To appear in Andronis, Mary, Christopher Ball, Heidi Elston and Sylvain Neuvel eds. CLS 37: The Main Session. Papers from the 37th Meeting of the Chicago Linguistic Society. Vol. 1. Chicago: Chicago Linguistic Society.
Manuscript revised May 16, 2001
Comments welcome; address correspondence to the author by email at: akg@alumni.indiana.edu

A stochastic OT approach to word order variation in Korlai Portuguese

Andrew J. Koontz-Garboden
Stanford University

1. Introduction

Boersma (1998, 2000) and Boersma and Hayes (1999, 2001) have recently proposed a stochastic version of optimality theory (OT; Prince and Smolensky 1993) along with a learning algorithm, the Gradual Learning Algorithm (GLA). This framework offers a great deal of promise in accounting for variable linguistic phenomena, which in general have received less attention in the theoretical literature. Boersma and Hayes exemplify stochastic OT with various empirical test-cases of phonological variation. Although OT in general has been increasingly applied in the realm of syntax (e.g. Aissen 1999, 2000; Bresnan 1999a,b, 2000a,b; Choi 1999; Grimshaw 1997, 2001; Sells 2001 among others), there have been relatively few applications of stochastic OT to syntactic phenomena (e.g. Asudeh 2001, Bresnan and Deo 2001, Bresnan, Dingare, and Manning in prep., Clark 2001, and Manning 2001). In what follows, the study of stochastic OT in the realm of syntax is furthered by applying this framework to a case of synchronic word order variation in Korlai Portuguese (KP; Clements 1990, 1996, 2000, 2001). The application of stochastic OT to this problem is argued to have at least two advantages. First, it provides a possible alternative to the competing grammars account of synchronic instability for synchronic stages where "mutually inconsistent" types of word order are observed to arise in the same system (cf. Kroch 1989, 1994). Second, it improves on Clements' (2001) parameter-based account of head-complement/head-adjunct order in Korlai Portuguese. Further, it is argued that the type of word order variation observed in KP is actually predicted to occur given the stochastic OT architecture, and given a small set of violable constraints. Crucially, the same constraints that constrain categorical phenomena also constrain variable phenomena. This idea has been stated quite clearly by Bresnan and Deo (2001: 40), who argue that under stochastic OT, "…individual variation samples the typological space of possible grammars."

Below, I first give a brief introduction to the stochastic OT framework. This is followed by a presentation of the types of word order variation found in KP. Also presented in this section are the frequency data collected by Clements (2001) on verb/object and verb/adjunct order in KP, which are at the core of the present study. Next, a stochastic OT account of these phenomena is developed by conducting a series of computer simulations using Clements' (2001) data and the *OTSoft* software package of Hayes, Tesar, and Zuraw (2000). Finally, I make some concluding remarks.

2. A brief introduction to Boersma and Hayes' stochastic OT framework

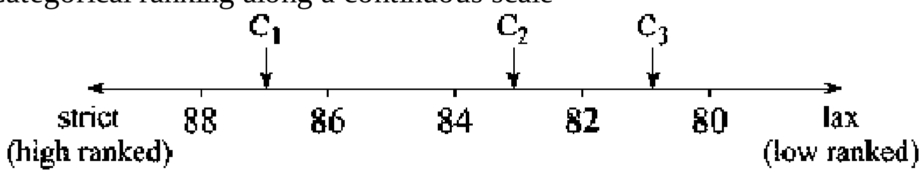
Under the standardly-held view of optimality theory (e.g. Prince and Smolensky 1993), grammars of individual languages are composed of categorically ranked constraints, which strictly dominate one another, as in (1).

(1) Hypothetical strict domination ranking in a language X
$C_1 >> C_2 >> C_3$

Under the view in (1), C_2 in language X, can never come to dominate C_1, or C_3 dominate C_2.[1] Thus, C_1 is said to strictly dominate C_2, while C_2 strictly dominates C_3 (with C_1 strictly dominating C_3 by transitivity).

Boersma and Hayes alter this view in two crucial ways.[2] First, they adopt a continuous linear ranking scale, as shown graphically in (2), which depicts a possible constraint ranking from (1).

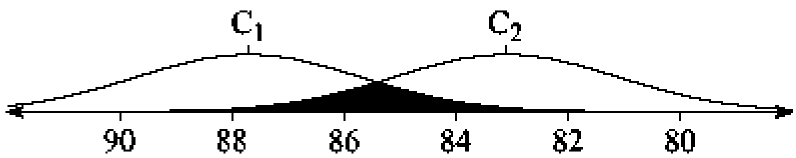
(2) Categorical ranking along a continuous scale[3]



On the scale in (2), higher numerical values correspond to a higher ranking. Thus, it might be said, for example, that C_1 outranks C_2 more than C_2 outranks C_3.

The next crucial alteration is stochastic candidate evaluation. Under this view, every time a candidate set is evaluated, the position of each constraint is perturbed by a random variable (either in a positive or negative direction). Boersma and Hayes (2001: 48) refer to the permanent value of a constraint as the ranking value, while a constraint's value on any given evaluation is called the selection point. Thus, each candidate has a range of selection points associated with its ranking value. These ranges are interpreted as normal probability distributions. This state of affairs is illustrated graphically in (3).

(3) Relative constraint rankings vary[4]



Under this view, categorical constraint ranking arises as a special case when two constraints have ranking values relatively far away from one another. When two constraints have relatively close ranking values, however, their ranges may overlap, as in the shaded portion of (3), allowing for selection points where on one evaluation $C_2 >> C_1$, while on others $C_1 >> C_2$.

Along with these changes, Boersma and Hayes adopt the GLA, developed by Boersma (1998, 2000), for the learning of stochastic OT grammars. In what follows, the GLA is used as part of Hayes et al.'s (2000) *OTSoft* software in the examination of KP word order variation.

3. Word order variation in Korlai Portuguese

KP, a Portuguese-based creole spoken on the western coast of India, is currently undergoing a typological shift as the result of contact with Marathi (Clements 1990, 1996, 2001). This shift, according to Clements, is resulting in variation in order within subordinate clauses (Clements 1996: 182-183), the order of nominal modification (Clements 1996: 166), adposition use (Clements 1996: 143-145), verb/object order, and verb/adjunct order. It is on the latter two areas of variation, for which substantial data collection has been undertaken, that we focus below.

According to Clements (1996: 152), "depending on the speaker, KP displays [verb/object and object/verb order] without any appreciable distinction in meaning, such as a focused vs. a non-focused reading." Thus, both (4) and (5) are grammatical in KP, and carry no distinction in meaning.

(4) Teru buk ulyan.
Teru book looking.
'Teru is looking at a/the book.' (Clements 1996: 152, (3b))

(5) Teru ulyan buk.
Teru looking book.
'Teru is looking at a/the book.' (Clements 1996: 152, (3b'))

The same is true of adjunct/verb order, so that both (6) and (7) are grammatical without any distinction in meaning.

(6) Teru tana amya.
Teru is going tomorrow.
'Teru is going tomorrow.'

(7) Teru amya tana.
Teru tomorrow is going.
'Teru is going tomorrow.'

The evidence for a typological shift in KP comes by way of Clements' (2001) quantitative examination of the order of object NPs (DOs and IOs) relative to the verb and the order of locative, goal, instrumental, and sentential adjuncts relative to the verb. Clements considered only declarative main clauses and direct quotes, so as not to include utterances carrying different pragmatic meaning. The source of the data was narratives from 16 KP native-speakers collected at the speakers' homes or at Clements' residence in the village. Clements' (2001: 5) analysis revealed that these speakers fall into four different groups, as in (8), with stage 1 representing archaic speech, and stage 4 representing speech of younger speakers. Stages 2 and 3 are intermediate to these two points.[5]

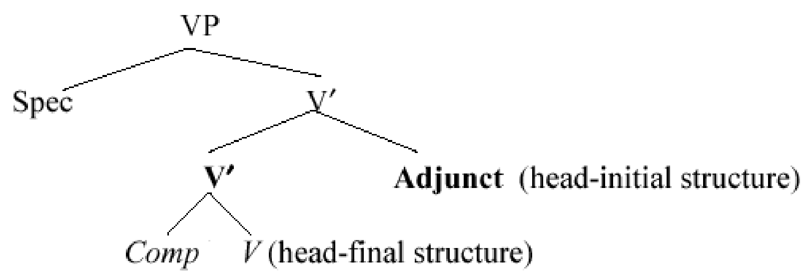
As can be seen in (8), whereas stage 1 speech has predominately VO constituent order in main clauses, stage 4 speech exhibits a tendency toward OV order in main clauses. Also evident in (8) is a shift from VAdjunct order to AdjunctV order.

(8) The four stages of VO –> OV and VAdjunct –> AdjunctV shift[6]

ORDERS	**VO**	**OV**	**VAdjunct**	**AdjunctV**
Stage 1	**80%** (101/127)	**20%** (26/127)	**90%** (44/49)	**10%** (5/49)
Stage 2	**51%** (83/163)	**49%** (80/163)	**92%** (99/108)	**8%** (9/108)
Stage 3	**42%** (127/300)	**58%** (173/300)	**56%** (75/133)	**44%** (58/133)
Stage 4	**27%** (209/767)	**73%** (558/767)	**53%** (361/676)	**47%** (315/676)

Two aspects of the shift made evident by the data in (8) are under investigation here. The primary concern is how to formally account for the indeterminate status of stage 2 and the largely indeterminate status of stage 3 with respect to verb/object ordering (i.e. competing grammars, or some other mechanism). A secondary concern is how to account for the very slow shift of verbs relative to adjuncts when compared to the shift in verb/complement order. Clements' (2001) solution to these problems is to posit, in addition to head-initial and head-final, a head-initial & head-final parameter setting, as shown in (9).

(9) Structure resulting from Clements head-initial & head-final parameter[7]



Even under this approach, however, stages such as stage 2 are still indeterminate with respect to all three parameters, since there appears to be nothing to motivate OV ordering over VO ordering and vice-versa.

One solution to this problem is the competing grammars approach of Kroch (1989, 1994), under which it might be said that KP has two separate but competing grammars: one head-initial grammar and one head-final grammar. While this is a possible explanation, an alternative analysis is available under stochastic OT. Indeed, stochastic OT offers the possibility of accounting for the indeterminacy problem, while at the same time avoiding the positing of two separate intra-speaker grammars.[8]

4. A stochastic OT account of KP word order variation

In this section, I first outline the OT constraints relevant for the analysis of verb/object variation in KP. This is followed by discussion of the constraints relevant for verb/adjunct ordering. I then turn to simulation of the KP grammar using these constraints.

4.1 Constraints on verb/object order

In order to account for the relative order of verbs with respect to objects in KP, I adopt the constraints on clause structure recently proposed by Grimshaw (2001). Although Grimshaw develops an analysis that allows for the absence of structure (e.g. the absence of heads, specs, etc.), since only very simple main clause declaratives (all of which have a verb and an object) are under consideration here, only the three constraints in (10) are relevant. It should also be made explicit that for the present purposes, a clause is assumed to be a VP.

(10) Constraints on clause structure (Grimshaw 2001)

SPEC-LFT	A specifier is leftmost in the projection which immediately dominates it.
HD-LFT	A head is leftmost in its projection.
HD-RT	A head is rightmost in its projection.

If we assume something along the lines of the VP-internal subject hypothesis (e.g. Kitagawa 1986) and note that subjects most often appear to the left of both the verb and the object in all varieties of KP (Clements p.c.), then it seems safe to assume that SPEC-LFT is highly ranked in KP, and is only rarely if ever violated, at least in main clause declaratives. The two crucial constraints for the purpose of examining verb/object order, then, are HD-LFT and HD-RT. The tableau in (11) illustrates how these constraints conflict with one another.

(11) HD-LFT and HD-RT in conflict

	SPEC-LFT	HD-LFT	HD-RT
a.[$_{VP}$ Teru [$_{V'}$ ulyan [$_{DP}$ buk]]. (SVO)		*	*
b.[$_{VP}$ Teru [$_{V'}$ [$_{DP}$ buk] ulyan]. (SOV)		**	

Under Grimshaw's (2001) analysis, a trace and an actual lexical item count equally in the computation of the head constraints. Assuming that the subject is either VP-internal, or leaves a trace in spec-VP after movement, there will always be either the actual subject or its trace in the left edge of the VP. Either way, it or its trace will always be at the left edge of the VP, due to the high-ranking of SPEC-LFT. For ease of presentation, it is assumed that the subject remains VP internal. This assumption, as explained, is not crucial, however. A VO candidate, such as candidate a, will incur one violation each of HD-LFT and HD-RT, since in this construction, the head of the VP (i.e. the verb) is situated between the specifier and the complement. An OV candidate, such as candidate b, however, will fully satisfy HD-RT, at the cost of incurring two violations of HD-LFT, since the head is two positions away from the left edge of the projection. Given these two constraints, then, within a strict domination approach to OT it would be predicted that cross-linguistically we would find head-initial languages and head final languages, which, of course, we do. These two types of languages are also predicted to exist under the stochastic OT framework; if HD-LFT is ranked sufficiently far away from HD-RT then the result will be a system that is categorically VO.[9] If HD-RT is ranked far enough away from HD-LFT, then the result will be a system that is categorically OV. Stochastic OT, however, predicts an additional type of system not predicted by a strict domination approach: systems for which HD-LFT and HD-RT are ranked very close to one another such that their ranges overlap, as in (3). It is this type of a system that we find in KP, as will be seen more clearly below.

4.2 Constraints on adjunction

As observed in (8), and as noted by Clements, the relative order of the adjunct with respect to the verb is somewhat independent of the relative order of the verb with respect to its object. It thus seems that additional constraints are needed to regulate adjunction. For present purposes, then, the two constraints in (12) are posited.

(12) Constraints on adjunction

- AJ-RT — An adjunct is rightmost in the projection that immediately dominates it.
- AJ-LFT — An adjunct is leftmost in the projection that immediately dominates it.

How these constraints interact with one another and with the constraints on heads is illustrated in (13), with evaluation parallel to that discussed for the HD constraints in (11).

(13) Adjunct/verb order

	SPEC-LFT	AJ-LFT	AJ-RT	HD-LFT	HD-RT
a.[$_{VP}$ Teru [$_{V'}$ tana [$_{AP}$ amya]]. (SVA)		**		*	*
b.[$_{VP}$ Teru [$_{V'}$ [$_{AP}$ amya] tana]. (SAV)		*	*	**	

One problem with independent constraints on adjunction, such as those in (12), as Julie Auger (p.c.) has noted, is that they fail to capture the fact that, although the two shifts seem to be independent of one another, verb/object order and verb/adjunct order are indeed moving in the same direction diachronically. It is hoped that in future research a suitable alternative to the constraints on adjunction may allow for the same generalizations to be captured, while at the same time also capturing the fact that the two orders do move diachronically in the same direction. Such work is beyond the scope of the present study, however, and the constraints on adjunction are therefore assumed despite their possible deficiencies.[10]

4.3 Simulating the KP grammar

In order to determine whether or not these constraints implemented in stochastic OT with the GLA could account for the word order variation observed throughout the various stages of KP, a series of computer simulations was performed using the *OTSoft* software package of Hayes et al. (2000). The OT constraints assumed are those discussed above: HD-LFT, HD-RT, AJ-LFT, and AJ-RT. SPEC-LFT is assumed to be highly-ranked, and is therefore not considered in the simulations

The simulations proceeded as follows.[11] Each of the four stages of the KP shift documented by Clements (2001) was conceived of as an individual grammar, i.e. as the grammar of an individual KP speaker. Experimentally, this was achieved by constructing an individual simulation for each of the four stages. Each simulation consists of two steps. The first step is the learning of a grammar using the GLA. This step was carried out by exposing the algorithm to one million input data. On the basis of these data, the GLA arrived at a steady-state grammar with a ranking value for each of the constraints at each of the four stages. The second step is the testing of that grammar. This is achieved by repetition of the process of stochastic evaluation and the comparison of the overall results to the relative frequencies that were actually observed by Clements. In this case, the process of stochastic evaluation was repeated two million times.[12]

With respect to step one, the output frequency for each stage independently observed by Clements was input into the program as if it were the input a child of the stage in question received as her primary linguistic data. This is a bit of an idealization, but a necessary and justifiable one. The problem is that if stage 1 speakers have 80% VO as their output, then how do stage 2 speakers have 51% VO as their primary linguistic data? What is being glossed over here is how one gets from point A to point B. It is assumed that in this particular case this involves social mechanisms having to do with language contact (Thomason and Kaufman 1988), thus making the issue largely irrelevant from a synchronic perspective. This constitutes a problem of "actuation" in the terminology of Weinreich, Labov, and Herzog (1968) and is beyond the scope of the present study.

As an example of how the simulations proceeded, consider stage 3 of the KP shift, characterized by the frequency data in (14) (repeated from (8)).

(14) Stage 3 frequencies

ORDER	**VO**	**OV**	**VAdjunct**	**AdjunctV**
Stage 3	**42%** (127/300)	**58%** (173/300)	**56%** (75/133)	**44%** (58/133)

The first step of the simulation was to have the GLA learn a grammar assuming as input (i.e. primary linguistic data) the data in (14). In order to do this, *OTSoft* was given two OT tableaux similar to those in (15), which evaluated object/verb order, and (16), which evaluated adjunct/verb order.

(15) Verb/object GLA input to stage 3

	AJ-LFT	AJ-RT	HD-LFT	HD-RT
42% a.[$_{VP}$ Teru [$_{V'}$ ulyan [$_{DP}$ buk]]. (SVO)			*	*
58% b.[$_{VP}$ Teru [$_{V'}$ [$_{DP}$ buk] ulyan]. (SOV)			**	

(16) Adjunct/verb GLA input to stage 3

	AJ-LFT	AJ-RT	HD-LFT	HD-RT
56% a.[$_{VP}$ Teru [$_{V'}$ tana [$_{AP}$ amya]]. (SVA)	**		*	*
44% b.[$_{VP}$ Teru [$_{V'}$ [$_{AP}$ amya] tana]. (SAV)	*	*	**	

Upon being exposed to the data in (15) and (16) one million times, the GLA arrived at the ranking values in (17), with each constraint having started from an arbitrary initial state of 100.

(17) Stage 3 constraint ranking

Constraint	Ranking Value
AJ-RT	100.708
HD-RT	100.228
HD-LFT	99.772
AJ-LFT	99.292

Following the learning of the grammar by the GLA, the stable grammar was tested to determine the extent to which it generates the variation as observed by Clements (2001). This testing was achieved, as described above, by repeating the process of stochastic evaluation two million times. When this was carried out with the stage 3 grammar, the frequencies in (18) and (19) were recorded as the output generated by the steady-state grammar.

(18) Verb/object output generated by stage 3 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SOV]	58%	56.4%	1127645
[SVO]	42%	43.6%	872355

(19) Verb/adjunct output generated by stage 3 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SVA]	56%	58.0%	1160947
[SAV]	44%	42.0%	839053

As can be seen in (18) and (19) by examining the input frequencies with respect to the generated frequencies, the grammar learned by the GLA generates output frequencies that are strikingly close to those actually observed by Clements. In other words, the grammar generates the type of variation observed for stage 3 KP speakers.

The table in (20) lays out the ranking values for each of the constraints at each of the four stages following exposure to the data in each of Clements' four stages and acquisition by the GLA.

(20) Constraint rankings across all four stages

Stage 1	Stage 2	Stage 3	Stage 4	Constraint
101.855	103.670	100.708	101.061	AJ-RT
98.145	96.330	99.292	98.939	AJ-LFT
101.092	99.930	99.772	99.088	HD-LFT
98.908	100.070	100.228	100.912	HD-RT

The tables in (21)-(24) give the frequencies of VO/OV order generated by the grammar in each of the four stages when the process of stochastic evaluation is repeated two million times (stage 3 in (18) is repeated as (23)).

(21) Verb/object output generated by stage 1 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SVO]	80%	78.0%	1559600
[SOV]	20%	22.0%	440400

(22) Verb/object output generated by stage 2 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SVO]	51%	48.0%	960775
[SOV]	49%	52.0%	1039225

(23) Verb/object output generated by stage 3 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SOV]	58%	56.4%	1127645
[SVO]	42%	43.6%	872355

(24) Verb/object output generated by stage 4 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SOV]	73%	74.0%	1480838
[SVO]	27%	26.0%	519162

As can be readily observed in (21)-(24), the output frequencies generated by the grammars learned by the GLA match very closely the actual frequencies observed by Clements (2001). The case is the same for verb/adjunct order, as is shown in (25)-(27) (with stage 3 reproduced from (19) as (27)).

(25) Verb/adjunct output generated by stage 1 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SVA]	90%	89.5%	1790114
[SAV]	10%	10.5%	209886

(26) Verb/adjunct output generated by stage 2 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SVA]	92%	91.2%	1824156
[SAV]	8%	8.8%	175844

(27) Verb/adjunct output generated by stage 3 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SVA]	56%	58.0%	1160947
[SAV]	44%	42.0%	839053

(28) Verb/adjunct output generated by stage 4 grammar

	Input Frequencies	Generated Frequencies	Number Generated
[SVA]	53%	52.4%	1048197
[SAV]	47%	47.6%	951803

With respect to the changes from one stage to another, it is worth pointing out that formally this is modeled as an increase or decrease in the relative distance of constraints with respect to one another. So, for example, a greater degree of VO or OV order at any given stage is the result of HD-RT and HD-LFT being ranked closer to or farther away from one another, as shown in (29). As the language goes from VO to OV, HD-RT goes from being ranked well below HD-LFT to being ranked well above it. Crucially, at the intermediate stages where order is free, the rankings of the constraints are closer to one another, as represented by the smaller numerical difference between the constraint rankings at those stages. The case with the adjunct constraints is similar.

(29) Differences between constraints at the four stages

Stage 1	Stage 2	Stage 3	Stage 4	Constraints
-2.184	-0.140	0.456	1.824	HEAD (HD-RT LESS HD-LFT)
3.710	7.340	1.416	2.122	ADJUNCT (AJ-RT LESS AJ-LFT)

5. Discussion and concluding remarks

Several important theoretical conclusions can be drawn from the results of this study.

First, they suggest that the stochastic OT framework developed by Boersma and Hayes (2001) using evidence from the realm of phonology is quite easily extended to the domain of syntactic variation. Indeed, given a set of well-motivated constraints on syntactic structure, such as those of Grimshaw (1997, 2001), the extension to syntax of their framework is straightforward.[13]

Also, the positing of separate constraints for adjunct/head order and head/comp order, while not without problems, allows us to capture Clements' observation that the shift in KP from VO to OV is somewhat independent of the shift in the order of adjuncts with respect to verbs. This obviates the need for a UG parameter such as Head-initial & Head-final (cf. Clements 2001); this result falls out of the constraint set. Further, the change in the relative frequencies of verb/object order and verb/adjunct at each of the four stages can be captured through small incremental shifts in the distance between the HD constraints for verb/object order and between the AJ constraints for verb/adjunct order.

Finally, whereas past studies on syntactic change (e.g. Kroch 1989, 1994) posit competing grammars to account for variation (e.g. stages 2 and 3 in (8)), a stochastic OT analysis actually predicts that such variation arises in a single grammar and is governed by the same linguistic constraints that govern categorical phenomena. Given the fact that cross-linguistically we find languages that are strictly VO and languages that are strictly OV, there can be no universal ranking of HD-LFT and HD-RT with respect to one another. If this is the case, then under stochastic OT it is predicted that cases will be found where not only are these constraints ranked very far away from one another, yielding categorical VO or OV ordering, but also where the constraints in question are ranked more closely to one another. In such a case, variation is expected to arise, since constraints cover normally distributed ranges and overlap when their ranking values are relatively close to one another. In KP we find precisely this type of system.

Thus, under stochastic OT, we may be able to account for phenomena currently attributed to competing grammars within a single stochastic OT grammar. Much more research, of course, is needed before such a conclusion can be definitively made. There are important outstanding questions, such as whether or not having overlapping ranges, such as in the case of HD-LFT and HD-RT above, can be diachronically stable, and if not, how to formally model this.[14] The discussion and findings reported here nonetheless demonstrate that stochastic OT merits further study as an alternative to the competing grammars account of linguistic systems which exhibit a great deal of seemingly typologically inconsistent variation, as well as for predicting the types of variable patterns commonly described in variationist studies.

Acknowledgements

During the preparation of this paper, I have benefited from discussion with Ash Asudeh, Joan Bresnan, Daniel Dinnsen, Ken de Jong, Ashwini Deo, Christopher Manning, and Elizabeth Traugott. I am especially grateful to Bruce Hayes for help with the *OTSoft* program, to Clancy Clements for sharing his knowledge of Korlai Portuguese grammar, and to Julie Auger, Paul

Boersma, Brady Clark, and Steven Franks for offering helpful comments during various stages of the paper's development. I also acknowledge useful feedback from members of the audience at CLS-37. I alone am responsible for the errors.

Throughout discussions with some of the people mentioned above, it has become apparent that the idea that stochastic OT might constitute an alternative to the competing grammars approach to syntactic change appears to have been reached independently by several researchers. In particular, Paul Boersma (p.c. 1/13/2001), Brady Clark (p.c. 3/8/2001, 2001), and Joan Bresnan and her associates (cf. Bresnan and Deo 2001, Bresnan, Dingare, and Manning in prep.) have also had ideas similar to those presented here.

Notes

[1] Of course, there might exist as a special case locally conjoined constraints which allow, e.g. C3 to dominate C1 under a special set of circumstances. See Smolensky (1995), Lubowicz (1998), and Itô and Mester (1999) on the use of local conjuction in phonology. See Aissen (1999, 2000) on the use of local conjunction in OT syntax.

[2] The very brief outline given here follows the discussion in Boersma and Hayes (2001: 47-50).

[3] The graphic in (2) is adapted from Boersma and Hayes (2001: 47), (1).

[4] The graphic in (3) is adapted from Boersma and Hayes (2001: 49), (6).

[5] To be more precise, stage 1 represents data collected from one older woman who memorized stories told to her as a child by her mother-in-law. Stage 2 data were collected from the narratives of a speaker who at the time of recording had been out of the Korlai village for 46 years. Stage 3 represents the narratives of three older KP speakers (65, 73, 63 at time of recording) who have resided in Korlai for their entire lives. Stage 4 data come from narratives of eleven native KP-speaking children, whose ages range from 7-16. Although Stages 3 and 4 are composed of multiple numbers of speakers, the averages given in (8) are representative of intra-speaker norms for those stages (Clements p.c.), which is what is being modeled with a stochastic OT grammar.

It is also worth making clear that the frequencies in (8) represent the frequency of occurrence of the type of clause in question out of the total collected. E.g. in stage 1 there were 127 VO or OV clauses collected, 101 of which were VO, and 26 of which were OV.

[6] The table in (8) is adapted from Clements (2001, Table 1). Clements' chi-square statistics on this table are the following: VO→OV: df 3, $\chi 2$= 144, p<0.001. VAdjunct→AdjunctV: df 3, $\chi 2$= 72.2, p<0.001.

[7] This graphic is adapted from Clements (2001: 11).

[8] In fact, it seems that a competing grammars account of the KP facts would require not two competing grammars, but more likely four, since the rate of change of the VO to OV shift appears to be quite different from the rate of change of the VAdjunct to AdjunctV shift. On the problem of independent variation leading to an increasing number of intra-speaker competing grammars, see Bresnan and Deo (2001: 40). For additional criticism of the competing grammars approach, especially as a response to register variation, see Newmeyer (2000: 48).

[9] The question of categoricity is slightly more complicated than this, since the values of the normal distribution never actually reach zero. For all practical purposes, however, categoricity can be said to arise. See Boersma and Hayes (2001: 50) for discussion of this point.

[10] Brady Clark (p.c. 5/8/2001) has suggested as an alternative to the present constraint set, one modeled on work by Peter Sells (2001), in which a constraint on adjunction is motivated. This constraint set would consist of ADJUNCT-LEFT, HEAD-LEFT, and OBJECT-LEFT, preserving portions of Kayne's (1994) antisymmetry proposal. In an implementation of the GLA using these three constraints and Boersma and Weenik's (2001) Praat program, Clark was able to model the frequency data in (8) in a manner similar to that carried out below. This constraint set, then, seems like a promising alternative to the one proposed here, and the choice between the two will ultimately come down to the predictions made by each set. Neither constraint set, however, solves the directionality problem mentioned above (i.e. that verb/object order and verb/adjunct order are

drifting in the same direction). Clark suggests that this might be accounted for by factors external to the grammar (e.g. Hawkins 1994).

[11] Throughout the simulations, evaluation noise was held constant at 2. Falling plasticities were used in the learning schedule with a beginning value of 2 and a final value of .002.

[12] A real life analogy of this computational exercise would be to ask a KP speaker to say the same sentence two million times, and then record the proportion of VO to OV utterances over the two million tokens.

[13] It should be noted, as Boersma and Hayes (2001: 54) have pointed out, that the stochastic OT framework is only as good as the inventory of constraints that is assumed; it crucially relies on a well-motivated set of constraints.

[14] Another important outstanding question, how to formally model Kroch's (1989) constant rate effect, is explored by Clark (2001).

References

ROA=http://ruccs.rutgers.edu/roa.html

Aissen, Judith. 1999. Markedness and subject choice in optimality theory. *Natural Language and Linguistic Theory* 17/4:673-711.

Aissen, Judith. 2000. Differential object marking: iconicity vs. economy. Ms., UC, Santa Cruz. Santa Cruz, CA.

Asudeh, Ash. 2001. Linking, optionality, and ambiguity in Marathi. Ms., Stanford University. Stanford, CA. [ROA-424]. To appear in *Formal and Empirical Issues in Optimality-theoretic Syntax*, ed. by Peter Sells. Stanford: CSLI.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Boersma, Paul. 2000. Learning a grammar in Functional Phonology. In Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (Eds.), 465-523.

Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32/1:45-86.

Boersma, Paul and David Weenik. 2001. Praat. Available on the www at http://www.praat.org.

Bresnan, Joan. 1999a. Explaining morphosyntactic competition. In *Handbook of Contemporary Syntactic Theory*, ed. by Mark Baltin and Chris Collins, 11-44. Oxford: Blackwell.

Bresnan, Joan. 1999b. The emergence of the unmarked pronoun. To appear in *Optimality-theoretic Syntax*, ed. by Géraldine Legendre, Sten Vikner, and Jane Grimshaw. Cambridge, MA: MIT Press.

Bresnan, Joan. 2000a. Optimal syntax. In Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (Eds), 334-385.

Bresnan, Joan. 2000b. Pidgin genesis in optimality theory. To appear in *Processes of Language Contact: Case Studies from Australia and the Pacific*, ed. by Jeff Siegal. Montreal: Les Editions Fides.

Bresnan, Joan and Ashwini Deo. 2001. 'Be' in the survey of English dialects: a stochastic OT account. Ms., Stanford University. Stanford, CA.

Bresnan, Joan, Shipra Dingare, and Christopher Manning. In preparation. Soft constraints mirror hard constraints: voice and person in English and Lummi. Ms., Stanford University. Stanford, CA.

Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford: CSLI.

Clark, Brady. 2001. Modeling the constant rate of change in stochastic optimality theory. Ms., Stanford University. Stanford, CA.

Clements, J. Clancy. 1990. Deletion as an indicator of SVO→SOV shift. *Language Variation and Change* 2:103-133.

Clements, J. Clancy. 1996. *The Genesis of a Language: the Formation and Development of Korlai Portuguese*. Creole Language Library 16. Amsterdam: John Benjamins.

Clements, J. Clancy. 2000. Korlai Creole Portuguese. Ms., Indiana University. To appear in *Comparative Creole Syntax*, ed. by John Holm and Peter Patrick. London: Battlebridge Publications.

Clements, J. Clancy. 2001. Word order shift and natural L2 acquisition in a Portuguese creole. To appear in *Romance Syntax, Semantics and L2 Acquisition*, ed. by Caroline Wiltshire and Joaquim Camps. Amsterdam: John Benjamins.

Dekkers, Joost, Frank van der Leeuw, and Jeroen van de Weijer (Eds.). 2000. *Optimality Theory: Phonology, Syntax, and Acquisition.* Oxford: Oxford University Press.

Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28/3:373-422.

Grimshaw, Jane. 2001. Clause structure. Plenary talk at the 2001 LSA annual meeting, Washington, D.C., January 5, 2001.

Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge, UK: Cambridge University Press.

Hayes, Bruce, Bruce Tesar, and Kie Zuraw. 2000. OTSoft. Software package with documentation, available on the WWW at http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Itô, Junko and Armin Mester. 1999. On the sources of opacity in OT: coda processes in German. Ms, UC, Santa Cruz. Santa Cruz, CA. [ROA-347].

Kayne, Richard S. 1994. *The Antisymmetry of Syntax.* Linguistic Inquiry monograph 26. Cambridge, MA: MIT Press.

Kitagawa, Yoshihisa. 1986. *Subjects in Japanese and English.* Ph.D. dissertation, University of Massachusetts, Amherst. Amherst, MA.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199-244.

Kroch, Anthony. 1994. Morphosyntactic variation. In *Papers from the 30th regional meeting of the Chicago Linguistic Society: parasession on variation and linguistic theory*, ed by Katharine Beals, Jeannette Denton, Robert Knippen, Lynette Melnar, Hisami Suzuki, and Erica Zeinfeld, 180-201.

Łubowicz, Anna. 1998. Derived environment effects in OT. Ms, University of Massachusetts, Amherst. Amherst, MA. [ROA-239].

Manning, Christopher. 2001. Probabilistic approaches to syntax. Paper presented at the Workshop on probability theory in linguistics held at the 2001 meeting of the LSA. Washington, DC, January 6, 2001.

Newmeyer, Frederick. 2000. Optimality and functionality: some critical remarks on OT syntax. Ms., University of Washington. Seattle, WA. [ROA-402].

Prince, Alan and Paul Smolensky. 1993. Optimality theory: constraint interaction in generative grammar. Ms. Rutgers University and University of Colorado.

Sells, Peter. 2001. *Structure, Alignment, and Optimality in Swedish.* Ms. Stanford University. To appear, Stanford: CSLI.

Smolensky, Paul. 1995. On the internal structure of the constraint component of UG. Handout of talk presented at UCLA, April 7, 1995. [ROA-86].

Thomason, Sarah and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics*, ed. by W.P. Lehmann and Yakov Malkiel, 95-188. Austin, TX: University of Texas Press.

Department of Linguistics
Margaret Jacks Hall, Building 460
Stanford University
Stanford, CA 94305-2150

andrewkg@stanford.edu
http://www.stanford.edu/~andrewkg/