Meaningful Variability: A Sociolinguistically-Grounded Approach to Variation in Optimality Theory

JUAN ANTONIO CUTILLAS ESPINOSA[*]
Universidad de Murcia

ABSTRACT

Most approaches to variability in Optimality Theory have attempted to make variation *possible* within the OT framework, i.e. to reformulate constraints and rankings to accommodate variable and gradient linguistic facts. Sociolinguists have attempted to apply these theoretical advances to the study of language variation, with an emphasis on language-internal variables (Auger 2001, Cardoso 2001). Little attention has been paid to the array of external factors that influence the patterning of variation. In this paper, we argue that some variation patterns —specially those that are socially meaningful— are actually the result of a three-grammar system. G_1 is the standard grammar, which has to be available to the speaker to obtain these variation patterns. G_3 is the vernacular grammar, which the speaker is likely to have acquired in his local community. Finally, G_2 is an *intergrammar*, which is used by the speaker as his 'default' constraint set. G_2 is a continuous ranking (Boersma & Hayes 2001) and domination relations are *consciously* altered by the speakers to shape the appropriate and variable linguistic output. We illustrate this model with analyses of English and Spanish.

KEYWORDS: Variability, Sociolinguistics, External Factors, Continuous Ranking, Three-Grammar System.

_____

* *Address for correspondence*: Juan Antonio Cutillas Espinosa, Departamento de Filología Inglesa, Facultad de Letras, Universidad de Murcia, Plaza de la Universidad s/n, 30.001 Murcia (Spain). Tel.: +34 968 363 416. Fax: +34 968 363 185. E-mail: jacuti@um.es

I. INTRODUCTION

Explaining language variation has been one of the main aims of Optimality Theory from its very origins. There is an essential motivation for such an endeavour. Labov's pioneering work (1966a, 1972) helped to establish two main principles: firstly, language is essentially variable; secondly, this variation is principled and should, therefore, be the subject of attention of linguistic theory. This is precisely why he suggested that sociolinguistics was nothing but *real* linguistics, pointing at the error of analysing language as an idealised, invariable model. Thus, if language is essentially variable, any acceptable grammatical model should be able to explain and produce variability. In other words, if Optimality Theory had proved to be unable to produce variable patterns, it would have lost part of its plausibility as an account of how grammar works.

There are two key issues in the discussion of variability. First of all, the implications for the standard OT conception of *constraints* and *rankings*. At the beginning, linguists working in OT had to decide how to make variation possible, that is to say, how to adapt the OT system to produce variable patterns. Once this has been achieved, a second question arises. It is not enough that OT grammars can produce variation. We also have to check whether these mechanisms of variability are consistent with what we know about language variation and change, basically through the findings of sociolinguistics.

II. IMPLICATIONS FOR CONSTRAINTS AND RANKINGS

Most OT analyses focus on situations where a well-established constraint ranking ($\mathcal{H}$) evaluates all the possible candidates submitted by GEN and chooses one single optimal candidate. Some of the constraints in that non-variable analysis may not be ranked, but this is non-crucial, i.e. it does not have consequences for the selection of the optimal candidate. This is illustrated in (1). Candidate *a* is selected by a ranking where $\mathbb{C}_1$ and $\mathbb{C}_2$ have equal importance.

(1)

/abc/	$\mathbb{C}_1$	$\mathbb{C}_2$	$\mathbb{C}_3$
☞ a. [abc]		*	*
b. [def]	*	*!	

If we establish a domination relationship of the type $\mathbb{C}_1 \gg \mathbb{C}_2$, there will be no change in the selected candidate (2). In this sense, non-ranking of constraints is just the expression of a lack of arguments favouring one domination relationship over the other, given that the resulting optimal candidate is going to be the same.

(2)

/abc/	$\mathbb{C}_1$	$\mathbb{C}_2$	$\mathbb{C}_3$
☞ a. [abc]		*	*
b. [def]	*!	*	

One of the obvious possibilities when trying to produce variability is to introduce mechanisms to obtain grammatical systems that do not choose just *one* single candidate or two candidates in equal proportions. Nagy & Reynolds (1996, 1997) suggest the concept of 'floating constraints'. They claim that a floating constraint should be allowed to move freely within its domain, even when the rest of constraints are hierarchically ranked. Thus, if we allow a constraint $\mathbb{C}_4$ to float freely in the hierarchically ranked domain $\{\mathbb{C}_1 \gg \mathbb{C}_2 \gg \mathbb{C}_3\}$ we would get all these different possible combinations:

(3) $\mathbb{C}_1 \gg \mathbb{C}_2 \gg \mathbb{C}_3 \gg \mathbb{C}_4$ $\qquad$ $\mathbb{C}_1 \gg \mathbb{C}_2 \gg \mathbb{C}_4 \gg \mathbb{C}_3$

$\qquad$ $\mathbb{C}_1 \gg \mathbb{C}_4 \gg \mathbb{C}_2 \gg \mathbb{C}_3$ $\qquad$ $\mathbb{C}_4 \gg \mathbb{C}_1 \gg \mathbb{C}_2 \gg \mathbb{C}_3$

Other researchers have considered alternatives to these freely floating constraints (Anttila 1997, Schütze 1997). Anttila (1997) suggests that we can cope with variability by making use of stratum-internal non-ranking. He divides constraints into sets "which are strictly ranked with respect to each other, but internally random except for universal rankings" (Anttila 1997: 15). He claims that "while mutually ranked, the sets are internally random reflecting the fact that the constraints are equally important [...]. This forms the probabilistic component of the grammar" (Anttila 1997: 21). Consequently, if we wanted to add a non-ranked constraint $\mathbb{C}_4$ to the strictly ranked $\mathbb{C}_1 \gg \mathbb{C}_2 \gg \mathbb{C}_3$, it would have to be included forming a set with *any* one of the ranked constraints —with $\mathbb{C}_3$ in the example presented in (4).

(4) $\mathbb{C}_1 \gg \mathbb{C}_2 \gg \{\mathbb{C}_3, \mathbb{C}_4\}$: $\qquad$ $\mathbb{C}_1 \gg \mathbb{C}_2 \gg \mathbb{C}_4 \gg \mathbb{C}_3$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\mathbb{C}_1 \gg \mathbb{C}_2 \gg \mathbb{C}_3 \gg \mathbb{C}_4$

This does not mean that Anttila's (1997) approach cannot produce the same variation patterns as the floating constraint model. It just implies a different theoretical approach that does not allow a constraint to be freely-ranked with respect to a set of ranked constraints. Anttila has subsequently moved on to propose a broader system of *partially-ordered* grammars, where the approach that we have just presented is only a special case (Anttila & Cho 1998; Anttila 2002 a, b). He claims that "it is not clear which of these approaches is correct, if any. It may turn out that some restricted version of the multiple grammars model [...] is sufficient. It may also be that

the full power of multiple grammars is necessary. Finally, it is possible that something quite different is needed" (Anttila 2002c: 231).

That 'something quite different' that Anttila refers to is a more drastic detachment from standard OT, represented by the continuous ranking scale (Boersma 1997, 2000; Hayes 2000; Boersma & Hayes 2001). The following quotation summarises the starting point of this approach:

> It would be useful in what follows to consider rankings not as simple arrangements of constraint pairs, but rather as the result of the constraints' each taking on a range of values on an abstract continuum [...[Within each band, [there is a] selection point, which is defined as the particular value of strictness taken on by a constraint on a given speaking occasion.
>
> *Hayes (2000: 89-90)*

Let us imagine a language where, as the result of an uncompleted sound change, all syllabic codas *tend* to be suppressed, except in a minority of cases where they are still retained. In most cases, the markedness constraint NOCODA ('avoid codas') would dominate the faithfulness constraint MAX ('elements in the input must have a correspondent in the output'). If we look at constraints as ranges of values which can vary minimally in each evaluation, we could explain the variable presence or absence of coda consonants as the result of NOCODA having a *ranking value* which is higher than MAX, but so close to it that overlapping may occur (see figure 1). All the details concerning how to calculate and use strictness values can be found in Boersma & Hayes (2001: 47-50) and will be further discussed in section IV of this paper.
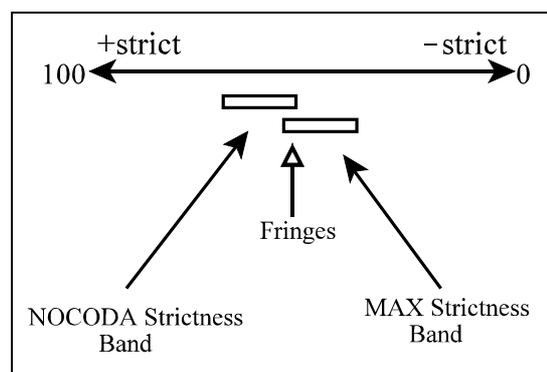
Figure 1. Strictness bands of NOCODA and MAX following Hayes' (2000) model, adapted from Cutillas (2003)

III. THE APPLICATION TO SOCIOLINGUISTIC STUDIES

III.1. What has been suggested so far: internal causes of variability

The theoretical developments proposed by Reynolds (1994), Nagy & Reynolds (1996, 1997) and Anttila (1997) have been used in sociolinguistic studies by Cardoso (2001) and Auger (2001). The application of these theories to sociolinguistic data has shown a potential to describe the linguistic behaviour of both the speaking community and the individual quite accurately, thus accounting for variable patterns.

In spite of this apparent success, there are some problems in these analyses. First of all, the data under study is quite unusual in quantitative sociolinguistics. The authors focus on a practically extinct language —Picard—, analysing written materials (Auger) or a mixture of written and spoken extracts (Cardoso). Given that the (relatively few) speakers of the language are older men living in rural, isolated areas, there is no chance for class or gender differentiation. Practically the same applies to stylistic variation. As Cardoso points out "Picard, as a dying (or recessive) language, is characterized by monostylism. [...] Languages in the process of decay are mostly used in a single formality style" (Cardoso 2001: 311). Stylistic effects are created by distinguishing three levels of formality (conversation, translation and written texts). Therefore, most of the discussion is limited to purely linguistic variability based on internal factors, with no reference to external causes of variation.

The reason why external factors are avoided may be found in OT literature. Anttila remarks that "it is not the business of grammatical theory to explain the effects of sex, age, style, register and social class" (Anttila 2002c: 212) before sketching two possible scenarios. Variability based on sex or style can be 'modular' and have no interaction with language-internal variation. Consequently, socially meaningful variability is just the result of a selection from a set of available grammars. The 'antimodular' approach suggests that internal and external constraints operate at the same level and can, therefore, be mixed.

III.2. What is left to say: external causes of variability

When we talk about variability, we cannot forget that variation related to linguistic factors is just one of the types that we may find. In fact, sociolinguists have been interested in many different kinds of extra-linguistic factors affecting linguistic performance. Considering language-external factors in an OT approach to variability is certainly challenging. Firstly, because the definition of these variables is sometimes difficult and controversial; secondly, because it introduces concepts such as language loyalty, prestige and identity building, which are quite outside the interests of most theoretical linguists.

Sociolinguistics has substantially advanced since its origins in the mid 1960s. More specifically, the way we look at independent variables has been modified quite significantly. As a matter of fact, these advances pose new challenges and problems to OT analyses of variation. Let us take the example of stylistic variation. In the very first sociolinguistic studies, it was seen

as the result of different levels of attention paid to one's speech (Labov 1966a, 1972; Trudgill 1974). Subsequent research has shown that informal styles are not necessarily the result of carelessness, but rather a sign of adequacy to context. In addition, 'formal' choices are not necessarily more faithful than informal ones. Most researchers (inside and outside OT) assume that different styles are the result of different grammars. In a way, they are reflecting a view of style that is very convenient for the purpose of linguistic theory, but quite contrary to developments in stylistic theory over the last twenty years.

Whether we consider style as a reflection of social variation (Bell 1984) or social variation as a reflection of style (Finegan & Biber 1991), we notice that stylistic variation is essentially gradual. This was first suggested by Boersma & Hayes (2001: 82-83). Let us consider the example of Couplands' study of the speech of a sales assistant in a travel agency (1980, 1984). Her use of non-standard forms depended quite heavily on the occupational status of her clients. With very few exceptions, the assistant's speech to members of each social class went parallel to the class members' own speech. The shift in her speech reached 55% in the direction of that of her clients, but that depended on the status difference between the client and the assistant. What we are describing is an essentially gradual phenomenon. We could, of course, suggest that the assistant has got six or seven formality grammars and chooses one of them depending on the perceived social and personal features of the interlocutor. But in that case, we would obtain a pattern of variation that is too *neat*. Once one grammar is assigned to one speaker, we should expect exactly the same percentage of variation along the whole conversational exchange. From research, we know that speakers can adapt and modulate their non-standardness depending on all sorts of external factors. Thus, we can start a conversation with an 80% of use of standard forms and move towards non-standardness gradually, making minimal changes that indicate continuity rather than an abrupt shift from one grammar to another[1]. In more recent approaches to stylistic variation, speakers are assumed to design their speech as a projection of a self-image (see Coupland 1985, 2001; Schilling-Estes 1999, 2002; or Traugott & Romaine 1985, for example). From these new perspectives, we are no longer talking about grammar selection based on the interlocutor's identity or a formal or casual context, but rather about a conscious process of identity building which is variable and orderly irregular.

Let us now consider the case of gender-based variablity. At the very beginning of sociolinguistic studies, gender was seen as an objective quality of the person. Individuals were socialised as either male or female, and this was in turn reflected by their linguistic behaviour. Thus, there would be 'male' constraint rankings and 'female' constraint rankings. Again, these two would be modulated depending on different external factors, such as context, addressee, age or social class. However, a drastically different viewpoint is suggested by Cameron (1998), based on the work of the philosopher Judith Butler (1990) and her concept of *performativity*. In agreement with what we commented on style, the linguistic behaviour of male or female speakers is seen as a projection, rather than as a reflection. Men or women do not speak the way they do

because of who they are; rather, they build their identity on the basis of their linguistic use. Consequently, "it is unhelpful for linguists to continue to use models of gendered speech which imply that masculinity and femininity are monolithic constructs, automatically giving rise to predictable (and utterly different) patterns of verbal interaction" (Cameron 1998: 282). Adapting Cameron's words to an OT perspective, it is unhelpful to attempt to explain gender-based variation using monolithic constraint rankings, automatically giving rise to predictable variable distributions. According to this new view of gender, speakers would be consciously altering their constraint rankings and linguistic outputs. It is not as simple as choosing from one of the available multiple grammars; what we see is a complex process of speech construction.

From facts about style and gender, we can infer that (*i*) variation is essentially *gradual* —in line with Boersma & Hayes— and (*ii*) it can perfectly be *conscious*. This is why we propose that we should approach variability from a gradual perspective, limiting the complexity of the multiple grammar model and leaving enough theoretical space for conscious choice to play a central part in the explanation of linguistic variability. We cannot agree with Anttila's (2002c) in considering external factors as something hardly relevant for an OT theory of variation. On the contrary, these play an essential role, providing us with arguments to select one variable grammar model over the others.

IV. CONTINUOUS RANKING AND THE THREE-GRAMMAR MODEL

We shall propose a model of analysis of external variation which is based on two main principles: *(i)* a continuous ranking approach to contraints (Boersma 1997, 2000; Hayes 2000; Boersma & Hayes 2001) and *(ii)* a limited number of available grammars.

According to this view, constraints are not monolithic entities that always select one single optimal candidate, or more than one if the ranking is arranged following one of the approaches discussed in section II. Rather, constraints are seen as *strictness bands*, with their relative importance being assessed and checked against the rest in every evaluation and thus, subject to (minimal, although sometimes crucial) variation.

Every time a candidate is evaluated, the grammar selects a particular point within the strictness bands of the different constraints. Under *stochastic evaluation*, this point varies whenever we choose a candidate. The extent to which this variation can be allowed in every evaluation is explicitly stated in the grammatical system, in the form of an *evaluation noise*. The mathematical implementation of this system is fairly simple. Strictness bands are seen as a range of values forming a normal distribution. Given a mean (the central *ranking value*) and a standard deviation (the *evaluation noise*), we can obtain all sorts of patterns of grammatical variation (for more details, see Boersma & Hayes 2001).
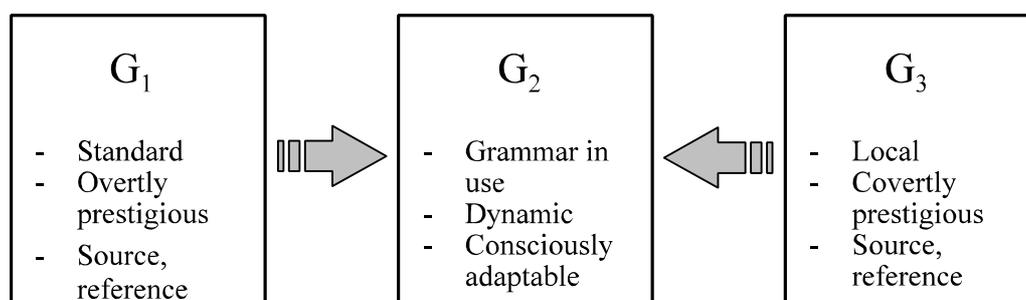
Let us consider the possibility that, apart from the evaluation noise —which generates variable selection points—, the system is also subject to meaningful variations of ranking values

consciously caused by the individual. The action of the speaker favouring the change in ranking values is the consequence of the availability of two different linguistic systems, which are seen as points of reference. The pattern is fairly straightforward. In many languages, we find that speakers use variable percentages of standard forms and non-standard, usually local forms. The exact percentage of use of each form depends on independent variables such as style, gender or social class. Let us assume that the speaker has three grammars available to him/her, which we shall call G_1, G_2 and G_3.

G_1 is the standard, prestigious variety associated with education and propriety. It provides speakers with an alternative to their 'native' G_3 forms, which can be chosen for a variety of purposes, either permanently or just in specific contexts. G_3 is the local grammar available to the community. The values expressed through this grammar are those of language loyalty and covert prestige, among others (see, for instance, Trudgill 1972). The system is acknowledged as part of the identity of the speech community, but it usually has connotations of non-standardness. Finally, we have G_2, which is an intermediate, personal grammatical system. In much the same way that the learner of a foreign language builds an intermediate system between her mother tongue and the target language (*interlanguage*, see Selinker 1972) or between two different dialects (*interdialect*, see Trudgill 1986: 63), it is reasonable to assume that the speaker in contact with two different grammars builds his/her own constraint ranking.

G_2 is the default grammar used by the speaker, both in everyday, informal contexts and in other more formal situations. In informal contexts, G_2 may or may not be identical to the local grammar (G_3), depending on a complex series of social variables.[2] In formal contexts, G_2 will be more or less similar to the standard grammar (G_1) depending on the availability of the standard forms and the conscious choice of the speaker.[3] The most defining feature of this grammar is that it is extremely dynamic, depending on the needs of the individual speaker. Taking G_1 and G_3 as references values, the individual *designs* his/her speech conveying different sorts of social and personal meaning, thus projecting a self-image and building a linguistic identity. We have attempted to summarise the links between these grammars in (5).

(5)

G_1		G_2		G_3
- Standard - Overtly prestigious - Source, reference	⟹	- Grammar in use - Dynamic - Consciously adaptable	⟸	- Local - Covertly prestigious - Source, reference

The system that we have just described is certainly identical in results to the 'style sensitivity' approach of Boersma & Hayes (2001: 82-83), as far as the behaviour of the speaker's current grammar is concerned (G_2). The difference is essentially theoretical. Instead of establishing two abstract points in a stylistic continuum, we suggest that we should assume two linguistic systems on their own right. Among other reasons, because these systems can be claimed to exist and are observable. By approaching one or the other, the speaker conveys the meanings that are associated with each one of the reference linguistic varieties. If the speaker approaches the standard —an accent as acceptable as any other—, he is probably attempting to convey the positive value judgments associated with the standard (education, propriety, formality, etc.). However, if the speaker approaches the vernacular —again, an accent like any other, not a 'casual' or 'careless' version of the standard—, he/she is probably trying to transmit the positive values associated with the local forms (language loyalty and group membership, among others). The obvious advantage is that this analysis can then be easily extended to other external variables.

IV.1. The three grammar system in English

In his New York City study, Labov (1966a) reports the case of Susan Salto, who showed striking stylistic shifts for variable (r). Simplifying the facts, in that variety of English (r) can be realised in two different ways: either it is pronounced as [ɹ], or it is suppressed. Salto used the standard variant [ɹ] 100% of the times in her most careful style and suppressed it in 98% of the cases in casual style (6).

(6)   Percentage of the [ɹ] variant in the speech of Susan Salto

MP	100%
WL	61%
RP	58%
IS	26%
CS[4]	2%

Let us now look at this data from the three-grammar perspective. The standard grammar (G_1) establishes that coda /r/ cannot be suppressed. Consequently, MAX must be high-ranked, making sure that candidates which are not faithful to the input are rejected. We also need a markedness constraint, militating against /r/ in coda position. We shall assume that this constraint is *CODA/r.[5] Thus, the constraint ranking in G_1 will be MAX » *CODA/r as shown in (7).

On the other hand, the local grammar (G_3) demands that /r/ in coda position has to be deleted. The markedness constraint *CODA/r is ranked over MAX, thus selecting the unfaithful candidate (8).

(7) G₁ grammar: [ɹ] kept in coda position in the word *guard*.

/gɔɹd/	MAX	*CODA/r
☞ a. [gɔɹd]		*
b. [gɔːd]	*!	

(8) G₃ grammar: [ɹ] deleted in coda position in the word *guard*.

/gɔɹd/	*CODA/r	MAX
a. [gɔɹd]	*!	
☞ b. [gɔːd]		*

     Susan's linguistic behaviour implies a total ranking reversal from minimal pair reading (MP) to casual style (CS) and a transition area which goes gradually from G₁ to G₃. One could argue that this can be explained by posing three more grammars (Word List grammar, Reading Passage grammar and Interview Style grammar). But there are good reasons against this logic. Firstly, these different styles are arbitrarily established. Why could not we have a distinction between an informal interview style and a job interview style? Or a casual style talking to friends as opposed to a casual style talking to colleagues? These distinctions are perfectly reasonable and we would be forced to propose some more grammars to explain these new chunks of the stylistic continuum. Secondly, Labov's styles are essentially artificial. We do not read minimal pair lists. We rarely, if ever, read aloud. However, the methodology followed by Labov is useful insofar as it help us to visualize the continuum of style shift. It is not reasonable to assume that we have dozens of grammars to explain a change, step by step, from MAX » *CODA/r (7) to *CODA/r » MAX (8). Rather, we can sketch a simple OT account of sociolinguistic facts by accepting continuous ranking scales and the three-grammar system that we outlined in the previous section.

     We base our calculations on the model presented in Boersma & Hayes (2001). We have to go from the percentages of variation that we have observed to the relevant ranking values, assuming a standard deviation of 2.[6] The resulting ranking values are presented in (9). Note that we have decided that the markedness constraint has a fixed ranking value of 80. Variation derives from a change in the ranking value of the faithfulness constraint MAX. Although, in principle, there are different possibilities to obtain the same result, we suggest the hypothesis that markedness constraints have a fixed place in the ranking —as a result of their 'physical grounding'— and faithfulness constraints are allowed to move.

IJES, vol. 4 (2), 2004, pp. 165-184

(9) Ranking values of MAX and *CODA/r and derived percentages of variation

Styles	MAX		*CODA/r	
	Ranking value	*Percentage of [ɹ]*	*Ranking value*	*Percentage of Ø*
MP	91'31	100%[7]	80	0%
WL	80'79	61%	80	39%
RP	80'57	58%	80	42%
IS	78'18	26%	80	74%
CS	68'69	0%	80	100%

From these results, we can establish the ranking values of MAX and *CODA/r in the two reference grammars, G_1 and G_3. In the standard grammar —G_1—, MAX would have a ranking value of 91'31, whereas *CODA/r would be assigned a ranking value of 80. Thus, coda /r/ would *always* be pronounced. In the local grammar —G_3—, MAX would have a ranking value of 68'69 and *CODA/r would always dominate MAX, so that coda /r/ would *never* be pronounced. In Susan Salto's personal and variable grammar (G_2), which takes G_1 and G_3 as the points of reference for the construction of social meaning, the values of MAX range from 91'31 to 68'69, being consciously modulated by the speaker.

The advantage of this model is that it does not assume the existence of multiple grammars. Rather, it suggests that there are two reference models and one personal grammar based on the principle of continuous ranking. Furthermore, it acknowledges the possibility that the speaker modulates his/her own constraint ranking to accommodate the extralinguistic context, to project a desired self-image or to build an identity. Grammar is no longer seen as a fully automated mechanism; personal and *meaningful* decisions are granted a place.

IV.2. The three grammar system in Spanish

Cutillas-Espinosa & Hernández-Campoy (2004) study the linguistic behaviour of a radio presenter at a local station in Murcia (Spain). More specifically, they compare his use of the standard and non-standard variants of variable (s), amongst some other prominent variables in the local dialect. In Murcia, (s) can be realised as either [s] —the standard, Castilian Spanish form— or as Ø, causing changes in the preceding vowels (for details, see Hernández-Campoy & Trudgill 2002).

Let us first define the two grammatical systems used as a *source*. The local variety of Murcian Spanish —G_3— is characterised by the practically total absence of consonant codas, except for the nasal /n/. The standard variety —Castilian Spanish, G_1— allows different types of one-member codas and a few complex codas of the type –consonant+/s/– as in *ins.truc.ción*

'instruction' or *bi.ceps*.[8] One of the most common codas in Castilian Spanish is /s/, which is systematically suppressed in Murcian Spanish. We shall propose two constraints, which are very similar to the ones in the preceding analysis: MAX (elements in the input must have a correspondent in the output) and NOCODA (codas are not allowed). The ranking of these two constraints in G_1 and G_3 are the following:

(10)   Rankings of MAX and NOCODA in G_1 and G_3

      G_1 (Standard Castilian Spanish): MAX » NOCODA
      G_3 (Local Murcian Variety): NOCODA » MAX

Given an input /pésas/ ('weights'), G_1 and G_3 would select the following optimal candidates (11 and 12).

(11) G_1 grammar: Coda consonant preserved.

/pésas/	MAX	NOCODA
☞ a. [pésas]		*
b. [pésæː]	*!	

(12) G_3 grammar: Coda consonant suppressed, causing vowel changes.

/pésas/	NOCODA	MAX
a. [pésas]	*!	
☞ b. [pésæː]		*

Cutillas-Espinosa & Hernández-Campoy (2004) study the differences in (s) deletion in the speech of the presenter in two different *performances*: (*i*) in broadcasting, when talking to a predominantly non-standard speaking audience on the phone; (*ii*) in a formal interview with the researchers. The results obtained are shown in (13).

(13)   Variant 1 (Standard): [s]   Variant 2 (Local): Ø

Broadcasting: 89%   11%
Interview: 1%   99%

IJES, vol. 4 (2), 2004, pp. 165-184

The presenter of the radio programme used an overwhelming majority of standard [s] forms when talking to audience members on the phone during broadcasting. In spite of the fact that audience members were non-standard speakers, he kept a high degree of standardisation. This was due, according to his own words, to an attempt to express 'respect' and awareness of being heard by an audience (G_1 being identified with 'respectability' and 'correctness'). The interview with the two researchers was *performed* by the presenter in a different way, in spite of the fact that an interview with a linguist is not precisely a 'casual' context. Two factors have to be considered. Firstly, the researchers used a local accent and the presenter shifted towards a practically complete absence of coda /s/ accordingly. Secondly, the presenter was aware of the linguists' curiosity for his high standardisation in broadcasting and, during the conversation, he insisted on a positive —although contradictory— view of the vernacular (G_3). It could be argued that he showed that he had no prejudice against Murcian Spanish by *using* G_3 forms extensively. The picture, therefore, is more complex than a purely formal *vs* casual contrast in one single grammar as suggested in the appendix to Boersma & Hayes (2001: 82-83).

Again, we shall interpret the data in terms of the three-grammar system and a continuous ranking. The value rankings of MAX and NOCODA in G_1 —Castilian Spanish grammar— are 91'31 and 80, respectively, making sure that /s/ in coda position is *always* pronounced. On the other hand, the value rankings of MAX and NOCODA in G_3 —the local grammar— are 68'69 and 80, respectively, making sure that /s/ in coda position is *never* pronounced. Taking those values as a reference, the radio presenter builds his own grammar —G_2—. Its observed values are calculated following the procedure described in the previous section. The results are presented in (14).

(14) Ranking values of MAX and NOCODA and derived percentages of variation.

Styles	MAX		NOCODA	
	Ranking value	*Percentage of [s]*	*Ranking value*	*Percentage of Ø*
Broadcasting	83'46	89%	80	11%
Interview	73'4	1%	80	99%

Thus, the observed ranking values of MAX in the presenter's grammar —G_2— range from 83'46 to 73'4. This variation in ranking values is meaningful and, again, it points towards a dynamic and continuous system delimited by the values of the two reference grammars, G_1 and G_3. For example, it is reasonable to assume that, when the presenter addresses the station director, the ranking value of MAX may be significantly lowered below the 83'46 ranking value of broadcasting, or raised above the 73'4 value selected for the interview. Thus, we admit that

the presenter can play with ranking values, which is an extremely effective way to describe the stylistic continuum.

IV.3. Three or more?

So far we have looked at an ideal —although commonly found— linguistic situation where the speaker is only exposed to *two* linguistic models, the standard and the local. However, we have to admit the possibility that a speaker may be exposed to more than one local grammar or, more rarely, more than one standard. It is also very likely that there is more than one standard or non-standard variant for each variable. This just adds some mathematical complexity to the system, but it remains fully operational. The fact that continuous ranking works with probability makes it the best approach to cope with the complexities of language variation and change.

IV.4. Going beyond style-shifting

At this point, it would be reasonable to wonder whether the three-grammar system is just applicable to style-shifting. We argue that it can be generalised to other sociolinguistic variables, such as gender (see discussion in section III.2) or social class. Such generalization is possible if we assume a *constructivist* view of language variability, whereby variation is not a reflection of personal characteristics, but rather a projection and construction of one's self-image. In OT terms, this view implies that the speaker is able to control ranking values *consciously*.

Let us take the case of social class. Labov (1966b) re-analysed his New York City study data taking into account how many of his informants could be regarded as socially 'upwardly mobile'.[9] In Labov's study, variable (dh) —like in *this, then, there*— has two variants: the standard one [ð] and the non-standard [d]. In (15) we show how patterns vary depending on whether working class and lower middle class members are upwardly mobile.

(15) Use of the non-standard [d] variant (*this* → [dɪs]) depending on class and social mobility.

	Working Class	Lower Middle Class
Upwardly Mobile:	27%	17%
Stable:	80%	50%

As we can see, people who move —or attempt to move— from one social class to another change their linguistic behaviour quite substantially, so that instead of preserving their grammar in a different situation, "upwardly mobile individuals *adjust* (emphasis mine) the frequency of certain linguistic variables in order to sound more like the class they are joining and less like the one they are leaving" (Chambers 1995: 55). This suggests that it is not enough to claim that there is a 'working class grammar' or a 'lower middle class' grammar, which are fixed and acquired

by the individual. Rather, there is a conscious process of identity building, via linguistic means. In other words, given two pieces of grammar extracted from two different grammatical systems (G_1, *this* → [ðɪs]; G_3, *this* → [dɪs]), the speaker is able to *adjust* his own grammar (G_2) in the direction of the standard (G_1) and convey a meaningful declaration of a particular social group membership.

Again, the fact that personal grammars are not monolithically preserved throughout a person's life favours the idea of continuous rankings and the three-grammar system. Under this approach, variation is essentially continuous, meaningful and conscious.

V. CONCLUSION

In this paper we have attempted to show how sociolinguistic theory can contribute to propose an OT model of variability that is *grounded* on more than four decades of extensive research. We have started by reviewing different approaches to variability in OT literature and their applications to sociolinguistic studies. The phonologists' concern has been to find a way to describe gradient and variable facts in an architecture which was designed to select just *one* optimal candidate, discarding the rest. Different solutions have been suggested. The sociolinguists' concern has been to apply these models of variability to *real* data, but there has been a tendency to limit this endeavour to the explanation of language-internal variability.

We have commented on the complexity of external factors. We disagree with Anttila's remark (2002c) that the explanation of these facts is not the business of grammatical descriptions. It is true that they add a lot of complexity to the picture, but we agree with Hayes' idea that "at present linguistics is *not difficult enough*" (Hayes 2000: 118). However, it is reasonable to assume —and this is probably close to what Anttila meant— that the study of these external facts is not *central* to phonological discussion.

The question is not only that external facts need to be explained, but also that they can provide us with valuable information for the choice of an OT model of variability. We cannot just claim that anything that produces variation is an acceptable solution. We also have to pay attention to the patterns of variability that are found in the real world, because these can be used as the criteria for the choice among the different OT models. We have concluded that the continuous ranking approach to variation (Boersma & Hayes 2001) seems to be the one that best fits sociolinguistic knowledge. It is capable of describing and producing both drastic and minimal adjustments to linguistic outputs.

In addition to continuous ranking, we propose a basic, three-grammar model of variability. There would be two reference grammars, G_1 —the standard— and G_3 —the local— and one intermediate, personal grammar —G_2— which may be closer to G_1 or G_3 depending on contextual, socio-demographic or just personal factors.

In spite of its apparent difficulty, this approach could prove useful. It may help to overcome the traditional gap between theoretical phonology —for sociolinguists, *armchair linguistics*— and real-world patterns, which have to be the centre of our business.

NOTES:

1. A similar idea is presented in Boersma & Hayes (2001: 82-83) where they assume a casual to formal continuum, with gradual stylistic variations based on an equation. The difference is that we do not refer to 'formal' and 'casual' within the same grammar; rather, we look at stylistic variation as the result of taking two different grammars on their own right as reference points. This analysis can then be extended to the analysis of other independent variables, or to the analysis of stylistic variation under audience or speaker design-based approaches.

2. The relative similarity of G_2 to one of the grammatical models —either the standard or the local— will depend on factors such as gender, social class, age and social network. The discussion of the effects of each specific factor is outside the scope of this study.

3. We are assuming that speakers have access to the standard grammar (G_1) and, of course, this is not necessarily the case. They may have a limited access to it, or they may have constructed a personal version of what they see as the standard grammar. In much the same way, speakers may have a limited access to the grammar of the local community (G_3), if we assume that their link to the community's social network and vernacular speech is very weak. In order to illustrate the theory, we shall assume that the speaker has full access to G_1 and G_3, although the possibility of having an individual version of these grammars cannot be discarded.

4. MP: Minimal Pairs; WL: Word List; RP: Reading Passage; IS: Interview Style; CS: Casual Style. The data have been extracted from Chambers (1995: 20).

5. This constraint is similar to the ones proposed by Hammond (1999) in his analysis of English phonotactics.

6. We carried out the following calculation to go from percentages to ranking values. Given that ranking values are based on an arbitrary scale, the essential point is to calculate the difference between MAX and *CODA/r's means (X), which justifies each specific percentage of variation. Let us see, for instance, what difference in ranking value has to separate MAX and *CODA/r to explain that Susan Salto uses variable [ɹ] 61% of the times when reading a word list:

If *CODA/r ~ N (μ_2, σ_2^2) and MAX ~ N (μ_1, σ_1^2) and both are independent, then

X = *CODA/r − MAX ~ N $(\mu_2 − \mu_1, \sigma_2^2 + \sigma_1^2)$

If X ≤ 0, then MAX » *CODA/r

$$P(X \leq 0) = P\left(\frac{X - (\mu_2 - \mu_1)}{\sqrt{8}} \leq \frac{0 - (\mu_2 - \mu_1)}{\sqrt{8}} \right) =$$

$$P\left(Z \leq \frac{-(\mu_2 - \mu_1)}{\sqrt{8}} \right) = 0'61 \text{ (which corresponds to an f(x) value of } 0'28)$$

$$\frac{-(\mu_2 - \mu_1)}{\sqrt{8}} = 0'28 \rightarrow -(\mu_2 - \mu_1) = 0'28 \times 2'83 \rightarrow -(\mu_2 - \mu_1) = 0'79 \rightarrow \mu_1 - \mu_2 = \underline{\mathbf{0'79}}$$

Thus, the ranking value of MAX will be the same as that of *CODA/r, plus 0'79, to generate 61% of the standard variant [ɹ]. We shall assume that *CODA/r has a ranking value of 80. Consequently, MAX must have a ranking value of 80'79.

7. In fact, a 100% probability cannot be obtained, as probability distributions tend to ∞. If MAX has a ranking value of 91'31, there is a 99'997% probability that it will dominate *CODA/r. The same applies to 0%.

8. There are few words in Castilian Spanish that have word-final complex codas, such as *biceps* or *fórceps*. They are usually loans that can be regarded as peripheral to the system. However, word-internal combinations of −consonant + s− are more common: *abs.tru.so* 'abstruse', *cons.tre.ñir* 'constraint', *ads.crip.ción* 'assignment'. In these cases, it is quite usual to drop the first consonant and simplify the coda to a single /s/ in pronunciation: *as.tru.so, cos.tre.ñir, as.crip.ción* (Alarcos 1994: 40).

9. For details, see Labov (1966b). A good summary of the facts can be found in Chambers (1995: 55-57).

REFERENCES

Alarcos-Llorach, Emilio (1994). Gramática de la lengua española. Real Academica Española, Madrid: Espasa-Calpe.

Anttila, Arto (1997). Deriving variation from grammar. In Frans Hinskens, Roeland van Hout & Leo Wetzels (Eds.), *Variation, Change and Phonological Theory*. Amsterdam: John Benjamins, 35-68. Also ROA-63.

Anttila, Arto & Cho, Young-mee (1998). Variation and change in Optimality Theory. *Lingua*, 104, 34-56.

Anttila, Arto (2002a). Derived environment effects in colloquial Helsinki Finnish. ROA-406.

Anttila, Arto (2002b). Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory,* 20:1, 1-42. Also ROA-425.

Anttila, Arto (2002c). Variation and phonological theory. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (Eds.), *The Handbook of Language Variation and Change*. Oxford: Blackwell, 206-243.

Auger, Julie (2001). Phonological variation and Optimality Theory: Evidence from word-initial vowel epenthesis in Vimeu Picard. *Language Variation and Change*, 13, 253-303.

Bell, Allan (1984). Language style as audience design. *Language in Society*, 13, 145-204.

Boersma, Paul (1997). How we learn variation, optionality and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, 21, 43-58. Also ROA-221.

Boersma, Paul (2000). Learning a grammar in Functional Phonology. In J. Dekkers, F. van der Leeuw & J. van de Weijer (Eds.), *Optimality Theory: Phonology, Syntax and Acquisition*. Oxford: Oxford University Press, 465-523.

Boersma, Paul & Hayes, Bruce (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, 32:1, 45-86.

Butler, Judith (1990). *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge.

Cameron, Deborah (1998). Performing gender identity: Young men's talk and the construction of heterosexual masculinity. In J. Coates (Ed.), *Language and Gender. A Reader*. Oxford: Blackwell, 270-284.

Cardoso, Walcir (2001). Variation patterns in across-word regressive assimilation in Picard: An Optimality Theoretic account. *Language Variation and Change*, 13, 305-341.

Chambers, Jack (1995). *Sociolinguistic Theory*. Oxford: Blackwell.

Coupland, Nikolas (1980). Style-shifting in a Cardiff work-setting. *Language in Society*, 9, 1-12.

Coupland, Nikolas (1984). Accommodation at work. Some phonological data and their implications. *International Journal of the Sociology of Language*, 46: 49-70.

Coupland, Nikolas. (1985). Hark, hark the lark: Social motivations for phonological style-shifting. *Language and Communication*, 5:3, 153-72.

Coupland, Nikolas. (2001). Dialect stylization in radio talk. *Language in Society*, 30: 3, 345-75.

Cutillas-Espinosa, Juan Antonio (2003). *Teoría Lingüística de la Optimidad. Fonología, Morfología y Aprendizaje*. Murcia: Servicio de Publicaciones de la Universidad de Murcia.

Cutillas-Espinosa, Juan Antonio & Hernández-Campoy, Juan Manuel (2004). Audience and speaker design in radio broadcasting. Attitudes to language as a starting script. Ms., Universidad de Murcia.

Finegan, Edward & Biber, Douglas (1994). Register and social dialect variation: An integrated approach. In Edward Finegan & Douglas Biber (Eds.), *Sociolinguistic Perspectives on Register*. Oxford: Oxford University Press.

Hammond, Michael (1999). *The Phonology of English. A Prosodic Optimality-Theoretic Approach*. Oxford: Oxford University Press.

Hayes, Bruce (2000). Gradient well-formedness in Optimality Theory. In J. Dekkers, F. van der Leeuw & J. van de Weijer (Eds.), *Optimality Theory: Phonology, Syntax and Acquisition*. Oxford: Oxford University Press, 88-120.

Hernández-Campoy, Juan Manuel & Trudgill, Peter (2002). Functional compensation and southern peninsular Spanish /s/ loss. *Folia Linguistica Historica*, XXIII, 141-167.

Labov, William (1966a). *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.

Labov, William (1966b). The effect of social mobility on linguistic behavior. *Sociological Inquiry*, 36, 186-203.

Labov, William (1972). *Sociolinguistic Patterns*. Oxford: Blackwell.

Nagy, Naomi & Reynolds, William T. (1996). Accounting for variable word-final deletion within Optimality Theory. In Jeniffer Arnold, Renée Blake, Brad Davidson, Scott Schwenter & Julie Solomon (Eds.), *Sociolinguistic variation. Data, theory, and analysis. Selected papers from NWAV23 at Stanford*. Stanford, California: CSLI Publications, 151-160.

Nagy, Naomi & Reynolds, William T. (1997). Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change*, 9:1, 37-55.

Reynolds, Bill (1994). *Variation and Phonological Theory*. Unpublished doctoral dissertation, University of Pennsylvania.

Schilling-Estes, Natalie (1999). Situated ethnicities: Constructing and reconstructing identity in the sociolinguistic interview. *University of Pennsylvania Working Papers in Linguistics*, 6:2 (Proceedings from NWAVE 27), 137-51.

Schilling-Estes, Natalie (2002). Investigating stylistic variation. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (Eds.), *The Handbook of Language Variation and Change*. Oxford: Blackwell, 375-401.

Schütze, Carson T. (1997). The prosodic structure of Serbo-Croatian function words: An argument for tied constraints. *MIT Working Papers in Linguistics*, 30, 355-367.

Selinker, Larry (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209-230.

Trudgill, Peter J. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, 1, 179-195.

Trudgill, Peter J. (1974). *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, Peter J. (1986). *Dialects in contact*. Oxford: Blackwell.