ON SUBCATEGORIZATION AND PRIORITY: EVIDENCE FROM WELSH ALLOMORPHY

Mykel Loren Brinkerhoff

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Linguistics.

Chapel Hill
2019

Approved by:

Jennifer L. Smith

Brian Hsu

Katya Pertsova

ABSTRACT

Mykel Loren Brinkerhoff: On Subcategorization and Priority: Evidence from Welsh Allomorphy
(Under the direction of Jennifer L. Smith)

This thesis examines the phonologically conditioned suppletive allomorphy (PCSA) of the definite article in Welsh and initial consonant mutations. The analysis of these patterns shows that Optimality Theory (Prince & Smolensky 2004 [1993]), with the addition of Lexical Selection's use of priority relationships and lexical subcategorization (Mascaró 2007; Bonet et al. 2007) and Prosodic Subcategorization (Inkelas 1990, 1993; Zec 2005; Bye 2007; Bennett et al. 2018; Tyler 2019), can account for the distribution of the definite article allomorphs and their interaction with the rest of the grammar as well as the behavior of Welsh initial consonant mutations contrary to Hannahs & Tallerman (2006). The analysis further argues for an expansion of prosodic subcategorization to include allomorph-specific subcategorization frames in light of the Welsh definite article. Additionally, this thesis makes the argument against a purely morphological subcategorization approach to phonologically conditioned allomorphy, contrary to the claims of Paster (2009, 2015).

Dedicated to Grant J. Wilson for teaching me to never give up and for teaching and encouraging me to always discover new things about the world.

ACKNOWLEDGEMENTS

This thesis would never have been possible without the help of a large number of people. To name all who have had an impact or who have contributed in any way to this thesis would fill a vast number of tomes. If I inadvertently leave someone out who is deserving of acknowledgment, know it is not out of spite or a slight against you.

First and foremost, I wish to express eternal gratitude to my wife Betsy, without whose love and support I would never have considered going to graduate school and would not have been able to complete this thesis. Additional thanks is owed to my parents, who pushed me to be better and for their support and encouragement.

I wish to express my gratitude and thanks for the countless hours of discussion, assistance, and advise that my advisor, Jennifer L. Smith, has given me during my time here at the University of North Carolina at Chapel Hill. Without her help, I would not be where I am today. I wish to also thank the rest of the members of my thesis committee, Brian Hsu and Katya Pertsova, for the assistance that they have given me as we discussed various aspects of the thesis.

I am indebted to the wonderful members of the UNC P-side research group for the feedback that they gave me as they heard various aspects and parts of this thesis.

I wish to thank the many friends and acquaintances that I have made during my time here: Steven Kramer, Mel Klein, Haley Boone, Brent Eisenbarth, Will Carter, Jen Boehm, Amy Reynolds, Simon Wolf, John Kidd, Stephen Kudlacek, Tayler Billings, and many more. You all have helped to make my experience here one of the best and helped to provide me with an outlet from the rigors of academic life.

I also wish to thank several of the other professors in the linguistics department that have helped me. Specifically, Mike Terry, for always being there with an open door when I needed to talk or had concerns and Paul Roberge, thank you for providing me with a final year full of stories from our prehistory to the northern lands.

Thank you one and all!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

AM	Aspirate mutation
ICM	Initial consonant mutation
LS	Lexical Selection
NM	Nasal Mutation
OT	Optimality Theory
PCSA	Phonologically Conditioned Suppletive Allomorphy
SM	Soft Mutation
x_o	Soft Mutation
x_n	Nasal Mutation
x_h	Aspirate Mutation

CHAPTER 1

Introduction

This thesis is concerned with the interaction of the phonological and morphological components of grammar within a constraint-based theory, specifically within Optimality Theory (Prince & Smolensky 2004 [1993]). This interaction is explored in the specific context of the Welsh language's definite article allomorphy. The selection of the correct definite article allomorph is conditioned by the syntactic, phonological, and morphological component at the intersection of these components.

I show that the interaction of both *prosodic subcategorization* (Inkelas 1990; Zec 2005; Bye 2007; Bennett et al. 2018; Tyler 2019) and Lexical Selection's constraints on morpheme and allomorph selection, in the form of PRIORITY and RESPECT (Mascaró 2007; Bonet et al. 2007; Brinkerhoff 2019), are necessary for giving a complete and adequate account of the Welsh definite article allomorphy. This is because of the clitic status of the Welsh definite article with its allomorphs exhibiting different cliticizing patterns. Additionally, this proposal is necessary to account for the phonologically non-optimizing nature of one of the definite article allomorphs.

This thesis provides an account for initial consonant mutations within Optimality Theory and Lexical Selection. This is accomplished by the listing of the mutation patterns into sets that the grammar selects through lexical subcategorization, which is accounted for by means of the morphological constraint RESPECT (Bonet et al. 2007).

This thesis contributes to our understanding of morpho-phonology by addressing the claims made by Paster that *morphological* subcategorization alone can account for all cases of allomorphy (Paster 2005, 2006, 2015). I demonstrate that this is not the case and that morphological constraints are needed to account for the definite article.

This thesis also extends the framework of prosodic subcategorization to include *allomorph-specific* subcategorization frames. This will be shown to be necessary for the definite article because its allomorphs are associated with different prosodifications, with one being enclitic and the others being proclitic. By allowing the different allomorphs of the definite article to have their prosodic subcategorizations, the facts of the definite article can be accounted for within Optimality Theory.

This thesis presents facts of the Welsh phonology, the definite article and its allomorphy, and initial consonant mutations in Chapter 2. Chapter 3 lays out the theoretical assumption and frameworks that are necessary for analysis. Chapter 4 presents the analysis of the definite article allomorphy using Lexical Selection and Prosodic Subcategorization. Chapter 5 presents an analysis of initial consonant mutations in general and an analysis of the interaction of the initial consonant mutations with the definite article allomorphy. Chapter 6 concludes the thesis.

CHAPTER 2

Data and Descriptive Generalizations

For any discussion of Welsh phonology or any part of the grammar interacting with the phonology, it is vital to note that there does not exist a unified spoken standard for the language. It is commonly known that the Welsh language is divided into two larger dialect areas: the Northern varieties and the Southern varieties (Morris-Jones 1913; Thorne 1993; King 2015; Borsley et al. 2007; Hannahs 2013b). These differences between the Northern and Southern varieties amount to changes in the lexicon, syntactic structures, morphology, and most noticeably phonological and phonetic differences. This chapter introduces first a summary of the Welsh phonology with an emphasis on the segment inventories, syllables, and phonotactics in Section 2.1. The data concerning the definite article allomorphy is introduced in Section 2.2. The fundamental aspects of initial consonant mutation in Welsh are introduced in Section 2.3. Finally, the data concerning the interaction of the definite article allomorphy and initial consonant mutations are presented in Section 2.4. This chapter concludes with a summary in Section 2.5.

2.1 Welsh phonology

As stated by Hannahs (2013b), "in the absence of a spoken standard language the description of the segment inventory needs to take into consideration a significant amount of regional variation." These differences—which can be noticeable between the North and the South—need acknowledgment in any account concerning Welsh phonology even if they do not pose an issue for an analysis involving the definite article, as is the case in this thesis. This section lays out the main facts concerning the segment inventory for Welsh and makes a note of the key differences between these two spoken norms. A discussion of the consonant inventories is given in Section 2.1.1. Section 2.1.2 is a discussion of the vowel inventory. A discussion of the orthography is given in Section 2.1.3. Finally, a discussion of the syllable structure and phonotactics is given in Section 2.1.4.

All information in this section is based on Hannahs (2013b), unless otherwise noted.

2.1.1 Consonants

Figures 2.1 and 2.2 below lay out the basic consonant inventories in the Northern and Southern varieties respectively. The main differences between the two varieties lie in their treatment of the glottal fricative. Southern varieties do not have this sound and additionally lack the voiceless nasals[1] and voiceless aspirated rhotic whereas the Northern varieties have them. Additionally, the Southern varieties have [z] which is only found in English loanwords such as *sw* [zuː] ‘zoo’.[2]

	Bilabial	Labio-dental	Dental	Alveolar	Lateral	Post-aveolar	Palatal	Velar	Uvular	Glottal
Stop	p b			t d				k g		
Nasal	m̥ m			n̥ n				ŋ̊ ŋ		
Affricate						tʃ dʒ				
Fricative		f v	θ ð	s	ɬ	ʃ			χ	h
Liquid				r̥ʰ r	l					
Glide	w						j	w		

Figure 2.1: Consonant inventory for Northern Welsh varieties

	Bilabial	Labio-dental	Dental	Alveolar	Lateral	Post-aveolar	Palatal	Velar	Uvular	Glottal
Stop	p b			t d				k g		
Nasal	m			n				ŋ		
Affricate						tʃ dʒ				
Fricative		f v	θ ð	s z	ɬ	ʃ			χ	
Liquid				r	l					
Glide	w						j	w		

Figure 2.2: Consonant inventory for Southern Welsh varieties

2.1.2 Vowels

The Welsh vowel inventory consists of between 11-13 monophthongs, depending on which variety is under consideration. Just as with the consonants, Southern varieties have a more restrictive vowel system than Northern varieties. Southern varieties, in general, do not contain the high central vowels

[1]This has led some researchers to the conclusion that all of these forms are better analyzed as consisting of a nasal followed by aspiration, [mʰ nʰ ŋʰ] or [m̥ʰ n̥ʰ ŋ̊ʰ].

[2]In the North these are all realized with [s].

which have merged with the high front vowels, whereas the Northern varieties have retained this sound. This difference is indicated in Figure 2.3 by the high central vowels being in parentheses.

	Front		Central		Back	
	Short	Long	Short	Long	Short	Long
High	ɪ	iː	(ɨ)	(ɨː)	ʊ	uː
Mid	ɛ	eː	ə		ɔ	oː
Low			a	aː		

Figure 2.3: Welsh Vowel Inventory

As can be seen in Figure 2.3, all Welsh vowels—except for schwa—make a distinction between long and short. Just as in English, the difference between long and short also corresponds to a change in vowel quality for the mid and high vowels ,with the long vowels being pronounced tensed and the short vowels lax. However, unlike English, the mid tense vowels are not diphthongs but monophthongs.

Welsh also has a series of diphthongs with different quantities between the North/South divide. In Southern varieties, there are a total of eight diphthongs which symmetrically fall into two categories with four diphthongs closing to the high front vowel and four diphthongs closing to the high back vowel.

(1) Southern diphthongs (Hannahs 2013b: 24)

a. [aɪ] [taɪ] *tai* 'houses'
   [ɔɪ] [trɔɪ] *troi* 'turn'
   [əɪ] [nəɪd] *neud* 'make, do'
   [ʊɪ] [mʊɪ] *mwy* 'more'

b. [aʊ] [ɬaʊ] *llaw* 'hand'
   [ɛʊ] [ɬɛʊ] *llew* 'lion'
   [əʊ] [ˈbəʊɪd] *bywyd* 'life'
   [ɪʊ] [ɬɪʊ] *lliw* 'colour'

As discussed above, Northern varieties retain the high central vowel which has coalesced with the high front vowel in Southern varieties. This vowel can also combine with other vowels into diphthongs; this is one reason for the disparity between the different quantities of diphthongs between the North/South divide. In the North, this results in a total of thirteen diphthongs which are divided into three categories: diphthongs that close to the high front vowel (2a), those that close to the high back vowel (2c), and those that close to the high central vowel (2b).

(2) Northern diphthongs (Hannahs 2013b: 25)

a. High front closing	b. High back closing	c. High central closing
[aɪ] [tʰaɪ] tai ‘houses’	[aʊ] [ɬaʊ] *llaw* ‘hand’	[aɨ] [kʰaɨ] *cau* ‘close’
[ɔɪ] [tʰr̥ɔɪ]] *troi* ‘turn’	[ɛʊ] [ɬɛʊ] *llew* ‘lion’	[ɑɨ] [kʰɑɨ] *cae* ‘field’[3]
[əɪ] [tʰəɪ] *tei* ‘tie’	[əʊ] [ˈbəʊɪd] *bywyd* ‘life’	[ɔɨ] [kʰɔɨd] *coed* ‘wood’
	[ɪʊ] [ɬɪʊ] *lliw* ‘colour’	[ʊɨ] [mʊɨ] *mwy* ‘more’
	[ɨʊ] [ɨʊ] *yw* ‘is’	[əɨ] [nəɨ] *neu* ‘or’

2.1.3 Orthography

Any discussion of the Welsh phonology is neglectful if it fails to mention the orthography of this language. One reason that this is necessary, is that almost all descriptions and accounts of various aspects of Welsh present the data solely in the orthography. This is done both because of the lack of a standardized spoken language and the differences in the pronunciation between the Northern and Southern varieties.

The Welsh orthography is relatively phonemic (Thorne 1993), which greatly aids in knowing the pronunciation of any given word, even across the North/South divide, as the differences between the two varieties are systematic. However, the correspondence is not always clear due to a large number of digraphs that are employed in Welsh. The following tables modified from Thorne (1993) present the phoneme and the grapheme associated with it for the consonants and vowels in Tables 2.1 and 2.2 respectively.

Welsh also employs a series of diacritics to encode other phonetic information (Thomas 1996). These include the circumflex (ˆ), grave accent (‘), acute accent (´), and diaeresis (¨). The circumflex indicates that the vowel is long or to differentiate between two homophones. However, not all long vowels have the circumflex because the length of the vowel can be predicted by the orthography. As noted by Williams (1983: 179), in stress syllables monophthongs are short if they are followed by /p t k m ŋ/ or if followed by two or more consonants, in all other cases the vowel is pronounced long.[4]

[3] The low back vowel is only in this diphthong and is only found in Northern varieties. This vowel is not anywhere else part of the vowel inventory.

[4] Hannahs (2013b: Chapter 3) attributes this to a word minimality requirement that words must be minimally bimoraic, which also extends to stressed syllables.

Table 2.1: Correspondence for consonants

/p/	p	/t/	t	/k/	c
/b/	b	/d/	d	/g/	g
/m/	m	/n/	n	/ŋ/	ng
/f/	ff, ph	/θ/	th	/s/	s
/ɬ/	ll	/ʃ/	si, sh	/χ/	ch
/h/	h	/v/	f	/ð/	dd
/l/	l	/r/	r	/r̥ʰ/	rh
/w/	w	/j/	i		

Table 2.2: Correspondence for vowels

/iː/	/ɪ/	i, y (in the South)			
/eː/	/ɛ/	e			
/aː/	/a/	a			
/oː/	/ɔ/	o			
/uː/	/ʊ/	w			
/ɨː/	/ɨ/	u, y (in the North)			
/ə/		y			
/əɪ/	ei	/ɪʊ/	iw, yw, uw	/əɨ/	eu
/aɪ/	ai	/ɛʊ/	ew	/aɨ/	au
/ɔɪ/	oi	/aʊ/	aw	/ɑɨ/	ae
/ʊɪ/	wy	/oʊ/	ow	/ɔɨ/	oe
		/ɨʊ/	yw, uw	/ʊɨ/	wy
		/əʊ/	yw		

The grave accent occasionally indicates that a vowel is short where a long vowel is expected from the orthography and is primarily found with foreign borrowings. The acute accent shows where stress is when it is different than the default stress placement. Both the grave and acute accent are commonly omitted in everyday communications. The last diacritic is the diaeresis, which indicates that two adjacent vowels are pronounced as monophthongs instead of as a diphthong (i.e., they are pronounced as belonging to two different syllables).

2.1.4 Syllables and phonotactics

As noted by Hannahs (2013b), there has not been much work on the syllable and its divisions beyond a few indications by Morris-Jones (1913) and Pilch (1957: 49–54).

Welsh allows syllables to lack onsets and codas. This means that a syllable is allowed to consist solely of a nucleus.

(3) Welsh syllable consisting of only a nucleus

a. [aː] *â* 'with'

b. [ʊɨ] *ŵy* 'egg'

Because the Welsh syllable finds onsets optional and does not explicitly forbid codas a monosyllable may consist of any combination of onset, nucleus, and coda (with the requirement that there has to be a nucleus).

(4) Welsh monosyllables with simple onsets and/or simple codas

a. [tʰɨː] *tŷ* ‘house’

b. [aʊχ] *awch* ‘edge’

c. [ɬaːð] *lladd* ‘kill’

Welsh also allows both complex onsets and codas (i.e., onsets and codas with more than one segment). For complex onsets, the maximum number of segments is limited to three. Thanks to the extensive initial consonant mutations found in Welsh there is a large number of two-member consonant clusters that are permissible in Welsh, see (5a). All of the permissible three-member complex onsets involve the phone [s] as the first member of the cluster, see (5b).

(5) Complex onsets in Welsh

a. Permissible two-member onset clusters, including both citation and mutation forms:

pl, pr, fl, fr, bl, br, m̥l, m̥r, vl, vr, ml, mr

tl, tr, θr, θl, dl, dr, n̥l, n̥r, ðr, nr

kl, kr, kn, χl, χr, χn, gl, gr, gn, ŋ̊l, ŋ̊r, ŋ̊n, ŋl, ŋr

gw, kw, χw

sb, st, sg

b. Permissible three member onset clusters:

sbr, sbl, str, sgl, sgr

Regarding coda clusters, Welsh allows two-segment coda clusters at maximum. These coda clusters strictly adhere to the sonority sequencing principle and because they do not take part in consonant mutations these clusters are not as extensive as the onset clusters. As noted by Hannahs (2013b), “[t]hese [clusters] may consist of an obstruent followed by an obstruent, *Pasg* [pask] ‘Easter’, *hollt* [hɔɬt] ‘split’, or a sonorant followed by a sonorant, e.g. *darn* [darn] ‘piece’, *talm* [talm] ‘portion’, or a sonorant followed by an obstruent, e.g. *barf* [barv] ‘beard’, *balch* [balχ] ‘proud’, *gardd* [garð] ‘garden’, *merch* [mɛrχ] ‘girl’.”

However, Welsh has a significant number of items that appear to violate the sonority sequencing principle if considering the orthography. These are all words that involve coda clusters that end in a liquid.

(6) Apparent sonority sequencing violations in Welsh *pobl* ‘people’

cwbl ‘all’

llestr ‘vessel’

ffenestr ‘window’

posibl ‘possible’

perygl ‘danger’

Welsh takes several different approaches in resolving these clusters which include svarabhakti (i.e., copy-epenthesis) of the preceding vowel, deletion of the liquid, or metathesis. These different processes all result in resolving violations of sonority sequencing. A detailed account is given in (Hannahs 2013b: Ch. 5) and is beyond the scope of this thesis.

In regards to the syllabification of words and phrases, syllabification occurs at the prosodic word and will not syllabify across prosodic word boundaries (Hannahs 2013b: 38ff). At the level of the prosodic word, syllables will be formed that will make the word most phonologically and phonotactically optimal.[5] For example, if we consider the phrase *gyda'r afon* ‘with the river’ the syllabification at the word level would be as follows (the # indicates a prosodic word boundary):

(7) Word level syllabification of *gyda'r afon*

[gɨ.dar # a.vɔn]

2.1.5 Summary

This section presents the basic descriptions of the phonology and syllable structure for Welsh. Although there is a lack of a spoken standard, there is still significant overlap between the Northern and Southern varieties which allows for their mutual discussion.

[5]Because words are rarely pronounced in isolation, there is some evidence that syllabification is fluid across word boundaries. Cardinaletti & Repetti (2009) claim that in Romance languages clitics and syllables occur at two different levels: one at the prosodic word and the other at the phonological phrase. Just as individual words combine in syntax to form constituents, words in those constituents are pronounced together in the phonological correspondent (i.e., phonological phrases). According to Cardinaletti & Repetti, it is in this higher level that syllabification of those syntactic constituents occurs. This means that the syllable structure in this higher phrase may or may not be identical to the syllabification at the prosodic word. According to Cardinaletti & Repetti it is these syllables that we encounter in language. It is possible that this type of syllabification is also present in Welsh. In order to ascertain if syllabification occurs at the word-level or phonological-phrase level experimentation or a phonetic analysis of Welsh is needed. However, Bennett (2018) raised objections to this higher-level syllabification on the grounds of Match Theory.

2.2 Definite Article allomorphy

The Welsh definite article has three phonologically conditioned allomorphic forms (*y/yr/'r*). The allomorph *y* appears when the following word begins with a consonant, while the *yr* allomorph is present when the following word begins with a vowel, glide, or the phone [h] (Hannahs & Tallerman 2006: 782).[6] The third allomorph of the definite article, *'r*, takes precedence over the other allomorphs whenever the preceding word is vowel-final regardless of the following context. This generalization is summarized below in (8). All examples in this section are from Hannahs & Tallerman (2006), unless otherwise noted. Welsh examples are primarily left in their orthographic forms because of a lack of a standardized pronunciation for Welsh as discussed above. IPA transcriptions will be used where necessary to help clarify a point.

(8) Welsh definite article by environment with examples.

a. *yr*/_V, h, G	b. *y*/_C	c. *'r*/V_
yr afon 'the river'	*y llyfr* 'the book'	*o'r afon* 'from the river'
		o'r llyfr 'from the book'

Based on the generalizations described in the preceding paragraph and in (8), the behavior appears to be a simple case of phonologically conditioned suppletive allomorphy (PCSA). PCSA is when the phonology triggers the allomorphy of a given lexeme but the choice of which allomorph is morpheme-specific and does not reflect language-wide phonological processes. Hannahs & Tallerman (2006), however, claim that this is not the case because the *'r* allomorph takes precedence over the other allomorphs whenever its environment is present, regardless if the environments for the other allomorphs are present or not.

[6] This is not entirely true as it has been remarked that in Modern Welsh the two glides behave differently in this regard. It was noted at least as far back as Morris-Jones (1913: 192) that the labio-velar glide patterns with the consonants in respect to the definite article allomorphy and the palatal glide patterns as a vowel, which is consistently transcribed by Morris-Jones as [i̯] instead of [j], which others follow (e.g. Iosad 2007, 2016, 2017). Others have noticed this pattern as well and frequently mentioned this pattern in contemporary grammatical accounts of the language (Thorne 1993; King 2015).

This also could be explained by Smith's (2012) account of glides in Korean and whether or not they are considered part of the consonantal-onset or rimal onglides. From Smith's account, it is seen that they function both as consonant- and vowel-like. This is owed to a definition of the constraint Onset to be head-based. This same logic can be applied to the Welsh glides, with the glide [w] being considered a consonant and being a member of the onset while the glide [j] is not, which would then be considered a type of rimal on-glide instead.

(9) The precedence of *'r*

a. ***o'r*** *llyfr/*o* ***y*** *llyfr (Preposition)*
from-the book

b. *Brynaist* ***ti'r*** *llyfr?/*Brynaist ti* ***y*** *llyfr? (Determiner)*
bought.2sg you-the book
'Did you buy the book?'

c. *Pwy* ***ydy'r*** *meddyg?/*Pwy ydy* ***y*** *meddyg? (Verb)*
Who is-the doctor
'Who is the doctor?'

d. ***rhieni'r*** *ysgolfeistr/*rhieni* ***yr*** *ysgolfeistr (Noun)*
parents-the schoolmaster
'the parents of the schoolmaster'

In Welsh, there exists a small class of morphemes which alternate between a vowel-final form and a consonant-final form, see (10). The consonant-final form of these morphemes surfaces only when the following word is vowel-initial, in all other cases the vowel-final form is observed.

(10) Functional morphemes that have vowel-/consonant-final alternations (Information collected from King 2015)[7]

gyda/gydag	'with'
â/ag	'with'
tua/tuag	'towards, about'
na/nag	'than'
a/ac	'and'
na/nac	'neither/nor'

When these morphemes (hereafter referred to as the *gyda*-like morphemes) are spoken in conjunction with the definite article two potential forms are logically possible. The first is with the enclitic definite article, *'r*, and the vowel final form of the *gyda*-like morphemes, such as *gyda*. The other possibility is the consonant-final form of the *gyda*-like morpheme, such as *gydag* and one of the proclitic allomorphs, *y/yr*. Of these two possible forms only one is ever encountered in the data; the vowel-final form of the *gyda*-like morphemes and the enclitic definite article allomorph.

[7] All of the morphemes that have this alternation are functional morphemes that end in a velar stop. Because they all end in a velar stop, it is possible that this stop has some role to play in the alternation between consonant- and vowel-final forms of the morphemes. However, this requires further investigation, which is beyond the scope of this thesis.

(11) The interaction of *gyda* and the definite article before a C-initial word

a. First possibility *gyda'r nod*

gyda'r nod [gɨ.da**r̲** nɔ**d̲**]

'with the aim'

b. Second possibility **gydag y nod*

gydag y nod* [*gɨ.dag ə.nɔd̲**]

'with the aim'

I show in Chapter 4 that the interaction of these two classes of morphemes is still phonologically conditioned. The choice of the different allomorphs is the result of the lexical listing of the allomorphs and the interaction of the prosodic structure of the definite article which has been demonstrated to be a clitic for over a hundred years (Morris-Jones 1913) and will be further discussed in Section 4.1. This is still the result of the phonological conditioning of the allomorphs.

In addition to the facts already presented, one further aspect of the definite article is commonly encountered in the data. As previously stated, the enclitic form of the definite article takes precedence whenever its environment is present. This is true for most instances, except in certain circumstances where the enclitic form is suppressed, which results in the proclitic forms of the definite article surfacing. In Welsh, the enclitic form is required to be pronounced in in conjunction with its phonological host and as such cannot be stressed. In circumstances where one wishes to stress the definite article for emphasis then one of the other allomorphs (*y/yr*) is present depending on the following phonological contexts. In the following examples, the environments and definite article are in italics.

(12) a. Caffi Morgan *ydy Y* lle i fynd!
cafe M. is the place to go
'Cafe Morgan is *the* place to go!'

b. Caffi Morgan *ydy YR* unig lle i fynd!
cafe M. is the only place to go
'Cafe Morgan is *the* only place to go!'

Additionally, the enclitic form is suppressed when giving citation forms (13a), when the definite article is part of a name or title (13b), or when there is a deliberate pause (13c).

(13) a. Citations.

Er	*mai*	‘*y*	pobloedd’	a	ddisgwylir
although	COMP	the	peoples	REL	expect:PRES:PSV

‘although *y pobloedd* is expected’

b. Part of a title or name[8]

Mae	*yr*	Arolwg	Ordnans	yn	adolygu	30,000	o	fapiau	y	flwyddyn.
is	the	survey	ordnance	PROG	review	30,000	of	maps	the	year

‘The Ordnance Survey revises 30,000 maps a year.’

c. Deliberate pause.

Ni	alwodd	neb	*yma* ‖	*y*	dydd	o’r blaen		‖ = pause
NEG	called	no-one	here	the	day	before		

‘No one called here the other day.’

According to (Hannahs & Tallerman 2006: 792f), the suppression of the enclitic form of the definite article is attributed to the presence of an intonational phrase boundary which, as noted by Nespor & Vogel (2007 [1986]), blocks various kinds of segmental rules from applying.[9]

2.2.1 Summary

We have seen in this section that the definite article has three distinct allomorphs which appear to have a skewed preference for one of the allomorphs, in this case, the allomorph *’r* [r], being chosen over the other two forms whenever its environment is present. This asymmetry between the allomorphs will be crucial to their analysis and will need to be accounted for in any analysis of the definite article.

2.3 Initial Consonant Mutation

Initial consonant mutation is a well-known process among the members of the modern Celtic languages (see Hannahs 2011 for a discussion given to the long history of initial consonant mutations

[8]In this example a further position that one expects to find the encliticized definite article is with *fapiau y flwyddyn.* However, because the phrase *y flwyddyn* is an adjunct to the verb phrase instead of forming a constituent with the preceding noun encliticization is blocked.

[9]Under Match Theory (Selkirk 2009, 2011), this cannot be caused by an intonational phrase boundary, because intonational phrases are only associated with CPs or ForcePs. Instead these boundaries would have to correspond to major phonological phrase boundaries.

in the Celtic language within linguistics and philology). In these languages, certain word-initial consonants undergo a series of systematic and phonetically distinct alternations which are conditioned by specific contexts and environments (Hannahs 2011). These contexts are generally understood to have originated historically as phonetic processes of external sandhi of the word-initial consonant in various phonologically conditioned environments which have undergone a process of grammaticalization followed by the loss of the phonological triggers of mutation.[10] This process of grammaticalization has resulted in specific lexical items governing the different mutations, with each of the former phonological processes now associated with the different mutation types.

2.3.1 Initial Consonant Mutation Patterns

In Modern Welsh, three different mutations can occur on the initial consonants which are referred to as *radicals*, which are specifically the non-mutated consonants and word forms.[11] These mutations are: Soft, Nasal, and Aspirate Mutation. These mutations only happen with a subset of the consonant inventory of the language. Those consonants that undergo mutation are summarized in Table 2.3 below. Cells that are left blank indicate that this mutation does not occur with the associated radical.

Table 2.3: Mutation table showing soft, nasal, and aspirate mutation (Borsley et al. 2007: 20)

Radical	*Soft*	*Nasal*	*Aspirate*
p [p]	b [b]	mh [m̥]	ph [f]
t [t]	d [d]	nh [n̥]	th [θ]
c [k]	g [g]	ngh [ŋ̊]	ch [χ]
b [b]	f [v]	m [m]	
d [d]	dd [ð]	n [n]	
g [g]	– zero	ng [ŋ]	
m [m]	f [v]		
ll [ɬ]	l [l]		
rh [r̥ʰ]	r [r]		

Of these mutations, the soft mutation is the only one in which all the radicals take part. Soft mutation is also the most varied in terms of the phonological changes that the radicals undergo. Under soft mutation, voiceless stops become voiced, and voiced stops become fricatives with the exception of /g/ which, descriptively speaking, deletes. Additionally for the remaining three sounds that participate in soft mutation, [m] becomes [v], [ɬ] becomes [l], and [r̥ʰ] becomes [r]. In nasal

[10] See Hannahs (2011) and the citations therein for a detailed discussion.

[11] As King (2015) defines it: “the basic, dictionary form of a Welsh word without any mutation to the initial letter”

mutation, each of the stops become nasal stops while preserving their voicing.[12] Finally, aspirate mutation causes the voiceless stops to become fricatives.

In (14), we see the mutations that the word *tad* 'father' undergoes with the different mutation triggers (with the radical—or citation—form given first). As shown, the soft mutation form of *tad* is shown to be triggered by the second person singular possessive pronoun (*dy*) resulting in *dad*. Next, we see that nasal mutation is triggered by the first person singular possessive (*fy*) resulting in *nhad* and finally we see aspirate mutation which is triggered by the third person feminine possessive pronoun (*ei*) and results in *thad*.

(14) Examples of Welsh initial consonant mutations on *tad* 'father' (Hannahs 2011: 2807)

radical	*tad*	[taːd]	'father'
soft mutation	*dy dad*	[də daːd] (2SG.POSS father)	'your father'
nasal mutation	*fy nhad*	[ə n̥aːd] (1SG.POSS father)	'my father'
aspirate mutation	*ei thad*	[i θaːd] (3FEM.SG.POSS father)	'her father'

We saw that each type of mutation is triggered by different sets of lexical items. There is only a small class of lexical triggers for each of the associated mutations with the sole exception of soft mutation, which will be discussed below. In (15), we see the specific lexical triggers associated with the nasal and aspirate mutations. These lexical triggers are commonly listed in Welsh grammars, and this list of some of the triggers are drawn from King (2015: Section 9) specifically. Each of these forms is also listed with the typical diacritic found in those same grammars and in most dictionaries to indicate what mutation is associated with those lexemes.

(15) Lexical triggers for nasal and aspirate mutation

a. Nasal Mutation

fy^n , *'(y)n*^{n}	'my'
yn^n	'in'

[12]This is only true of the Northern varieties of Welsh. In Southern varieties all nasal mutations are voiced regardless of the original voicing of the radical.

b. Aspirate Mutation

a/ac^h	'and'
â/ac^h	'with'
chwe^h	'six'
ei^h	'her [poss.]'
gyda/gydag^h	'with'
tri^h	'three (m)'
tua/tuag^h	'towards, about'

2.3.2 Soft Mutation triggers

Soft mutation is by far the most widespread of the different types of mutation found in Welsh. In addition to a list of over 20 lexical triggers, this mutation also has at least three purely morphosyntactic contexts.

(16) Some of the lexical triggers for soft mutation

a^o	[relative]
dacw^o	'(over) there is/are...'
dau^o	'two (m)'
ei^o	'his'
rhy^o	'too'
neu^o	'or'

One of these morphosyntactic contexts is observed when feminine singular nouns follow the definite article. This context presents some unique issues related to the distribution of the definite article allomorphs, which will be discussed further in Section 2.4.

Another instance of a purely morphosyntactic trigger is with adjectives. Adjectives in Welsh are typically post-nominal (i.e., occur after the noun they modify like in Spanish or French), but Welsh does allow all adjectives—except for comparatives and superlatives (Borsley et al. 2007: 179f)—to be pre-nominal. Pre-nominal adjectives cause soft mutation on whatever word immediately follows (typically a noun), see (17a), while feminine nouns cause soft mutation on all adjectives that follow them, see (17b).

(17) Morphosyntactic contexts for SM (Borsley et al. 2007)

a. Pre-nominal adjective trigger of SM on nouns

hen gath (*cath* > *gath*)

'old cat'

b. Feminine singular noun trigger of SM on all following adjectives

merch dal gerf (*tal* > *dal*, *cerf* > *gerf*)

girl.FEM.SG tall strong

'(a) tall, strong girl'

An additional context that soft mutation occurs in is called "syntactic soft mutation." According to many authors, this mutation is purely syntactic, hence the name (see Tallerman 2006 and Borsley et al. 2007 for a detailed discussion and review of syntactic soft mutation and previous accounts). According to Borsley et al. (2007: 223-226), there has been extensive debate over the last twenty-five years as to what exactly causes or triggers "syntactic soft mutation." The examples in (18) show the range of syntactic positions and categories in which this mutation occurs, including objects (18a), VPs (18b), embedded clauses (18c), and subjects (18d) as possible targets for syntactic soft mutation.

(18) Examples of syntactic soft mutations which are marked with an underline (taken from Borsley et al. 2007: 224)

a. Direct object *beic*

Prynodd y ddynes <u>feic</u>.
buy.PAST.3S the woman bike

'The woman bought a bike.'

b. VP with initial word *gwerthu*

Gwnaeth y dyn [<u>werthu</u> beic].
do.PAST.3S the man sell.INF bike

'The man sold a bike.'

c. Initial word in embedded clause *golchi*

Dechreuodd Huw [<u>olchi</u> 'r llestri].
begin.PAST.3S Huw wash.INF the dishes

'Huw began to wash the dishes.'

d. Subject of the sentence *ci*

Mae	*yn*	*yr*	*ardd*	*gi.*
be.PRES.3S	in	the	garden	dog

'There's a dog in the the garden.'

This mutation and its conditioning are beyond the scope of this thesis and will not be discussed here. Instead, I refer the interested reader to Borsley et al. (2007: Ch 7) and Tallerman (2006) for a detailed discussion concerning this mutation.

2.3.3 Exceptions to mutation

One issue that remains to be discussed concerning mutation is the lexical exceptions that do not undergo mutation or that only undergo a subset of the mutations. It has been noted by many that there exists a series of lexical items that are impervious to the effects of mutation (e.g., Thorne 1993; Green 2006, 2007; Hannahs 2011; King 2015). According to King, these exceptions for mutation are: words that are already mutated such as *beth?* 'what' (from *peth* 'thing') or *dros* 'across, over' (originally *tros*), miscellaneous words such as *dy*° 'your', *byth* 'ever/never', and *lle* 'where'; non-Welsh place-names, personal names, and some foreign words (especially those that begin with [g]). This exceptionality will prove crucial to the analysis of initial consonant mutations presented in Chapter 5.

2.4 Interaction of the Definite Article and Initial Consonant Mutation

As was mentioned in the previous section, feminine singular nouns undergo mutation when there is a definite article directly proceeding them in the sentence.

(19) *y* *ddynes* (*dynes* > *ddynes*)
'the woman.FEM.SG'

In most cases, this does not present an issue to the selection of the correct definite article allomorph, but there is one case that does. As was mentioned previously, words that begin with the voiced velar stop in the radical appear to "delete" this consonant under soft mutation. This presents a unique situation for the definite article and these feminine singular nouns that undergo mutation. In all instances, these /g/-initial radicals could present the definite article with either a vowel- or consonant-initial word depending on what follows the elided segment. For example, if the Welsh word begins with a consonant cluster when soft mutation occurs, the resultant form will be consonant-initial. If, however, the word begins with the radical /g-/ only, then it would result in a vowel-initial form after soft mutation, see (20). These forms then become the basis for which allomorph will appear for the definite article.

(20) Definite article trigger of SM on feminine singular nouns

a. *y lasog* (*glasog* > *lasog*)
'the gizzard.FEM.SG'

b. *yr ardd* (*gardd* > *ardd*)
'the garden.FEM.SG'

As stated by Hannahs & Tallerman (2006): "We seem then to have a paradox, a classic chicken and egg situation. The lenition on the noun…is triggered when the noun follows the article, but the correct form of the article (*y* or *yr*) cannot be inserted unless the initial segment of the noun is known, which depends on whether or not lenition has applied." This is indeed the case if we consider this from a Distributed Morphology (Halle & Marantz 1993) perspective and the order in which lexical items are inserted into the derivation. However, this is not a problem from a parallel insertion of the lexical material into the morpho-phonological component of the grammar like the one presented in this thesis in Chapter 5.

2.5 Summary

This chapter has laid out the basics of the Welsh phonology as well as information regarding the syllable structure. Most importantly this chapter has presented the facts concerning the definite article and the initial consonant mutations that occur in Welsh. These facts will prove vital for further discussion throughout this thesis; especially the allomorphy of the Welsh definite article and its phonological environments. I demonstrate in Chapter 4 that the clitic status of the definite article will prove vital to the analyses throughout this thesis. Clitics will require the use of prosodic subcategorization to explain their effects on the rest of the prosodic elements surrounding it, because of their phonological dependence on a host.

This chapter has also presented facts concerning initial consonant mutations in Welsh. Even though these initial consonant mutations appear to be phonologically driven—or at least have phonological reflexes—they appear to be primarily triggered by lexical item as well as various morphosyntactic structures. In chapter 5 of this thesis, I discuss an Optimality Theory approach to initial consonant mutations. I argue that initial consonant mutations are part of the mental lexical entry for each of the radicals and that the selection of the different mutation forms resides outside of the phonological component of the grammar. The selection of the correct mutation is argued to be the result of lexical subcategorization which is accounted for within Optimality Theory by means of the morphological constraints Priority and Respect.

CHAPTER 3

Theoretical Background

This chapter lays out the theoretical tools that I will be implementing for my analysis of the Welsh definite article allomorphy and initial consonant mutations. This chapter lays out the necessary information to understand how these theories function before proceeding to the analyses that will make use of these theories in the following chapters. This chapter will first describe two theoretical additions to the constraint-based grammar framework of Optimality Theory (Prince & Smolensky 2004 [1993]): Lexical Selection and Prosodic Subcategorization. Lexical Selection is discussed in Section 3.1. This theoretical device is primarily concerned with the capturing of phonologically conditioned allomorphy—especially in instances of suppletive allomorphy—through the introduction of ordered allomorph sets and the constraint PRIORITY (Mascaró 2007; McCarvel 2016). In addition to phonologically conditioned allomorphy, Lexical Selection is concerned with the modeling of lexical subcategorization, where lexical items encode various phonological or morphological information that they require to surface in the output. This is accounted for through the constraint RESPECT (Bonet et al. 2007). A brief discussion of clitics is given in Section 3.2 to provide background to *prosodic subcategorization* (Inkelas 1990, 1993; Zec 2005; Bye 2007; Bennett et al. 2018; Tyler 2019) which is discussed in Section 3.3. This framework is concerned with capturing the prosodic organization of affixes and clitics. These discussions will be followed by a general summary that will lay out the main points of each theoretical device and make reference to how they will be vital for the analyses in the following chapters.

3.1 Lexical Selection

Lexical Selection, as a theory of the morphology/phonology interface, is concerned with how the phonological grammar can handle the idiosyncratic nature of some allomorphs, in other words, exceptional allomorphy (Bonet et al. 2007; Mascaró 2007; McCarvel 2016). The theory functions by restricting the way in which morphology interacts with the phonology. The interactions that the morphology is restricted to are the lexical listing of allomorphs, the prioritization of an allomorph or subset of allomorphs when needed, the use of the constraint PRIORITY to regulate the prioritization, and the accounting for lexical subcategorization by means of the constraint RESPECT.

3.1.1 Lexical listing

Within Lexical Selection, a lexical listing of all the allomorphs of a given morpheme is in the input to the phonological grammar. The lexical listing of the allomorphs is necessary when there are cases of arbitrary, unmotivated instances that are not phonologically natural. In other words, if there is an alternation between the allomorphs that appears to be idiosyncratic to the rest of the grammar (e.g., the alternation between *a* and *an* in English is determined by the morpheme avoiding vowel hiatus, which is tolerated elsewhere in the language) then the allomorphs are lexically listed into a set that forms the input to an OT grammar. In order to incorporate and explain this behavior, Mascaró (2007) formalized a correspondence relationship for these lexical listing in (21) below.

(21) a. The set of allomorphs of a morpheme M (m_1, m_2, ..., m_n) can be represented as a partially ordered set.

b. For M = /m_1, m_2, ... , m_n/, GEN(/m_1, m_2, ..., m_n/) = GEN(m_1) $\cup$ GEN(m_2) ... $\cup$ GEN(m_n). (Given a set of allomorphs, the candidate set is the collection of the individual candidate sets of each allomorph.)

c. Each candidate morph in b. stands in a correspondence relation to one of the underlying allomorphs (i.e., if $\text{cand}_1 \in$ GEN(/m_j/), then $\text{cand}_1 \Re m_j$)

d. Under input allomorphy, candidate faithfulness violations are computed with respect to the candidate's corresponding underlying allomorph.

From this we are able to account for the variation that is observed in phonetically conditioned allomorphy. This is illustrated with the relative simple cases of the English indefinite article *a/an* and the allomorphy of the infinitival affix in Baix Empordà Catalan in Section 3.1.2 are employed.

According to 21b, each allomorph of the morpheme is listed in the input as an allomorph set {ə, ən}, in no particular order. The interaction with the constraints during EVAL will always select the correct output as can be seen in (22) below.

(22) OT/LS analysis of *a/an* allomorphy in English (Mascaró 2008: 517)

a. *an* [ən]

{$ə_1$, $ən_2$} *impossible*	MAX	DEP	ONSET	NOCODA
☞ a. ə.n_2*impossible*			*	
b. $ə_1$.*impossible*			**!	

b. *a* [ə]

{$ə_1$, $ən_2$} *possible*	MAX	DEP	ONSET	NOCODA
a. $ən_2$.*possible*			*	*!
☞ b. $ə_1$.*possible*			*	

Because the allomorphs in the input stand in direct correspondence to an allomorph in the output, allomorphs only receive faithfulness violations when there is a change in phonological structure to their correspondent. Because there is no change to the phonological structure between the correspondents for the allomorphs, there are no correspondence violations in either MAX or DEP. This is in agreement with how faithfulness constraints interact with allomorphs as outlined in (21d) and the correspondence relationships expressed in (21c). This results in the markedness violations determining the optimum with the least phonologically marked candidate winning.

The only instance of a faithfulness violation would be when there is a phonological mismatch between the corresponding allomorphs in the input and output, see (21d).

(23) Faithfulness violations in OT/LS (Mascaró 2007: 721)

ilun-{tu_1, du_2}	IDENT(voice)	*NÇ	*VOICEDOBST
a. ilun-tu_1		*!	
☞ b. ilun-du_2			*
c. ilun-du_1	*!		*
d. ilun-tu_2	*!	*	

Candidates (23.c-d) each receive a single Ident violation as the voicing of the corresponding allomorphs in the input and output do not agree in voicing. This then allows the two phonologically faithful candidates to compete against one another in markedness violations.

In summary, the use of allomorph sets allows for the grammar to select the correct allomorph for each of the phonological environments. These allomorphs in the input stand in direct correspondence to an allomorph in the output. Because of these correspondence relationships, allomorphs only receive faithfulness violations when there is a change in phonological structure to their correspondent. However, these allomorph sets cannot account for the selection of an apparently non-optimizing output of the grammar.

3.1.2 Prioritization

Per (21a), the listing of allomorphs can be ordered when the language prefers one or more allomorphs (Mascaró 2007: 716). This is represented in the formalism by the symbol "$>$" and represents a dominance relation between the allomorphs or, in other words, a prioritization of one allomorph or a subset of allomorphs over the other allomorphs. If we have a set of allomorphs $\{m_1 > m_2 > m_3\}$ then $m_1$dominates m_2, which in turn dominates m_3.

An example of prioritization in allomorph sets is found in Mascaró (2007). Mascaró discusses the allomorphy of the verbal infinitival affix from Baix Empordà Catalan. In this variety of Catalan, the infinitival morpheme for verbs is generally [ɾ], which assimilates to the following pronominal enclitic's initial consonant but remains as [ɾ] in all other environments see (24).

(24) a. Nonassimilation of [ɾ]

pɔza**ɾ**-u	'to put it'
pɔza**ɾ**-i	'to put there'
pɔza**ɾ**	'to put'

b. Assimilation of [ɾ]

pɔza**m**-mə	'to put me'
pɔza**l**-lə	'to put it-fem'
pɔza**l**-li	'to put him/her-dat'
pɔza**s**-sə	'to put oneself'
pɔza**t**-tə	'to put you'
pɔza**l**-ləs	'to put them-fem'
pɔza**n**-nə	'to put some'

Based on this data, Mascaró (2007) proposes that all of these allomorphs are lexically listed with the [ɾ] being given priority because it appears in the least restrictive environment. This yields the allomorph set in (25).

(25) Allomorph set for the infinitival morpheme

$\{ɾ_1 > (n_2, l_3, t_4, s_5, m_6)\}$

In order to account for the allomorph prioritization during EVAL Mascaró (2007: 726) introduces a morphological faithfulness constraint PRIORITY.

(26) PRIORITY

Respect lexical priority (ordering) of allomorphs.

"Given an input containing allomorphs m_1, m_2, ..., m_n, and a candidate m_i', where m_i' is in correspondence with m_i, PRIORITY assigns as many violation marks as the depth of ordering between m_i and the highest dominating morph(s)."

This constraint will take an allomorph set like the one above in (25) and it will assign violations for not using the highest ranked allomorph. Additionally, as stated by this definition there can be multiple rankings within the allomorph sets. If a set of allomorphs like $\{ m_1 > m_2 > m_3 \}$ existed in which m_1 is ranked above m_2 which in turn is ranked above m_3 than Mascaró's (2007) PRIORITY constraint will assign no violations for m_1, it will assign one violation for m_2, and m_3 will receive two violations.

Subsequent work by Wolf (2008, 2015) in phonologically conditioned suppletive allomorphy (PCSA) shows that this use of allomorph ordering is too powerful as there is no limit to the number of dominance relations. Wolf also claims that PRIORITY is evaluated and assigns violations gradiently which is generally an undesirable trait of any phonological constraint. Building on Wolf and data from Jèrriais, McCarvel (2016) restricts the use of these dominance relations to a single allomorph having priority over one or more other allomorphs which would then mean that the new set would consist of $\{ m_1 > m_2, m_3\}$ where m_1 is prioritized over both m_2 and m_3. This new understanding of the lexical listing and allomorph sets leads McCarvel to reformulate this constraint to be a markedness constraint see (27), where there is only ever an alternation between a prioritized allomorph, or in her words a default, and non-prioritized allomorphs.

(27) PRIORITY: (McCarvel 2016)

Assign one violation mark for the use of any allomorph other than the default allomorph.

This constraint means that there can only ever be a single allomorph that is given prominence over one or more allomorphs. McCarvel gives the following allomorph set for the Jèrriais plural definite article

(28) Allomorph set for Jèrriais's plural definite article according to McCarvel (2016)

{lei, (lz > leiz)}

In this set—according to McCarvel (2016)—there is only a ranking between the allomorphs [lz] and [leiz]. This is illustrated by having the two allomorphs that are ranked in parentheses. In this ranking [lz] is ordered above [leiz] and the remaining allomorph is left unordered to both. What this means during evaluation with the constraint PRIORITY only the allomorph [leiz] receives a violation, and the other allomorphs do not result in any violations.

This redefinition is not wholly adequate for the behavior of PRIORITY as more than one allomorph can be prioritized over one or more allomorphs which is a better way to analyze the allomorphy in Jèrriais because only the allomorph [leiz] receives a violation of the constraint PRIORITY. This then means that both of the other allomorphs should be ordered over this allomorph instead of saying that one never receives a violation because it is unordered and that the other does not receive one because it is ordered over another allomorph. This unnecessarily complicates the evaluation which runs counter to Occam's Razor. Because of this fact, I revised the constraint PRIORITY in Brinkerhoff (2019) to better reflect this observation.

(29) PRIORITY: (Brinkerhoff 2019)

Assign one violation mark for the use of any allomorph other than the prioritized allomorph(s).

This constraint and rethinking of lexical ordering now allows for a single ranking to be placed anywhere in the allomorph set. Suppose there is a set of allomorphs: { m_1, m_2, m_3, m_4 }. This set—according to this definition—could have a single ranking anywhere there is a comma resulting in one, two or three allomorphs being ranked above three, two, or one other allomorph. This constraint will then assign a violation for the use of any allomorph that is dominated. For example in the allomorph set {m_1, $m_2 > m_3$ }, candidates m_1' and m_2' will be faithful to PRIORITY and m_3' would incur one violation.

Returning to the Baix Empordà Catalan data, the constraint PRIORITY is required, according to Mascaró, because the alternative would result in any of the allomorphs satisfying the grammar when there is a clitic that is vowel-initial, as illustrated in (30).

(30) Exceptional over-assimilation in Baix Empordà Catalan. Infinitive = /ɾ > (n, l, t, s, m)/ (Mascaró 2007: 726)[1]

a. Prevocalic

pɔzá-{$ɾ_1$ > (n_2, l_3, t_4, s_5, m_6)}-u	IDENT(F)	AGREE/C	PRIORITY
☞ a. puzá-$ɾ_1$-u			
b. puzá-n_2-u			∗!
c. puzá-l_3-u			∗!
d. puzá-t_4-u			∗!
e. puzá-s_5-u			∗!
f. puzá-m_6-u			∗!

b. Preconsonantal

pɔzá-{$ɾ_1$ > (n_2, l_3, t_4, s_5, m_6)}-lə	IDENT(F)	AGREE/C	PRIORITY
a. puzá-$ɾ_1$-lə		∗!	
☞ b. puzá-l_3-lə			∗

When there is a clitic beginning with a consonant than there is place agreement between the infinitival morpheme and the clitic's word initial consonant which is what the grammar in (13) correctly predicts.

In summary, Lexical Selection restricts the role of morphology in a grammar to three points. These points are the use of allomorph sets, the prioritization of the allomorphs within a set, and the introduction of the constraint PRIORITY.[2] These points collectively restrict grammars to generate some patterns but not others.

[1]There is a process in Catalan of vowel reduction that changes the unstressed vowels to [i u a ə] or [ə] depending on the variety which will not be addressed. The stressed vowel is marked with an acute accent on the vowel (i.e., á).

[2]There is also the possibility that when there are multiple morphemes that each have an ordered allomorph set then there are PRIORITY constraints for each allomorph (Brinkerhoff 2019). This requires further research to determine if this indeed exists.

3.1.3 Lexical Prespecification

It was proposed in Bonet et al. (2007), that in addition to the use of PRIORITY and lexical ordering of allomorphs there exist "idiosyncratic lexical specifications" which also have to bear on the choice of which allomorph is allowed to appear in a given context. This involves a morpheme determining the exponence of another morpheme (i.e., the lexical entry of that lexeme includes subcategorization for what allomorph is allowed to occur with said lexeme). This involves a morpheme determining the exponence of another morpheme.

An example of this comes again from Catalan. Like other Romance languages, Catalan makes use of specific gender morphology in order to indicate the grammatical gender of the word. Masculine words typically do not inflect for gender; however, some masculine words are found to be indicated by [u] or [ə]. For feminine words, the allomorphs indicating gender are [ə] or is not marked.

(31) Examples showing Catalan gender allomorphy (Bonet et al. 2007: 916)

a. Masculine
   i. /gɔ́t-∅/: [gɔ́t] 'glass'
   ii. /mós-u/: [mósu] 'lad'
   iii. /páɾ-ə/: [páɾə] 'father'

b. Feminine
   i. /mós-ə/: [mósə] 'lass'
   ii. /sál-∅/: [sál] 'salt'

Based on this evidence we can conclude that these allomorphs each form two sets that distinguish whether a noun is masculine or feminine. Additionally, it is claimed that these allomorph sets are ordered with the least morphologically marked allomorph ranked highest over the marked and most marked.[3]. Their results are summarized in Table 3.1 below.

[3]The evidence used by Bonet et al. (2007) to decide whether or not an allomorph was unmarked or marked was based on the statistical distribution of the different allomorphs. Based on their results they found that the masculine gendered allomorphs were ordered with the null allomorph being considered the morphologically unmarked form of the morpheme, because it had the highest statistical distribution being the unmarked and the allomorph with the lowest statistical distribution being the most marked.

Bonet et al. (2007) probably base this on the distribution of the allomorphs in the lexicon, with the allomorph with the highest proportion being considered the unmarked and the allomorph with the lowest proportion being considered the most marked. However, there is a problem with this as this does not work cross-linguistically.

A good example of this comes from German, where the default plural allomorph is not the statistically greatest proportion. However, is the default because when speakers are presented with new words that is the plural form that chosen over the other forms (Schuler et al. 2016)

Table 3.1: Table of the gender allomorphs in Catalan taken from Bonet et al. (2007: 917)

	Unmarked	Marked	Most marked
Masculine	∅	u	ə
Feminine	ə	∅	

(32) Allomorph sets for the masculine and feminine gender allomorphs

a. Masculine:

{ ∅ > u > ə }

b. Feminine:

{ ə > ∅ }

In order to account for the fact that some lexical items can only appear with specific gender allomorphs, the authors claim that these lexical items come specified from the mental lexicon with which gender allomorph it can take. These preferences are represented with a subscript of that allomorph on the underlying form of the lexical item and are to be understood as a subcategorization requirement.

(33) Lexical specification of atypical gendered nouns in Catalan.

a. Masculine:

/mos$_{\mathrm{u}}$/, /paɾ$_{\mathrm{ə}}$/

b. Feminine:

/sal$_{\varnothing}$/

From this claim the authors also posit the addition of a new constraint that govern faithfulness to these lexical subcategorization requirements. This constraint is given as Respect and receives violations for every subcategorization that is not met. In their paper the definition was not the most informative according to the criteria laid out by McCarthy (2008), where definitions need to make reference to how violations are assigned. Following this criteria we can give the formal definition of this constraint as the following.

(34) Respect:

Assign one violation for every mismatch between the lexical subcategorization in the input and the output.

By ranking this new constraint above that of Priority, we are assured that the lexical specifications are upheld in the course of Eval. If this were not so the lexical specification would not be satisfied and the choice of which allomorph would be left to violations of general faithfulness and markedness constraints. In the next two examples, we see that when there is no lexical specification the constraint Priority will ensure that the unmarked gender allomorph is selected (Note that this is using Mascaró's (2007) original constraint definition and evaluation for Priority and multiple allomorph orderings).

(35) Evaluation of words without lexical specification

a. Masculine *pas* [pás] 'step'

pás-{ ∅ > u > ə }	Respect	Priority
☞ a. pás		
b. pásu		∗
c. pásə		∗!∗

b. Feminine *taca* [tákə] 'stain'

ták-{ ə > ∅ }	Respect	Priority
☞ a. tákə		
b. ták		∗!

In words that do have an underlying lexical specification the constraint Respect ensures that the subcategorization is found in the output. We see in the following examples that even with a violation of Priority for the use of the non-prioritized allomorph the higher ranked Respect will eliminate any candidate that does not respect the lexical specification of the base.

(36) Evaluation of words with lexical specification

a. Masculine *mosso* [mósu] 'lad'

mós_{u}-{ ∅ > u > ə }	Respect	Priority
a. mós	∗!	
☞ b. mósu		∗
c. mósə	∗!	∗∗

b. Feminine *sal* [sál] 'salt'

$\text{sál}_{\varnothing}$-{ ə > ∅ }	Respect	Priority
a. sálə	∗!	
☞ b. sál		∗

3.1.4 Summary

Lexical Selection has the power needed to ensure that asymmetrical relationships between allomorphs can be accounted for in constraint-based grammars. This is accomplished through the use of allomorph sets and rankings within those sets to show the asymmetrical relationships when they exist. The grammar evaluates these ranked allomorph sets utilizing the constraint Priority. We will see in the following chapters that this is important for the evaluation of the definite article allomorph and the preference of the allomorph *'r* in Welsh as well as which initial consonant mutation is chosen. In addition to these aspects of Lexical Selection, the addition of lexical subcategorization in the grammar allows lexical items to select material that is required to appear in conjunction with them. This subcategorization is regulated by means of the constraint Respect. This will prove important for the evaluation of initial consonant mutations, which analysis is given in Chapter 5.

3.2 Clitics

In order to obtain an accurate picture of the definite article and its allomorphy, a discussion of clitics is necessary. As will be illustrated in Section 4.1, the definite article's allomorphs are clitics both enclitic and proclitic.[4]

Clitics are lexical items that are syntactically free (i.e., they form the head of a syntactic phrase) but form a phonological constituent with either the proceeding or the following word thus giving us enclitics and proclitics respectively. For example, in English, the clitic *='s* is an independent syntactic element that is the reduced form of the verb *is*.

(37) Examples of the clitic *='s* (Inkelas 1990: 81)

a. The [pilot's] in the cockpit

b. The pilot I [met's] in the cockpit

This word stands on its own syntactically but must be pronounced in conjunction with the preceding word. This results in the clitic combining with the prosodic word in some structure in the vein of Selkirk (1981, 1984), and Nespor & Vogel (2007 [1986]).

Clitics are very similar to affixes in their behavior, but there are a few key differences that differentiate a lexical item from being a clitic or an affix. Zwicky & Pullum claim that several tests are used to differentiate between clitics and affixes. A summary of these tests are given below in (38) which is taken from (Anderson 2005: 33).

(38) Summary of Clitic tests from Zwicky & Pullum (1983) summarized by Anderson (2005)

a. Clitics have a low degree of selection with respect to their hosts; affixes a high degree of selection.

[4] Or as Tyler (2019) calls them "right-branching" for enclitics and "left-branching" for proclitics.

b. Affixed words are more likely to have accidental or paradigmatic gaps than host+clitic combinations.
c. Affixed words are more likely to have idiosyncratic shapes than host+clitic combinations.
d. Affixed words are more likely to have idiosyncratic semantics than host+clitic combinations.
e. Syntactic rules can affect affixed words, but not groups of host + clitic(s).
f. Clitics, but not affixes, can be attached to material already containing clitics.

These tests also reveal that clitics are often associated with functional morphemes and do not bear stress. This is the primary criteria that led researchers to declare that the definite article is a clitic in Welsh (Morris-Jones 1913; Hannahs 2013b). In the following section, I discuss how the prosodic status of clitics is accounted for within this thesis.

3.3 Prosodic Subcategorization

Prosodic subcategorization was developed as a tool to capture the dependence relationships of affixes and clitics (Inkelas 1990, 1993). Inkelas claims that all affixes and clitics have as part of their lexical entries *prosodic subcategorization frames* which focus on the prosodic relationships and are actual pieces of the prosodic structure for these lexical items. This is in contrast to *morphological subcategorization frames* as proposed by Lieber (1980) which focus on—as Inkelas (1990: 78) puts it—"passively constraining the immediate context of the bound morpheme".

One of the motivations for the development of *prosodic subcategorization* over the continued use of *morphological subcategorization* was the unique place of clitics within the grammar. As Inkelas (1990: 80f) says:

> "But unlike roots and affixes, clitics do not require the presence of a morphological sister in order to be morphologically well-formed. Clitics are in fact independent syntactic terminals, while affixes and roots are not...If clitics are not morphologically attached to the preceding word, then how may we characterize their dependence on their host? I suggest that clitics are prosodically bound elements, which do not form independent prosodic

units of their own but must be pronounced in conjunction with another word to produce a well-formed prosodic constituent."

In short, what Inkelas means by this is that clitics form prosodic constituents with their host. These prosodic constituents are represented by prosodic subcategorization frames, which encode the obligatory dependence of clitics on other prosodic constituents.

I follow recent work by Bennett et al. (2018) and Tyler (2019) in assuming the syntactic and prosodic structure correspondence found in Match Theory (Selkirk 2009, 2011). Match Theory is a theory of the correspondence between the syntactic structure and prosodic structure. Match Theory further assumes that Complementizer Phrases (CP) and Force Phrases (ForceP) in the syntactic structure correspond to intonational phrases (ι) in the the prosodic hierarchy. All other syntactic phrases (XP) correspond to phonological phrases (ϕ). Lastly, syntactic heads (X^0) correspond to prosodic words (ω).[5] These basic correspondence relationships are summarized in Table 3.2.

Table 3.2: Correspondence table between the syntactic and prosodic structure (Tyler 2019: 5)

Syntactic (type α)	Prosodic (type π)
CP/ForceP	ι
XP	ϕ
X^0	ω

Each of these relationships is enforced by a correspondence constraint which is evaluated within an OT-grammar. These constraints come in two types which describe either correspondence from the syntactic input structure to the output prosodic structure (i.e., Input-Output correspondence) or the correspondence between the output prosodic structure to the syntactic input structure (i.e., Output-Input correspondence). These constraints are given in Selkirk (2011) which are repeated in (39) below. In these constraints, α represents a syntactic element and π a prosodic element.

[5]Bennett et al. (2016) propose syntactic phrases that are both maximal and minimum (i.e., syntactic phrases that consist of the syntactic head only) correspond to prosodic words instead of phonological phrases, universally. They attribute this to a BINARY constraint that forces phonological phrases to be binary branching. A syntactic phrase that is both maximal and minimal will necessarily violate this if it is a phonological phrase instead of a prosodic word.

(39) MATCH constraints (Selkirk 2011)

a. MATCH(α,π) [= S-P faithfulness]

The left and right edges of a constituent of type α in the input syntactic representation must correspond to the left and right edges of a constituent of type π in the output phonological representation.

b. MATCH(π,α) [= P-S faithfulness]

The left and right edges of a constituent of type π in the output phonological representation must correspond to the left and right edges of a constituent of type α in the input syntactic representation.

These constraints then interact with a series of prosodic markedness constraints such as the constraint BINMIN—which assigns a violation for every instance of a phonological phrase (ϕ) that does not contain at least two prosodic words (ω). By ranking these prosodic markedness constraints above the MATCH constraints the non-correspondence between the syntax and prosodic structures is able to be resolved. This can be seen in the tableau below with a simple syntactic clause as the input.

(40) Interaction of BINMIN and MATCH(XP,ϕ)

$_{\text{clause}}$[[verb [noun]$_{\text{NP}}$]$_{\text{VP}}$]$_{\text{clause}}$	BINMIN	MATCH(XP,ϕ)
a. $_\iota$($_\phi$(verb $_\phi$(noun)$_\phi$)$_\phi$)$_\iota$	$*$!W	L
☞ b. $_\iota$($_\phi$(verb noun)$_\phi$)$_\iota$		$*$

In this tableau, the constraint BINMIN will assign a violation to candidate (a) because the ϕ that contains the noun does not contain at least two prosodic words. By having it ranked above the MATCH constraint this will select as the optimal candidate the candidate that violates this constraint—by "deleting" a set of phrase boundaries.

3.3.1 Subcategorization frames

In recent work by Bennett et al. (2018) and Tyler (2019) the prosodic subcategorization frames from Inkelas (1990) have been incorporated into Match Theory. As was mentioned above, prosodic subcategorization was developed to explain the behavior of affixes and—most importantly for this thesis—clitics with the prosodic structure. In order to capture these relationships Inkelas (1990) proposed prosodic subcategorization. Prosodic subcategorization functions by declaring the prosodic relationships that affixes and clitics have with the rest of the prosodic structure. The following example from Inkelas (1990); Bennett et al. (2018) and Tyler (2019) illustrates what a prosodic subcategorization frame is and how such a frame functions within OT-based frameworks.

In English, adjectives primarily have ante-penultimate stress which may or may not shift depending on what is affixed to the base. For example, the affix *un-* does not cause stress shift to occur on the base. This is captured by the fact that the lexical entry of this affix in the mental grammar includes a subcategorization frame which is read as "the affix *un-* requires that its mother and sister nodes be of the prosodic category of ω, and *un-* must be the left branch." This is shown in (41). It is generally understood that prosodic words already have stress assigned and as such by stating that *un-* takes a prosodic word as its prosodic sister then stress shift cannot happen, which is indeed the case.

(41) Subcategorization frame for *un-*: [$_\omega$ *un-* [$_\omega$ ...]]

  a. [$_\omega$ *un-* [$_\omega$ fínished]]

  b. *[$_\omega$ *ún-* finished]

In contrast, the affix *in-* does cause stress shift to occur when joined to its adjectival base. This affix corresponds to a slightly different subcategorization frame which states: "the affix *in-* requires that its mother node is of category ω and *in-* must be the left branch."

(42) Subcategorization frame for *in-*: [$_\omega$ *in-* [...]]

  a. *[$_\omega$ *in-* [$_\omega$ fínite]]

  b. [$_\omega$ *ín-* finite]

These subcategorization frames have been adopted into constraint-based grammars in the form of the constraint SubCat(X).

(43) SubCat(X): (Tyler 2019: 9)

  Assign one violation for every instance of morpheme X where X's prosodic subcategorization frame is not satisfied.

This constraint assigns violations for every non-compliance to the subcategorization frame for a specific morpheme. These constraints follow McCarthy & Prince (1993a) and Pater (2010) in that constraints can be morpheme-specific.

To illustrate how these constraints function let us consider the prepositional phrase *to Andy*. Following Tyler (2019), the preposition *to* corresponds to the subcategorization frame in (44) which shows that—in English at least—prepositions are associated with subcategorization frames. This frame is read as "*to* requires its mother node to be category ω, and it requires a sister node of any category on its right." This constraint is evaluated in (45).

(44) Subcategorization frame for *to*: [$_\omega$ *to* [...]]

(45) Illustration of SubCat in an OT-grammar (Tyler 2019)

[$_{PP}$ to [$_{DP}$ Andy]]	SubCat(*to*)	MW	MP
a. ($_\phi$ ($_\omega$ to)($_\omega$ Andy))	*!		
b. ($_\phi$ to ($_\omega$ Andy))	*!	*	
c. ($_\omega$ ($_\omega$ to)($_\omega$ Andy))	*!	*	*
d. ($_\omega$ to Andy)		***!	*
☞ e. ($_\omega$ to ($_\omega$ Andy))		**	*

Because SubCat is ranked higher than Match Word (MW) and Match Phrase (MP), any prosodic structure that is not in compliance with the subcategorization frame for *to* will lose even if they fare better in MW or MP. This is what occurs with candidates (a-c) in the tableau in (45). Because the remaining two candidates (d-e) do not violate SubCat the violations in MW and MP are left to determine the optimum leaving candidate (e) as the optimum because of the fewer violations in MW than candidate (d).

3.3.2 Summary

These prosodic subcategorization frames allow for the modeling and capturing of the prosodic relationships that clitics have with their prosodic constituents. This will prove vital to the discussion of the Welsh definite article which is a clitic (Morris-Jones 1913; Hannahs & Tallerman 2006; Hannahs 2013b) in Chapter 4.

3.4 Summary

This chapter has laid out the details of the theoretical frameworks that will be implemented in this thesis. Through Lexical Selection we will be able to capture the definite article and its phonological conditioned allomorphy. This will be thanks to partially-ordered allomorph sets and the constraint PRIORITY.

The addition of lexical subcategorization into Lexical Selection allows for the grammar to account for lexemes to come prespecified to select for various lexical items. These lexical subcategorizations are regulated in the grammar through the morphological constraint RESPECT. Lexical subcategorization will prove vital for accounting for the relationship that Initial Consonant Mutations have with the rest of the grammar.

The use of prosodic subcategorization accounts for the prosodic relationships that exist between clitics and the prosodic structures to which they are bound. This will prove itself essential to capture the relationship that the definite article has with the surrounding lexical items.

CHAPTER 4

Definite Article Analysis

This chapter presents an analysis of the definite article. I will show that in order to account for the behavior of the allomorphs the use of prosodic subcategorization *and* the constraint Priority are required. The use of prosodic subcategorization and its corresponding SubCat constraints are required because of the clitic status of the definite article and its allomorphs. This allows for the prosodic structure that is associated with the allomorphs to be considered during allomorph selection, which has been an issue of previous accounts which has ignored the prosodic structure of the clitic. By allowing the enclitic allomorph to be ranked above the other allomorphs the constraint Priority allows for the enclitic allomorph to surface during allomorph selection.

In Section 4.1, I present my analysis of the definite article making special reference to prosodic structure and the preferential treatment of the enclitic and show that by considering the definite article's prosodic status the analysis is possible within OT. Section 4.2 describes several alternative explanations of the definite article allomorphy with the analysis from Hannahs & Tallerman (2006) in Section 4.2.1. My previous account in Brinkerhoff (2019) which uses only Priority is given in Section 4.2.2. Section 4.2.3 is a consideration of using only morphological subcategorization to explain the definite article allomorphy as proposed by Paster (2015) for all cases of phonologically conditioned allomorphy. A summary of the chapter and key points for further analysis are given in Section 4.3.

4.1 Main proposal

In providing an analysis for the definite article, prosodic subcategorization frames are required in order to account for the clitic-status of the definite article in Welsh. In general, the definite article is proclitic, meaning that it cliticizes onto the immediately following prosodic word (Morris-Jones 1913; Hannahs 2013b). Evidence for this comes from the distribution of schwa in Welsh. As noted by Hannahs (2013b: 38), schwa is never allowed to appear in word-final syllables in Welsh. This is true of all Welsh content morphemes; however, when we consider the functional morphemes in the language, we observe several monosyllabic morphemes which contain a schwa.

(46) Some monosyllabic function morphemes that have schwa (Hannahs 2013b: 38)

dy [də] 'your'

fy [(v)ə] 'my'

yr [ər] 'the'

These morphemes seem to be at odds with the observation made by Hannahs. However, these morphemes are still consistent with the Welsh avoidance of schwa in word-final syllables, because of their status as clitics. These forms were observed as far back as Morris-Jones (1913) as being clitics, which was noted by Morris-Jones (1913) that "many frequently-recurring monosyllables bear no stress, but are pronounced in conjunction with another word. These are proclitics, which precede the accented word, and enclitics, which follow it." This behavior and observation for these function morphemes by Morris-Jones is consistent with the definition of clitics as laid out by Zwicky (1977) and elaborated on by subsequent work on clitics (see Anderson 2005).

According to (Morris-Jones 1913: 48) and Hannahs (2013b),[1] the definite article is a proclitic and forms a prosodic constituent with the following word. We can reasonably conclude that the prosodic subcategorization frame that captures this behavior is in (47).

(47) The prosodic subcategorization frame for the definite article morpheme:

$[_\omega$ Def $[_\omega$ ...]]

This subcategorization frame models the behavior of the definite article morpheme, as a whole. This subcategorization frame states that the definite article takes as its prosodic sister a prosodic word, which has to be on its immediate right, and is a daughter to a prosodic word. The sister to the definite article needs to be a prosodic word because of the lack of stress shift that occurs with these clitics when they are attached to their prosodic hosts as reported by Hannahs (2013b) and in keeping with the discussion of prosodic subcategorization as discussed in Section 3.3. This behavior is similar to that of the English affix *un-* which also does not cause stress shift to occur on the adjectives to which it is is affixed.

[1] Hannahs does mention that at least one of the allomorphs is enclitic rather than proclitic.

(48) Prosodic structure of the Welsh determiner

DP → D° NP: yr afon $\Longrightarrow$ ω_{max} → σ ω_{min}: yr afon

We see that the clitic is associated with a syllable instead of a prosodic word, which was argued for by Itô & Mester (2009a,b) and adopted by Tyler (2019) for monosyllabic English prepositions, auxiliaries and determiners which form recursive prosodic words with their prosodic complement and which likewise applies to Welsh. In the above example for the Welsh DP *yr afon* 'the river' we see that the definite article is associated with a bare syllable and forms a recursive prosodic word with its complement.

As discussed in Section 2.1.4, this structure's syllabification occurs at the maximal prosodic word. This results in the definite article being pronounced as part of its prosodic host as seen in (49).

(49) Syllabification for *yr afon*

[ə.ra.vɔn]

In keeping with the work of incorporating prosodic subcategorization into constraint-based frameworks, the subcategorization frame given above in (47) is associated with the following SubCat constraint:

(50) SubCat(def):

Assign one violation for every instance of the definite article morpheme where the definite article morphemes's prosodic subcategorization frame is not satisfied.

This constraint ensures that the subcategorization frame for the definite article morpheme is satisfied. It will assign violations for any use of the definite article morpheme, regardless of allomorph, that is not proclitic. However, this constraint is only satisfied by two of the three allomorphs of the definite article which are *proclitic*. As was discussed in Section 2.2, the allomorph *'r* is an *enclitic* rather than a *proclitic* and will never satisfy this constraint.

I propose that the solution to the enclitic definite article, *'r*, is the addition of another subcategorization frame and its corresponding SubCat constraint that is specific to the enclitic allomorph of the definite article. This is the first time that *allomorph-specific* prosodic subcategorization has been proposed. The notion that allomorphs are associated with separate subcategorization frames is not a new notion and is, in fact, common in *morphological* subcategorization frameworks (Paster 2015). Within morphological subcategorization, each allomorph is associated with different subcategorization frames. However, as claimed by Paster (2005, 2006, 2009, 2015) this is purely for *affixes*, not clitics. Under this analysis, the enclitic allomorph has a different subcategorization than that of the proclitics, like in (51).

(51) Subcategorization frame for the enclitic allomorph:

[$_\omega$ [...] *'r*]

(52) SubCat('r):

Assign one violation for every instance of the allomorph *'r* where its prosodic subcategorization frame is not satisfied.

Unlike the previous subcategorization frame, this prosodic subcategorization frame states that the enclitic allomorph attaches to the prosodic constituent to its immediate left and together forms a prosodic word. The prosodic constituent to its left is not associated with a specific prosodic category because it has the potential of cliticizing with proclitics that are immediately preceding it (e.g., *a* 'and' which is a proclitic according to Morris-Jones 1913: 48). The prosodic object that the definite article could potentially encliticize to could be of any category.

Unlike the proclitic definite article allomorphs, such as *yr*, shown in (48) and (49), which attach as syllables to their hosts, the enclitic definite article *'r* is what (Tyler 2019: 12) would classify as a "very reduced" clitic. These clitics are not syllables but are instead non-syllabic consonants which attach to the preceding phonological host and become codas in the resulting prosodic word. This means that *o'r afon* would have the phonological structure shown in (53a), and the syllabic structure in (53b).

(53) a. Phonological structure of *o'r afon*[2]

[$_\omega$ [o] 'r] afon

b. Syllabification for *o'r afon*

[ɔr # a.vɔn]

[2]The prosodic status of *o* 'from' has intentionally been left undefined as its prosodic status has not been determined. Some authors claim that most prepositions in Welsh, which *o* is a member, are proclitics (Morris-Jones 1913; Hannahs 2013b).

Thanks to the notions laid out by work by Selkirk (1984); Nespor & Vogel (2007 [1986]) and many others since then in the *Strict-Layering Hypothesis* and more recent approaches to the correspondence between syntax and prosody (see Bennett 2018 and the references given therein for a summary) a given node can only dominate something equal to or lower than itself in the prosodic hierarchy. This then means for the subcategorization frame above in (51), the enclitic can only attach to prosodic words or lower (i.e., feet and syllables).

If we apply these two constraints to the DP *yr afon* 'the river' and *o'r afon* 'from the river' we get the proper allomorphs surfacing with the addition of the constraints Hiatus and NoCoda, which will be illustrated below in (56) and (57). These constraints—especially Hiatus—are required because of the distribution of the definite article's allomorphs which, as was discussed in 2.2, will always avoid instances of vowel hiatus. The constraint NoCoda will be needed when we consider the constructions involving the *gyda*-like morphemes.

In this chapter and subsequent chapters, I will be representing maximal prosodic word boundaries with the symbol # to help illustrate how the allomorphs are prosodically attached. If the # is to the left of the definite article allomorph, it shows that the allomorph is proclitic and a # on the right of the definite allomorph shows that it is enclitic. In keeping with Section 2.1.4 syllabification happens with each maximal prosodic word. Additionally, the definite article is in bold and underlined to facilitate readability.

One additional aspect of the definite article that needs to be addressed is that the definite article is associated with an allomorph set. Allomorph sets for the definite article are required because of the morpheme-specific avoidance of marked phonological structures of hiatus.

As stated by Mascaró (2007) and Bonet et al. (2007), the use of the allomorph sets allows for the distribution of these allomorphs to be derived by universal markedness constraints. This allows the correct form of the allomorph to surface. This also allows the rest of the grammar to interact with these constraints while preserving the input occurrences of hiatus and codas because of higher ranked faithfulness constraints. This is also true for the *gyda*-like morphemes, repeated here as (54).

(54) Functional morphemes that have vowel-/consonant-final alternations (Information collected from King 2015)

gyda/gydag	'with'
â/ag	'with'
tua/tuag	'towards, about'
na/nag	'than'
a/ac	'and'
na/nac	'neither/nor'

As can be seen in the selection of the appropriate allomorph of the morpheme *gyda*, the choice of which allomorph is entirely determined by markedness constraints. No faithfulness violations occur with the use of allomorph sets in keeping with Lexical Selections tenet that "candidate faithfulness violations are computed with respect to the candidate's corresponding underlying allomorph," see (21d) and Section 3.1.

(55) OT/LS output for *gyda/gydag*.

a. *Gyda* in isolation or before a consonant-initial word *gwên* 'smile'

/{gɨda$_1$, gɨdag$_2$} (gwe:n)/	Faith	Hiatus	NoCoda
☞ a. gɨ.da$_1$ (gwe:n)			
b. gɨ.dag$_2$ (gwe:n)			∗W

b. *Gydag* followed by vowel-initial word *eraill* 'others'

/gɨda$_1$, gɨdag$_2$} ɛraɪɬ/	Faith	Hiatus	NoCoda
a. gɨ.da$_1$ ɛ.raɪɬ		∗W	∗
☞ b. gɨ.dag$_2$ ɛ.raɪɬ			∗∗

In accounting for the definite article allomorphy, the selection of the correct allomorph will be dictated by the interaction of SubCat and Hiatus, which ensures that the proper prosodic structure and the avoidance of hiatus are chosen by the grammar.

(56) a. Definite article before vowel-initial words: *yr afon*

/{#ə, #ər, r#} avɔn/	SubCat('r)	Hiatus	NoCoda	SubCat(def)
☞ a. # **ə.r**a.vɔn			∗	
b. # **ə**.a.vɔn		∗!W	∗	
c. **r** # a.vɔn	∗!W		∗	∗W

b. Definite article between two vowels: *o'r afon*

/ɔ {#ə, #ər, r#} avɔn/	SubCat('r)	Hiatus	NoCoda	SubCat(def)
a. ɔ # **ə.r**a.vɔn		∗!W	∗L	L
b. ɔ # **ə**.a.vɔn		∗!∗W	∗L	L
☞ c. ɔ**r** # a.vɔn			∗∗	∗

In tableau (56a), the constraint SubCat('r) eliminates candidate (c) because the enclitic form of the definite article is not obeying its subcategorization frame because it does not have a host to encliticize to. With candidate (c) out of the evaluation, the markedness constraint Hiatus will eliminate candidate (b) for the instance of vowel hiatus. In tableau (56b) we get the motivation for constraint ranking of Hiatus ≫ SubCat(def). With Hiatus ranked above SubCat(def) candidates (a) and (b) both are eliminated. In both tableaux candidate (c) receives a single violation of the constraint SubCat(def), which will assign a violation when the definite article morpheme is not proclitic.

This ranking will not, however, produce the same results when we consider the interaction of consonant-initial words such as the phrase *o'r llyfr* 'from the book.' The reason for this lies in the enclitic *'r*, see tableau (57). In the evaluation of this input with this grammar the enclitic will produce another violation of NoCoda when it is present in a valid environment (i.e., when following the preposition *o* 'from'). In order for the enclitic form to win, the constraint Hiatus has to outrank both of the constraints NoCoda and SubCat(def).

(57) Definite article between a vowel and consonant: *o'r llyfr*

/ɔ {#ə, #ər, r#} ɬɨvr/	SubCat('r)	Hiatus	NoCoda	SubCat(def)
a. ɔ # **ər**.ɬɨ.vɨr		∗!W	∗∗	L
b. ɔ # **ə**.ɬɨ.vɨr		∗!W	∗L	L
☞ c. ɔ**r** # ɬɨ.vɨr			∗∗	∗

If this was not the case than candidate (b) would win because it performs better than candidate (a) in NoCoda violations. Additionally, Hiatus has to be ranked above SubCat(def) because any instance of the enclitic allomorph will automatically result in a violation of this constraint. See Section 2.1.4 for a discussion of why *llyfr* has an additional vowel which is lacking in the input.

In turning to the evaluation of the definite article following the *gyda*-like morphemes, which are also associated with allomorph sets, the selection of the appropriate definite article allomorph is due to the alternation between consonant-final and vowel-final allomorphs in the *gyda*-like morphemes, which is not motivated by a language-wide process but is idiosyncratic to these morphemes, which is the prime criteria for the use of allomorph sets.

In order to get the correct output (i.e., the enclitic form of the definite article) we need to incorporate partially-ordered sets and the constraint Priority into our grammar. The reason for incorporating these into our grammar is to be able to capture the preference that Welsh has for the enclitic definite article as discussed above in 2.2. This preference results in the enclitic definite article allomorph being ranked above the other definite article allomorphs in the partially ordered set as in (58) below.

(58) Priority ranking for the definite article:

{r# $>$ #ə, #ər}

This will result in the enclitic form being selected thanks to the enforcement of this ranking by the constraint Priority during any evaluation of the grammar.

In addition to this allomorph ranking and Priority, a phonotactic constraint is needed. This constraint will assign violations for any syllable-final consonant-cluster involving the Welsh liquids as the final element. This constraint is motivated by the general avoidance of these types of consonant-clusters by the language as discussed above in Section 2.1.4.

(59) $^*CL]_\sigma$:

Assign one violation for every instance of a consonant followed by a liquid in the syllabic coda.

We see in the following tableau that through the incorporation of the constraints Priority and $^*CL]_\sigma$ the correct output is selected by the grammar. This occurs if the constraint Priority is ranked above NoCoda and SubCat(def). This is the only constraint ranking that can be motivated in this tableau. The ranking of $^*CL]_\sigma$ cannot be determined; all that matters is that the constraint is present in the grammar.

(60) Tableau for *gyda'r nod* 'with the aim'

/{gɨda, gɨdag} { r# > #ə, #ər} nɔd/	SubCat('r)	Hiatus	*CL]$_\sigma$	Priority	NoCoda	SubCat(def)
a. gɨ.da # **ər**.nɔd		∗!W		∗W	∗∗	L
b. gɨ.da # ə̲.nɔd		∗!W		∗W	∗L	L
☞ c. gɨ.dar̲ # nɔd					∗∗	∗
d. gɨ.dag # ə̲.nɔd				∗!W	∗∗	L
e. gɨ.dag # **ər**.nɔd				∗!W	∗∗∗W	L
f. gɨ.dagr̲ # nɔd			∗!W		∗∗	∗

In this tableau, the constraints Hiatus and Priority eliminate candidates (a) and (b) because of both the occurrence of hiatus and the use of the non-prioritized allomorphs of the definite article. Additionally, the constraint Priority will eliminate candidates (d) and (e) for the use of the non-prioritized allomorphs. Candidate (f) is eliminated because of the marked phonotactics of the CL cluster that is the result of encliticizing the definite article to the allomorph *gydag*.

The ordering of the allomorphs and the constraint Priority are also needed to produce the correct output of the grammar. Without these additions, the grammar would select candidate (d) as the optimum, which is the other potentially correct form according to Hannahs & Tallerman (2006). This can be seen in the following tableau.

(61) Selection of **gydag y nod* 'with the aim'

/{gɨda, gɨdag} { r#, #ə, #ər} nɔd/	SubCat('r)	Hiatus	*CL]$_\sigma$	NoCoda	SubCat(def)
a. gɨ.da # **ər**.nɔd		∗!W		∗∗	
b. gɨ.da # ə̲.nɔd		∗!W		∗L	
(☞) c. gɨ.dar̲ # nɔd				∗∗	∗W!
☹ d. gɨ.dag # ə̲.nɔd				∗∗	
e. gɨ.dag # **ər**.nɔd				∗∗∗!W	
f. gɨ.dagr̲ # nɔd			∗!W	∗∗	∗W

In this evaluation the subcategorization frame and its corresponding SubCat constraint for the definite article in general will eliminate the correct candidate from winning, which allows the two markedness constraints, Hiatus and NoCoda, to determine the optimum. Candidate (d), which always hamonically bounds the remaining candidate, is selected by the grammar during the evaluation of SubCat(def).

If the ranking that includes Priority is indeed the correct ranking it should still produce the correct output for the forms in (56) and (57), which is indeed the case as can be seen in the following tableaux for *afon*. The tableau in (62a) also serves to show that Priority is dominated by the constraint SubCat('r). By ranking SubCat('r) above Priority this will result in the enclitic form losing during the evaluation even with the prioritization of the enclitic over the other forms. Without this ranking, the enclitic form of the definite article would win. The Hasse diagram showing the complete constraint ranking is given in (63).

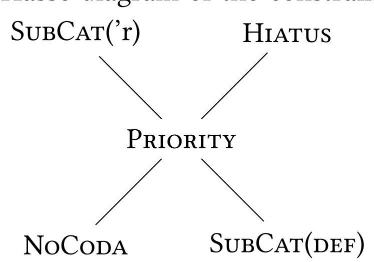
(62) a. Tableau for *yr afon*

/{ r# > #ə, #ər} avɔn/	SubCat('r)	Hiatus	*CL]$_\sigma$	Priority	NoCoda	SubCat(def)
☞ a. # **ə.r**a.vɔn				*	*	
b. # **ə**.a.vɔn		*!W		*	*	
c. **r** # a.vɔn	*!W			L	*	*

b. Tableau for *o'r afon*

/ɔ { r# > #ə, #ər} avɔn/	SubCat('r)	Hiatus	*CL]$_\sigma$	Priority	NoCoda	SubCat(def)
a. ɔ. # **ə.r**a.vɔn		*!W		*W	*	L
b. ɔ # **ə**.a.vɔn		*!*W		*W	*	L
☞ c. ɔ**r** # a.vɔn					**	*

(63) Hasse diagram of the constraint rankings



As was discussed in Section 2.2, there exist environments that suppress the selection of the enclitic definite article allomorph. Hannahs & Tallerman (2006) attribute to suppression of the enclitic allomorph to the presence of an intonational phrase boundary. This boundary does not pose an issue for the analysis presented in this thesis. Following Tyler's (2019) treatment of English enclitics that fail to encliticize, this is due to the lack of any phonological material present within its phonological domain, taken by Tyler to be within the same phonological phrase, that it can select. Because the enclitic is unable to select this material because of the major prosodic boundary, SubCat correctly predicts that the enclitic allomorph will not surface, which then leaves the proclitic allomorphs of the definite article being the only forms that the grammar is able to select from.

In this section, it was demonstrated that the use of subcategorization frames can account for the definite article allomorphy. It was also demonstrated that the interaction of the definite article with the *gyda*-like morphemes can be accounted for through the incorporation of ordered allomorph sets and the constraint Priority and prosodic subcategorization.

4.2 Alternative approaches

The account just presented is one of several potential accounts. My account differs from and is superior to these other accounts in addressing the clitic status of the definite article and the prosodic structure that it is associated with which the other accounts fail to do. The first of these alternative accounts is the account presented by Hannahs & Tallerman (2006). The second account is one that is based on using Priority alone to account for the selection of the correct definite article allomorph. The last account address the use of subcategorization frames in the vein of Paster (2005, 2006, 2009, 2015).

4.2.1 Hannahs & Tallerman's (2006) account

The principle alternative to the account given in this thesis is the one presented by Hannahs & Tallerman (2006). In their account Hannahs & Tallerman take the approach that the definite article is not derived from a single underlying form but instead consists of multiple forms. These forms are then chosen by the grammar based on the phonological environments into which the definite article is inserted. This is similar to the approach taken in this thesis with multiple inputs to the grammar.

(Hannahs & Tallerman 2006: 786) take the view that any purely phonological account is unable to account for the selection of the correct definite article allomorph and that if an optimality theoretic approach is to be taken "something non-phonological" has to be added. This is where there is agreement between our two accounts. Because the selection of the definite article cannot be chosen purely by phonological means on account of the morpho-syntactic considerations at work on the selection of the correct allomorph for the definite article, there has to be "something non-phonological" added to account for this. This "non-phonological" component is the constraint Priority.

Where the divergence between our two accounts occurs is in the implementation of the other grammatical components to account for the definite article allomorphy. In Hannahs & Tallerman, they analyze the selection of the definite article from the Distributed Morphology (Halle & Marantz 1993) framework with serial insertion of lexical items and application of phonological rules at each stage of the derivation. The account that I employ is based on Optimality Theory (Prince & Smolensky 2004 [1993]) and its parallel analysis of the allomorphy with the addition of morphological and prosodic constraints in the form of Priority and SubCat respectively.

Hannahs & Tallerman propose that the selection of the correct allomorph is the result of the staggered insertion of certain lexical classes during the syntactic derivation following the method proposed by Emonds (2002). However, this type of lexical insertion is not the standard approach as exemplified by Embick & Marantz (2008). In the standard approach, lexical insertion occurs from the inside out meaning that the lexical insertion first targets the most deeply embedded nodes in the tree structure, with subsequent lexical items being inserted as the derivation progresses out of the phrase. Emonds (2002), however, proposes that insertion does not occur inside out but in stages based on the semantic features of the lexeme. The first inserted items are any lexical item that is associated with a purely semantic feature f. These items primarily correspond to content words. The second insertion takes lexical items that contain no purely semantic feature, but nevertheless have some "cognitive syntactic feature F i " that assists with interpretation in the Logical Form. The last insertion takes all remaining lexical items which do not contribute to the Logical Form interpretation and inserts them in the Phonological Form. These inserted lexemes primarily correspond to functional words. These three levels of lexical insertion are summarized by (Hannahs & Tallerman 2006: 804) and repeated here as (64).

(64) Emonds's (2002) Stages of Lexical Insertion

a. First level of insertion (Deep Insertion): involves lexical items associated with a purely semantic feature *f*, i.e., open-class items and idioms, which must be inserted at the outset of derivations. This includes nouns, verbs, adjectives and possibly, some prepositions.

b. Second level of insertion (Syntactic Insertion): if an item contains no purely semantic feature *f*, but has ''cognitive syntactic features F_i'' which contribute to interpretation, e.g., *to* as a PATH marker, it is inserted during the syntactic derivation leading to LF.

c. Third level of insertion (Late Insertion): this involves items which lack both purely semantic features and also interpretable instances of F, or else which serve purely as a placeholder for predictable values of F. These are inserted at PF and are ‘‘absent during the derivation from underlying structure to LF’’. An example would be case-assigning *of* in *destruction of the city.*

The reason for Hannahs & Tallerman's choice in Emonds's theory of lexical insertion lies in the interaction of the definite article and the *gyda*-like morphemes. As discussed in 2.2, the definite article always surfaces in its enclitic form while the *gyda*-like morpheme is in the vowel-final form. According to Hannahs & Tallerman, this is due to the order in which these lexical items are inserted during the derivation to phonological form.

To illustrate the derivation Hannahs & Tallerman use the prepositional phrase *gyda'r nod* ‘with the aim’ as seen in (65).

(65) Proposed derivation for *gyda'r nod* ‘with the aim’

Syntactic structure		[P	D	N]
i)	insertion of content word			nod
ii)	Insertion of preposition	gyda	__	nod
iii)	insertion of the article	gyda	'r	nod

According to Hannahs & Tallerman, the first inserted elements are content words comprising nouns and verbs, like *nod* ‘aim’. The next elements are inserted during “syntactic insertion” and correspond to certain prepositions. Hannahs & Tallerman take this to be the point that the *gyda*-like morphemes are inserted into the derivation. Because the immediately following slot for the determiner is at this stage unfilled, the vowel-final form of these morphemes, like *gyda*, are inserted into the preposition slot. The last element to be inserted is the determiner. Because the immediately preceding and following phonological contexts are present the correct allomorph for the definite article is able to selected by the grammar through the elsewhere condition (Kiparsky 1973, 1982), in this case the enclitic allomorph.

The account provided by Hannahs & Tallerman does appear to account for the data concerning the definite article. However,their account still has the issue that it relies on Emonds's (2002) theory of a three stage lexical insertion. The use of Emonds's (2002) theory of a three stage lexical insertion runs contrary to the significant evidence for a view of lexical insertion that does not refer to category distinctions within Distributed Morphology such as Embick & Noyer (2006).

4.2.2 PRIORITY alone

In Brinkerhoff (2019), I argue that the definite article allomorphy can be accounted by relying solely on Lexical Selection and the use of the constraint PRIORITY.

I argued that the selection of the correct allomorph of the definite article and the *gyda*-like morphemes, repeated here as (66) from (10) above , is easily accounted through the mechanism of Lexical Selection in OT.

(66) Functional morphemes that have vowel-/consonant-final alternations (King 2015)

gyda/gydag	'with'
â/ag	'with'
tua/tuag	'towards, about'
na/nag	'than'
a/ac	'and'
na/nac	'neither/nor'

Lexical Selection—as was mentioned in Section 3.1—allows for each allomorph to compete in its own right during EVAL through lexically listing each of the allomorphs into a set. These sets then allow each morpheme to pattern idiosyncratically with respect to the rest of the grammar through markedness violations of the resulting outputs (see further discussion in Bonet et al. 2007; Mascaró 2007, 2008; McCarvel 2016).

As can be seen in the tableaux in (67), which form of *gyda/gydag* is chosen by the grammar is the one that is the least phonologically marked.

(67) Correct OT output for *gyda/gydag*.

a. *Gyda* in isolation or before a consonant

/{*gyda*$_1$, *gydag*$_2$} (C)/	FAITH	HIATUS	NOCODA
☞ a. *gy.da*$_1$ (C)			
b. *gy.dag*$_2$ (C)			∗W

b. *Gydag* followed by vowel-initial word

/{*gyda*$_1$, *gydag*$_2$} *eraill*/	FAITH	HIATUS	NOCODA
a. *gy.da*$_1$*.e.raill*		∗W	∗
☞ b. *gy.da.g*$_2$*e.raill*			∗

Tableau (67a) has *gyda* outputted in order to satisfy NoCoda. In tableau (67b), we see that *gydag* is the output in order to satisfy Hiatus.

The ordering that I took to be present for the definite article is based on the evaluation of simple determiner phrases. I argued that the set given below for the definite article was required in order to get the two proclitic allomorphs to surface when the definite article was pronounced as part of a simple determiner phrase as seen the tableaux below.

(68) Ordering needed for the definite article.

{(*y*, *yr*) > *'r*}

(69) a. *yr* output

/{(y_1, yr_2) > 'r_3} *afon*/	Hiatus	Priority	NoCoda
a. *y_1.a.fon*	∗W		∗
☞ b. *y.r_2a.fon*			∗
c. *'r_3a.fon*		∗W	∗

b. *'r* output

/*o* {(y_1, yr_2) > 'r_3} *afon*/	Hiatus	Priority	NoCoda
a. *o.y_1.a.fon*	∗∗W	L	∗
b. *o.y.r_2a.fon*	∗W	L	∗
☞ c. *o.'r_3a.fon*		∗	∗

This, however, was a mistake on my part as I ignored the clitic status that the definite article had. As was illustrated above in Section 2.2, the prosodic subcategorization associated with the allomorphs will automatically ensure that the two proclitics are chosen whenever the environment for the enclitic is not present.

In attempting to account for the interactions of *gyda*-like morphemes and the definite article, I argued that the morpheme *gyda/gydag* needs to have a revised allomorph set with *gyda* prioritized over *gydag*. Because I encountered the same problem as above in (61) with the harmonic bounding of the other possible output from Hannahs & Tallerman (2006), **gydag y nod*, the grammar was unable to produce the correct output.

(70)

/{gyda$_1$,gydag$_2$} {(y_1, yr_2) > 'r_3} nod/	Hiatus	Priority	NoCoda
a. *gy.da$_1$.y_1.nod*	∗W		∗
b. *gy.da$_1$.yr$_2$.nod*	∗W		∗
(☞) c. *gy.da$_1$r$_3$.nod*		∗W	∗∗W
💣 d. *gy.da.g$_2$y$_1$.nod*			∗

By adding this allomorph ranking to *gyda/gydag*, I was able to get the correct output after adding a second Priority constraint specified to *gyda*.

(71) Correct output for *gyda/gydag*

/{gyda$_1$>gydag$_2$} {(y$_1$, yr$_2$) > 'r$_3$} nod/	HIATUS	PRIORITY$_{gyda}$	PRIORITY$_{yr}$	NOCODA
a. *gy.da$_1$.y$_1$.nod*	*W		L	*L
b. *gy.da$_1$.yr$_2$.nod*	*W		L	*L
☞ c. *gy.da$_1$r$_3$.nod*			*	**
d. *gy.da.g$_2$y$_1$.nod*		*W	L	*L

This analysis was successful to a point but still failed to address the clitic status of the definite article. By ignoring this fact of the definite article and its resultant prosodic structure I produced the wrong results and was not able to present a full analysis of the facts concerning the definite article. This is also the same problem that plagued Hannahs & Tallerman's (2006) analysis. By addressing the clitic status of the definite article in the analysis presented in Section 2.2 above all of the facts of the definite article can be accounted instead of focusing on only a few aspects and ignoring the rest as I did in Brinkerhoff (2019).

4.2.3 Subcategorization

Paster (2006, 2009, 2015) argues that all instances of PCSA reside solely in the morphological component of the grammar, because any phonological effects that are observed are the result of the morphology selecting phonological elements in the stem. She further argues—based on a survey of 67 languages—that PCSA is best explained through *morphological* subcategorization frames (Lieber 1980) instead of a model of OT where P(honological) constraints are ranked above M(orphological) constraints as proposed by McCarthy & Prince (1993a,b). Paster styles this approach as "P $\gg$ M" and explicitly states that all constraint-based accounts falls under this heading regardless of their approach to morphology.

In comparing these two approaches, Paster outlines four predictions that each of these systems make. These predictions are laid out in Table 4.1 with the results of the two approaches, subcategorization and "P$\gg$M".[3] These same predications are evaluated against the approach taken in thesis. It will be shown that this approach conforms to the "P$\gg$M" predictions and offers support for this approach instead of subcategorization only.

[3]The results listed in the table for subcategorization are based on her survey of 67 languages (Paster 2009, 2015).

Table 4.1: Predictions and results of the two approaches to PCSA

Prediction	**P≫M**	**Subcategorization**
a. PCSA is always optimizing	Yes	No
b. Input or Output- conditioned	Output	Input
c. Directionality of conditioning	Any	Inside-out only
d. Adjacency required	No	Yes

According to (Paster 2015: 221f), "P≫M" makes the prediction that all cases of PCSA are optimizing. This is taken by Paster as meaning that universal phonological constraints optimize words with respect to a given phonological constraint. Whenever there are cases that do not appear to be "obviously optimizing" then the allomorphy is derived from the interaction of different morphological constraints with those phonological constraints. This thesis's approach agrees with the "P≫M" approach, because the selection of the correct allomorph is determined by the interaction of the phonological constraints during Eval, which results in an optimal syllable structure.[4] The only instance when a morphological constraint—which is represented by the constraint Priority[5]—is required in interaction with the *gyda*-like morphemes. Without the constraint Priority, the incorrect form is selected by the grammar as can be seen in the tableau in (61). This is in agreement with the first predication made by the "P≫M" model.

This in turn also supports the "P≫M" model's second prediction on output-conditioning. Because the selection of the correct DA allomorph is determined by a competition between different output forms (i.e., potential surface forms), the correct output candidate is the most optimal according to the OT grammar. This is in contrast to the subcategorization approach in which all forms of PCSA are entirely determined in the *morphology* prior to being input to the phonological component where the regular phonological operations are then performed (Paster 2015: 224).

The third prediction for the "P≫M" model states that the directionality of the allomorph selection can occur in any direction, which is what is observed in the case of the allomorphy for the Welsh definite article. The facts of the Welsh definite article clearly show that this is *not* inside-out conditioned. Inside-out conditioning means that the allomorphy of an element is conditioned by the

[4] This is noted by Mascaró (2007) to be one of the hallmarks of Lexical Selection.

[5] Priority is a departure from the traditional account of morpheme-specific constraints, where constraints introduce morphs (e.g., Kager's (1996) Genitive = /-n/ constraint for Djabugay). This is noted because that approach has the consequence of excluding any allomorphic variation of the morph, because it forces the phonetic form that is found in the constraint (Mascaró 2007: footnote 14).

element that is more deeply embedded within the syntactic phrase structure. For example if there is a prepositional phrase that contains a determiner phrase which in turn contains a noun phrase then the noun phrase is the element that conditions the allomorphy of the determiner which then would condition the allomorphy of the preposition. Considering just stems and affixes, inside-out conditioning assumes that the stem conditions the affix.

If the allomorphy of the definite article was inside-out then it would result in the wrong allomorphs surfacing for the definite article and the *gyda*-like morphemes. This would result in the definite article not undergoing encliticization with the *gyda*-like morphemes which in turn would have the consonant-final form.

(72) Inside-out conditioning stages: **gydag y nod*

Syntactic structure		[P	D	N]
i)	insertion of content word			nod
ii)	conditioning of the article		y	nod
iii)	conditioning of the preposition	gydag	y	nod

Because the actual form is not *gydag y nod* but *gyda'r nod* it cannot be inside-out conditioned. At the same time it is not outside-in conditioned, which is the exact inverse of inside-out conditioning, because the most external element also does not condition the allomorphy of the internal elements. In the case of Welsh the conditioning happens in parallel, where both the internal and external elements condition the allomorphy of the definite article.

The final prediction for the "P≫M" approach is that the conditioning for allomorphy can come from anywhere in the word and the conditioning is not necessarily required to be adjacent (e.g., an element on the left-edge of a word can condition the allomorphy of a suffix). According to Paster (2009), these long-distance conditioning effects are not found in any example from her survey where all examples are conditioned by immediately adjacent elements. She further claims that these long-distance effects do not actually exist and that "the P≫M approach overgenerates." From the Welsh data, the conditioning for the definite article allomorphy is conditioned on the immediately preceding and following contexts which shows support for the subcategorization account. However, this does not mean that all PCSA accounts must be adjacent. Other accounts could potentially make use of long-distance conditioning.[6]

Table 4.2: Welsh Data against the two approaches to PCSA

Prediction	**P≫M**	**Subcategorization**	**Welsh Data favors**
a. PCSA is always optimizing	Yes	No	P≫M only
b. Input or Output- conditioned	Output	Input	P≫M only
c. Directionality of conditioning	Any	Inside-out only	P≫M only
d. Adjacency required	No	Yes	Both

The results of the the Welsh data are summarized in Table 4.2. The Welsh data offers support for the "P≫M" approach and is a counter example to most of the claims made by Paster (2005, 2006, 2009, 2015) that PCSA is best analyzed in a morphological subcategorization approach, because the data and analysis presented in this thesis matches the "P≫M" approach in three of the four predictions. Regarding the last prediction, the Welsh data is compatible with both approaches, on this criteria. However, this is only due to the fact that the selection of the definite article is dependent on the immediately preceding and following contexts. Lexical Selection additionally does not claim that the selection of the correct allomorph is conditioned by the immediate context but instead by the interaction of phonological constraints.

4.3 Summary

This chapter has shown that the allomorphy of the definite article can be modeled through the use of prosodic subcategorization frames and ordered allomorph sets. This is possible because of two reasons; first, the clitic status of the definite article requires the consideration of the prosodic structure associated with these lexical items. Second, through the use of ordered allomorph sets the preferential treatment of the enclitic allomorph of the definite article is accounted for in those situations that would otherwise seem to warrant a more phonologically well-formed output of the grammar. This chapter also introduces the expansion of prosodic subcategorization to include *allomorph-specific* prosodic subcategorizations for clitics, like in the case of Welsh where one of the allomorphs behaves differently than the other allomorphs in their cliticization. This analysis of the definite article will

[6] An example of a potential long distance conditioning in the allomorphy of the suffix *-alis/-aris* in Latin where the choice of which allomorph is conditioned by the presence of a lateral anywhere in the stem. For example, *lunaris* 'of the moon', *militaris* 'of a soldier' vs. *dualis* 'of the number two', *regalis* 'of the king'.

become crucial to later discussion of the interaction of the definite article and initial consonant mutations in Chapter 5.

CHAPTER 5

Initial Consonant Mutation

This chapter explores initial consonant mutation within the framework of Optimality Theory (OT; Prince & Smolensky 2004 [1993]). This is accomplished through a synthesis of two other accounts: Green's (2006) lexical listing approach which accounts for initial consonant mutations within OT and Hannahs's (2013a) account of pattern extraction which models the broader language-wide generalizations and patterns that are observed in the different consonant mutations. This chapter will show that through the mechanisms found in Lexical Selection and its lexical listing with allomorph ordering and lexical prespecification (Bonet et al. 2007; Mascaró 2007) the selection of the correct output candidate will be possible. Through lexical listing of the the mutation patterns of each radical, the correct form is selected through morphological constraint interactions. The use of lexical prespecification is required to allow the lexical triggers to select for the mutation specified as part of their lexical entry. My analysis of initial consonant mutation is presented in Section 5.1 followed by an analysis of the interaction of initial consonant mutation and the definite article allomorphy in Section 5.2. These analyses are followed by a discussion of several alternative analyses for initial consonant mutations and one alternative analysis for the interaction of initial consonant mutations and the definite article allomorphy in Section 5.3. A summary of the chapter is given in Section 5.4.

5.1 Initial Consonant Mutation Analysis

As stated in Chapter 2, initial consonant mutations involve a subset of the consonants in Welsh. Because each consonant is associated with a specific set of mutation consonants and patterns, they can be considered to form a set from which the correct form is selected. For example, all words that begin with /t/ follow the same pattern of phonological reflexes for the different types of mutation with soft mutation represented by [d], aspirate mutation by [θ], and nasal mutation as [n̥].[1] An example of this pattern is seen in Figure 5.1 with the word *tad* 'father'.

According to Hannahs & Tallerman (2006), this pattern is extracted and stored within the lexicon of the speakers for each sound, see Figure 5.2. The correct form is extracted by means of subcategorization relationships. In adopting this behavior into an OT grammar we can assume that each of

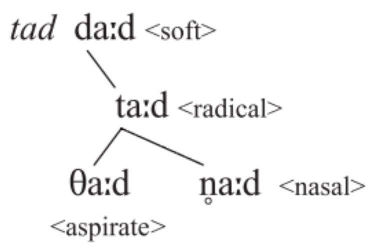


Figure 5.1: *t*-initial pattern association (morpheme)

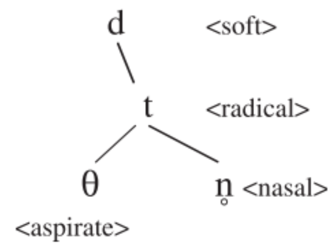


Figure 5.2: *t*-initial pattern extraction (segment)

these patterns is represented in the lexicon of a Welsh speaker as a set similar to the allomorph sets of Lexical Selection. These pattern sets consist of each radical consonant and its mutated forms. The radical form would receive priority, in keeping with Lexical Selection, to ensure that it is chosen by the grammar when the conditions for the different mutations are not present to force the selection of one form over another. These pattern sets then exist for each of the different mutations in Welsh, as shown in Table 2.3, which is repeated here for ease of reference. An example of how these pattern

Table 5.1: Welsh Mutations table

Radical	Soft	Nasal	Aspirate
p [p]	b [b]	mh [m̥]	ph [f]
t [t]	d [d]	nh [n̥]	th [θ]
c [k]	g [g]	ngh [ŋ̊]	ch [χ]
b [b]	f [v]	m [m]	
d [d]	dd [ð]	n [n]	
g [g]	– zero	ng [ŋ]	
m [m]	f [v]		
ll [ɬ]	l [l]		
rh [r̥ʰ]	r [r]		

Table 5.2: Mutation Subscripts

Mutation Type	Subscript
Radical	–
Soft Mutation	o
Nasal Mutation	n
Aspirate Mutation	h

sets are represented is given below with each of the different mutations marked with a subscript, as summarized in Table 5.2.

[1] In Southern Welsh varieties that lack [h], the voiceless nasals are not present, but are instead represented by their voiced nasal counterparts. This is, however, not reflected in the orthography.

(73) Pattern set for /t/

$\{t > d_o, \text{n̥}_n, \theta_h\}$

This set is then present as the first segment of the lexical entry for any word that begins with /t/ in the radical as can be seen in (74) for the Welsh word *tad* 'father'.

(74) Lexical entry for *tad* [tad] 'father'

$\{t > d_o, \text{n̥}_n, \theta_h\}$ad

Because the triggers for the different mutations are now various lexical items instead of phonological environments those lexemes must have a feature or a lexical prespecification (see Section 3.1.3) which lexically subcategorizes for the mutation it governs. As mentioned above, each mutation is indicated by a subscript which was summarized in Table 5.2. These subscripts are found on both the mutations as well as the mutation trigger within this formalism to illustrate which mutation is found with each of the different lexical triggers. For example, the Welsh word *dy* [də] 'your' will have a lexical prespecification for soft mutation which will then subcategorize for that given mutation.

(75) Prespecification input to an OT grammar for *dy* [də] 'your'

/ də_o /

This input will then require some morphological constraint that enforces the subcategorization found with this lexeme. This is accomplished through Bonet et al.'s (2007) constraint Respect, which was given above in Section 3.1.3 and is repeated here.

(76) Respect:

Assign one violation for every mismatch between the lexical subcategorization in the input and the output.

The constraint Respect ensures that any lexical subcategorization that is specified in the input, represent by subscripts on the lexeme, will be present in the output. This is accomplished by assigning violations whenever the subcategorization for a lexeme is not respected or in other words when the subcategorization is not matching between the lexical trigger and the different mutated forms in the output. We can see how this constraint assigns violations and its interaction with the constraint Priority in the tableaux in (77).

(77) Tableaux showing the constraint Respect interacting with the grammar.

a. Soft mutation specification *dy dad* 'your father'

/də$_o$ { t > d$_o$, n̥$_n$, θ$_h$ }ad/	Faith	Respect	Priority
a. də$_o$ tad		∗!W	L
☞ b. də$_o$ dad$_o$			∗
c. də$_o$ n̥ad$_n$		∗!W	∗
d. də$_o$ θad$_h$		∗!W	∗

b. No specification *y tad* 'the father' (behavior of the definite article simplified)

/ə { t > d$_o$, n̥$_n$, θ$_h$ }ad/	Faith	Respect	Priority
☞ a. ə tad			
b. ə dad$_o$			∗!W
c. ə n̥ad$_n$			∗!W
d. ə θad$_h$			∗!W

In tableau (77a) we see that a higher ranked Respect constraint insures that the subcategorization specified by the lexeme /də$_o$/ occurs. When no such subcategorization is specified Priority forces the allomorph *tad* to appear as the conditioning for the other forms are not met and when there is no higher ranked phonological markedness constraint. If however Priority is ranked above Respect any specification in the input would be erased.

Similar to the selection of the definite article, the candidates do not receive faithfulness violations, because violations are only computed with respect to the candidate's corresponding underlying allomorph.

As noted above in Section 2.3, there exists lexical items that are impervious to mutation. This is attributed to the lack of any of these mutation pattern sets as part of their lexical entry. We can see this in the word *gêm* 'game' which is resistant to mutation.

(78) *y gêm* 'the game'

/ə$_o$ gem/	Faith	Respect	Priority
☞ a. ə$_o$ geːm		∗	
b. ər$_o$ eːm	∗!W	∗	

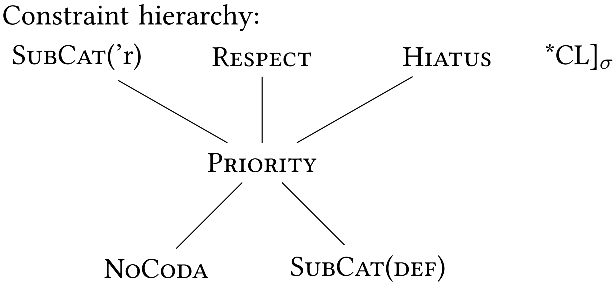
Because the word *gêm* is not associated with any lexical listing then any change will automatically violate a faithfulness constraint. In tableau (78) it is impossible to satisfy the constraint Respect because there is no pattern set associated with this lexical entry which would allow for the lexical subcategorization that the definite article is able to select. If the input where to show something analogous to soft mutation than the input would need to undergo deletion of the initial /g/ which would receive a violation of Faith. Because of the faithfulness constraints being ranking above that of Respect and Priority and the lack of the pattern set being listed of the mutated forms these words are resistant to initial consonant mutation.

5.2 Initial Consonant Mutation and the Definite Article Analysis

As previously mentioned, this analysis is a combination of the analyses in the previous chapter and the analysis above. As was previously discussed in Section 2.4, the definite article causes soft mutation to occur on feminine singular nouns only. In order to specify that mutation occurs with feminine singular nouns, we have several options for the lexical prespecification on the definite article. Either there exists in the lexicon a separate lexical item of the definite article that is specific for feminine singular nouns, which is entirely distinct from the definite article for masculine and plural nouns or the specification on the definite article has specific reference to feminine as well as soft mutation. I will not take a position on these options at this time; for simplicity I will show a specification for soft mutation and assume it is a separate lexical item that is the feminine definite article.

Because several constraints that were not present in Chapter 4 were added in this chapter they need to be incorporated into the previous constraint hierarchy that was needed to account for the definite article allomorphy. This will produce the new hierarchy as found in (79) below.

(79) Constraint hierarchy:



This constraint ranking produces the following tableaux in (80) and shows that the correct optima are selected by the grammar. The constraint *CL]$_\sigma$ was removed from the tableaux for readability and because it does not play a role in the selection of the correct optimum.

(80) Behavior of the Definite Article and feminine nouns

a. Tableau for *yr ardd* ‘the garden’

/{ r# > #ə, #ər}$_{o.fem}$ { g > $\emptyset_o$, ŋ$_n$}arð/	SC(’r)	Hiatus	Respect	Priority	*Coda	SC(def)
a. # ə$_o$.garð			*!W	*	*	L
b. # ə$_o$.arð$_o$		*!W		*	*	L
☞ c. # ə.r$_o$arð$_o$				*	*	
d. r$_o$ # arð$_o$	*!W			*	*	*W
e. # ə$_o$.ŋarð$_n$			*!W	*	*	L

b. Tableau for *y lasog* ‘the gizzard’

/{r# > #ə, #ər}$_{o.fem}$ { g > $\emptyset_o$, ŋ$_n$}lasɔg /	SC(’r)	Hiatus	Respect	Priority	*Coda	SC(def)
a. # ə$_o$.gla.sɔg			*!W	*	*	
☞ b. # ə$_o$.la.sɔg$_o$				*	*	
c. # ər$_o$.la.sɔg$_o$				*	**!W	
d. r$_o$ # la.sɔg$_o$	*!W			*	*	*W
e. # ə$_o$.ŋlasɔg$_n$			*!W	*	*	

In both tableaux, the constraint SubCat(’r) will eliminate the candidates that have the enclitic definite article allomorph because their prosodic subcategorization is not honored, in this case candidate (d) in both tableaux. In tableau (80a), Hiatus eliminates candidate (b) from the evaluation. This leaves the constraint Respect to eliminate candidates (a) and (e) for not respecting the lexical specification found on the definite article. This then leaves candidate (c) to become the optimum, which is indeed the correct output in Welsh.

The behavior of the constraint Respect in tableau (80b) also eliminates candidates (a) and (e) for the same reasons as mentioned above for tableau (80a)’s candidates, leaving only the candidates that have soft mutation. This results in the candidates (b) and (c) competing against one another. Candidate (c) is eliminated from the evaluation for the higher number of violations in NoCoda than the optimum.

This analysis shows that the interaction of definite article allomorphy with initial consonant mutation can be accounted for by incorporating the analyses from the previous chapter.

5.3 Alternative Analyses

Because initial consonant mutations have had a long history of various accounts that attempted to account for their patterns, I will focus on only three alternative accounts, which are relevant for the account that I present in this thesis. I refer interested readers to Hannahs (2011) and his discussions of earlier accounts of initial consonant mutations that are not discussed in this Section.

The first alternative that I discuss is Kibre's (1995) floating feature approach in Section 5.3.1. This is followed by Green's (2006) lexical listing approach in Section 5.3.2. In Section 5.3.3, Hannahs's (2013a) pattern extraction approach is discussed. These accounts are concerned with how to model initial consonant mutations. Section 5.3.4 offers Hannahs & Tallerman's (2006) account of the interaction of the definite article with initial consonant mutations.

5.3.1 Kibre's (1995) Floating Features

Kibre's (1995) account of initial consonant mutation is concerned with accounting for initial consonant mutation within the framework of Lexical Phonology especially with morphological rule-triggering as proposed by Mohanan (1982, 1986). Kibre also explores the possibility of using Optimality Theory with floating segments to account for Welsh initial consonant mutations.

Under this approach, the different mutations are governed by the interaction of IDENT and PARSE-FEATURE constraints interacting with floating features in the input to the OT grammar following the use of latent features and segments as proposed by Zoll (1996). This floating feature in the input then interacts with three different constraints that are defined below in (81).

(81) Constraints for Kibre's OT grammar:

a. IDENT(-voi, cont):

Assign a violation for every segment that belongs to the natural class [+voi], which has a different underlying and surface specification for [continuant].

b. IDENT((+cont, -son, -lat), voi):

Assign a violation for every segment that belongs to the natural class [+cont, -son, -lat], which has a different underlying and surface specification for [voi].

c. Parse-Feat:

Assign a violation for every floating feature that is not realized.

For example, under Kibre's proposal the word *ei* 'his' is represented in the input with the phone /i/ and the floating features <+voi, +cont>, because these are the features that are associated with soft mutation according to Kibre. This input then interacts with the constraints defined above as seen in the following tableaux which are reproduced here from Kibre (1995).

(82) a. Tableau for *ei ben* 'his head' (Kibre 1995: 39)

/i-<+voi, +cont> + pen/	Ident(-voi,cont)	Ident(+cont..., voi)	Parse-Feat
a. i <+voi, +cont> pen			**!W
☞ b. i <+cont> ben			*
c. i ven	*!W		L

b. Tableau for *ei ffrind* 'his friend' (Kibre 1995: 39)

/i-<+voi, +cont> + frind/	Ident(-voi,cont)	Ident(+cont..., voi)	Parse-Feat
☞ a. i <+voi> frind			*
b. i vrind		*!W	L

In tableau (82a), the constraint Ident(-voi,cont) will eliminate candidate (c) which has fully parsed the floating features, because the underlying /p/ changes its continuant feature. This then leaves candidates (a) and (b) which do not fully parse the floating features. Candidate (b) ultimately wins because it has parsed the [+voi] feature from the input while the fully faithful candidate loses on violations of Parse-Feat. Tableau (82b) illustrates what occurs when there is no mutation. The constraint Ident((+cont, -son, -lat), voi) will assign a violation to candidate (b) because the /f/ has become voiced from the parsing of the floating feature [+voi]. Because the underlying /f/ is already [+cont] that floating feature automatically is parsed into the segment.

The primary issue with Kibre's account is his evaluation of initial consonant mutations from a purely phonological perspective. This is not feasible, especially for the soft mutation which requires different types of changes for different classes (see Section 2.3 for a review of the types of changes). This is because of the behavior of certain lexical exceptions to initial consonant mutation for which Kibre does not provide an account. If Kibre's account of initial consonant mutation is adopted as is, then these lexical exceptions should undergo assimilation to these floating features. However, this is not what is observed in Welsh. In contrast, lexical exceptions to mutation are predicted in my approach because Priority and Respect are inherently morpheme-specific constraints

Another issue is that Kibre rejects this proposal in favor of a ruled-based account in Lexical Phonology. According to Kibre this is because the OT account is a "less direct way of expressing the facts of mutation than rules…[and] this model has little to recommend it over rule-based approaches."

The difference between my approach and Kibre's lies in which component of the grammar is responsible for the initial consonant mutations. For Kibre, the mutations are part of the phonological component whereas my account assumes that the mutations are part of the morphological component instead. This is because mutations reside in the lexicon. Evidence for this comes from the the presence of the lexical exceptions as mentioned previously. If initial consonant mutations were purely phonological than these lexical exceptions should not exist. In fact this is the approach that many authors have taken in recent years. Two of these accounts are are given in Sections 5.3.2 and 5.3.3 of this thesis.

5.3.2 Green's (2006) Lexical Listing

Another OT approach to initial consonant mutation is the one presented by Green (2006, 2007). The approach that Green takes is similar to the one that I present in this thesis. Similar to my approach, Green claims that the initial consonant mutations in the Celtic languages are not purely phonological but instead belong to the morphological component of the grammar. Green claims that initial consonant mutations are regulated by grammatical subcategorization relations found in the lexical entry for each lexeme that selects for a specific type of mutation (i.e., the lexemes are lexically prespecified for the type of mutation that is present with the lexeme). However, contrary to the approach taken in this thesis, Green also assumes that all lexemes have these grammatical subcategorizations whether or not they are associated with a specific mutation. If the word does not select for a specific type of mutation, then it has a subcategorization for the radical form.

(83) Lexical specification in the masculine and feminine definite article in Welsh (Green 2006)

a. Masculine definite article

$ə_{+\textsc{rad}}$

b. Feminine definite article

$ə_{+\textsc{len}}$

The biggest differences between our approaches is how Green accounts for the initial consonant mutations. Green takes the approach that every lexeme in the lexicon of a speaker is associated with the full lexical listing of the lexeme with each of its mutation forms in a set that contains the full phonetic specification for each form. An example of his sets is found in the following example.

(84) Lexical listing according to Green (2006) for *tad*.

{ tad_r, dad_o, n̥ad_n, θad_h }

In order to regulate the choice of which form is chosen Green makes use of the constraint MUTATIONAGREE which functions identically to the constraint RESPECT as proposed by Bonet et al. (2007) and used above in this thesis. Instead of respecting all subcategorizations like the constraint RESPECT, Green's MUTATIONAGREE constraint is only concerned with respecting those subcategorizations associated with initial consonant mutations. This constraint is given below.

(85) MUTATIONAGREEMENT (MUTAGREE):

Assign one violation mark for every instance of the mutation required by the trigger not agreeing with the mutation of the target

This constraint will then assign violations for any mutation or radical that does not match the subcategorization found on the lexical trigger. An example of this is given below. In (86), Green treats the definite article as two separate lexical entries: one is associated with feminine singular nouns and a second is associated with the masculine singular and plural nouns regardless of gender. These two entries then each have different subcategorizations with the feminine singular definite article subcategorizing for soft mutation and the other definite article that subcategorizes for radicals.

(86) Tableaux showing initial consonant mutation (Green 2006: 1977f)

a. Tableau for the masculine and the radical: *y bardd* 'the bard'

/$\text{ə}_{+\text{RAD}}$ { barð_{RAD}, varð_{SM}}/	MUTAGREE	IDENT(cont)
☞ a. ə barð		
b. ə varð	∗!	

b. Tableau for the feminine and soft mutation: *y faner* 'the flag'

/$ə_{+SM}$ { $baner_{RAD}$, $vaner_{SM}$}/	MutAgree	Ident(cont)
☞ a. ə barð	*!	
b. ə varð		

However, Green's approach makes too many assumptions about these lexical prespecifications. This is particularly true in that he assumes that all words include these prespecifications. These prespecifications also enforce the selection of the radical. My approach is superior through the use of Priority and the ordering of the radical over the other mutated forms. By ranking the radical over the mutated forms, the grammar can select the correct form when there is no lexical prespecification. Additionally, Green's approach places too much emphasis on the lexicon by assuming that full lexical listing occurs for every entry. This misses a crucial generalization that each mutation is associated with a broader language-wide pattern for each mutation. My approach captures this broader pattern by placing each mutation pattern into a set. This allows for a simplification in the lexicon of the speaker over Green's approach.

5.3.3 Hannahs's (2013a) Pattern Extraction

The last alternative concerning the modeling of just the initial consonant mutation that will be considered is the one put forward by Hannahs (2013a). This approach is based on the assumption that initial consonant mutations in all of the Celtic languages are represented in the mental lexicon of speakers as a series of associations which relates the radical consonant to its associated mutations. Which of the mutated forms of the radical is allowed to surface is determined through subcategorization, and then that form is extracted from the associations based on the patterns in the language in the morphological component of the grammar only. According to Hannahs this approach is superior to full lexical listing as it only requires the specific patterns for each radical consonant to be stored in the mental lexicon of the speaker, whereas full lexical listing would require that each word would be listed with each allomorphic form listed in the lexicon with its full phonetic specification like the approach taken by Green (2006, 2007).

In regard to subcategorization, Hannahs (2013a) follows the morphological subcategorization schema as proposed by Lieber (1980) and used by subsequent researchers like Paster (2015). For example, each lexical entry according to Lieber's morphological subcategorization is associated with specific subcategorization frames that specify the type of material it selects. This material could include information that is syntactic, morphological, or phonological. In Hannahs's account, he provides an example of how the lexical entry for a word that selects for a specific type of mutation would appear. This is given below.

(87) Lexical entry fot the preposition *â* 'with' (Hannahs 2013a: 13)

â: Prep, [___ $\underset{\text{<aspirate>}}{\text{X}}$]$_{\text{PP}}$

The lexical entry for this word consists of the word itself, its lexical category and finally its subcategorization frame which also specifies what type of syntactic constituent is created. This word will then case all complements to show aspirated mutation. The native speaker would then extract from those patterns that are stored in their lexicon for each of the radicals and their associations the correct form.

This alternative theory does seem to present an improvement over other approaches in reducing the load that is stored in the mental lexicon to these patterns instead of the full lexical listing of the mutation forms. This approach is the basis of the representational aspects of the account that I present in this chapter. However, I assume that these patterns are attached to each lexical listing and those lexemes that lack mutation do not have these pattern sets, like *gêm*. This is contrary to Hannahs's approach that does not provide an account for these lexical exceptions except to state that they exist and that they are resistant to mutation. My account does address the issue that these lexical exceptions represent for a model of pattern extraction. This is accomplished by not allowing the mutation pattern sets to exist as part of the lexical entry for those lexemes that lack initial consonant mutations. This then allows the grammar, which is modeled within OT in my account, to correctly predict and model the behavior of the lexical exceptions and those lexemes that undergo mutation through lexical subcategorization.

5.3.4 Hannahs & Tallerman (2006)

The principal alternative analysis for the interaction of the definite article and the initial consonant mutation is the analysis presented in Hannahs & Tallerman (2006). As was mentioned in Section 4.2.1, their approach to the situation relies on Distributed Morphology and its use of lexical insertion.

However, they are only able to explain what is occurring through the reliance on Emonds's (2002) approach which states that the insertion of the lexical items is staggered, with content morphemes being inserted first followed by the insertion of functional morphemes that serve some "cognitive syntactic" purpose. Only after the lexical insertion of these two previous stages is the definite article inserted during the final stage of insertion.

In order to account for the fact that the definite article needs to know the final phonological shape of its complement in order to select the correct allomorph Hannahs & Tallerman propose that initial consonant mutation occurs as a rule during one of the steps before the definite article is inserted. The rule that Hannahs & Tallerman propose is given below in (88).

(88) Mutation trigger by an empty D

$[\_\_]_D$ [gardd $]_N \rightarrow [\_\_]_D$ [ardd $]_N$

This rule assumes that some point during the derivation there is an empty determiner position followed by the noun and the mutation is triggered by the unfilled determiner. This then will result in mutation occurring on the noun which then becomes the phonological conditioning for which allomorph is inserted into the derivation. The definite article will then select the correct allomorph based on the elsewhere principle from Kiparsky (1973, 1982) based on a series of rules similar to the phonological environments given in (8) and then the correct form of the definite article is inserted as in (89) below.

(89) Proposed derivation for *yr ardd* 'the garden'

Syntactic structure		[D°	N°]
i)	insertion of content word	__	gardd
ii)	triggering of mutation	__	ardd
iii)	insertion of the article	yr	ardd

In this example of the proposed derivation for *yr ardd* 'the garden,' the content word is inserted first into the syntactic structure. This then will cause the rule given in (88) to apply because of the noun is following an empty determiner head. The rule will cause the word *gardd* to undergo soft mutation resulting in the soft mutated form *ardd.* This then is sent to step three of the derivation which is the insertion of the definite article. At this point, the definite article will consider which of its allomorphs is most appropriate according to the elsewhere principle. This analysis still runs into the same problem as was previously mentioned in Section 4.2.1 which is the use of Emonds's (2002) three levels of lexical insertion.

5.4 Summary

In this chapter, I illustrated that through lexical listing of the mutation patterns and lexical subcategorization, the facts of initial consonant mutation, in addition to its interaction with the allomorphy of the definite article, can be modeled within Optimality Theory. The lexical listing of the different mutation patterns into sets allows lexical triggers for mutation to select the correct mutation. Additionally, lexical ordering allows for the radical to be selected when there is not a lexical trigger present to select for a given mutation. This analysis also provides an account for the lexical exceptions to initial consonant mutations. This is due to those lexical items lacking the mutation pattern sets in their lexical entry, which is then input to the OT grammar resulting in the fully faithful candidate being selected by the grammar. Any change to these lexical exceptions, to make them appear as if they underwent initial consonant mutation, results in faithfulness violations. This is in contrast to most accounts of initial consonant mutations where these lexical exceptions are ignored or relegated to a footnote without an attempt for analysis.

CHAPTER 6

Conclusion

This thesis presented data concerning the definite article and initial consonant mutation in the Welsh language in Chapter 2. In Chapter 4, the allomorphy of the Welsh definite article was analyzed by considering the clitic status of the definite article by means of prosodic subcategorization.

This thesis provided support for the expansion of prosodic subcategorization frames to allomorphs of clitics. This was necessary for the definite article because its allomorphs are associated with different cliticization patterns. This expansion allows for constraint interactions in the grammar to explain the various prosodifications of multiple allomorphs, which provided an advantage over other accounts on the definite article by addressing the prosodification.

This thesis also explored the patterns of initial consonant mutations in Welsh in Section 2.3. Even though these initial consonant mutations appear to be phonologically driven, initial consonant mutations are caused by lexical item triggers and morpho-syntactic structures. In chapter 5 of this thesis, I presented an Optimality Theoretical account for Welsh initial consonant mutations. I argued that initial consonant mutations are part of the mental lexical entry for each radical and can be explained through lexical subcategorization.

Chapter 3 explained the details of the theoretical frameworks that were implemented in this thesis. Through Lexical Selection, the definite article and its phonologically conditioned allomorphy can be accounted. This was due to partially-ordered allomorph sets and the constraint Priority. The addition of lexical subcategorization from Bonet et al. (2007) proved vital to capturing the relationships that initial consonant mutations have with the rest of the grammar.

Chapter 4 showed that the allomorphy of the definite article was accounted for through the use of prosodic subcategorization frames and ordered allomorph sets. This was possible for two reasons: First, the clitic status of the definite article required the consideration of the prosodic structure associated with these lexical items. If the prosodic structure that these forms are associated with is ignored critical aspects of the distribution are missed and will lead to erroneous analyses from the improper applications of tools not designed to account for these forms. Second, through the use of ordered allomorph sets the preferential treatment of the enclitic allomorph was modeled.

Chapter 5 illustrated that the through lexical listing and lexical prespecification the initial consonant mutation can be accounted for by the grammar. The lexical listing of each of the different mutations allowed for the lexical prespecification of the mutation triggers to select the correct mutation. Additionally, lexical ordering allowed for the radical to be selected when there was not a trigger. This analysis also provided an account for the lexical exceptions to initial consonant mutations. This is due to those lexical items lacking the mutated forms in the input. Any change to these lexical exceptions resulted in faithfulness violations. Chapter 5 also showed that an account of the interaction of the definite article and initial consonant mutation was possible within Optimality Theory.

This thesis offered a solution to the interaction of morphology and phonology within Optimality Theory as illustrated by the Welsh definite article allomorphy. This thesis also offered evidence in the form of the Welsh definite article allomorphy against a morphological subcategorization approach to phonologically conditioned allomorphy, as proposed by Paster (2015) and showed that the "P$\gg$M" model of allomorphy (McCarthy & Prince 1993a,b) is better suited to account for phonology conditioned allomorphy.

REFERENCES

Anderson, Stephen. 2005. *Aspects of the Theory of Clitics.* Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199279906.001.0001. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199279906.001.0001/acprof-9780199279906.

Bennett, Ryan. 2018. Recursive prosodic words in Kaqchikel (Mayan). *Glossa: a journal of general linguistics* 3(1). 1–33.

Bennett, Ryan, Emily Elfner & James McCloskey. 2016. Lightest to the Right: An Apparently Anomalous Displacement in Irish. *Linguistic Inquiry* 47(2). 169–234. http://dx.doi.org/10.1162/LING_a_00209. http://www.mitpressjournals.org/doi/10.1162/LING{_}a{_}00209.

Bennett, Ryan, Boris Harizanov & Robert Henderson. 2018. Prosodic Smothering in Macedonian and Kaqchikel. *Linguistic Inquiry* 49(2). 195–246. https://muse.jhu.edu/article/690043.

Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró. 2007. Allomorph selection and lexical preferences: Two case studies. *Lingua* 117(6). 903–927. http://dx.doi.org/10.1016/j.lingua.2006.04.009. http://linkinghub.elsevier.com/retrieve/pii/S0024384106001100.

Borsley, Robert D, Maggie Tallerman & David Willis. 2007. *The Syntax of Welsh* Cambridge Syntax Guides. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511486227. http://ebooks.cambridge.org/ref/id/CBO9780511486227.

Brinkerhoff, Mykel Loren. 2019. Optimality Theory with Lexical Selection needs multiple PRIORITIES: A case study in Welsh allomorphy. *Proceedings of the Linguistic Society of America* 4(1). 25. http://dx.doi.org/10.3765/plsa.v4i1.4528. https://journals.linguisticsociety.org/proceedings/index.php/PLSA/article/view/4528.

Bye, Patrik. 2007. Allomorphy – selection, not optimization. In Sylvia Blaho, Patrik Bye & Martin Krämer (eds.), *Freedom of analysis?*, 63–92. Berlin: Walter de Gruyter. http://dx.doi.org/10.1515/9783110198591.63. https://www.degruyter.com/view/books/9783110198591/9783110198591.63/9783110198591.63.xml.

Cardinaletti, Anna & Lori Repetti. 2009. Phrase-level and word-level syllables: Resyllabification and prosodization of clitics. *Phonological Domains: Universals and Deviations* 16. 77—-102.

Embick, David & Alec Marantz. 2008. Architecture and Blocking. *Linguistic Inquiry* 39(1). 1–53.

Embick, David & Rolf Noyer. 2006. Distributed morphology and the syntax-semantics interface. *The Oxford Handbook of Linguistic Interfaces* (1994). 289–324.

Emonds, Joseph. 2002. A common basis for syntax and morphology: tri-level lexical insertion. In Paul Boucher (ed.), *Many morphologies*, 235–262. Somerville, MA: Cascadilla Press.

Green, Antony D. 2006. The independence of phonology and morphology: The Celtic mutations. *Lingua* 116(11). 1946–1985. http://dx.doi.org/10.1016/j.lingua.2004.09.002. http://linkinghub.elsevier.com/retrieve/pii/S0024384105000781.

Green, Antony D. 2007. *Phonology limited.* Potsdam: Universitätsverlag Potsdam.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20*, 111–176. Cambridge, MA: MIT Press.

Hannahs, S. J. 2011. Celtic Mutations. In *The blackwell companion to phonology*, 2807–2830. Oxford, UK: John Wiley & Sons, Ltd. http://dx.doi.org/10.1002/9781444335262.wbctp0117. `http://doi.wiley.com/10.1002/9781444335262.wbctp0117`.

Hannahs, S. J. 2013a. Celtic initial mutation: pattern extraction and subcategorisation. *Word Structure* 6(1). 1–20. http://dx.doi.org/10.3366/word.2013.0033. `http://www.euppublishing.com/doi/abs/10.3366/word.2013.0033`.

Hannahs, S. J. 2013b. *The phonology of Welsh*. Oxford: Oxford University Press.

Hannahs, S. J. & Maggie Tallerman. 2006. At the interface: selection of the Welsh definite article. *Linguistics* 44(4). 781–816. http://dx.doi.org/10.1515/LING.2006.025. `https://www.degruyter.com/view/j/ling.2006.44.issue-4/ling.2006.025/ling.2006.025.xml`.

Inkelas, Sharon. 1990. *Prosodic Constituency in the Lexicon*. New York, London: Routledge. http://dx.doi.org/10.4324/9780429455209. `https://www.taylorfrancis.com/books/9780429455209`.

Inkelas, Sharon. 1993. Deriving Cyclicity. In *Studies in lexical phonology*, 75–110. Elsevier. http://dx.doi.org/10.1016/B978-0-12-325071-1.50009-1. `https://linkinghub.elsevier.com/retrieve/pii/B9780123250711500091`.

Iosad, Pavel. 2007. Phonological processes as lexical insertion: more evidence from {Welsh} and elsewhere.

Iosad, Pavel. 2016. *Prosodic structure and suprasegmental features*, vol. 19 3. Springer Netherlands. http://dx.doi.org/10.1007/s10828-016-9083-8. `http://link.springer.com/10.1007/s10828-016-9083-8`.

Iosad, Pavel. 2017. Welsh svarabhakti as stem allomorphy. *Transactions of the Philological Society* 115(2). 141–175. http://dx.doi.org/10.1111/1467-968X.12085. `http://doi.wiley.com/10.1111/1467-968X.12085`.

Itô, Junko & Armin Mester. 2009a. The Extended Prosodic Word. In Janet Grijzenhout & Bariş Kabak (eds.), *Phonological domains: Universals and deviations* 1, 135–194. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110217100.2.135.

Itô, Junko & Armin Mester. 2009b. The onset of the prosodic word. In Steve Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*, 227–260. London: Equinox.

Kager, René. 1996. On affix allomorphy and syllable counting. In Ursula Kleinhenz (ed.), *Interfaces in phonology*, 155–171. Berlin: Akademie Verlag.

Kibre, N J. 1995. *A model of mutation in Welsh*: UCSB Master's thesis.

King, Gareth. 2015. *Modern Welsh: A Comprehensive Grammar*. Routledge third edit edn.

Kiparsky, Paul. 1973. 'Elsewhere' in phonology. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for morris halle*, 93–106. New York: Holt, Rinehart, and Winston.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, 131–175. Dordrecht: Foris.

Lieber, Rochelle. 1980. *On the orginatization of the lexicon*: Massachusetts Institute of Technology Doctoral dissertation.

Mascaró, Joan. 2007. External Allomorphy and Lexical Representation. *Linguistic Inquiry* 38(4). 715–735. http://dx.doi.org/10.1162/ling.2007.38.4.715. `http://www.mitpressjournals.org/doi/10.1162/ling.2007.38.4.715`.

Mascaró, Joan. 2008. External Allomorphy as Emergence of the Unmarked. In *Optimality theory in phonology*, 513–522. Oxford, UK: Blackwell Publishing Ltd. http://dx.doi.org/10.1002/9780470756171.ch28. `http://doi.wiley.com/10.1002/9780470756171.ch28`.

McCarthy, John J. 2008. *Doing Optimality Theory: Applying theory to data*. Oxford: Blackwell.

McCarthy, John J. & Alan Prince. 1993a. Generalized Alignment. In G. Booij & J. Van Marle (eds.), *Yearbook of morphology*, vol. 1993, 79–153. http://dx.doi.org/10.1007/978-94-017-3712-8_4. `http://link.springer.com/10.1007/978-94-017-3712-8{}4`.

McCarthy, John J. & Alan Prince. 1993b. Prosodic Morphology I: Constraint interaction and satisfication.

McCarvel, Miranda Kelly. 2016. *Harmonic Serialism with Lexical Selection: Evidence from Jèrriais allomorphy*: University of Utah Doctoral dissertation.

Mohanan, Karuvannur Puthanveettil. 1982. *Lexical Phonology*: MIT Doctoral dissertation.

Mohanan, Karuvannur Puthanveettil. 1986. *The theory of lexical phonology* (Studies in natural language and linguistic theory v. 6). Dordrecht ; Boston : Norwell, MA: D. Reidel Pub. Co.

Morris-Jones, J. 1913. *A Welsh Grammar: Historical and Comparative*. Oxford: Clarendon Press.

Nespor, Marina & Irene Vogel. 2007 [1986]. Prosodic phonology :with a new foreword.

Paster, Mary. 2005. Subcategorization vs. Output Optimization in Syllable-Counting Allomorphy. In John Alderete, Chung-hye Han & Alexei Kochetov (eds.), *Proceedings of the 24th west coast conference on formal linguistics*, 326–333. Somerville, MA: Cascadilla Proceedings Project.

Paster, Mary. 2006. *Phonological Conditions on Affixation*: University of California, Berkeley Doctoral dissertation.

Paster, Mary. 2009. Explaining phonological conditions on affixation: Evidence from suppletive allomorphy and affix ordering 1. *Word Structure* 2(1). 18–37. http://dx.doi.org/10.3366/E1750124509000282. `http://www.euppublishing.com/doi/abs/10.3366/E1750124509000282`.

Paster, Mary. 2015. Phonologically Conditioned Suppletive Allomorphy: Cross-linguistic Results and Theoretical Consequences. In Eulàlia Bonet, Maria-Rosa Lloret & Joan Mascaró (eds.), *Understanding allomorphy: Perspectives from optimality theory*, 218–253. London: Equinox Publishing Ltd. http://dx.doi.org/10.1558/equinox.25219. `https://journals.equinoxpub.com/index.php/books/article/view/25219`.

Pater, Joe. 2010. Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Steve Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*, London: Equinox.

Pilch, Hebbert. 1957. Das kymrische Lautsystem. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 75(1/2). 24–57. https://www.jstor.org/stable/40848002.

Prince, Alan & Paul Smolensky. 2004 [1993]. *Optimality Theory: Constraint Interaction in Generative Grammar.* Oxford, UK: Blackwell Publishing Ltd. http://dx.doi.org/10.1002/9780470759400. http://doi.wiley.com/10.1002/9780470759400.

Schuler, Kathryn D., Charles Yang & Elissa L. Newport. 2016. Testing the Tolerance Principle : Children form productive rules when it is more computationally efficient to do so. In A. Papafragou, D. Grodner, D. Mirman & J.C. Trueswell (eds.), *Proceedings of the 38th annual conference of the cognitive science society*, vol. 5 1958, 37082. Austin, TX: Cognitive Science Society.

Selkirk, Elisabeth. 1981. On prosodic structure and its relation to syntactic structure. In Thorstein Fretheim (ed.), *Nordic prosody ii: Papers from a symposium*, 111 – 140. Trondheim: TAPIR.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The relation between sound and structure.* Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 2009. On Clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu* 2009(136). 35–73.

Selkirk, Elisabeth. 2011. The Syntax-Phonology Interface. In *The handbook of phonological theory*, 435–484. Oxford, UK: Wiley-Blackwell. http://dx.doi.org/10.1002/9781444343069.ch14. http://doi.wiley.com/10.1002/9781444343069.ch14.

Smith, Jennifer L. 2012. The formal definition of the ONSET constraint and implications for Korean syllable structure. *Prosody matters: Essays in honor of Elisabeth Selkirk* (2). 73–108.

Tallerman, Maggie. 2006. The syntax of Welsh "direct object mutation" revisited. *Lingua* 116(11). 1750–1776. http://dx.doi.org/10.1016/j.lingua.2004.10.004. https://linkinghub.elsevier.com/retrieve/pii/S0024384105000859.

Thomas, Peter Wynn. 1996. *Gramadeg y Gymraeg.* Caerdydd: Gwasg Prifysgol Cymru.

Thorne, David. 1993. *A Comprehensive Welsh Grammar/Gramadeg Cymraeg Cynhwysfawr.* Oxford: Blackwell.

Tyler, Matthew. 2019. Simplifying MATCH WORD: Evidence from English functional categories. *Glossa: a journal of general linguistics* 4(1). 1–30. http://dx.doi.org/10.5334/gjgl.631. http://www.glossa-journal.org/articles/10.5334/gjgl.631/.

Williams, Briony Jane. 1983. *Stress in modern Welsh.* Cambridge: Universty of Cambridge Doctoral dissertation. http://dx.doi.org/10.17863/cam.16507.

Wolf, Matthew Adam. 2008. *Optimal interleaving: Serial phonology-morphology interaction in a constraint-based model*: University of Massachusetts Amherst Doctoral dissertation.

Wolf, Matthew Adam. 2015. Lexical Insertion occurs in the phonological component. In Eulàlia Bonet, Maria-Rosa Lloret & Joan Mascaró (eds.), *Understanding allomorphy: Perspectives from optimality theory*, 361–407. http://dx.doi.org/10.1558/equinox.25223.

Zec, Draga. 2005. Prosodic differences among function words. *Phonology* 22(1). 77–112. http://dx.doi.org/10.1017/S0952675705000448.

Zoll, Cheryl Cydney. 1996. *Parsing below the Segment in a Constraint Based Framework*: University of California, Berkeley Doctoral dissertation. http://dx.doi.org/10.7282/T37080BS. https://rucore.libraries.rutgers.edu/rutgers-lib/37655/.

Zwicky, Arnold M. 1977. On clitics. *Phonologica 1976* 23–39.

Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English N'T. *Langauge* 59(3). 502–513. http://dx.doi.org/10.2307/413900. https://www.jstor.org/stable/413900?origin=crossref.