Linguistic Research 32(3), 533-571
DOI: 10.17250/khisli.32.3.201512.002

A reanalysis of the great English vowel shift under contrast preservation theory*

Julien Fulcrand
(Université Charles-de-Gaulle - Lille 3)

Fulcrand, Julien. 2015. A reanalysis of the great English vowel shift under contrast preservation theory. *Linguistic Research* 32(3), 533-571. The goal of this paper is to present a reanalysis of the Great English Vowel Shift in terms of Contrast Preservation Theory (Lubowicz 2003, 2012). Chain shifts like the Great English Vowel Shift pose a challenge for constraint-based theories such as Optimality Theory because they are an instance of opacity. In a system with only two levels of representation, it is impossible to both forbid a sound and allow it to surface in the same contexts. The current paper proposes to evaluate the adequacy of Contrast Preservation Theory by applying it to diachronic data, specifically the Great English Vowel Shift. We will show how a model developed for synchronic data can be applied to diachronic sound change. In addition, we will claim that the application of modern linguistic theory to diachronic development can offer insights into how language change occurs. Finally, our analysis provides further support for Contrast Preservation Theory. **(Université Charles-de-Gaulle - Lille 3)**

Keywords Diachronic phonology, phonological theory, Contrast Preservation Theory, Great English Vowel

1. Introduction

Chain shifts have always been problematic for phonological theories. They are hard to model in both input-driven derivational theory and output-driven Optimality Theory. The first goal of this paper will be to prove that point. It will be shown that when traditional generative phonological theories try to account for historical chain shifts, the links between the various stages of the chain are broken. As for

* I would like to thank the two anonymous reviewers who provided very insightful remarks and enabled this paper to gain in accuracy and analytical depth. I would also like to thank my PhD directors, Roland Noske and Kathleen O'Connor, as well as my fellow PhD student Joshua Albair for all their help and comments.

Optimality Theory, its two-level representational structure will be proved to be the main weakness of this theory, as far as chain shifts are concerned. Then, after reviewing three unsatisfactory output-driven analyses − one of which is actually a bit more satisfactory than the other two − of the Great English Vowel Shift (henceforth GEVS), a more satisfactory analysis of GEVS under Contrast Preservation Theory (henceforth CPT) will be given. CPT (Łubowicz 2003, 2012) is an output-driven theory that was originally devised for synchronic chain shifts. Nevertheless, in this paper, it will be shown that CPT can also account for diachronic chain shifts. Therefore this paper will serve the additional goal of providing evidence regarding the validity of CPT for diachronic shifts.

A chain shift can be defined as a complex phonological process involving a series of interlinked changes. A general representation of a chain shift is given in (1):

(1) /A/ → [B] > /B/ → [C] > /C/ → [D]

For reasons of clarity, here in (1), slashes and brackets have a different use from in synchronic phonology. Here, slashes characterise pre-shift forms and brackets characterise post-shift forms. (1) represents a chain shift with three stages. First, /A/ moves to the phonetic position of /B/ and becomes [B]. Then, during the second stage, under systemic pressure − i.e. pressure in order to preserve contrast −, here represented by '>', /B/ moves to the phonetic position of /C/ and becomes [C]. Finally, in the last stage, still under systemic pressure, /C/ becomes [D]. It is important to bear in mind that the output of one stage and the input of the next stage are of a different nature. For example, in the first and the second stages of the chain shift described in (1), the output [B] of the first stage is not the direct input of the second stage. /B/ is the actual input of the second stage and existed in the considered system before the beginning of the first stage of the chain shift. As a whole, in (1), before the chain shift, the theoretical phonological system is /A, B, C/. At the end of the chain shift, there is this new system [B, C, D]. A first observation is to see that the number of contrasts is the same in the system before and after the chain shift. Before the chain shift, there are two levels of contrast: one between /A/ and /B/ and another one between /B/ and /C/. After the chain shift, there still are two levels of contrast: one between [B] and [C] and another one between [C] and [D]. The fact that the number of contrasts remains the same before and after the chain shift will refer to what will be seen in section 5 with Contrast Preservation Theory. Here, GEVS is used to illustrate (1).

The term GEVS was first coined by Jespersen (1909). It affected the long vowels in the evolution from Middle English to Early Modern English. It is the main cause of the spelling peculiarities found in Present-Day English. Consider (2) and (3)[1] where phonemic transcriptions of words before and after GEVS are compared. Front vowels shift is illustrated in (2)[2] and likewise for the back vowels in (3).

(2)

Pre-GEVS	Post-GEVS
maken /m**a**ː kən/	make /m**ɛ**ː k/
meat /m**ɛ**ː t/	meat /m**e**ː t/
feet /f**e**ː t/	feet /f**i**ː t/
tide /t**i**ː də/	tide /t**aɪ**d/

(3)

Pre-GEVS	Post-GEVS
goat /g**ɔ**ː t/	goat /g**o**ː t/
food /f**o**ː d/	food /f**u**ː d/
hous /h**u**ː s/	hous /h**au**s/

Therefore, based on the representation of a chain shift given in (1) and the data of (2) and (3), the two vocalic chain shifts shown in (4) and (5) can be inferred:

(4) /aː/ → [ɛː] > /ɛː/ → [eː] > /eː/ → [iː] > /iː/ → [aɪ]
(5) /ɔː/ → [oː] > /oː/ → [uː] > /uː/ → [au]

The paper is organised as follows: in the next section, it will be shown that classical generative theories like SPE (Chomsky & Halle 1968) have great difficulties in accounting for chain shifts. Then, in section 3, it will be shown that this is also true for output-driven theories like OT (Prince & Smolensky 1993). In section 4, three case studies of GEVS will be reviewed. These studies used modified versions of OT in order to tackle the issues described in section 3. Nevertheless, despite the modifications, the weaknesses and the limitations of these approaches will be pointed out. All in all, these approaches will be proved to be unsatisfactory − or limited for one of them − to account for GEVS. This will lead to section 5, in which the analysis of GEVS under CPT will be presented. It will be proved that, with CPT, the difficulties found in previous analyses are not encountered.

---

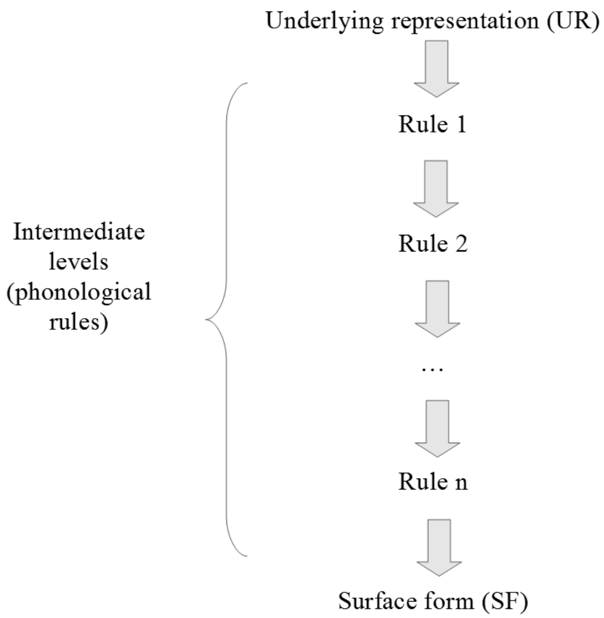
1 The description of GEVS is from Baugh & Cable (1993, 2002:238).

2 In the case of the verb 'make', there is also an infinitival inflection shift, which is outside the scope of the present article.

2. Chain shifts and classical generative theory

This theoretical framework based on phonological rules dates back to Chomsky & Halle (1968). SPE assumes a sequential application of rules as shown in (6):

(6)



Such a sequential organisation, or derivation, is already problematic for historical chain shifts. In a traditional chain shift as depicted in (1), if a sequential rule application as shown in (6) is applied, then one would end up with the /A/ → [D] mapping. However, such an analysis is problematic because in the end there are several surface forms, [B], [C] and [D], not [D] alone. This is shown in (7):

(7) /A/ → [B] > /B/ → [C] > /C/ → [D] (push chain)

UR	/A/	/B/	/C/
/A/ → [B]	B		
/B/ → [C]	C	C	
/C/ → [D]	D	D	D
SF	*[D]	*[D]	[D]

/A/ should map to [B] and /B/ should map to [C] but under this derivation, both /A/ and /B/ map to [D]. In (7), which illustrates derivational theory, once the /A/ → [B] shift has occurred, there is nothing to block the /B/ → [C] shift afterwards and so on. The same principle applies for the /B/ → [C] shift: once the latter has occurred, there is nothing to block the /C/ → [D] shift. That is why chain shifts are a clear instantiation of opacity.

The notion of opacity has played an important role in the debate of phonological theories. It was defined by Kiparsky (1968:79) and shed some light on very problematic cases where one found a form which contradicted a particular surface generalisation. In (7), for example, the global merger to [D] is a surface generalisation predicted by the derivational theory illustrated in (7). This generalisation is not what is observed in the chain shift in (1). Four types of rule order were defined: feeding, bleeding, counterfeeding and counterbleeding.[3] Opacity can be found in counterfeeding and counterbleeding. In this paper, the focus will be on the counterfeeding order. The latter is described below in (8):

[3] In this article, the focus will be on the counterfeeding order. Nevertheless, the other types of rule order are briefly described as follows. Consider two rules A and B:

- In a feeding order, A creates a context in which B is applicable. A feeds B.
- In a bleeding order, A destroys a context in which B is applicable. A bleeds B.
- The counterbleeding order, as the term suggests it, is the opposite of the bleeding order. It means that A still destroys a context in which B is applicable, but, in such a rule order, B is ordered before A and thus can be applied since its context of application is not destroyed by Rule A yet.

(8) Counterfeeding:

Consider two rules A and B. In a feeding order, A creates a context in which B is applicable. However, in a counterfeeding order – the opposite of the feeding order as suggested by the term counterfeeding –, rule B is ordered before rule A and thus rule B cannot be applied since its context of application has not been created by A yet. This will be described in the schematic example below:

Consider the following rules to account for the surface from [DBE] derived from /DAE/:
Rule 1: /A/ → [B] / D__E
Rule 2: /B/ → [C] / D__E

UR	/DAE/
Rule 2: /B/ → [C] / D__E	
Rule 1: /A/ → [B] / D__E	DBE
SF	DBE

In (8), the surface form [DBE] is considered opaque because it contains the context which makes Rule 2 applicable and yet the latter is not applied. Rule 1 can feed Rule 2 but the outcome will be the problematic output *[DCE]. Here, in order to get the correct output,
Rule 2 must be ordered before Rule 1, hence the counterfeeding order.

This can be applied to chain shifts. As seen in (7), the correct outputs were not obtained at the end of the derivation. In (9), the order of the rules is reversed, just like in the counterfeeding derivation in (8), and then the correct outputs are obtained:

(9) /A/ → [B] > /B/ → [C] > /C/ → [D] (push chain)

UR	/A/	/B/	/C/
/C/ → [D]			D
/B/ → [C]		C	
/A/ → [B]	B		
SF	[B]	[C]	[D]

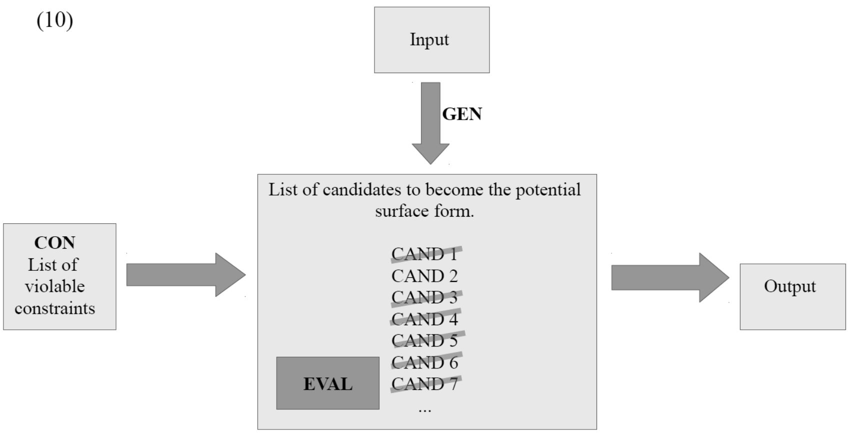
In (7), the derivation followed the order of the considered chain shift and the correct outputs were not obtained. In (9), the rules are reversed in order to obtain the correct outputs. Even though the correct outputs are obtained, the links between the different stages of the chain shift are broken. Because of the broken links, the explanatory power is lost; the three changes are now isolated facts.

Despite all its merits, derivational theory has proved unsatisfactory to account for chain shifts because in order to get the correct outputs, the links between the different stages must be destroyed. In the next section, we will see if Optimality Theory (OT) is a better alternative to account for chain shifts. It will be shown that it is not.

3. Optimality theory in chain shifts

OT is an example of an output-driven theory. The standard version was proposed by Prince & Smolensky (1993). OT was a major and sudden change in phonological theories. The focus now is on the output. In such a theory, there is no derivation. Instead, universal and violable constraints are applied on a set of candidates generated from the input. As just mentioned, the constraints are universal and in any OT analysis, they are ranked. The ranking is language-specific. The theory states that candidates generated from the input will be blocked during evaluation due to constraint violation. After the evaluation of the candidates, one of them is designated as the optimal form, or output. A brief sketch of OT is given in (10):

(10)



We start from an input. Next, the component GEN (=generator) generates a set of candidates. The component CON (=constraints) is a set of violable constraints, ranked for the language in question. The constraints are violable as well as universal. The ranking is language-specific. The component EVAL (=evaluation) applies the constraints to the set of candidates. Some candidates will violate high-ranked constraints in the hierarchy and these candidates will be blocked from becoming the optimal candidate. The evaluation goes per level and the satisfaction of a low-ranked constraint does not compensate the violation of a high-ranked constraint. Every time a candidate violates a constraint, the latter assigns a violation mark to the candidate. A candidate might violate a constraint several times, receiving then several violation marks. However, it is important to remember that the relevance of the number of violations is relative. If, at the end of the evaluation, there are still two competing candidates, if this competition is on the same constraint then the number of violations will be relevant. The candidate with the lowest number of violations will win. If this final competition concerns two different constraints in the hierarchy, then the ranking will decide the outcome of the competition and not the number of violations. The candidate which violates the higher-ranked constraint will lose, even if it has less violation marks than the other candidate. After the evaluation, the candidate which violates low-ranked constraints is designated as the optimal candidate. A schematic example of an OT analysis is presented in (11):

(11)

Input	CON 1	CON 2	CON 3	CON 4	CON 5
CAND 1	*!				
☞ CAND 2			*	*	
CAND 3		*!	*		
CAND 4			*	**!	*
CAND 5		*!	*	*	*

* = violation mark
! = fatal violation
☞ = optimal candidate chosen

In (11), CAND 1 is blocked, for it violates the highest constraint in the hierarchy, CON 1. Next, at the level of CON 2, CAND 3 and CAND 5 violate this constraint and thus are both blocked. At this stage of the analysis, there are two candidates left, CAND 2 and CAND 4. At the level of CON 3, both CAND 2 and CAND 4 violate this constraint. If the number of violation marks is different, then CON 3 will be decisive for the evaluation. However, here, it is not, since both candidates have one violation mark. CON 4 is decisive. Both CAND 2 and CAND 4 violate this constraint but this time, the number of violations is different. CAND 4 has more violation marks than CAND 2 and therefore is blocked. After the evaluation, CAND 2 is designated as the optimal candidate.

Several weaknesses of OT have been described.[4] The main weakness concerns the opacity issue. Opacity is not satisfactorily dealt with within OT. In derivational theories, opacity is represented with intermediate levels. However in OT, there are only two levels of representation, the input level and the output level. As illustrated in (10), OT has several components − GEN, CON and EVAL − but these components must not be mistaken for equivalents of the intermediate levels. Contrary to a rule-based derivation, in OT the application of the components does not create intermediate representations. All the candidates are evaluated at the same time. Consequently, even a simple chain shift like /A/ → [B] > /B/ → [C] is problematic for OT. B cannot be both optimal as it has to be in /A/ → [B] and non-optimal as it has to be in /B/ → [C]. This is shown in (12):

(12) /A/ → [B] > /B/ → [C]

a. /A/ → [B]

/A/	*/A/	*/B/	*/C/
[A]	*!		
[B]		*!	
☹ [C]			*

b. /B/ → [C]

/A/	*/A/	*/B/	*/C/
[A]	*!		
[B]		*!	
☞ [C]			*

In (12b), the correct output is obtained, but not in (12a), as indicated by the symbol ☹. Prince & Smolensky's (1993) version of OT cannot block the /A/ → [C]

[4] See *The Linguistic Review*, vol.17:2-4 for some articles about these weaknesses.

mapping. As seen before, in derivational theories, the opacity issue can be tackled by modifying the rule order but it leads to a highly unsatisfactory analysis. The question raised is whether this issue can be tackled here by modifying the hierarchy of the constraints. This is shown in (13):

(13) /A/ → [B] > /B/ → [C]

a. /A/ → [B]

/A/	*/C/	*/A/	*/B/
[A]		*!	
☞ [B]			*
[C]	*!		

b. /B/ → [C]

/A/	*/C/	*/A/	*/B/
[A]		*!	
☹ [B]			*
[C]	*!		

With the modified constraint hierarchy, [B] is now the optimal candidate in (13a). However, in (13b), [C] is not the optimal candidate anymore under this new constraint hierarchy. Moreover, in (12) and (13), as previously observed in derivational theories, the links between the stages of the chain shift are broken. The two stages of the chain shift are analysed separately. In order to have a relevant analysis of chain shifts in output-driven theories, one must come up with a unified analysis of a chain shift. It suggests that the whole chain shift has to be analysed in the same tableau. If an analysis of GEVS followed what was seen in (12), we would end up with nonsensical mappings like /aː/ → [ai] or /ɔː/ → [au]. As a whole, with respect to chain shifts, OT faces the same difficulties as derivational theories, at least in the case of a counterfeeding order. Some modifications have since been proposed in order to tackle the issue of opacity. Some studies of GEVS were made to illustrate the supposed efficiency of these modifications. In the next section, three studies of GEVS in modified OT will be reviewed. These studies will be proved not to be wholly satisfactory in accounting for GEVS and consequently for chain shifts in general.

4. GEVS analyses in modified OT

This section presents three analyses which tried to modify OT to tackle the weaknesses described in the previous section. Section 4.1 presents Miglio & Moren's (2003) analysis of GEVS using *complex constraints*. Section 4.2 presents Lee's (2004) analysis using the concept of *distantial faithfulness*. Section 4.3 presents Ahn's (2002) analysis based on *Dispersion Theory* and '*Maintain Contrast*' constraint. It will be explained why these analyses prove unsatisfactory in accounting for GEVS. As for Ahn's analysis, it will be shown that it is not totally unsatisfactory, nevertheless some improvements can be done.

4.1 Miglio & Moren (2003): A first step in OT analysis of GEVS

In their analysis, Miglio & Moren (2003) distinguish three stages, the second of which deals with the chain shift observed in GEVS. There are two things to note in their analysis: (1) they do not include the /aː/ → [ɛː] shift, but rather start with the /ɛː/ → [eː] shift; (2) their underlying representations are always short vowels. In their analysis, there is no /ɛː/ → [eː] shift but rather one involving /ɛ/ → [eː], and this is the same for the other vowel shifts of GEVS. A last thing to mention before starting with Miglio & Moren's analysis is the way they consider vowels in terms of features. They use privative features, while, for this paper, binary values are applied. For example, under Miglio & Moren's analysis, high vowels are specified as [high], mid vowels as [high, low], and low vowels as [low].[5] Consequently, for section only, the privative approach to features will be used to specify the vowels.

4.1.1 Miglio & Moren's analysis of the /ɛː/ → [eː] stage[6]

Starting with the analysis of /ɛː/ → [eː], they use the following constraints described in (14):

---

5 With binary values, for instance, high vowels would be specified as [+high, −low], the presence and the absence of the features are indicated by the symbols '+' and '−'. However, under the privative approach to features, only the present features are indicated, that is to say those which are associated with the symbol '+' under the equipollent approach. Consequently, that's why, under the privative approach, high vowels are specified as [high].

6 According to Miglio & Moren, /e/ and /ɛ/ are [high, low].

(14)

FootBinary (FtBin)	Feet must be binary at either the mora or syllable level (Prince & Smolensky 1993)
DepLink-Mora [SEG]	Do not add morae [7] to a particular type of segment that it did not have underlyingly.
Iden[RTR]	Each segment of the output must share the feature [RTR] with their corresponding segment in the input.

FtBin must be high in the hierarchy because it blocks the transformation of a long vowel into a short vowel. Between the two DepLink-Mora constraints, DepLink-Mora [RTR, LOW, HIGH] must be higher than DepLink-Mora [HIGH, LOW] to block candidate [ɛː]. Iden[RTR] is designated as the lowest constraint for the analysis of the /ɛː/ → [eː] stage given in (15a):

(15) a.

/ɛ/	FtBin	DepLink-Mora [RTR, LOW, HIGH]	DepLink-Mora [HIGH, LOW]	Iden[RTR]
[ɛː]		*!		
☞ [eː]			*	*
[ɛ]	*!			
[e]	*!			*

Both [e] and [ɛ] are blocked by the highest ranked constraint since they are short vowels. [ɛː] receives a violation mark from DepLink-Mora [RTR, LOW, HIGH] because [ɛː] and /ɛ/ share the same features specified within the brackets: they are both [RTR, high, low], and furthermore there is a moraic difference between these two vowels, /ɛ/ is short while [ɛː] is long. For the same reasons, [eː] receives a violation mark from DepLink-Mora [HIGH, LOW] since, like /ɛ/, [eː] is [high, low], and there is once again a moraic difference. [eː] receives its first violation mark on the third constraint in the hierarchy. However, since the other candidates violate higher ranked constraints, [eː] is the optimal candidate.

In this analysis, the problem is to account for the fact that /ɛ/ does not become [iː]. This is really problematic because if the candidate [iː] is added in (15), the undesired output is obtained as shown in (15b):

7 A mora is a weight element of the syllable. One mora is equivalent to a short syllable.

(15) b.

/ɛ/	FtBin	DepLink-Mora [RTR, LOW, HIGH]	DepLink-Mora [HIGH, LOW]	Iden[RTR]
[ɛː]		*!		
[eː]			*!	*
☹ [iː]				*
[ɛ]	*!			
[e]	*!			*

If these constraints are kept, [iː] would become the optimal candidate. However, at this stage of GEVS, [iː] is not the correct output. A possible solution would be to add a constraint DepLink-Mora [HIGH] in order to target [iː]. The problem is that DepLink-Mora [HIGH] is lower ranked than DepLink-Mora [HIGH, LOW] in the hierarchy. This would lead to the analysis in (16):

(16)

/ɛ/	FtBin	DepLink-Mora [RTR, LOW, HIGH]	DepLink-Mora [HIGH, LOW]	DepLink-Mora [HIGH]	Iden[RTR]
[ɛː]		*!			
[e]			*!		*
☞ [iː]				*	*
[ɛ]	*!				
[e]	*!				*

[iː] receives another violation mark but it has no consequences on the output selection of the analysis. The violated added constraint is low in the hierarchy. [iː] remains the optimal candidate. Miglio (1999) proposes a local conjunction constraint, Iden[RTR] & Iden[LOW] [8] to tackle this issue. Such a complex constraint is violated when both of its constraints are violated by the same segment of a candidate. Adding this complex constraint and consequently also Iden[LOW], Miglio sets the tableau in (17):

[8] Miglio & Moren define /ɛ/ as [low].

(17)

/ɛ/	FtBin	Iden[RTR] & Iden[LOW]	DepLink-Mora [RTR, LOW, HIGH]	DepLink-Mora [HIGH, LOW]	DepLink-Mora [HIGH]	Iden [RTR]	Iden [LOW]
[ɛː]			*!				
☞ [eː]				*		*	
[iː]		*!			*	*	*
[ɛ]	*!						
[e]	*!					*	

Candidate [iː] violates both of the constraints composing the complex constraint: /ɛ/ is [RTR, high, low] but [iː] is [high] and since Iden[RTR] & Iden[LOW] is higher ranked than DepLink-Mora [RTR, LOW, HIGH], [iː] is blocked and [eː] is the optimal candidate. Given the analysis of this stage of the chain shift, we can move onto the next stage.

4.1.2 Miglio & Moren's analysis of the /eː/ → [iː] stage

Miglio & Moren maintain the constraints and ranking seen in the previously and thus provide the analysis in (18a):

(18) a.

/e/	FtBin	DepLink-Mora [HIGH, LOW]	DepLink-Mora [HIGH]
[eː]		*!	
☞ [iː]			*

Next, they say that they have to find a constraint to block a reverse movement of the shift, namely the fact that /e/ → [aː]. They propose a constraint Iden[HIGH].[9] Adding the candidate [aː], the analysis shown in (18b) is obtained:

9 Miglio & Moren define /e/ as [high, low].

(18) b.

/e/	FtBin	Iden[HIGH]	DepLink-Mora [HIGH, LOW]	DepLink-Mora [HIGH]
[a ː]		*!		
[e ː]			*!	
☞ [i ː]				*

In (18b) [i ː] is chosen as the optimal candidate. Now, the last stage /i ː/ → [aɪ] will be examined.

4.1.3 Miglio & Moren's analysis of the /i ː/ → [aɪ] stage

In their analysis of this stage of the shift, Miglio & Moren add the faithfulness constraint defined in (19):

(19)

Integrity	"No Breaking" = No element of S1 has multiple correspondents in S2. As a whole, the role of this constraint is to block diphthongisation.

Their analysis with this new constraint is given in (20):

(20)

/i/	FtBin	DepLink-Mora [HIGH]	Integrity [HIGH]
[i ː]		*!	
☞ [aɪ]			*

The constraint Integrity[HIGH] adds more precision to the analysis: it has no real consequences for the analysis. It simply states that the optimal candidate violates a faithfulness constraint lower ranked than the other constraints violated by other candidates. Now that Miglio & Moren's analysis has been summarised, its main drawback can be examined.

Each stage of the chain shift is analysed separately and the links between the different parts of the chain shift are lost. This destroys the explanatory power of the analysis. All in all, this analysis does not shed any new light on the opacity issue in OT but it has the merit of providing one of the first OT analyses of a historical chain shift. In the next section, another OT analysis of GEVS, Lee (2004), will be reviewed. It will be proved that once again, the opacity issue is not satisfactorily dealt with, but the analysis provides another view of the forces that drove GEVS.

4.2 Lee (2004): A different view of GEVS

Lee's analysis of GEVS is based on the use of distantial faithfulness constraints. To fully understand what 'distantial faithfulness' means, one must look to Kirchner (1995, 1996). Kirchner provides a study of the Western Basque Hiatus Raising and attributes this phenomenon to the following constraint given in (21):

(21) Hiatus Raising: In a hiatus V_1V_2, maximise height of V_1.

He then shows that the feature system can provide the four-way height distinction found in Etxarri Basque. He says that the relation “higher than” can be evaluated over the values of the height features. So violations of Hiatus Raising can be addressed in a scalar manner. This is presented in (22):

(22) “>” = higher than
+raised > −raised
+high > −high
−low > +low

	low	high	raised
i^y, u^w	−	+	+
i, u	−	+	−
e, o	−	−	−
a	+	−	−

If a phonetic scale based on features is assumed, the output cannot be more than a certain “distance” from its input value along that scale. In his study of Etxarri Basque, Kirchner uses the distantial constraint given in (23):

(23) V-Height Distance ≤ 1 (initial formulation): The output may not be a distance > 1 from the input value with respect to vowel height.

Since this constraint seems to work well for the Basque data, Lee (2004:104) tested it on GEVS. Lee shows that without the distantial faithfulness constraint, the correct output cannot be obtained. This first analysis is given in (24) as presented in the article (apart from the symbols used to designate the chosen candidate):

(24)

	FtBin	$\text{Parse}_{\text{low}}$	Hiatus Raising	$\text{Parse}_{\text{high}}$	$\text{Parse}_{\text{raised}}$
ɛ → aː		*!	***		
ɛ → ɛː			**!		
ɛ → eː			*!		
☹ ɛ → iː				*	*
e → aː		*!	***		
e → ɛː			**!		
e → eː			*!		
☞ e → iː				*	*

Parse constraints are used to block a candidate that shares the feature specified in these constraints. So $\text{Parse}_{\text{low}}$ blocks 'ɛ → aː' because [aː] is [+low]. Hiatus Raising, which is at the core of this analysis, blocks the other two less-optimal candidates. In order to understand how this constraint functions, one has to see what "maximising the height of the output" - seen in (21) - means. There is a four-level vocalic height system, /aː/ the lowest, followed by /ɛː/, then /eː/ and finally /iː/, the highest in this system. Given Hiatus Raising, a candidate receives a violation mark for every vocalic height level higher than its own vocalic height level. [eː] receives only one violation mark because there is only one higher level, the /iː/ level. [ɛː] receives two violation marks because there are two higher levels, /eː/ and /iː/. [aː] receives three violation marks because there are three higher levels, /ɛː/, /eː/ and /iː/. [iː] does not receive any violation mark because in the considered system, it cannot go higher.

Without the distantial faithfulness constraint, even though the correct output is obtained for the second part of the analysis, the optimal candidate should have been 'ε → eː', and not 'ε → iː'.

In (25), the distantial faithfulness constraint is added to the analysis and the Parse constraints are replaced by Iden constraints.

(25)

	FtBin	Iden [LOW]	V-Height Distance ≤ 1	Hiatus Raising	Iden [HIGH]	Iden [RAISED]
ε → aː		*!		***		
ε → ɛː				**!		
☞ ε → eː				*		
ε → iː			*!		*	*
e → aː		*!	*	***		
e → ɛː				**!		
e → eː				*!		
☞ e → iː					*	*

The distantial faithfulness constraint seems to be a good solution. 'ε → eː' is now the optimal candidate in the first part of the analysis. 'ε → iː' is blocked by this constraint because /ε/ and [iː] differ in more than one feature: /ε/ is [−high, −low, −tense] and [iː] is [+high, −low, +tense]. In the second part of the analysis, 'e → aː' receives one violation mark for the same reasons: /e/ is [−high, −low, +tense] and [aː] is [−high, +low, −tense].

Lee's analysis is more unified than Miglio & Moren's since it relies on one notion, distantial faithfulness, and one ranking of constraints, at least for part of the front vowel chain shift. Nevertheless, distantial faithfulness has some limitations. It seems to be applicable more for vocalic chain shifts because vowels can be organised on a scale and so follow the distantial faithfulness principle. However, distantial faithfulness might face some problems with consonants since the latter are more categorical than scalar. Consonants can be organised on a scale, for example the sonority scale, but the latter is harder to handle for distantial faithfulness than the scales of height and backness on which the vowels are organised. Another point is that, as presented in the article, Lee's analysis is not complete since it only deals with part of GEVS. The analysis does not deal with the whole chain shift of front vowels and there is nothing said about the back vowels.

However, the concept of distantial faithfulness remains interesting. It is based on the proximity of the input and the output in terms of features. It provides a more insightful way to analyse chain shifts than Miglio & Moren's approach. This will be linked to what will be seen in section 5, where CPT is treated and provides a more satisfactory analysis of GEVS. But, before moving on to the section dealing with CPT approach and GEVS analysis under this framework, a last GEVS OT-analysis must be considered, for this one has some connections with what will be discussed in section 5.

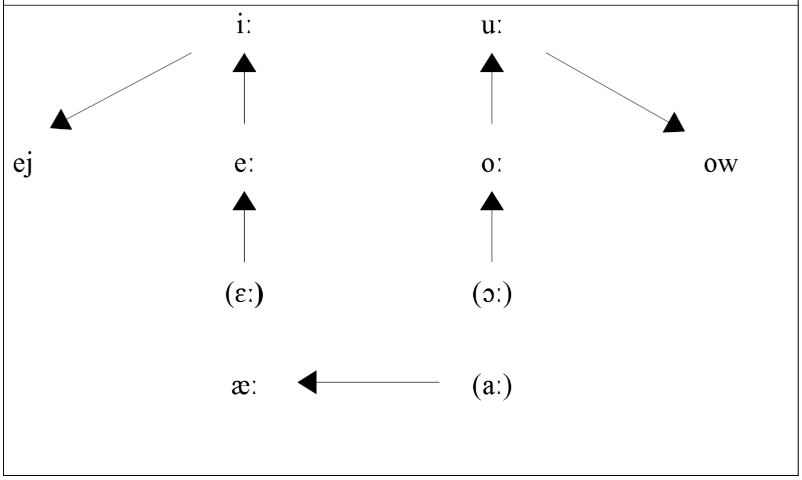
4.3 Ahn (2002): A first step into a contrast preservation approach[10]

Even if this analysis is older than the other two, which had been discussed in sections and , this is the closest to what will be dealt with in section . This analysis is based on *Dispersion Theory* (Flemming 1995, 1996) and OT. These two different frameworks are associated in order to show that "[...] the well-formedness of the vowel system cannot be evaluated in isolation, the overall result is obtained by the pattern evaluation of the adjacent vowels." (Ahn 2002:156). We will see that this analysis offers a very insightful way in how to account for diachronic chain shifts. Nevertheless, it will also be suggested that we can obtain a more thorough analysis of GEVS through CPT.

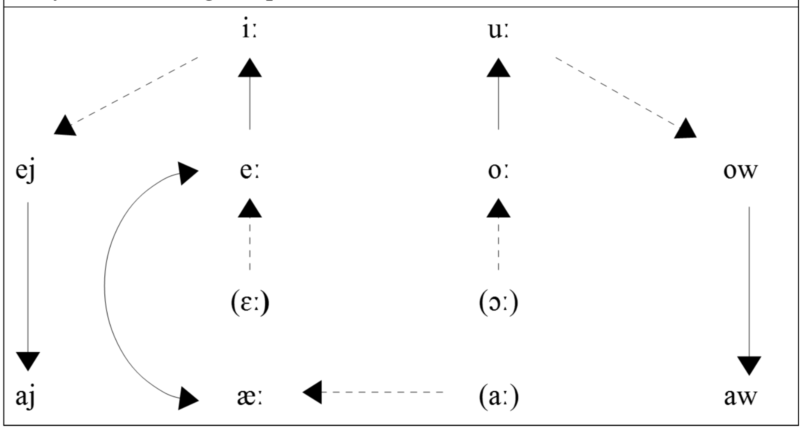
In Ahn's analysis, GEVS is split into two main periods. There is a series of changes that occurred during the Shakespearean era and another one during the Early Modern English period, after the former. (26a) and (26b) are adapted from Ahn's (2002:159, 167) description of the two phenomena. The dotted arrows in (26b) indicate the changes that had already occurred during the Shakespearean era.

10 I would like to thank the anonymous reviewer who gave me the reference of this article. I did not come across it during my research and I did find it quite insightful.

(26a) Shakespearean era [11]



(26b) Early Modern English period [12]



Ahn starts with the analysis of the chain shift during the Shakespearean era. Based on universal phonetic tendencies, the raising of mid-vowels [ɛː] and [ɔː] is assumed to be the triggering event of the chain shift. Through a standard OT analysis, Ahn (2002:159-162) accounts for the raising of [ɔː]. (28) gives the final tableau of this analysis and (27) the constraints used in it.

---

11 Ahn uses this symbol to refer to the low back vowel.

12 The way Ahn represents the diphthongs will be kept as such only for this section.

(27) *Long[lax]: A long vowel may not be lax.

*aː: The long low back vowel is not allowed.

Ident(back): The backness of the input vowel is identical to that of the output.

Ident(high): The height of the input vowel is identical to that of the output.

(28)

/ɔː/	*Long[lax]	*aː	Ident(back)	Ident(high)
ɔː	*!			
aː		*!		*
☞ oː				*
uː				**!
eː			*!	*

Every constraint has a specific objective, everything aiming at the same goal, which is the proper raising of [ɔː]. *Long[lax] triggers the change, since, according to Ahn, [ɔː] is a lax vowel. That's why the candidate [ɔː] is blocked. *aː has two functions: it enables the [aː]-fronting and blocks [ɔː] from lowering to [aː]. The latter function explains why the candidate [aː] is blocked. Ident(back) prevent [ɔː] from fronting to [eː], thus accounting for the fact that the candidate [eː] is blocked. Ident(high) is here to limit the raising of [ɔː]. In (28), both of candidates [oː] and [uː] violate this constraint. Nevertheless, there is a difference in the violation degree. Candidate [oː] violates Ident(high) only once since there is only one height-level difference between /ɔː/ and [oː]. However, candidate [uː] violates Ident(high) twice since there are two height-level differences between /ɔː/ and [uː]. In the end, [oː] is designated as the optimal candidate. (28) shows that the raising of [ɔː], the assumed triggering event, can be accounted for through a ranking of violable constraints. Nevertheless, Ahn shows that this ranking cannot be applied to the front vowels. This is shown in (29):

(29)

/ɛː/	*Long[lax]	*aː	Ident(back)	Ident(high)
ɛː	*!			
aː		*!	*	*
iː				**!
? eː				*
? æː				*

If the same ranking as used in (28) is applied in (29), there is a tie between two candidates. Normally, the correct output would be [eː]. Based on this observation, Ahn (2002:162) suggests that "we [⋯] need to consider the whole paradigm of vowel shift, rather than each individual vowel". As a consequence, Ahn proposes a new constraint that would imply input-output correspondence in the considered vocalic system. This constraint is described in (30).

(30) Maintain Contrast: Maintain input contrasts between adjacent vowels in the output.

This 'Maintain Contrast' constraint aims at maintaining any input contrast in the output. Actually, its main function is to avoid any possible neutralisation. So, if one candidate displays input contrast neutralisation, for every occurrence of it, the candidate will receive one violation mark. Ahn (2002:163) adds this new constraint to the ranking applied before and shows the issues previously encountered with the front vowels is solved. This is shown in (31).

(31)

/eː, ɛː, aː/	*Long[lax]	Maintain Contrast	*aː	Ident(back)	Ident(high)
eː, ɛː, aː	*!		*		
eː, æː, aː			*!		*
eː, eː, æː		*!			*
☞iː, eː, æː				*	**
eː, æː, æː		*!		*	*

As shown in (31), candidate 'eː, eː, æː' is blocked because the contrast between /eː/ and /ɛː/ is neutralised. More importantly, candidate 'eː, æː, æː' is now blocked as well. Once again, we do realise the importance of the 'Maintain Contrast' constraint. Without it, the analysis would have ended up with a wrong output, which is 'eː, eː, æː'. In such a situation, the supposedly optimal candidate 'iː, eː, æː' would have been blocked by the Ident(back) constraint. Ahn (2002:164) also shows that the ranking used in (31) can be applied to the back vowels.

The last elements that need to be included in this analysis to have the whole GEVS are the diphthongs. In order to do so, Ahn (2002:164,165) introduces two other constraints. The latter are given in (32):

(32) Ident-IO(μ): The mora count of the input should remain identical in the output.

Minimum Distance (MinDis): The difference in height between the input vowel and the output one should be kept minimum, i.e., less than 2 steps: MinDis $\leq$ 1.

These two constraints are mainly focused on the input hight vowel change. Ident-IO(μ) makes sure that the long high vowel will not shorten. MinDis is focused on the diphthongisation process by limiting [iː] transformation. In other words, MinDis makes sure that, during its diphthongisation, [iː] will not go too low in the vocalic system. (33) adapts Ahn's analysis of GEVS for the front vowels, including in the ranking the two new constraints seen in (32).

(33)

/iː, eː, ɛː, aː/	*Long[lax]	Ident-IO (μ)	MinDis	Maintain Contrast	*aː	Ident (back)	Ident (high)
iː, eː, ɛː, aː	*!				*		
i, eː, æː, æː		*!		*		*	*
aj, iː, eː, æː			*!			*	***
☞ej, iː, eː, æː						*	***

Candidate 'i, eː, æː, æː' is blocked by Ident-IO(μ) because /iː/ is shortened to [i]. Candidate 'aj, iː, eː, æː' is blocked by MinDis because /iː/ goes to far. It goes through the mid-close and mid-open vocalic level to go directly to the low vocalic level and becomes the diphthong [aj]. However, candidate 'ej, iː, eː, æː' does not violate MinDis, since, in its diphthongisation, /iː/ only goes one vocalic level down, that is to say to the mid-close vocalic level, ending up as the diphthong [ej]. Using this ranking, Ahn managed to account for GEVS for the front vowels. This ranking can also be applied to back vowels.

Then, Ahn moves on to the analysis of GEVS during the second period established at the beginning the article, the Early Modern English period. Recalling what was seen in (26b), one of the main differences with the Shakespearean era is the diphthongisation, which now produces [aj] and [aw]. For this part of the analysis, Ahn introduces a new constraint *Long[low] described in (34).

(34) *Long[low]: Long vowels may not be low.[13]

The objective of this constraint is to prevent high vowels from becoming low vowels, and thus leaving the path to the diphthongisation. So there will not be /iː/ → [æː] or /uː/ → [aː], but /iː/ → [aj] and /uː/ → [aw]. In order to account for this new stage in GEVS, Ahn makes some modifications in the constraint ranking that was applied before. (35) recalls it and (36) shows the modifications made.

(35)

Constraint ranking for the Shakespearean era:
*Long[lax] >> Ident-IO(μ) >> MinDis >> Maintain Contrast >> *aː >> Ident(back) >> Ident(high)

(36)

Constraint ranking for the Early Modern English era:
*Long[lax] >> Ident-IO(μ) >> **Maintain Contrast >> *Long[low] >> MinDis** >> *aː >> Ident(back) >> Ident(high)

[13] The constraint *aː is changed into this more generalised constraint *Long[low] in order to account for both front and back vowels.

Based on (36), there are two main modifications. 'Maintain Contrast' now dominates MinDis and the new constraint *Long[low] is inserted between the latter. With this new ranking, Ahn analyses GEVS for the front vowels. Ahn's analysis is adapted in (37):

(37)

/iː, eː, æː/	*Long [lax]	Ident-IO (µ)	Maintain Contrast	*Long [low]	MinDis	Ident (back)	Ident (high)
iː, eː, ɛː	*!						*
i, iː, eː		*!	*				*
æː, iː, eː				*!	*		***
☞ aj, iː, eː					*	*	***
ej, iː, eː			*!				***

Candidate 'iː, eː, ɛː' is blocked because there is − as considered by Ahn − the lax vowel [ɛː]. Ident-IO(µ) blocks candidate 'i, iː, eː' because /iː/ is shortened to [i]. This time, candidate 'ej, iː, eː' is blocked by 'Maintain Contrast' since the contrast between /iː/ and /æː/ in neutralised.[14] Finally, candidate 'æː, iː, eː' is blocked by *Long[low] because [æː] is [+low]. So, in the end, candidate 'aj, iː, eː' is designated as the optimal candidate. Once again, this ranking can also be applied to the back vowels.

Ahn's analysis offers something quite original with the 'Maintain Contrast' constraint. It gives some insights in what operates GEVS from the inside. Just like Jespersen (1909:232) wrote (as quoted in Ahn 2002:163): “the changes of the single vowels cannot be considered separately; they are evidently parts of one great linguistic movement, which affected all words containing a long vowel in ME [Middle English]”. Indeed, Ahn shows that a certain preservation of contrast is at the core of GEVS. Definitely, Ahn's analysis is more satisfactory than Miglio & Moren's for it has some true explanatory force. Furthermore, Ahn's analysis is more satisfactory than Lee's because, contrary to what we suggested about the notion of *distantial faithfulness*, Ahn's approach can be applied to consonants. For instance, Ahn (2003:7-16) shows that the 'Maintain Contrast' constraint is also relevant in accounting for consonantal chain shifts.

14 Ahn (2002:168) explains this neutralisation as such: “Note that the vowel /e/ in /ej/ is not distinctive to the output /eː/ (< /æː/). In other words, the input vowel contrast in /iː, æː/ is neutralised as [eː] (in the output /ej/ and /eː/).”

Nevertheless, some minor limits can also be pinpointed. For example, under this approach, Ahn needs to separate the GEVS into two different periods. This surely justifies the fact that two different constraint rankings are needed. Nonetheless, this damages a little the unity of the analysis. We will see in the next section that we only need one constraint ranking in order to account for the whole GEVS, both for the front and back vowels. Secondly, the fact that maintaining the contrast between the vowels in the output is definitely something most relevant when one deals with chain shifts. Nevertheless, Ahn does not go deeply enough into this notion of preservation of contrast. The analysis does not say much about how this preservation of contrast functions nor how it influences the proceedings of the chain shift. In the next section, Łubowicz's CPT will be shown to be a more detailed and thorough approach to contrast preservation. It will be shown how contrast preservation actually functions and how it can be considered to be the engine of the whole chain shift.

5. GEVS analysis under Contrast Preservation Theory (CPT)

5.1 Introduction to CPT

An alternative for the analysis of chain shifts was developed by Łubowicz (2003, 2012) with *Contrast Preservation Theory*.[15] Łubowicz used this approach to account for synchronic chain shift. In this section, CPT will be applied to a diachronic chain shift, GEVS. This will constitute another test of CPT on diachronic data along Montreuil (2006) and Noske (2012). CPT revolves around three main concepts: a finite list of *scenarios*, the concept of *contrast* and *preserve contrast constraints*. Before providing an analysis of GEVS under CPT, a description of these three concepts will be given, starting with *scenarios*.

In CPT, there are no individual candidates. Instead, *scenarios* are evaluated. There are four possible *scenarios* (Łubowicz 2003:7-9; Montreuil 2006:112-113) described with schematic examples in (38):

---

15 As far as I know, CPT has only been applied to opaque phenomena. In her PhD thesis, Łubowicz (2003) deals with opaque synchronic chain shifts. Later, Łubowicz (2012) deals with some other opaque phenomena that can be observed in phonological and morphological contrasts.

(38) Identity:

In an identity scenario, there is no contrast difference between the input and the output. All the contrasts are preserved.

/A, B, C/ → [A, B, C]

Transparency:

In a transparency scenario, we lose one of the possible contrasts.

/A, B, C/ → [A, A, C] or [B, B, C]

Here we lose the contrast between A and B and consequently there are two possible outputs.

/A, B, C/ → [A, B, B] or [A, C, C]

Here we lose the contrast between B and C and consequently there are two possible outputs.

Fusion:

In a fusion scenario, all the contrasts are lost.

/A, B, C/ → [C, C, C]

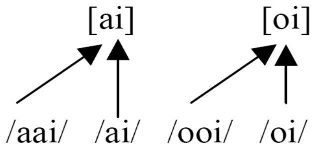
Opaque:

In an opaque scenario, we find the chain effect reaction of underapplication or overapplication. As we will see in (41), minimal contrast is preserved between the input and the output. Furthermore, the same number of contrasts exists before and after but the substance is different.

/A, B, C/ → [B, C, D]

In the identity scenario, there are no contrast differences between the input and the output. All levels of contrast observed in the input are kept identical in the output, without any change of the substance. In the transparency and opaque scenarios, there are contrast differences between the input and the output. In the transparency scenario, as seen in (38), the input-output contrast differences are based on contrast loss through mergers. For example, there is a merger between two neighbouring elements of the input that map onto the same output. This leads to a loss of contrast. It is important to bear in mind that in every transparency scenario, there is only one lost contrast. But, it is equally important to note that such contrast loss can lead to several mergers. Łubowicz (2003:8) gives the transparent scenario from Finnish reproduced in (39):

(39) Finnish transparent scenario:



In (39), only the length contrast is lost but this leads to two mergers. This demonstrates that in a transparent scenario, only one contrast is lost but this loss can lead to several mergers. The opaque scenario is different from the transparency scenario in the sense that there is no real loss of contrast. As discussed above, in such a scenario, one is more likely to observe contrast transformation as described in the Grimm's Law example in (41). The same number of contrast levels is maintained, but every level of contrast will move to another feature. Finally, there is the fusion scenario in which all the contrasts are lost. There is a total merger of all elements in the input to a single output. Among these four scenarios, the opaque scenario is the most interesting for this study since it is the scenario associated with chain shifts. For example, the schematic example given in (38) with this opaque scenario resembles what was presented in (1). Given this introduction of the concept of *scenarios*, now the concept of *contrast* will be described.

Łubowicz (2003:18) gives the following definition of contrast: “a pair of inputs, in_a and in_b, contrast in P [a phonological property], when corresponding segments in those inputs, seg_a and seg_b, are such that seg_a has P and seg_b lacks P.” For example, /d/ and /t/ are contrastive on the [voice] feature since /d/ is [+voice] and /t/ is [−voice].

To illustrate this, we will use the example of Grimm's Law [16] as described in (40):

16 As for Grimm's Law, the traditional view of this phenomenon, with plains plosives for Proto-Indo-European, is chosen for this article (see Brugmann & Delbrück 1906-1917). There are however other approaches such as the glottalic theory (see Hopper 1973 and Gamkrelidze & Ivanov 1995, among others) in which the Proto-Indo-European plosives are not plain but ejective.

(40) Grimm's Law
D^h → D > D → T > T → Þ [17]

a. D^h → D = /bh, dh, gh, g^{wh}/ → [b, d, g, g^w]
b. D → T = /b, d, g, g^w/ → [p, t, k, k^w]
c. T → Þ = /p, t, k, k^w/ → [f, Θ, x, x^w]

Grimm's Law was a phonological process that affected the plosives in Proto-Germanic, more precisely during the transition between Proto-Indo-European and Proto-Germanic. It was a three-stage consonantal chain shift. During the first stage, the aspirated voiced plosives became unaspirated voiced plosives (40a). During the second stage, the voiced plosives became voiceless plosives (40b). Finally, during the last stage, the voiceless plosives became voiceless fricatives (40c).

In terms of CPT, 'preservation of contrast' means that the number of contrasts is maintained before and after the shift. In the particular case of chain shifts, Łubowicz is talking about 'contrast transformation'. It means that at every stage of a chain shift, the input and the output will always preserve a minimal contrast based on one different feature. This is similar to what was seen previously with distantial faithfulness. In (41), we take again Grimm's Law shown in (40) and we describe the preservation of contrast:[18]

17 Abstracting their place of articulation, D^h designates all the voiced aspirated plosives, D all the voiced plosives, T all the voiceless plosives and Þ all the voiceless fricatives.
18 Noske (2012) gives an analysis of Grimm's Law under CPT.

(41) Grimm's Law

$D^h \rightarrow D$	>	$D \rightarrow T$	>	$T \rightarrow$ Þ
aspirated voiced plosives /bh, dh, gh, g^{wh}/				
⇩				
unaspirated voiced plosives [b, d, g, g^w]	>	unaspirated **voiced** plosives /b, d, g, g^w/		
		⇩		
		unaspirated **voiceless** plosives [p, t, k, k^w]	>	unaspirated voiceless **plosives** /p, t, k, k^w/
				⇩
				unaspirated voiceless **fricatives** [f, Θ, x, x^w]

At every stage of the chain shift, the contrast is preserved and is associated with a different feature. For the first stage, the contrast is on the aspiration feature. In the second stage, the contrast moves to the feature [voice]. In the last stage, the contrast moves to the feature [cont]. The notion of contrast is the core of CPT.

The last concept needed to be described before starting GEVS analysis is the Preserve Contrast constraints (= PC constraints). Łubowicz distinguishes three categories of PC constraints: input-oriented PC (PC_{IN}(P)), output-oriented PC (PC_{OUT}(P)), and relational PC (PC_{REL}(P)); P stands for a phonological property. Łubowicz's (2003:18, 20) definitions of these PC constraints are given in (42):

(42) Input-oriented PC (PC_{IN}(P)):
"If inputs are distinct in P, they need to remain distinct."

A transparency scenario like /A, B, C/ → [A, A, C] would receive one violation mark because we lose the contrast between /A/ and /B/.
A fusion scenario like /A, B, C/ → [C, C, C] would receive two violation marks because we lose the contrast between /A/ and /B/ and the contrast between /B/ and /C/.

Output-oriented PC (PC_{OUT}(P)):
"Avoid outputs ambiguous in P property."

A transparency scenario like /A, B, C/ → [A, A, C] would receive one violation mark because we have one ambiguous output, [A], which is the result of the merger between /A/ and /B/. [A] can correspond to /A/ or /B/.
A fusion scenario like /A, B, C/ → [C, C, C] would receive one violation mark because we have one ambiguous output, [C]. There is a merger of the three elements in the input, /A, B, C/, to [C].

Relational PC (PC_{REL}(P)):
"If a pair of outputs are minimally distinct in P, then a pair of inputs must be distinct in P as well."[19]

An opaque scenario like /A, B, C/ → [B, C, D] will receive one violation mark because we lose the contrast between /A/ and /B/ but we acquire a new contrast between [C] and [D]. The contrast between /B/ and /C/ of the input is preserved in the output between [B] and [C].

These three types of constraints evaluate contrast. Their aim is to see if the levels of contrast present in the input are preserved, transformed or lost in the output. As with

19 My brief definition of relational PC.

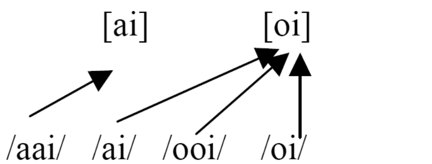
the constraints in an OT analysis, PC constraints are ranked alongside markedness constraints and faithfulness constraints. Every PC constraint has a specific role.

PC_{IN} minimises the loss of contrasts if mergers occur in a scenario. It is important to bear in mind that PC_{IN} admits contrast transformation since in such a case there is no real merger. That is why, as shown in (42), in a transparency scenario like /A, B, C/ → [A, A, C] PC_{IN} gives one violation because there is a merger between /A/ and /B/ to [A]. However, in an opaque scenario like /A, B, C/ → [B, C, D], PC_{IN} gives no violation marks because there is no merger − so no contrast losses − but contrast transformation. The contrast between /A/ and /B/ is transformed into the contrast between [B] and [C].

PC_{OUT} ensures that if mergers occur in a scenario, they affect one particular element of the output rather than several. As such, every ambiguous output results in one violation mark for this constraint. In the transparent scenario presented in (39) for example, a PC_{OUT} based on a length contrast would assign two violation marks because there are two ambiguous outputs for the length contrast, [ai] and [oi].

However, care must be taken, as illustrated by the opaque scenario in (43) (see Łubowicz 2003:8):

(43) <u>Finnish chain-shift scenario</u>:



While this scenario still focuses on the length contrast, PC_{OUT} assigns only one violation mark because the length contrast between /ooi/ and /oi/ is lost since they merged to [oi]. On the other hand, the length contrast between /aai/ and /ai/ is not lost because there is no merger. Instead, the length contrast is transformed to a rounding contrast between [ai] and [oi]. As such, PC_{OUT}(length) does not assign a second violation mark to this scenario.

PC_{REL} ensures the recoverability of the input contrast levels in the output contrast levels. In other words, it prevents contrast transformation, so PC_{REL} targets only opaque scenarios.

Finally, under CPT, markedness constraints are used to block the identity scenario. In an identity scenario like /A, B, C/ → [A, B, C], it suffices to have a markedness constraint like *[A] to block the identity scenario. Now that the required elements of CPT for our analysis have been introduced, GEVS can be analysed under this approach.

5.2 Analysis of GEVS under CPT

Recall that there are two chain shifts:

/aː/ → [ɛː] > /ɛː/ → [eː] > /eː/ → [iː] > /iː→ [aɪ]
and
/ɔː/ → [oː] > /oː/ → [uː] > /uː/ → [au]

In order to establish the PC constraints for this analysis, it is necessary to see which phonological property they will be associated with. In (44), the front and back pre-shift vowels are described in terms of features.[20] As illustrated in (41) for Grimm's Law, the preservation of contrast is also noticeable in (44). For each pair of vowels − /i/ and /e/, /e/ and /ɛ/, /ɛ/ and /a/, /u/ and /o/, /o/ and /ɔ/ −, the two elements differ from one another by only one feature.

[20] In this analysis, we assume that [ɜ, a, ɔ] are [−tense] and [e, o] are [+tense]. There are some – fairly justified – disagreements about the relevance of this feature [tense] in Present-Day English (see Durand 2005 for example). Today, in English – at least in RP-English – there is no longer phonemic contrasts like /e~ɛ/ or /o~ɔ/. But such contrasts did exist in former historical stages of English, as here in Middle English/Early Modern English. Furthermore, in contemporary Dutch, which vowel system can be related to the one of Middle English, [ɛ] and [ɔ] are seen as lax. If one does not adopt the feature [tense] for the considered vowels, this feature can be easily replaced by another one, like [ATR] for example. This kind of modification would not have any consequences on the analysis given later in this section. In such a case, rewriting some constraints would be necessary but, once again, it would not damage the analysis.

(44) The front vowels:

	/i/	/e/	/ɛ/	/a/
[high]	+	-	-	-
[tense]	+	+	-	-
[low]	-	-	-	+

The back vowels:

	/u/	/o/	/ɔ/
[high]	+	-	-
[tense]	+	+	-
[low]	-	-	-

As such, the PC constraints used in this analysis are associated with three features: [high], [tense] (=[tns]) and [low]. Nevertheless, it is not necessary to include every possible constraint involving these three features when evaluating for the optimal scenario.

The constraints used to block the non-optimal scenarios are given in (45):

(45) *V[+low]: no V[+low] vowels in the outputs.

*V[−tns]: no V[−tns] vowels in the outputs.

PC_{IN}(low): if two inputs contrasting in [low] map onto the same output, assign one violation mark.

PC_{IN}(tns): if two inputs contrasting in [tns] map onto the same output, assign one violation mark.

PC_{IN}(high): if two inputs contrasting in [high] map onto the same output, assign one violation mark.

INTEG: no diphthongs in the outputs.

(46) shows the analysis for the front vowels:

(46)	/aː, ɛː, eː, iː/	*V [+low]	PC_{IN} (low)	PC_{IN} (tns)	PC_{IN} (high)	*V [−tns]	INTEG
ID	Identity: [aː, ɛː, eː, iː]	*!				*	
T1	Transparency: [ɛː, ɛː, eː, iː]		*!			**	
T2	Transparency: [aː, eː, eː, iː]	*!		*			
T3	Transparency: [a:, ɛ:, i:, i:]	*!			*	*	
FU	Fusion: [ai, ai, ai, ai]		*!	*	*		****
OP	☞ Opaque: [ɛː, eː, iː, ai]					*	*

ID, T2 and T3 are blocked by *V[+low] because these scenarios have a [+low] vowel, [aː]. T1 and FU are blocked by PC_{IN} (low) because the contrast is lost between /aː/ and /ɛː/ and they both map onto the same output, [ɛː]. In the end, OP is chosen as the most optimal scenario. OP does not violate *V[+low]. The primary aim of this constraint is to block the identity scenario ID, and indirectly it enables the triggering of the first stage of the chain shift /aː/ → [ɛː] since it stipulates that there must not be [aː] in the output. Due to contrast transformation during the chain shift, there are no mergers, and the PC_{IN} constraints are not violated.

As a whole, CPT and its different tools give a unified analysis of the front vowel chain shift. Furthermore, the explanatory force is present through this analysis because of the notion of contrast preservation. (46) demonstrates that the CPT analysis seems to work on the front vowels involved in GEVS. In order to see if CPT is completely satisfactory in accounting for GEVS, it must also be able to account for the back vowel chain shift, which is presented in (47).

(47)	/ɔː, oː, uː/	*V [+low]	PC_{IN} (low)	PC_{IN} (tns)	PC_{IN} (high)	*V [−tns]	INTEG
ID	Identity: [ɔː, oː, uː]					*!	
T1	Transparency: [oː, oː, uː]			*!			
T2	Transparency: [ɔː, uː, uː]				*!	*	
FU	Fusion: [au, au, au]			*!	*		***
OP	☞ Opaque: [oː, uː, au]						*

PC_{IN}(tns) blocks T1 and FU. In these scenarios, the contrast is lost between /ɔː/ and /oː/. PC_{IN}(high) blocks T2, since the contrast is lost between /oː/ and /uː/. In ID, there is a [−tns] vowel in the outputs, [ɔː]. So this scenario is blocked by *V[−tns]. After the evaluation, OP is the optimal scenario. Once again, OP does not violate PC_{IN} constraints since, in this scenario, there are no mergers but contrast transformation. In (47), the triggering of the chain shift, this time, is ensured by the markedness constraint *V[−tns]. All these analogies with the previous analysis in (46) show that CPT is a fully satisfactory approach to account for chain shifts. Another argument for this conclusion is the fact that for both analyses, (46) and (47), the same hierarchy of constraints is used.

These examples show that CPT can both account for chain shifts and offer some insights about the mechanism of sound change with the notion of contrast preservation. Even more importantly, in the case of GEVS and Grimm's Law among others,[21] it has been shown that CPT can also account for historical chain shifts. In the conclusion, some of the implicatures of the account are considered.

21 For example, Montreuil (2006) gives an analysis of a vocalic chain shift in Gallo – a French dialect - under CPT.

6. Conclusions

This paper provides a new account of GEVS under CPT. It was shown that CPT provides a unified analysis of GEVS. Applying the concept of scenarios to GEVS allows all the stages of the chain shift to be accounted for using a single constraint ranking. There is a single candidate in which the whole chain shift is evaluated. Therefore, it is no longer necessary to provide different analyses of the different stages. This analysis thus is unified and insightful, by providing explanatory force using the concept of contrast preservation. Furthermore, it does not have the limitations of distantial faithfulness. CPT can account for both vocalic and consonantal chain shifts, both driven by the need for contrast preservation.

Moreover, it has been shown that even though CPT was initially designed to account for synchronic chain shifts, it can also explain diachronic chain shifts.

This analysis has provided evidence that diachronic data can be used as a secondary test for the validity of modern linguistic theory. Furthermore, it has demonstrated that a theory developed to account for synchronic phenomena can also provide insights on historical sound changes. CPT could also further our understanding of other types of sound changes. For example, it could be interesting in regards to comprehension of sound change in the chain shifts observed in the acquisition of language.

Finally, one interesting consequence of CPT as pointed out by Łubowicz is that pull shifts are not accounted for by CPT. This suggests two things: 1) what appears to be pull shifts do not exist; or 2) pull shifts exist but are driven by other forces. This calls for a deeper study of uncontroversial examples of pull shifts to see what forces are behind them.

References

Ahn, Sang-Cheol. 2002. An optimality approach to the great vowel shift. *Korean Journal of Linguistics* 27(2): 153-170.

Ahn, Sang-Cheol. 2003. Towards an optimal account of diachronic chain shifts: Part I (Grimm's Law). In Jennifer Spenader, Anders Eriksson, and Östen Dahl (eds.), *Variation within optimality theory*, 7-16. Department of Linguistics, Stockholm

University, Sweden.

Baković, Eric. 2011. Opacity and ordering. In John A. Goldsmith, Jason Riggle, and Alan C. L. Yu (eds.), *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishing.

Baugh, Albert C., and Thomas Cable. 2002. *A history of the English language*. Taylor and Francis.

Brugmann, Karl, and Berthold Delbrück. 1906. *Grundriss der vergleichenden grammatik der Indogermanischen sprachen*. Trübner. 5 vols. Strasbourg.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. NewYork; Evanston; London: Harper & Row.

Clements, George N. 2000. In defense of serialism. *The Linguistic Review* 17(2-4): 181-198.

Durand, Jacques. 2005. Tense/lax, the vowel system of English and phonological theory. In Philip Carr, Jacques Durand, and Colin J. Ewen (eds.), *Headhood, elements, specification and contrastivity. Phonological papers in honour of John Anderson*, 77-90. Amsterdam: John Benjamins.

Flemming, Edward. 1996. Evidence for constraints on contrast: The dispersion theory of contrast. *UCLA Working Papers in Phonology* 1: 86 – 106.

Jespersen, Otto. 1909. *A modern English grammar on historical principles*. George Allen & Unwin Ltd.

Kager, René. 1999. *Optimality theory*. Cambridge textbooks in linguistics. Cambridge; New York, NY: Cambridge university Press.

Kenstowicz, Michael J., and Charles W. Kisseberth. 1971. Unmarked bleeding orders. *Studies in the Linguistic Sciences* 1: 8-28.

Kenstowicz, Michael J., and Charles W. Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.

Kenstowicz, Michael J., and Charles W. Kisseberth. 1979. *Generative phonology: Description and theory*. New York: Academic Press.

Kiparsky, Paul. 1982. How abstract is phonology? *Explanation in phonology*, 119-164. Foris Publications, Holland.

Kirchner, Robert. 1996. Synchronic chain shifts in optimality theory. *Linguistic Inquiry* 27(2): 341-350.

Lee, Ju Yong. 2004. The great vowel shift in optimality theory. *Language & Information Society* 6: 84-106.

Lin, Yen-Hwei. 2011. Affricates. In Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice (eds.), *The blackwell companion to phonology*, 1: 367-390. Oxford: Blackwell.

Łubowicz, Anna. 2003. *Contrast preservation in phonological mappings*. PhD Dissertation, Amherst, MA: University of Massachusetts. http://roa.rutgers.edu/files/554-1002/554-1002-LUBOWICZ-10-0.PDF.

Łubowicz, Anna. 2012. *The phonology of contrast*. Equinox Publishing.

McCarthy, John J. 1999. Sympathy and phonological opacity. *Phonology* 16: 331-399.

McCarthy, John J. 2007. *Hidden generalizations: Phonological opacity in optimality theory*. London: Equinox.

Miglio, Viola. 1999. Interactions between markedness and faithfulness constraints in vowel systems. PhD Dissertation, Baltimore: University of Maryland.

Miglio, Viola, and Bruce Moren. 2003. Merger avoidance and lexical reconstruction: An optimality-theoretic model of the great vowel shift. In D. Eric Holt (eds.), *Optimality theory and language change*, 191-228. Dordrecht: Kluwer.

Montreuil, Jean-Pierre. 2006. Contrast preservation theory and historical change. In Randall S. Gess and Deborah Arteaga (eds.), *Historical Romance linguistics: Retrospective and perspectives*, 111-129.

Newton, B. E. 1971. Ordering paradoxes in phonology. *Journal of Linguistics* 7: 31-53.

Noske, Roland. 2012. The Grimm-Verner push chain and contrast preservation theory. In Bert Botma and Roland Noske (eds.), *Phonological explorations: Empirical, theoretical and diachronic issues*.

Prince, Alan, and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar*. Rutgers University: New Brunswick: New Jersey. http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF.

Julien Fulcrand
STL (UMR 8163)
Université Charles-de-Gaulle - Lille 3
3 Rue du Barreau, 59650 Villeneuve-d'Ascq, France
Email: jf.ganon@hotmail.fr

Received: 2015. 09. 11.
Revised: 2015. 12. 14.
Accepted: 2015. 12. 14.