

Optimality Theory and language death¹

Darcie Blainey

Tulane University
dwillia7@tulane.edu

1 Introduction

The literature that uses an optimality-theoretic framework to examine phonological variation in endangered language varieties is limited. Besides Auger and Steele (1999) and Auger (2001), little work has been done mapping raw frequencies from dying languages into optimality-theoretic constraint hierarchies, even though the inter- and intra-speaker variation and multiple acceptable surface forms characteristic of dying languages can provide crucial testing ground for extensions of Optimality Theory (Prince and Smolensky, 1993; McCarthy and Prince, 1994). This article therefore endeavours to present the case of Louisiana Regional French as an example of a dying language variety that can offer intriguing questions to phonological frameworks such as Optimality Theory.

The paper begins with a brief overview of Optimality Theory. Section 3 explores different proposals that have been made, such as crucially unranked constraints (Anttila and Cho, 1998; Anttila, 2002), floating constraints (Reynolds, 1994; Nagy and Reynolds, 1997), Linear Optimality Theory (Keller, 2000, 2006; Sorace and Keller, 2005), and Stochastic Optimality Theory (Boersma, 1998; Boersma and Hayes, 2001), in order to account for types of linguistic variation while maintaining the general tenets of Optimality Theory. The linguistic characteristics of moribund languages are considered in section 4, followed by an overview in section 5 of dying language research that has employed Optimality Theory to explain phonological variation and change. Section 6 narrows in on Louisiana Regional French, an enclave dialect that is the focus of study in section 7, and section 8 provides several tentative conclusions and hypotheses.

2 Optimality Theory

Optimality Theory (henceforth OT) moves away from rule-based explanations of a speaker's internal grammar, turning instead to the idea of a hierarchical ranking system of universal and violable faithfulness and markedness constraints (Prince and Smolensky, 1993; McCarthy and Prince, 1994). Functions involved in this conceptualisation of a speaker's internal grammar are CON, which ranks the constraints relative to one another; GEN, which generates a potentially infinite set of output candidates for a given underlying form (the input); and EVAL, which evaluates the candidates relative to the constraint hierarchy. The candidate judged most harmonic in terms of the constraints it violates is deemed the winner, marked by the symbol φ . Therefore, in Table 1, Candidate₂ is the optimal output form chosen by the grammar.

/input/	Constraint ₁	Constraint ₂	Constraint ₃
Candidate ₁	*!		
φ Candidate ₂			*
Candidate ₃		*	*!

Table 1: Theoretical OT tableau

Constraint rankings are expressed in a left-to-right fashion, with high-ranking, often undominated constraints to the left. Solid lines in the tableau divide crucially ranked constraints (in Table 1, Constraint₁ dominates Constraint₂), while a dotted line indicates that two constraints are unranked with respect to one another (Constraint₂ and Constraint₃ have the same ranking). Each violation of a constraint is marked with an asterisk. A fatal violation excludes a candidate from consideration, shown by an exclamation point. Shaded cells indicate that the associated constraints do not play a role in that candidate's evaluation, due to crucial violation of a higher constraint. Thus, Candidate₁ fails by crucially violating high-ranking Constraint₁, and Candidate₃ fails because it violates both Constraint₂ and Constraint₃. Candidate₂ is left as the most harmonic (least bad) candidate, with only one violation of a lower-ranked constraint, and is selected as the surface form.

In OT, the phonological behaviour of a given language is determined by the relative rank it affords each constraint. As a result, this theory is useful for explaining inter-linguistic and inter-dialectal variation and change, which “have not been among the leading issues in generative linguistics” (Anttila and Cho, 1998: 31). However, classic OT has serious limitations in terms of mapping variation, since all constraints must be strictly ranked in a given grammar. This has encouraged scholars to posit different extensions of OT that can cope with variable outputs; these theories are explored in the following section.

3 Models of variation in OT

In order to account for surface variation, the multiple grammars theory for phonology proposes that constraint re-ranking allows a speaker to have many underlying grammars at his or her disposal for a given language (Kiparsky, 1993, 1994). This theory affords the advantage of easily explaining any number of instances of intra-speaker variation, but it also comes with two major problems. First, there is no limit on the number of grammars a speaker may possess for a single language, which hardly seems a reasonable model of human language processing and production (Anttila, 2002 : 220; Reynolds, 1994: 75). Second, if every potential case of surface variation warrants a new hierarchical ranking, it is not clear how the theory of multiple grammars can exclude unattested or unacceptable variants (Anttila, 2002 : 220). Anttila and Cho (1998 : 48) and Anttila (2002 : 221-224) show that partial ordering of OT constraints in a stratified hierarchy is a better alternative to the too-powerful theory of multiple grammars, because the former can limit how many grammars are predicted. Specifically, Anttila and Cho (1998 : 36) propose that language grammars have a “partial order” because they do not exhibit connectedness, one of the four properties of the relation between constraints originally proposed in OT. This stratified grammar “consists of internally unranked strata of constraints which are strictly ranked with respect to each other” (Anttila, 2002 : 230), and can be used to explain attested frequencies of different variants. A second modified approach to the idea of multiple grammars is that of floating constraints (Reynolds, 1994; Reynolds and Nagy, 1997). It is shown that while some constraints are strictly ranked, others are able to “float” within a defined domain; this domain can be as small as one other constraint or larger (Reynolds and Nagy, 1997 : 43).

Variation can also be considered from the point of view of perception, and what forms are considered as more acceptable or less acceptable. Keller (2000, 2006) proposes Linear Optimality Theory (LOT) as “a framework designed to account for gradient judgment data” (Keller, 2006 : 270). Keller (2006 : 271) explains that LOT “relies on the assumption that constraint ranks are represented as sets of numeric weights, instead of as partial orders [...and] the grammaticality of a given structure is proportional to the sum of the weights of the constraints it violates”. LOT is attractive because constraint weights can be calculated as results of linear equations using Gaussian Elimination or Least Square Estimation (LSE) (Keller, 2006 : 278-9). LOT works well in accounting for gradience in syntax, since it posits that there are “hard constraints (that trigger categorical linguistic judgments) and soft constraints (that trigger gradient judgments)” (Sorace and Keller, 2005 : 1502). Another interesting feature advanced in LOT is cumulative violations, so that lower-ranked constraints can “gang up against [a] higher ranked constraint” (Sorace and Keller, 2005:1507) to eliminate a candidate. This behaviour is “predicted to be impossible under the assumption of *strict domination* of constraints (as in Optimality Theory)” (Sorace and Keller, 2005 : 1508).

There are, however, limitations to LOT. First, the authors admit that “LOT is not intended as a model of human language acquisition [...rather, its purpose is] to determine an optimal set of constraint weights for a given data set” (Sorace and Keller, 2005 : 1519). As a result, it is unclear when and how LOT comes into play. Also, while Keller (2006 : 283-6) claims that LOT is better able to account for gradient data than models such as Stochastic OT (Boersma, 1998; Boersma and Hayes, 2001) examined below, further phonological studies are needed to confirm that this assertion holds, and to assess whether or not it applies to variation in general.

Stochastic OT, instead of employing a discrete ranking hierarchy, places constraints on an infinite linear scale, such that constraints farther to the left still dominate those farther to the right, but constraints placed close together have a “less fixed” relative ranking than those farther apart (Boersma and Hayes, 2001 : 47). Each constraint is actually a normal probability distribution of “selection points” to be chosen from at the moment of evaluation for the output; the mean of the distribution, a real number, is the “ranking value” of the constraint on the linear scale (Boersma and Hayes, 2001 : 47-8). Thus, if constraints are close together on the linear scale, their probability distributions may overlap, causing (in rare cases) reversal of the normal ranking (Boersma and Hayes, 2001:49). It is important to note that since all constraints have the same standard deviation, and all are fixed on the linear scale, constraints cannot move or float in this theory (Boersma and Hayes, 2001:50). The error-driven Gradual Learning Algorithm (GLA) positions constraints on the linear scale (Boersma and Hayes, 2001 : 51). The GLA is more conservative than Constraint Demotion (Tesar and Smolensky, 2000), since the latter “responds to input data with a radical change, namely, whole-sale reranking” (Boersma and Hayes, 2001:66), which means that the former is less likely to respond to speech errors, it can model multiple acceptable surface forms, and it can handle gradient judgment data (Boersma and Hayes, 2001 : 46). The fact that the GLA models language acquisition, while the mechanism for LOT does not, is a further advantage to adopting this theory. The algorithm is also accessible, since researchers can implement a GLA analysis of their data by using Praat software (Boersma and Weenick, 2008) as laid out in Boersma (1999). These factors seem to make Stochastic OT an ideal analytical tool for variable pronunciations in dying language data.

In the following sections, first, the structural characteristics of endangered languages are presented (section 4); secondly, existing research on OT, dying languages, and phonology are examined (section 5).

4 Linguistic attributes of dying languages

Bullock and Gerfen (2004 : 95) remark that “most work on the role of convergence in attrition has focused on aspects of the morpho-syntactic restructuring of the attriting language”, such as Austin (2006), King (2005), Morimoto (2002), and Rottet (1998) among others. Accordingly, Mougeon et al. (2005) lay out a four-step methodology for determining whether contact-induced linguistic changes are due to inter-systemic transfer, whether this transfer be overt (syntactic) or covert (morphological or semantic), or intra-systemic transfer (Mougeon et al., 2005 : 102-104). Examples of morphological and syntactic effects of language contact include calques, simplification of gender distinctions, and verb positioning (Bullock and Gerfen, 2004:99) as well as analogical leveling (Dorian, 1977 : 29). Additionally, in a situation of language death, speech registers contract and categorical rules can become variable (Dressler, 1972 : 454-5; Dorian, 1977:29; Dressler and Wodak-Leodolter, 1977 : 37).

Despite this evidence, Cook (1995 : 218) claims that “convergence rarely or never occurs in a dying language”, an argument that is refuted by numerous phonological studies. Notably, Ola-Orie (2009) examines attritional changes in Trinidad Yoruba phonology, and shows that some “result from convergence and [others from] language acquisition reversal” (Orie, 2009), but not from incomplete acquisition of the language. Also, Bullock and Gerfen (2004) analyse Frenchville French, noting that “this linguistic variety is undergoing a perceptually striking process of phonetic convergence [...] motivated by the auditory and acoustic similarity between a subset of vowels in the contact languages” (Bullock and Gerfen, 2004 : 95). The authors go on to say that “bilingual phonologies may [thus] become particularly permeable to inter-linguistic influence precisely where they are acoustically and perceptually unstable, and where they are already congruent to some degree” (Bullock and Gerfen, 2004 : 95). These

findings present interesting possibilities to the phonological study of dying languages, since Orié (2009) supports Dressler's (1972 : 450) hypothesis comparing convergence to a reversal of language acquisition, and Bullock and Gerfen (2004) point to a phonetics-phonology mismatch as a factor.

The variability presented by speakers of dying languages is difficult to account for, since language death is accompanied by a linguistic continuum of fluency, with older fluent speakers at one end and semi-speakers at the other (Rottet, 2005 : 243). As a result, it is very important for researchers to carefully assess the fluency of each language user considered in an analysis. Moreover, instances of language convergence and contraction do not share a common degree of phonetic variation. Connell (2002) identifies numerous studies to show that language contraction can lead to significantly high levels of variation, some variation indicative of convergence, or little variation or reduction at all, and that these different degrees of variation are not necessarily reported (Connell, 2000 : 168-169). In terms of his own work, the researcher reports very high levels of phonetic variation in a moribund language spoken daily in Cameroon by thirty people, despite the sociolinguistic homogeneity of the speech community; he concludes that extensive variation can also reflect the relaxed language norms of small, isolated speech communities (Connell, 2002 : 181). Phonetic and phonological variation can also be caused by lexical fading (Sabino, 1994 : 514), whereby "a rule is slowly dropped by successively withdrawing words from its input" (Dressler, 1972 : 452). Finally, the same phenomenon (for example, loss of front rounded vowels in French) can be internally motivated in some language varieties and both internally and externally motivated in others (Bullock and Gerfen, 2004 : 103). For instance, older fluent bilingual speakers of Cajun French² and English have been shown to introduce attritional changes in Cajun French that are then adopted by both fluent and restricted speakers of the next generation, so that such changes accumulate as they pass down through generations (Dubois and Noetzel, 2005 : 132). These sociolinguistically-conditioned changes are internal, but contact with the dominant language (English) represents an external influence that motivates additional attritional changes (Dubois and Noetzel, 2005 : 140). The researcher must therefore take these factors into account when examining variable surface forms in an endangered language.

5 Studies of dying languages and OT

Very few phonological analyses of dying languages employ OT, and different varieties of OT are adopted in each study, which are examined here in chronological order. Green (1997) explores the prosody of Irish, Scots Gaelic and Manx, with the idea that free variation due to sound change is best explained by "multiple grammars with varying constraint rankings", so that different sociolinguistic groups have different grammars, rather than "as a result of unranked constraints", since this second option "has undesirable sociolinguistic and methodological implications" (Green, 1997 : 23). Green does not report intra-speaker variation, instead attributing free variation to dialect groups and noting that "any individual speaker will presumably prefer either [one or the other] variant, depending on various social circumstances" (Green, 1997:94). In a different approach, Nevin (1998) studies the Pit River language, Achumawi, and uses the theory of floating constraints (Nagy, 1994; Nagy and Reynolds, 1997) to explain surface variability. Nevin does acknowledge variable surface forms not treated in his thesis, "[assuming] this reflects attrition" (Nevin, 1998 : 37), but does not give attested frequencies of each surface variant. In a study of Mundurukú, Picanço (1999) concedes that classic OT cannot account for the multiple surface forms attested in the language, but the author proposes that this variation simply reflects ongoing changes in the constraint rankings of the language (Picanço, 1999 : 131) and "the underlying representations of alternating morphemes" (Picanço, 1999 : 309). Kim (2003) supports her OT analysis with grounded phonology (Archangeli and Pulleyblank, 1994) to explain dialectal, morphological and prosodic variation in Nuu-chah-nulth; however, this researcher does not mention any kind of intra-speaker variation. Finally, Bradley (2007) focuses on sonority and syllable constraints controlling metathesis in Judeo-Spanish (JS). His analysis of variation is in accordance with Anttila and Cho (1998), asserting that "the variability of metathesis results from a partially-ordered hierarchy in which the relationship between some of the constraints is in flux" (Bradley, 2007 : 175).

Unlike the examinations of dying language phonological processes considered thus far, Auger's (2001) study of vowel epenthesis in Vimeu Picard is distinguished by the fact that the OT analysis attempts to predict observed frequencies of this phenomenon (see also Auger and Steele, 1999). As in Anttila (2002) and Bradley (2007), certain constraints are crucially unranked in Auger's analysis in order to account for variable patterns of word-initial vowel epenthesis (Auger, 2001 : 274). All possible rankings and the winning candidate in each case are calculated and tallied, giving expected frequencies of different surface forms, which are then compared for statistical probability with the observed surface forms (Auger, 2001 : 277). Another interesting aspect of Auger's work is support for the idea of individual grammars in Optimality Theory (Auger, 2001 : 253), following Reynolds (1994). Specifically, Auger (2001 : 286) "propose[s] that variable OT grammars allow us to implement the intuition that individual grammars are more restricted versions of the community grammar by specifying some of the rankings left unspecified in the community grammar", so that crucially unranked constraints at the community level can be strictly ranked at the individual level in order to explain inter-speaker variation. This stance seems to be supported in the moribund language of Cameroon, Cambap, where "each speaker shows an individual pattern" for many of the attested allophonic consonant variations (Connell, 2002 : 176).

6 Louisiana Regional French

One type of speaker has not been discussed as of yet: the fluent speaker of a dying language who no longer regularly uses his or her native tongue in a speech community or Community of Practice (CofP) (Eckert and McConnell-Ginet, 1999). These speakers still distinguish between formal and informal speech registers (Blainey, 2009), but add a layer of complication to the variability of surface realisations of underlying forms, since it seems that their ranking hierarchy is beginning to disintegrate as their memory fades. They thus appear to forget strict rankings sometimes, and either generalise by analogy or variably apply phonological rules, posing unique challenges to OT as a framework for explaining variability.

Such speakers can be found in the Louisiana recordings (Klingler and LaFleur, 2007) that form part of *la phonologie du français contemporain* (PFC, Durand, Laks and Lyche, 2002, 2005). This international French phonology corpus aims to examine liaison and schwa phenomena in contemporary varieties of spoken French, but also provides comparable French language corpora for other phonological inquiries. There are many areas around the world where French is endangered; in North America alone, Rottet (2005 : 243) identifies numerous areas where French is in a linguistic minority situation with English, such as Newfoundland, New England, Ontario, New Brunswick, and Louisiana. The latter location and variety forms the focus of the remainder of the paper.

Louisiana French is an enclave dialect, because it no longer has linguistic contact with the French spoken in France; more specifically, this variety "constitute[s] a collection of subdialects, all of which are being endangered simultaneously" (Picone, 1997 : 117). Indeed, Louisiana French will likely die out in the near future, because it is no longer transmitted to children (Picone, 1997 : 125; Rottet, 2005 : 243), even though a 1990 CODOFIL and the University of Southwestern Louisiana telephone survey reveal that most self-reported French communication (66.9%) occurs between relatives (Picone, 1997 : 130). Decline in language use is also due in part to language attitudes, although evidence is mixed concerning just what those attitudes are. Dubois and Melançon (1997) find that for people living in the French Triangle of Louisiana, while "the most important criterion for Cajun identity is ancestry [...] self-identification is ultimately linked to ability in Cajun French for all age groups" (Dubois and Melançon, 1997 : 84). This leads the authors to conclude that "the notion of a Cajun speech community remains valid despite the precarious state of the language in Louisiana" (Dubois and Melançon, 1997 : 84). However, Dubois and Horvath (1998) report a somewhat different scenario: "it is now socially advantageous to show that you are Cajun. However, for the young generation, if Cajun identity is to be signaled by language, then it is to be done in English" (Dubois and Horvath, 1998 : 257).

It is important to remember that the term 'Cajun' used in these studies is problematic as a language label, as is the identifier 'Creole', because these words are also used to designate ethnic groups in Louisiana. As

a result, speakers will use the label for their ethnic group ('Cajun' or 'Creole') to identify the language they speak, even if this does not correspond, by linguistic definition, to the variety of French they use (Klingler, 2005 : 354-356). 'Cajun' has not always been a positive word, either. Trépanier (1991) suggests that institutional interference changed language label use in the late 1960s. The author explains that the Council for the Development of French in Louisiana (CODOFIL) decided to use the term 'Cajun' to identify and unify French-speaking Louisiana, a "choice [that] can only be interpreted as the desire for the French Louisiana elite to assure for the region a 'white' identity" (Trépanier, 1991 : 164). In order to transform the word "Cajun" from an insult to a badge of honour, the link between the Acadians and the Cajuns was emphasized by several state institutions (Trépanier, 1991 : 164). In fact, Klingler (2009) considers Acadian influences on Louisiana varieties of French, in terms of settlement and linguistic characteristics, and demonstrates that the link between 'Acadian' and 'Cajun' in Louisiana has been largely overstated. Nevertheless, the misperception and over-exaggeration of the link between the Acadians and the Cajuns is extensive in academic literature on French-speaking Louisiana, where 'Cajun' and 'Acadian' are used interchangeably (Klingler, 2003 : 77; Picone and Valdman, 2005 : 145; Trépanier, 1991 : 161; Valdman, 1979, quoted in Bullock and Gerfen, 2004 : 97). For these reasons, the present study adopts the term 'Louisiana Regional French' (LRF) in contrast with 'Louisiana Creole' (Klingler, 2003 : 77).

7 Preliminary research with the LRF corpus

The two male (M1, M2) and two female (F1, F2) speakers in the Ville Platte PFC corpus are all over 60 years of age. It is difficult to test sociolinguistic variables with so few speakers and such a limited age range; this situation will improve as more speakers are added to the corpus. Preliminary research suggests that the corpus exhibits a great deal of inter-speaker (Russell, 2009) as well as intra-speaker (Blainey, 2009) variation. For instance, a pilot study of schwa epenthesis in obstruent-rhotic (OR) clusters and in word-initial position shows that in formal speech, schwa is inserted 59% of the time, whereas in an informal register schwa epenthesis occurs 37% of the time (Blainey, 2009). However, if the OR cluster data are considered alone, then the individual rates of schwa epenthesis, shown in Figures 1 and 2 below, demonstrate that the pattern is not always so clear.

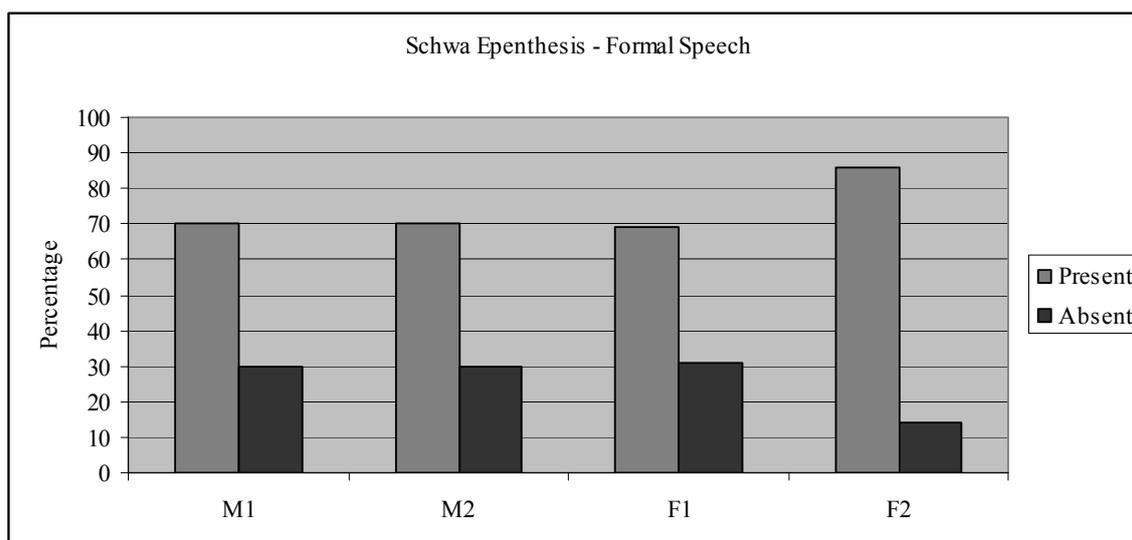


Figure 1: Rate of schwa epenthesis by speaker in formal speech (translation exercises).

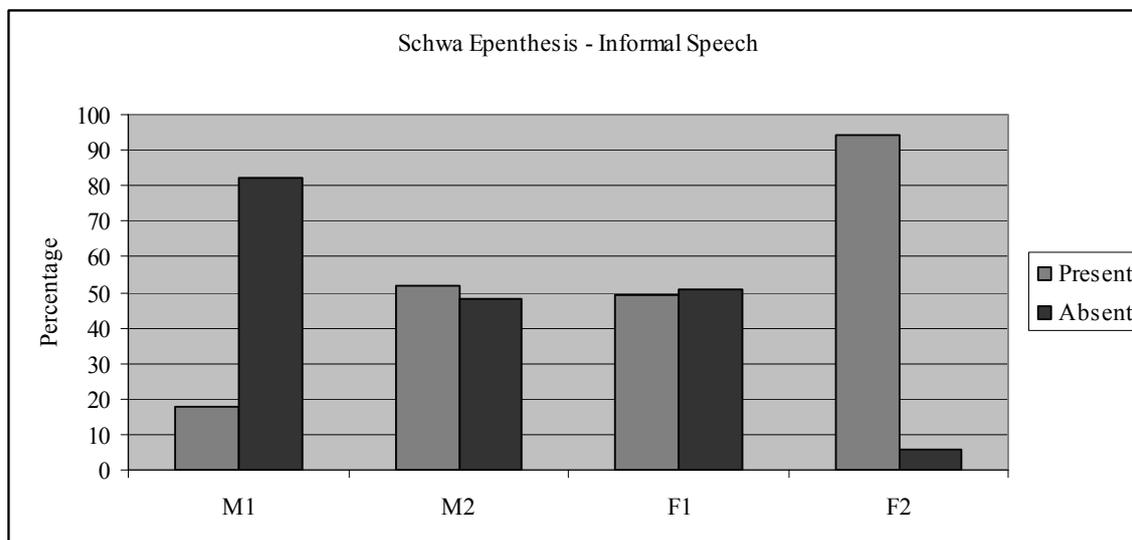


Figure 2: Rate of schwa epenthesis by speaker in informal speech (free conversation).

In Figure 1, the formal register exhibits a consistent pattern of schwa epenthesis for all speakers (between 69% and 86%). Conversely, Figure 2 displays varying rates of schwa epenthesis. While M1 simplifies OR clusters by inserting schwa only 18% of the time (13 of 72 tokens), M2 and F1 use schwa insertion roughly half the time (36 of 69 tokens, and 23 of 47 tokens, respectively). It should be noted that F2's exceptional patterning may be the result of a lack of tokens; as has been observed in previous work (Russell, 2009 : 45), F2 is not as talkative as the other speakers in the corpus. As a result, there are only 16 OR clusters to consider for F2 in the informal register, only one of which does not undergo simplification via schwa epenthesis. It is also possible that these data demonstrate that F2 does not change her rate of schwa epenthesis based on the formality of the speech event. At any rate, M1's non-conformity with the other two speakers' rates of schwa epenthesis is not due to a relative paucity of tokens, nor can it be linked to the pattern of schwa epenthesis he adopts in the formal register. Therefore, the lack of a clear trend in the informal register may indicate that when speakers feel more at ease, they are more likely to vary in their application of what is otherwise a consistent pattern of schwa epenthesis. The addition and analysis of more speakers will confirm whether or not these results support the hypotheses of Auger and Steele (1999) and Auger (2001), who promote individual grammars to account for individual patterns.

The qualities of the obstruent itself, such as voicing, may be responsible at least in part for the rate of schwa insertion in OR clusters in LRF (Blainey, 2009). This link has been noted in other varieties of French where OR cluster schwa insertion occurs (Colantoni and Steele, 2007; Chow and Poiré, 2007). The authors of the former study remark that in varieties of French spoken in Québec and in France, "[schwa] epenthesis is quasi-categorical in voiced [OR] clusters, whereas there is virtually no epenthesis in voiceless clusters [...while] in Spanish, [schwa] epenthesis is categorical [in all OR clusters]" (Colantoni and Steele, 2007 : 119) because of the similarity between the obstruent and the Spanish tap. In LRF, /R/ can surface as any of [r, ɾ, ʀ] (Lyche, 1996 : 34), the most common of which is [r]. Thus, LRF exhibits a similar epenthetic pattern to Québec and France French, but a similar inventory of rhotic allophones to the Spanish varieties in Colantoni and Steele (2007 : 118). However, the Ville Platte corpus data do not exhibit a pattern that clearly follows either language's schwa epenthesis behaviour, as demonstrated in Table 2. While the labial and labiodental phonemes follow a general trend of higher rates of epenthesis following voiced obstruents and lower rates following voiceless obstruents, as seen in other varieties of French, this pattern is by no means categorical. Furthermore, while at least 50% of the tokens for each [-labial] obstruent exhibit epenthesis, this tendency does not apply to all phonemes and all clusters, as it would in varieties of Spanish.

Obstruent	Schwa		Obstruent	Schwa	
	Present	Absent		Present	Absent
/p/	12	49	/t/	64	40
/b/	35	5	/d/	11	4
/f/	9	28	/k/	33	27
/v/	10	1	/g/	92	23

Table 2: Presence or absence of schwa in OR clusters by obstruent

Since the Ville Platte corpus has been coded according to the PFC protocol, coding does not include phonetic information: all rhotics are coded as [R]. In a study of the PFC Windsor French corpus, Chow and Poiré (2007 : 1024) identify a link between the type of rhotic and the rate of schwa epenthesis. It is therefore essential that future development of this analysis consider the type of rhotic as a factor in OR schwa epenthesis patterns. Just the same, these frequencies clearly distinguish the patterns of LRF schwa epenthesis from those of the more robust varieties of France and Québec French.

At the level of the lexeme, it becomes even more difficult to predict rates of schwa epenthesis. Some words in the corpus always contain an epenthetic schwa, such as *brun(e)* ‘dark-skinned, brown’. It is possible that in these cases, the weak vowel is now part of the word’s underlying representation. Other lexemes seem conditioned to nearly always prohibit schwa insertion (*français* ‘French’). The constraint *VCE∩FRIC_(C→), “avoid voiced fricatives post-consonantly” (Colantoni and Steele, 2007 : 122) works to explain patterns of schwa epenthesis in Québec and France French, since speakers of these varieties produce the dorsal fricative [ɣ] (Colantoni and Steele, 2007 : 114). This constraint will not suffice for LRF data, due to the variable nature of /R/ explored above.

At present, the following constraints can be used to give a rough estimation of epenthesis frequencies in the Ville Platte corpus:

- | | |
|---------------------------------|---|
| (1) *RHOTIC _(C[VD]→) | avoid rhotics after voiced consonants |
| (2) *SCHWA | avoid schwa |
| (3) DEP-IO-V | every output vowel must appear in the input |
| (4) *COMPLEX-ONSET | avoid onsets with more than one segment |

A word such as *grand* ‘big’, attested 17 times in the LRF corpus, has a variable pronunciation. Schwa epenthesis occurs 15 times, but the lexeme is also said twice without schwa. Table 3 is able to express that the form containing an epenthetic schwa is the winner, since [gRã] violates higher-ranking *RHOT_(C[VD]→), but it cannot capture the fact that both candidates are attested.

/gRã/	*RHOTIC _(C[VD]→)	*SCHWA	DEP-IO-V	*COMPLEX-ONSET
☞ [gəRã]		*	*	
[gRã]	*!			*

Table 3: OT Tableau of *grand* ‘big’

By the same token, the lexical item *prends* ‘take (2SG.PRES)’ has two possible pronunciations in LRF, though the more faithful output form, [pRã], occurs five out of the six times the word is said in the corpus. As can be expected from traditional OT, only a categorical relationship is expressed in Table 4.

/pRã/	*RHOTIC _(C[VD]→)	*SCHWA	DEP-IO-V	*COMPLEX-ONSET
☞ [pRã]				*
[pəRã]		*!	*	

Table 4: OT Tableau of *prends* ‘take (2SG.PRES)’

These cases are almost categorical in their treatment of epenthetic schwa, but many other lexical items do not make such a clear divide. For example, five tokens of the word *brin* ‘(a little) bit’ feature schwa insertion ([bəRɛ̃]), while three do not ([bRɛ̃]), a ratio that is quite different from that of *grand* ‘big’ in Table 3. Further examples of lexical items in the Ville Platte corpus are given below in Table 5. This list is by no means exhaustive. Rather, it gives an idea of the scope of the phenomenon in terms of lexical items, grammatical categories and syllable structure; it also illustrates the complexity and variability of schwa epenthesis patterns in the French of Ville Platte.

Lexeme	Gloss	Transcription ³	Obstruent	Schwa	
				Present	Absent
emprunter	to learn	[ãpRɛ̃te]	/p/	5	3
printemps	Spring	[pRɛ̃tã]	/p/	0	2
brun	dark-skinned, brown (m.)	[bRœ̃]	/b/	12	0
brune	dark-skinned, brown (f.)	[bRœ̃n]	/b/	9	0
frais	fresh (m.)	[fRe]	/f/	3	2
froment	flour	[fRamõ]	/f/	1	3
avril	April	[avRi]	/v/	4	0
février	February	[fɛvRije]	/v/	5	0
étriller	to currycomb a horse	[etRije]	/t/	4	5
l'étrier	stirrup	[letRije]	/t/	7	1
trop	too much	[tRo]	/t/	2	5
drôle	odd	[dRøl]	/d/	4	1
crapaud	toad	[kRapo]	/k/	3	3
croquesignole	doughnut	[kRœksijøl]	/k/	2	7
graisse	fat	[gRɛs]	/g/	10	1
gratons	cracklings	[gRatõ]	/g/	5	1
gros	big (m.)	[gRo]	/g/	11	39
grosse	big (f.)	[gRos]	/g/	9	5

Table 5: Partial list of schwa epenthesis frequencies in OR clusters

A full analysis of schwa epenthesis by speaker in LRF is outside the limits of this article, but the lexical items examined above give a beginning indication of the complicated nature of the data. Though individual grammars may be a possibility for explaining the LRF schwa epenthesis data, it is not clear that the concept of multiple underlying grammars will be a necessary analytical tool. Floating constraints do not seem to offer a solution, either (Naomi Nagy, personal communication), though partial ordering of certain constraints may be able to explain a certain amount of variability. An extension of OT that employs a linear scale, such as LOT or Stochastic OT, is left as the most likely way to accurately account for the high levels of variation in the LRF schwa epenthesis data. Once more tokens and speakers are available for analysis, these expansions of OT can be compared in order to determine which of the two most economically and correctly predicts rates of phonological variation in moribund languages. If lower-ranking constraints are shown to gang up on higher-ranking constraints, LOT will be a viable theory for explaining the LRF data; however, the advantages outlined in section 3 favour Stochastic OT as a model of language variation.

Based on the information presently available, it seems that as speakers begin to forget their native language, the probability curves of formerly separate constraints begin to overlap, causing surface variation. This trend of constraint overlap is more pronounced and variable in informal conversation. For some lexical items, one of the new surface variants is reanalysed as the underlying form, as may be the case with *brun(e)* ‘dark-skinned, brown’. For more variable forms, constraint movements occur in favour of markedness constraints, promoting forms that are less difficult to articulate and more perceptually salient. Further analysis of surface variation in LRF will test these hypotheses, as will research in other endangered varieties of French and dying languages around the world. If confirmed in a wide array of language varieties and families, these tendencies may encourage the formulation of a reverse GLA for language death.

8 Conclusion

This paper has sought to underline the importance of the variable nature of dying languages to theories attempting to predict and account for phonological variation. Specifically, it has examined the literature on variation and OT in order to show that a classic optimality-theoretic perspective is insufficient for variable data, since it insists on strict constraint domination and a discrete ranking system. Though several of the proposed extensions to OT are attractive and account for certain data sets, Stochastic OT has the added advantage of coming equipped with a language learning algorithm, and accurately mapping raw count data into a formal phonological framework. Such an algorithm would certainly benefit from research in dying languages, since language death has been proposed to be a kind of reverse language acquisition. Furthermore, if such a theory were able to properly describe the process of language death and the unique variability such a process entails, that success would make the theory even more attractive. Studies in dying languages are therefore central to the advancement of OT, but most existing OT studies involving endangered languages are primarily concerned with morpho-syntax. Phonological corpora such as the PFC provide large and easily comparable databases which researchers can access, a luxury not often available to those studying endangered language varieties. The case of French, a language spoken all over the world that nonetheless comprises numerous enclave dialects, is thus an excellent starting point for such research.

Naturally, the results from the preliminary analyses of LRF sound patterns considered in this work are not conclusive, and the Ville Platte corpus must be expanded in terms of speakers and tokens in order to test the rates of schwa epenthesis against predictions made in the different extensions of OT, especially LOT, Stochastic OT and the GLA. Furthermore, perceptual tests will be necessary in order to determine whether or not an epenthetic schwa has been reanalysed as part of the underlying form for certain lexical items in LRF. Finally, the rhotic allophones of LRF must be coded and analysed as a potential factor in rates of schwa epenthesis. Nevertheless, this pilot study does suggest that the process of OR cluster simplification through schwa epenthesis in LRF is distinct from other varieties of French, such as Québec French and France French. In addition, introductory examinations of LRF phonology put forward interesting questions, challenges and implications for OT. Moreover, any tentative conclusions that can be

drawn from preliminary OT research in certain endangered varieties of French demand confirmation from other enclave dialects of French, and unrelated dying languages as well. Therefore, it is imperative that this small body of literature grow, in order to explore and explain the perplexing and fascinating nature of endangered language phonology.

References

- Anttila, A. (2002). Variation and phonological theory. In Chambers, J., P. Trudgill and N. Schilling-Estes (eds.), *Handbook of Language Variation and Change*, Oxford and Malden: Blackwell, 206–243.
- Anttila, A. and Y. Cho (1998). Variation and change in Optimality Theory. *Lingua*, 104, 31–56.
- Archangeli, D. and D. Pulleyblank (1994). *Grounded phonology*. Cambridge: MIT Press.
- Auger, J. (2001). Phonological variation and Optimality Theory: Evidence from word-initial vowel epenthesis in Vimeu Picard. *Language Variation and Change*, 13(3), 253–303.
- Auger, J. and J. Steele (1999). A case study in microparametric variation: Vowel epenthesis in Vimeu Picard. *10th International Conference on Methods in Dialectology (Methods X), St. John's, NL, Aug. 1–6*.
- Austin, P. (2006). Content questions in Sasak, Eastern Indonesia: An optimality theoretic syntax account. In Schulze, F. and H. Warnk (eds.), *Insular Southeast Asia. Linguistic and Cultural Studies in Honour of Bernd Nothofer*, Wiesbaden: Harrassowitz Verlag, 1–12.
- Blainey, D. (2009). Schwa behaviour in formal and informal speech in the French of Ville Platte, Louisiana. *Poster presented at the Canadian Linguistics Association/L'association canadienne de linguistique (CLA/ACL), Carleton University, Ottawa, ON, May 23–25*.
- Boersma, P. (1999). Optimality-Theoretic learning in the Praat program. *IFA proceedings*, 23, 17–35.
- Boersma, P. and B. Hayes (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, 32(1), 45–86.
- Boersma, P. and D. Weenink (2008). Praat: doing phonetics by computer (Version 5.0.42). Computer program. Retrieved September 01, 2008, <http://www.praat.org/>.
- Bradley, T. (2007). Constraints on the metathesis of sonorant consonants in Judeo-Spanish. *Probus*, 19(2), 171–207.
- Bullock, B. and C. Gerfen (2004). Phonological convergence in a contracting language variety. *Bilingualism: Language and Cognition*, 7(2), 95–104.
- Chow, I. and F. Poiré (2007). Rhotic variation and schwa epenthesis in Windsor French. In van Hamme, H. and R.J.J.H. van Son (eds.), *Proceedings of INTERSPEECH-2007*, 1022–1025.
- Colantoni, L. and J. Steele (2007). Voicing-dependent cluster simplification asymmetries in Spanish and French. In Prieto, P., M.J. Solé and J. Mascaró (eds.), *Segmental and prosodic issues in Romance linguistics*, Amsterdam: John Benjamins, 109–129.
- Connell, B. (2002). Phonetic/phonological variation and language contraction. *International Journal of the Sociology of Language*, 2002(157), 167–185.
- Cook, E. (1995). Is there convergence in language death? Evidence from Chipewyan and Stoney. *Journal of Linguistic Anthropology*, 5(2), 217–231.
- Dorian, N. (1977). The problem of the semi-speaker in language death. *International Journal of the Sociology of Language*, 12, 23–32.
- Dressler, W. (1972). On the phonology of language death. *Chicago Linguistics Society*, 8, 448–457.
- Dressler, W. and R. Wodak-Leodolter (1977). Language preservation and language death in Brittany. *Linguistics*, 191, 33–44.
- Dubois, S. and B. Horvath (1998). Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change*, 10, 245–261.

- Dubois, S. and M. Melançon (1997). Cajun is dead – long live Cajun: Shifting from a linguistic to a cultural community. *Journal of Sociolinguistics*, 1(1), 63–93.
- Dubois, S. and S. Noetzel (2005). Intergenerational pattern of interference and internally-motivated changes in Cajun French. *Bilingualism: Language and Cognition*, 8(2), 131–143.
- Durand, J., B. Laks and C. Lyche (2002). La phonologie du français contemporain: usages, variétés et structure. In Pusch, C. and W. Raible (eds.), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*, Tübingen: Gunter Narr Verlag, 93–106.
- Durand, J., B. Laks and C. Lyche (2005). Un corpus numérisé pour la phonologie du français. In G. Williams (ed.), *La linguistique de corpus*, Rennes: Presses Universitaires de Rennes, 205–217.
- Eckert, P. and S. McConnell-Ginet (1999). New generalizations and explanations in language and gender research. *Language in Society*, 28, 185–201.
- Green, A. (1997). The prosodic structure of Irish, Scots Gaelic, and Manx. *Doctoral dissertation, Cornell University*.
- Keller, F. (2006). Linear Optimality Theory as a model of gradience in grammar. In Fanselow, G., C. Féry, R. Vogel and M. Schlesewsky (eds.), *Gradience in Grammar: Generative Perspectives*, New York: Oxford University Press, 270–288.
- Keller, F. (2000). Gradience in grammar : Experimental and computational aspects of degrees of grammaticality. *Doctoral dissertation, University of Edinburgh*.
- Kim, E. (2003). Theoretical issues in Nuu-chah-nulth phonology and morphology. *Doctoral dissertation, The University of British Columbia*.
- King, R. (2005). Crossing grammatical borders: Tracing the path of contact-induced linguistic change. In Filppula, M., J. Klemola, M. Palander and E. Penttilä (eds.), *Dialects Across Borders: Selected Papers from the 11th International Conference on Methods in Dialectology (Methods XI), Joensuu, August 2002*, Amsterdam and Philadelphia: John Benjamins, 233–251.
- Kiparsky, P. (1993). Variable rules. *Handout, Rutgers Optimality Workshop (ROW-1), New Brunswick, NJ, Oct. 22–24*.
- Kiparsky, P. (1994). An optimality-theoretic perspective on variable rules. *Paper presented at NWAVE 23, Stanford University*.
- Klingler, T. (2003). Language labels and language use among Cajuns and Creoles in Louisiana. *Working papers in linguistics*, 9, 77–90.
- Klingler, T. (2005). Le problème de la démarcation des variétés de langues en Louisiane : étiquettes et usages linguistiques. In Valdman, A., J. Auger and D. Piston-Hatlen (eds.), *Le français en Amérique du Nord: état présent*, Québec: Les Presses de l'Université Laval, 349–367.
- Klingler, T. (2009). How Much Acadian is there in Cajun? In Mathis-Moser, U. and G. Bischof (eds.), *Acadians and Cajuns: The Politics and Culture of French Minorities in North America*, Innsbruck: Innsbruck University Press, 91–103.
- Klingler, T. and A. LaFleur (2007). L'enquête PFC en Louisiane. In Detey, S. and D. Nouveau (eds.), *Bulletin PFC N. 7 : PFC: enjeux descriptifs, théoriques et didactiques 2007*, <http://www.projet-pfc.net/?pfc-rc:bibliopfc:pfc7>, 331–40.
- Lyche, C. (1996). Genèse et traits caractéristiques du français cadien : un aperçu phonologique. *Revue romane*, 31(1), 29–49.
- McCarthy, J. and A. Prince (1994). The emergence of the unmarked: Optimality in prosodic morphology. In M. González, M. (ed.), *Proceedings of the 24th North East Linguistics Society*, Amherst, Massachusetts: Graduate Linguistic Student Association, 333–379.
- Morimoto, Y. (2002). Antisymmetry and optimality : Positional variation in negation. *Doctoral dissertation, Universität Düsseldorf*.
- Mougeon, R., T. Nadasdi and K. Rehner (2005). Contact-induced linguistic innovations on the continuum of language use: The case of French in Ontario. *Bilingualism: Language and Cognition*, 8(2), 99–115.

- Nagy, N. and W. Reynolds (1997). Optimality Theory and variable word-final deletion in Faetar. *Language variation and change*, 9(1), 37–55.
- Nevin, B. (1998). Aspects of Pit River phonology. *Doctoral dissertation, University of Pennsylvania*.
- Ola-Orie, O. (2009). Trinidad Yoruba phonological attrition : a case of convergence and language acquisition reversal. *Paper presented at the South Eastern Conference on Linguistics (SECOL-LXXVI), Tulane University, Apr. 8-10*.
- Picanço, G. (1999). Mundurukú : Phonetics, phonology, synchrony, diachrony. *Doctoral dissertation, Federal University of Pará*.
- Picone, M. (1997). Enclave dialect contraction: an external overview of Louisiana French. *American Speech*, 72(2), 117–153.
- Picone, M. and A. Valdman (2005). La situation du français en Louisiane. In Valdman, A., J. Auger and D. Piston-Hatlen (eds.), *Le français en Amérique du Nord: état présent*, Québec: Les Presses de l'Université Laval, 143–165.
- Prince, A. and P. Smolensky (1993). Optimality Theory: Constraint interaction in generative grammar. *Ms., Rutgers University and University of Colorado at Boulder*.
- Reynolds, W. (1994). Variation and phonological theory. *Doctoral dissertation, University of Pennsylvania*.
- Rottet, K. (1998). Clause subordination structures in language decline. *Journal of French language studies*, 8(1), 63–95.
- Rottet, K. (2001). Language shift in the coastal marshes of Louisiana. *Studies in Ethnolinguistics* 8. New York: Peter Lang Publishers.
- Rottet, K. (2005). Variation et étoilement en français cadien. In Valdman, A., J. Auger and D. Piston-Hatlen (eds.), *Le français en Amérique du Nord: état présent*, Québec: Les Presses de l'Université Laval, 243–260.
- Russell, A. (2009). L'affrication à la Ville Platte: la variation dans la réalisation des occlusives en français louisianais. *Master's thesis, University of Western Ontario*.
- Sabino, R. (1994). They just fade away: Language death and the loss of phonological variation. *Language in Society*, 23(4), 495–526.
- Sorace, A. and F. Keller (2005). Gradience in linguistic data. *Lingua*, 115(11), 1497–1524.
- Tesar, B. and P. Smolensky (2000). *Learnability in Optimality Theory*. Cambridge: MIT Press.
- Trépanier, C. (1991). The Cajunization of French Louisiana: Forging a regional identity. *The Geographical Journal*, 157(2), 161–171.
- Valdman, A. (1979). Créolisation, français populaire et le parler des isolats francophones d'Amérique du Nord. In Valdman, A. (ed.), *Le français hors de France*, Paris: Champion, 181–197.

¹ Research supported by SSHRC Doctoral Fellowship # 752-07-1462.

² For a discussion of the term 'Cajun French', see section 6 of this paper.

³ Some lexical items have multiple attested pronunciations; for clarity and simplicity, the transcription given here selects the most common attested pronunciation, without the epenthetic schwa.