YOU?
[affirmative wh-question]
North: 2-91. West: 2193-2268. South: 4488-4580. East: 6832-6916.

VIII.9.5 We drink water when WE ARE/I AM/SHE IS/THEY ARE thirsty.
[affirmative declarative, adjectival predicate]
North:92-389. West: 2269-2602. South: 4581-4881. East: 6917-7253.

IX.7.1 To find out whether you are right, you ask quite simply AM I right?
[affirmative y/n question, adjectival predicate]
North: 390-464. West: 2603-2676. South: 4882-4956. East: 7254-7338.

IX.7.2 ARE YOU/IS SHE/ARE THEY married?
[affirmative y/n question, adjectival predicate]
North: 465-679. West: 2677-2893. South: 4957-5187. East: 7339-7589.

IX.7.3 But AREN'T YOU/ISN'T SHE/AREN'T THEY married?
[negative y/n question, adjectival predicate]
North: 680-898. West: 2894-3110. South: 5188-5411. East: 7590-7843.

IX.7.4 And if it was you, you'd say to yourself AREN'T I lucky?
[negative y/n question, adjectival predicate]
North: 899-972. West: 3111-3183. South: 5412-5484. East: 7844-7926.

IX.7.5 He's alright there ISN'T HE/AREN'T I/AREN'T YOU/AREN'T THEY?
[negative tag question]
North: 973-1258. West: 3184-3479. South: 5485-5781. East: 7927-8252.

IX.7.7 Which of you is English here? you could answer I AM/YOU ARE/SHE IS/THEY ARE.
[affirmative declarative, predicate ellipsis]
North: 1259-1526. West: 3480-3767. South: 5782-6082. East: 8253-8587.

IX.7.9 Oh yes WE ARE/I AM/YOU ARE/SHE IS (English).
[affirmative declarative, predicate ellipsis]
North: 1527-1794. West: 3768-4055. South: 6083-6378. East: 8588-8909.

IX.7.10 Oh no I'M NOT/SHE ISN'T/THEY AREN'T (drunk).
[negative declarative, predicate ellipsis]
North: 1795-1988. West: 4056-4271. South: 6379-6605. East: 8910-9160.

IX.7.11 Get away, I'M NOT drunk.
[negative declarative, adjectival predicate]
North: 1989-2063. West: 4272-4347. South: 6606-6680. East: 9161-9245.

IX.9.2 You see a dog chasing your sheep, and you know it's not yours, so you wonder WHOSE IT IS.
[affirmative wh declarative]
North: 2064-2140. West: 4348-4417. South: 6681-6750. East: 9246-9330.

IX.9.4 WHO ARE your parents?
[affirmative wh-question]
North: 2141-2193. West: 4418-4487. South: 6751-6831. East: 9331-9417.

Appendix B: Abbreviations for region names

Bd = Bedfordshire; Bk = Buckinghamshire; Brk = Berkshire; C = Cambridgeshire; Ch = Cheshire; Co = Cornwall; Cu = Cumberland; D = Devon; Db = Derbyshire; Do = Dorset; Du = Durham; Ess = Essex;

Ha = Hampshire; He = Herefordshire; Hrt = Hertfordshire; Hu = Huntingdonshire; Gl = Gloucestershire; K = Kent; L = Lincolnshire; La = Lancashire; Lei = Leicestershire; M = Monmouthshire; Man = Isle of Man; MxL = Middlesex and London; Nb = Northumberland; Nf = Norfolk; Nt = Nottinghamshire; Nth = Northamptonshire; O = Oxfordshire; R = Rutland; Sa = Shropshire; Sf = Suffolk; So = Somerset; Sr = Surrey; St = Staffordshire; Sx = Sussex; W = Wiltshire; Wa = Warwickshire; We = Westmoreland; Wo = Worcestershire; Y = Yorkshire.

References

Aissen, J. (1999). Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* **17**: 673–711.

Anderwald, L. (2001). Was/Were-variation in non-standard British English today. *English World-Wide* **22**: 1–21.

Anderwald, L. (2002). *Negation in non-standard British English: gaps, regularizations and asymmetrics*. London: Routledge.

Anderwald, L. (2003). Non-standard English and typological principles: the case of negation. In Rohdenburg, G. & B. Mondorf, (eds.), *Determinants of grammatical variation in English*. Berlin: Mouton de Gruyter. 507–29.

Anderwald, L. & B. Kortmann (2002). Typology and dialectology: a programmatic sketch. In van Marle, J. & J. Berns (eds.), *Present day dialectology*, vol. I: *Problems and discussions.* Berlin: Mouton de Gruyter. 159–71.

Anttila, A. (1997). Variation in Finnish phonology and morphology. PhD thesis, Stanford University.

Anttila, A. (2002). Variation and phonological theory. In Chambers, J. K., P. Trudgill & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 206–43.

Bell, A. (1984). Language style as audience design. *Language in Society* **13**: 145–204.

Benua, L. (1995). Identity effects in morphological truncation. In Beckman, J., L. Dickey, & S. Urbanczyk (eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers 18. Amherst: University of Massachusetts. 77–136.

Boersma, P. (1997). How we learn variation, optionality, and probability. In Caron, B. (ed.), *Proceedings of the Institute of Phonetic Sciences 21*. Paris. 43–58. Rutgers Optimality Archive: http://ruccs.rutgers.edu/roa.html.

Boersma, P. (1998). Functional phonology: formalizing the interaction between articulatory and perceptual drives. The Hague: Holland Academic Graphics. PhD thesis, University of Amsterdam.

Boersma, P. (1999a). Optimality-theoretic learning in the Praat program. *IFA Proceedings* **23**: 17–35.

Boersma, P. (1999b). Review of Arto Anttila: *Variation in Finnish phonology and morphology*. *GLOT International* **5**: 33–40.

Boersma, P. (2000). Learning a grammar in Functional Phonology. In Dekkers, J., F. van der Leeuw, & J. van de Weijer (eds.), *Optimality theory: phonology, syntax and acquisition*. New York: Oxford University Press. 465–523.

Boersma, P. & B. Hayes (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* **32**: 45–86.

Boersma, P. & D. Weenink (2000). Praat computer program. Institute of Phonetic Sciences, University of Amsterdam: http://www.fon.hum.uva.nl/praat/.

Bonet, E. (1995). Feature structure of Romance clitics. *Natural Language & Linguistic Theory* **13**: 607–47.

Börjars, K. & C. Chapman (1998). Agreement and pro-drop in some dialects of English. *Linguistics* **36**: 71–98.

Bresnan, J. (2000). Explaining morphosyntactic competition. In Baltin, M. & C. Collins (eds.), *The handbook of contemporary syntactic theory*. Oxford: Blackwell. 11–44.
Bresnan, J. (2001a). *Lexical-functional syntax*. Oxford: Blackwell.
Bresnan, J. (2001b). Optimal syntax. In Dekkers, J., F. van der Leeuw & J. van de Weijer (eds.), *Optimality theory: phonology, syntax and acquisition*. Oxford: Oxford University Press. 334–85.
Bresnan, J. (2001c). The emergence of the unmarked pronoun. In Legendre, G., J. Grimshaw & S. Vikner (eds.), *Optimality-theoretic syntax*. Cambridge, MA: MIT Press. 113–42.
Bresnan, J. (2002). The lexicon in Optimality Theory. In Stevenson, S. & P. Merlo (eds.), *The lexical basis of syntactic processing: formal, computational, and experimental issues*. Amsterdam: John Benjamins. 39–58.
Bresnan, J. & A. Deo (2001). Grammatical constraints on variation: 'be' in the *Survey of English Dialects* and (Stochastic) Optimality Theory. Stanford University: http://www.stanford.edu/∼bresnan/download.html.
Britain, D. (2002). Diffusion, levelling, simplification and reallocation in past tense BE in the English Fens. *Journal of Sociolinguistics* **6**: 16–43.
Brown, P. & S. Levinson (1987). *Politeness: some universals in language usage*. 2nd edition. Cambridge: Cambridge University Press.
Carstairs-McCarthy, A. (1987). *Allomorphy in inflexion*. London: Croom Helm.
Cheshire, J. (1991). Variation in the use of 'ain't' in an urban British English dialect. In Trudgill, P. & J. Chambers (eds.), *Dialects of English: studies in grammatical variation*. London: Longman. 54–73.
Cheshire, J. (1996). Syntactic variation and the concept of prominence. In Klemola, J., M. Kytö & M. Rissanen (eds.), *Speech past and present: essays in English dialectology in memory of Ossi Ihalainen*. Frankfurt: Peter Lang. 1–17.
Cheshire, J., V. Edwards & P. Whittle (1993). Non-standard English and dialect levelling. In Milroy, J. & L. Milroy (eds.), *Real English: the grammar of English in the British Isles*. London: Longman. 53–96.
Clark, B. (2004). A Stochastic Optimality Theory approach to syntactic change. PhD thesis, Stanford University.
Cornips, L. (2006). Intermediate syntactic variants in a dialect–standard speech repertoire and relative acceptability. In Fanselow, G., C. Féry, M. Schlesewsky & R. Vogel (eds.), *Gradience in grammar*. Oxford: Oxford University Press. 85–105.
Coupland, N. (1988). *Dialect in use: sociolinguistic variation in Cardiff English*. Cardiff: University of Wales Press.
Croft, W. (2003). *Typology and universals*. Cambridge: Cambridge University Press, 2nd edn.
de Lacy, P. (2002). The formal expression of markedness. PhD thesis, University of Massachusetts at Amherst.
Deo, A. & D. Sharma (2006). Typological variation in the ergative marking of Indo-Aryan languages. *Linguistic Typology* **10**: 369–419.
Foley, W. (1986). *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
Givón, T. (1978). Negation in language: pragmatics, function, ontology. In Cole, P. (ed.), *Syntax and Semantics 9: Pragmatics*. New York: Academic Press. 69–112.
Goldwater, S. & M. Johnson (2003). Learning OT constraint rankings using a maximum entropy model. In Spenader, J., A. Eriksson & Ö. Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm University. 111–20.
Greenberg, J. (1966). *Language universals, with special reference to feature hierarchies*. Janua Linguarum, Series Minor 59. The Hague: Mouton de Gruyter.
Grimshaw, J. (1997a). Projection, heads, and optimality. *Linguistic Inquiry*, **28**: 373–422.

Grimshaw, J. (1997b). The best clitic: constraint conflict in morphosyntax. In Haegeman, L. (ed.), *Elements of grammar*. Dordrecht: Kluwer Academic Publishers. 169–96.
Grimshaw, J. (2001). Optimal clitic positions and the lexicon in Romance clitic systems. In Legendre, G., J. Grimshaw & S. Vikner (eds.), *Optimality-theoretic syntax*. Cambridge, MA: MIT Press. 113–42.
Hayes, B., B. Tesar & K. Zuraw (2000). OTSoft 2.1 software package. UCLA: http://www.linguistics.-ucla.-edu/people/hayes/ot-soft/.
Hibiya, J. (2000). Variation studies and linguistic theory. Paper presented at the Eighteenth National Conference of the English Linguistic Society of Japan, 18–19 November, 2000, Konan University, Kobe, Japan.
Hughes, A. & P. Trudgill (1987). *English accents and dialects: an introduction to social and regional varieties of British English*. 2nd edition. London: Arnold.
Ihalainen, O. (1991). On grammatical diffusion in Somerset folk speech. In Trudgill, P. & J. Chambers (eds.), *Dialects of English: studies in grammatical variation*. London: Longman. 104–19.
Jäger, G. (in press). Maximum entropy models and stochastic Optimality Theory. In Grimshaw, J., J. Maling, C. Manning, J. Simpson & A. Zaenen (eds.), *Architectures, rules, and preferences: a festschrift for Joan Bresnan*. Stanford: CSLI Publications.
Jäger, G. & A. Rosenbach (2006). The winner takes it all – almost. *Linguistics*. **44**: 937–71.
Johnson, M. (1988). *Attribute-value logic and the theory of grammar*. Stanford, CA: CSLI Publications.
Jones, M. (2001). *Jersey Norman French: a linguistic study of an obsolescent dialect*. Oxford: Blackwell.
Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
Keller, F. & A. Asudeh (2002). Probabilistic learning algorithms and Optimality Theory. *Linguistic Inquiry* **33**: 225–54.
König, E. (1988). Concessive connectives and concessive sentences: cross linguistic regularities and pragmatic principles. In Hawkins, J. (ed.), *Explaining language universals*. Oxford: Blackwell. 145–85.
Kortmann, B. (1999). Typology and dialectology. In Caron, B. (ed.), *Proceedings of the 16th International Congress of Linguists, Paris 1997*. Amsterdam: Elsevier Science. CD-ROM.
Kroch, A. (2000). Syntactic change. In Baltin, M. & C. Collins (eds.), *The handbook of contemporary syntactic theory*. Oxford: Blackwell. 699–729.
Kuhn, J. (1999). Towards a simple architecture for the structure-function mapping. In Butt, M. & T. H. King (eds.), *Proceedings of the LFG '99 Conference, University of Manchester*. http://csli-publications.stanford.edu/LFG/4/lfg99.html.
Kuhn, J. (2000). Faithfulness violations and bidirectional optimization. In Butt, M. & T. H. King (eds.), *Proceedings of the LFG 2000 Conference, Berkeley*. 161–81. http://csli-publications.stanford.edu/LFG/5/lfg00.html.
Kuhn, J. (2001). Generation and parsing in Optimality Theoretic syntax. In Sells, P. (ed.), *Formal and empirical issues in Optimality-Theoretic syntax*. Stanford, CA: CSLI Publications.
Kuhn, J. (2002). OT syntax – decidability of generation-based optimization. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL02)*. Philadelphia. 48–55.
Kuhn, J. (2003). *Optimality-theoretic syntax – a declarative approach*. Stanford, CA: CSLI Publications.
Kusters, W. (2000). Morphological simplification: more than erosion? In Gilbers, D. G., J. Nerbonne, & J. Schaeken (eds.), *Languages in contact*. Studies in Slavic and General Linguistics. 28, Amsterdam: Rodopi. 225–30.

Kusters, W. (2003). *Linguistic complexity: the influence of social change on verbal inflection*. Utrecht: LOT.
Lee, H. (2002). Referential accessibility and stylistic variation in OT: a corpus study. *Chicago Linguistic Society (CLS)* **38**(1): 361–78.
Lee, H. (2003). Parallel optimization in case systems. In Butt, M. & T. H. King (eds.), *Nominals: inside and out*. Stanford: CSLI Publications. 15–58.
Lee, H. (2006). Parallel optimization in case systems: evidence from case ellipsis in Korean. *Journal of East Asian Languages* **15**: 69–96.
Legendre, G., W. Raymond & P. Smolensky (1993). An Optimality-theoretic typology of case and grammatical voice systems. In *Proceedings of the 19th Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA. 464–78.
LePage, R. & A. Tabouret-Keller (1985). *Acts of identity: creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
Nagy, N. & D. Heap (1998). Francoprovençal null subject and constraint interaction. *Chicago Linguistic Society* (*CLS*) **34**: 151–66.
Nordlinger, R. (1998). *Constructive case: evidence from Australian languages*. Stanford, CA: CSLI Publishers.
Orton, H. et al. (1962–71). *Survey of English dialects*. Leeds: University of Leeds. published for the University of Leeds by E. J. Arnold.
Payne, J. R. (1985). Negation. In Shopen, T. (ed.), *Language typology and syntactic description*, vol. I: *Clause structure*. Cambridge: Cambridge University Press. 197–242.
Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
Prince, A. & P. Smolensky (2004). *Optimality Theory: constraint interaction in generative grammar*. Oxford: Blackwell.
Schilling-Estes, N. & W. Wolfram (1994). Convergent explanation and alternative regularization patterns: 'were/n't' leveling in a vernacular variety. *Language Variation and Change* **6**: 273–302.
Smolensky, P. (1995). On the structure of the constraint component Con of UG. Talk handout, UCLA. Rutgers Optimality Archive: http://ruccs.rutgers.edu/roa.html.
Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.
Trudgill, P. (1999). *The dialects of England*. 2nd edition. Oxford: Blackwell.
Trudgill, P., and J. Chambers (1991). Verb systems in English dialects. In Trudgill, P. & J. Chambers (eds.), *Dialects of English: studies in grammatical variation*. London: Longman. 49–53.
Urbanczyk, S. (1995). Double reduplications in parallel. In Beckman, J., L. W. Dickey & S. Urbanczyk (eds.), *Papers in Optimality Theory* 18. University of Massachusetts Occasional Papers in Linguistics. Amherst: GLSA, UMASS. 499–531.
van Oostendorp, M. (1997). Style levels in conflict resolution. In Hinskens, F., R. van Hout & W. Leo Wetzels (eds.), *Variation, change and phonological theory*. Amsterdam: John Benjamins. 207–29.
Wolfram, W. (2000). Issues in reconstructing earlier African-American English. *World Englishes* **19**: 39–58.
Zwicky, A. & G. Pullum (1983). Cliticization versus inflection: English 'nt'. *Language* **59**: 502–13.