

Variable and Exceptional Assimilation of the Definite Article [l-] in Moroccan Arabic

Ali Nirheche
anirheche@linguist.umass.edu

Abstract

This paper investigates the morpheme-specific phonological behavior of the Moroccan Arabic definite article [l-] and its variable assimilation process with nouns that begin with the palatal fricative [ʒ]. By analyzing of a corpus of these nouns, I identify one major factor that influences this variability: the context following [ʒ]. The results of a nonce word experiment I conducted show that speakers follow the lexical tendencies when generalizing to [ʒ]-initial nonce words, which does not align with the assumptions proposed in the literature (Harrell, 1962; Heath, 1987, 1989; Freeman, 2016). I will show that a learnability approach that uses Maximum Entropy Grammar (Goldwater and Johnson, 2003) with Lexically-indexed Constraints (Pater, 2000, 2009) successfully accounts for both the categorical behavior of known [ʒ]-initial words as well as the variability predicted for nonce words.

1 Introduction

The investigation of variable and exceptional phonological patterns has become an area of growing interest in phonological theory. Research has focused on the representation and modeling of phenomena involving both variable and exceptional patterns (Hayes and Londe, 2006; Pater et al., 2012; Linzen et al., 2013; Shih, 2018; Hughto et al., 2019), emphasizing the complexity of phonological systems and the necessity for theoretical mechanisms capable of capturing this complexity. This paper contributes to this research area by investigating the morpheme-specific assimilation patterns of the Moroccan Arabic definite article [l-], specifically focusing on the variable assimilation observed in [ʒ]-initial words. A comprehensive corpus analysis of these words reveals that this variability is phonologically-conditioned (e.g. more assimilation when a consonant follows [ʒ]). This study argues against previous proposals that assume categorical assimilation behavior for [ʒ]-initial nonce words. I show that assimilation for nonce words follow the lexical frequencies, i.e. FREQUENCY MATCHING. It will also be shown that a Maximum Entropy (Goldwater and Johnson, 2003, MaxEnt) model incorporating Lexically-indexed

Constraints (Pater, 2000, 2009, LICs), similar to the one proposed by Moore-Cantwell and Pater (2016), can effectively learn a grammar that accurately predicts both the variation observed for [ʒ]-initial nonce words as well as the fixed categorical behavior of real words.

As in Modern Standard Arabic (MSA) and other Arabic varieties, the definite article [l-] in Moroccan Arabic exhibits morpheme-specific behavior, in the form of total assimilation (gemination), when attached to a coronal-initial word (Harrell, 1962; Heath, 1987, 1989; Maas and Procházka, 2022), as in [l-dar] → [ddar] ‘house’. Uniquely, in Moroccan Arabic, assimilation does not uniformly occur with the palatal fricative [ʒ], leading to variation in assimilation with some [ʒ]-initial words triggering it (e.g. [ʒʒar] ‘DEF-neighbor’), while other words resisting it (e.g. [lʒil] ‘DEF-generation’). This paper challenges the binary distinction argued for by previous studies, which categorized the assimilated [ʒ]-initial words as default and non-assimilated ones as exceptions (Harrell, 1962; Heath, 1987; Freeman, 2016), demonstrating, instead, that these assimilation patterns are gradient. A corpus study reveals that the assimilation of [ʒ]-initial words is phonologically-conditioned by the sound following [ʒ], with words where a consonant follows [ʒ] are more likely to assimilate than those where a schwa follows [ʒ], which, in turn, are more likely to assimilate than those where a vowel follows [ʒ].

A critical question that this paper addresses is how Moroccan Arabic speakers generalize the assimilation patterns to [ʒ]-initial nonce words. Two possible hypotheses are proposed based on previous literature: [ʒ]-initial nonce words either categorically assimilate or categorically do not. This paper argues against both hypotheses, showing that a nonce word experiment supports an alternative hypothesis where assimilation patterns for nonce words align with the distributions observed across the lexicon. Previous studies on the productivity of morphophonological alternations have shown that, when speakers generalize to nonce forms, they tend to follow the lexical statistics (Zuraw, 2000; Ernestus and Baayen, 2003; Hayes and Londe, 2006; Hayes, 2009; Linzen et al., 2013; Becker and Gouskova, 2016). The experimental results presented in this paper align with these studies.

One challenge that learning theories face is accounting for both the stochastic behavior of nonce words and the fixed pronunciations of real words. While some previous studies, such as Hayes and Wilson (2008), offered solutions that were not applicable to alternations, other studies offered more promising solutions to this problem (Becker and Gouskova, 2016; Moore-Cantwell and Pater, 2016; Hughto et al., 2019). The definite article assimilation patterns in Moroccan Arabic provide an ideal test case for learning models given the complex exceptionality and variation patterns it embodies. In this paper, I implement a MaxEnt model with LICs, which accounts for exceptionality by the use of constraints indexed to exceptional lexical items. I will show that the proposed model can successfully predict variation for [ʒ]-initial nonce words as well as a fixed categorical behavior for real words.

The rest of the paper is organized as follows: Section 2 illustrates the patterns observed within definite article assimilation in Moroccan Arabic with a particular focus on the behavior of [ʒ]-initial words, presenting a corpus study to identify factors influencing

assimilation for these words. Section 3 presents an analysis of these assimilation patterns using MaxEnt with LICs that accounts for the two possible scenarios proposed in the literature: the one assuming categorical assimilation and the other assuming categorical non-assimilation. Section 4 proposes a learnability approach based on MaxEnt with LICs that accurately predicts both the variation predicted for [ʒ]-initial nonce words and fixed behavior of real words. Section 5 presents the results of a nonce word experiment that support frequency marching, indicating that the assimilation of [ʒ]-initial nonce words varies and aligns with lexical regularities. Section 6 concludes.

2 The Exceptional and Variable Patterns of Definite Article Assimilation

The definite article in Moroccan Arabic exhibits exceptional and variable assimilation patterns when attached to nouns that begin with a coronal consonant. This section delves into these patterns, focusing on the factors that trigger assimilation in [ʒ]-initial words. It will be shown that, while other varieties of Arabic exhibit a consistent assimilation pattern in [ʒ]-initial words, Moroccan Arabic show variation in the application of assimilation in these words. This section will provide a review of the previous literature (Harrell, 1962; Heath, 1987; Freeman, 2016) and show that their categorization of [ʒ]-initial words is not accurate. To understand this variability, I will provide a corpus analysis of [ʒ]-initial words. It will be shown that the context following [ʒ] influences the likelihood of assimilation in [ʒ]-initial words.

2.1 In Moroccan Arabic

In Moroccan Arabic, definite nouns are formed by attaching the prefix [l-] to a given noun (1). When the noun begins with a CV sequence where C is a non-coronal consonant, the definite article is attached faithfully (1a). When the noun begins with a CC sequence where the initial C is a non-coronal consonant, a schwa is inserted between the definite article and the noun (1b). When the noun begins with a coronal consonant (either in a CC or CV sequence), however, attaching [l-] results in the total assimilation of the definite article to the initial coronal consonant, forming a geminate sound¹ (Harrell, 1962; Heath, 1987, 1989; Maas and Procházka, 2022) as seen in (1c).

¹I assume that geminates are single sounds. This has also been proposed for Tashlhiyt Berber, spoken in Morocco, where a geminate sound belongs to the onset of the same syllable ([llan] ‘they exist’), whereas non-identical CC sequences are parsed into two separate syllables ([g.ru] ‘glean’) (Ridouane, 2016).

(1)	Noun	Definite Noun	Gloss
a.	bənt	l.bən.t	‘girl’
	kora	l.ko.ra	‘ball’
b.	fraʃ	ləf.raʃ	‘bed sheets’
	kbal	lək.bal	‘popcorn’
c.	dar	ddar	‘house’
	nhar	nn.har	‘day’
	ʃəmʃ	ʃʃəm.ʃ	‘sun’
	lil	llil	‘night’

The assimilation pattern seen in (1c) is morpheme specific behavior, since it is not triggered in the presence of the preposition prefix [l-] when attached to nouns to form prepositional phrases, although it is identical to the definite article (2).

(2)	/l/ + /dar/ + /o/	ldaro	*ddaro	‘to-house-POSS.3SG’
	/l/ + /nhar/	lənhar ssəbt	*nnhar ssəbt	‘to-day Saturday’

Moreover, assimilation does not occur stem-internally (3a) or across morpheme boundaries when [l] is the final consonant of the stem (3b).

(3)	a.	/wəld/	wəld	*wədd	‘boy’
		/ldida/	ldida	*ddida	‘delicious’
	b.	/gəl/ + /ti/	gəlti	*gətti	‘say.PAST-2SG’
		/qfəl/ + /na/	qfəlna	*qfənna	‘lock.PAST-1PL’

Assimilation is not applied categorically to all [ʒ]-initial words when the word-initial consonant is a coronal. Instead, the application of assimilation is variable in such words (Heath, 1987, 1989; Maas and Procházka, 2022). The palatal fricative [ʒ] triggers assimilation in some words (4a) and does not in others (4b). This is an unusual exception given the fact that [ʃ], which only differs from [ʒ] in voicing, categorically triggers assimilation.

(4)	Noun	Definite Noun	Gloss
a.	ʒoqa	ʒʒoqa	‘gathering’
	ʒməl	ʒʒməl	‘camel’
	ʒuʃ	ʒʒuʃ	‘hunger’
b.	ʒomhor	lʒomhor	‘audience’
	ʒəlsa	lʒəlsa	‘court session’

2.2 In other varieties of Arabic

While the morpheme specific behavior of the definite article is present in most Arabic varieties, the assimilation patterns of [ʒ]-initial words is different among these varieties. In MSA, assimilation occurs when the definite article [ʔal] is attached to coronal-initial nouns (5b) except for those that begin with [ʒ] (5c). This failure to trigger assimilation is due to diachronic change. Watson (2002) pointed out that [ʒ] originated from the pre-Classical Arabic velar plosive [g], a non-coronal consonant that does not trigger assimilation (Blanc, 1969). Apparently, this remained the case after the diachronic change to [ʒ] in MSA.

(5)	Noun	Definite Noun	Gloss
a.	bint	ʔal-bint	‘girl’
	firaaʃ	ʔal-firaaʃ	‘bed sheets’
b.	daar	ʔad-daar	‘house’
	nahaar	ʔan-nahaar	‘day’
	ʃams	ʔaʃ-ʃams	‘sun’
c.	ʒamal	ʔal-ʒamal	‘the camel’
	ʒamiʃ	ʔal-ʒamiʃ	‘the mosque’
	ʒundiyy	ʔal-ʒundiyy	‘the soldier’

Freeman (2016) investigated the historical development of [ʒ] in Arabic varieties and suggested a phonological explanation for its exceptional behavior. He argues that there are three different grammars for present-day Arabic dialects:

(6)

- Grammar I: dialects in which all coronals trigger assimilation except [ʒ] (MSA; Peninsular Arabian Dialects).
- Grammar II: dialects in which all coronals trigger assimilation including [ʒ] (North African Varieties; Levantine Arabic; Iraqi Arabic).
- Grammar III: dialects in which all coronals trigger assimilation and velars optionally trigger assimilation (Cairene Arabic).

Freeman (2016), following McCarthy (1994), argues that the exceptional behavior of [ʒ] in Grammar I dialects can be explained in representational terms by the adoption of the phonotactic restriction known as Root Co-occurrence. The latter was proposed by McCarthy (1994) to account for a restriction observed in MSA roots: a trilateral root in MSA cannot have two adjacent consonants that have similar manner and place of articulation. McCarthy claims that [ʒ] patterns underlyingly as the velar [g], which explains its failure to trigger assimilation in MSA, a fact that is supported by the historical proposal that [ʒ]

originated from the pre-Classical Arabic velar plosive [g]. Freeman (2016) adopted McCarthy's root co-occurrence restriction, but proposed that [ʒ] underlyingly corresponds to the palatal plosive [j], not the velar [g], and that the place feature of [j] is dorsal rather than coronal. Based on this approach, Freeman argues that Grammar I dialects have underlying [j], while Grammar II dialects, in which all coronals trigger assimilation, have underlying [ʒ].

2.3 The source of variation in [ʒ] initial words

What's of interest to us in this paper is the case of Moroccan Arabic, which is considered by Freeman as a Grammar II dialect, although assimilation is not categorically applied to all [ʒ]-initial nouns as seen in (4). This raises questions about Freeman's proposal which stems from the common view that is held in the literature and that classifies cases like (4b) as exceptions to the categorical assimilation rule. According to Harrell (1962), [ʒ]-initial words that do not undergo assimilation are those that belong to religious terminology. Freeman (2016, p. 177) claims that the failure of [ʒ] to trigger assimilation in some Moroccan Arabic words is "the result of diglossic interference from the standard or classical language". Heath (1987) argues that assimilation of [ʒ]-initial words is default for inherent Moroccan Arabic words, while all MSA loans beginning with [ʒ] do not trigger assimilation. However, this could also mean that non-assimilation is default if we assume that the default behavior is that of new forms. In Moroccan Arabic, novel words are mostly borrowed MSA words. Since the latter do not assimilate, this suggests that assimilation do not occur by default. Given these two options, two possible hypotheses about the generalization of assimilation patterns of [ʒ]-initial words can be proposed:

(7)

- **Categorical Assimilation Hypothesis:** assimilation is default to [ʒ]-initial words (similar to all coronal-initial words), but some [ʒ]-initial words in the lexicon exceptionally resist assimilation. Nonce words are expected to assimilate categorically. This hypothesis is aligned with the proposals made by Harrell (1962), Heath (1987), and Freeman (2016).
- **Categorical Non-assimilation Hypothesis:** non-assimilation is default for all [ʒ]-initial words (similar to the patterns seen in MSA), but an exceptional class of [ʒ]-initial words in the lexicon triggers assimilation. Nonce words are expected not to assimilate.

There are two main issues with these previous proposals. First, there are words that are borrowed from other languages or that do not belong to religious/political register, but still fail to trigger assimilation (8a). Second, the issue with Heath's proposal is identifying what these "MSA loans" are, which is by no means a straightforward task. In fact, even what one might call "inherent" Moroccan Arabic words are historically derived from

Classical Arabic. What's more, MSA words are being borrowed continuously, so it's not an easy task to specify how old a word should be in order to be an inherent Moroccan Arabic word. It can be seen in (8b) that some words that do not seem to be recent borrowings from MSA do not assimilate. It is also unclear how the learner would identify these borrowed words to treat them exceptionally.

(8)	Noun	Definite Noun	Gloss
a.	ʒø	l-ʒø	'game'
	ʒõʕ	l-ʒõʕ	'genre'
b.	ʒuw	ʒʒuw	'weather'
	ʒud	l-ʒud	'generosity'

The next section presents the results of a corpus study showing that assimilation of [ʒ]-initial words is phonologically determined: the sound that follows [ʒ] is the major factor determining if an [ʒ]-initial word is more likely to assimilate or not.

2.4 Assimilation of [ʒ]-initial words is phonologically determined

A possibility that was not discussed in the literature is for assimilation to be phonologically-conditioned. One factor that is worth examining is the following context, i.e. the sound following [ʒ]. There are three relevant categories of sounds: consonants, full vowels, and the schwa. The expected behavior of [ʒ]-initial words is for the words with a consonant following [ʒ] to assimilate more than the ones with a schwa following [ʒ] which are, in turn, expected to assimilate more than the words with a full vowel following [ʒ].

The reason that assimilation is more likely to occur when a consonant follows [ʒ] is the restriction against having three adjacent consonants. The definite form of the noun [ʒmil] is [ʒʒmil] because the non-assimilated version [lʒmil] begins with a sequence of three adjacent consonants. When a vowel follows [ʒ], however, assimilation is less likely to occur given the absence of a CCC sequence. Such sequence is avoided in non-coronal initial contexts through schwa epenthesis. The non-assimilated definite form version of [ʒil], which is [lʒil], begins with a sequence of two consonants, which is acceptable in Moroccan Arabic. When a schwa follows [ʒ], assimilation is expected to have an intermediate rate of assimilation given the ambiguous status of the schwa in Moroccan Arabic (Benhallam, 1980; Al Ghadi, 1990; Boudlal, 2001; among others). The Moroccan Arabic schwa is often argued to be an epenthetic vowel only inserted for syllabification purposes, and whose status in the phonological grammar is different from that of a full vowel.

In order to investigate this factor, a corpus of 120 Moroccan Arabic [ʒ]-initial words was created and examined. The corpus is representative of the Moroccan Speaker's knowledge of [ʒ]-initial words given the various sources used for data collection. These sources include the author's knowledge as a native speaker of Moroccan Arabic, consulting other

native speakers, a large online corpus of Moroccan Arabic: Darija Open Dataset (Outchakoucht and Es-Samaali, 2021), previous work that discussed this phenomenon (Harrell, 1962; Freeman, 2016; Maas and Procházka, 2022), and a Moroccan Arabic-English dictionary (Harrell and Sobelman, 1966). Overall, 76 out of the 120 words in the corpus exhibit assimilation. It can be seen in Table 1 that, as was predicted, assimilation is more likely to occur when a consonant follows [ʒ], is less likely to occur when a full vowel follows [ʒ], and has an intermediate rate of assimilation when a schwa follows [ʒ].

Following Context	Count	Assimilation (%)
Consonant	26	96%
Schwa	38	81%
Full Vowel	57	37%
total:	120	63%

Table 1: Corpus Statistics about the Assimilation Patterns based on the following context

The results in Table 1 were supported by the logistic regression analysis with custom contrasts that was conducted to determine if the following context significantly influences assimilation of [ʒ]-initial words. For coding this factor, Helmert coding was applied, contrasting the schwa with consonant, and vowel with the average of schwa and consonant contexts. The analysis revealed a statistically significant effect for the contrast "Vowel vs. Average of Schwa and Consonant" ($\beta = -2.49$, $p < 0.00001$), indicating a strong negative effect of a full vowel on assimilation. The contrast "Schwa vs. Consonant" was not statistically significant ($\beta = -0.9968$, $p = 0.23$).

2.5 Conclusion

In this section, I showed that there are two possible hypotheses about the behavior of [ʒ]-initial words based on the literature: categorical assimilation and categorical non-assimilation. The corpus study, however, shows that assimilation is phonologically-determined: the context that follows [ʒ] determines the likelihood of assimilation. The next section presents an analysis using MaxEnt with LICs that accounts for the definite article assimilation patterns in Moroccan Arabic. The following section will lay out a proposed learnability approach that examines if frequency matching can be predicted using a learning model with unsupervised indexation.

3 A MaxEnt Analysis with Lexically-Indexed Constraints

3.1 Introduction

This section delves into the analysis of definite article assimilation patterns in Moroccan Arabic, focusing on behavior of [ʒ]-initial words. Previous work shows that assimilation is a morpheme-specific phenomenon, only being triggered in the presence of the definite article. In Moroccan Arabic, specifically, some [ʒ]-initial words trigger assimilation, while others resist it, exhibiting a pattern of exceptionality and variation that requires a highly complex analysis. This section is devoted to analyses that are based on previous proposals where either assimilation or non-assimilation is categorical.

It has been shown above that assimilation in definite nouns is a morpheme-specific phenomenon. It has also been shown that [ʒ]-initial words vary, with some words triggering assimilation and others resisting it. To account for this exceptionality and variation, I use Maximum Entropy grammar (Goldwater and Johnson, 2003) together with lexically-indexed constraints (Pater, 2000, 2009). MaxEnt is a probabilistic model that captures categorical and variable patterns in phonology. It assigns probabilities to different output candidates based on weighted constraints. Lexically-indexed constraints will also be proposed. Lexical indexation explains exceptionality by allowing constraints to be lexically specific, i.e. they apply only to certain lexical items or morphemes, not across the grammar. In our case, general constraints will account for the default behavior of the sequence of [l] followed by a coronal, where assimilation does not occur, and lexically-indexed constraints will account for the exceptionality of the definite article and the variable behavior of [ʒ]-initial words.

In this section, I will show that this approach can explain the exceptional behavior of the definite article in triggering assimilation. Moreover, the categorical assimilation/non-assimilation hypotheses proposed in the literature for [ʒ]-initial words will also be accounted for. I will also show that the predictions of both analyses for the behavior of [ʒ]-initial nonce words do not match the lexical frequencies.

3.2 The Default Behavior of l-coronal

In the absence of the definite article, the sequence of [l] followed by a coronal does not trigger assimilation as discussed in section 2.1. To achieve this outcome, an analysis in which the faithfulness constraint MAX(lat) interacts with the markedness constraint *l[cor] is proposed. MAX(lat), which requires identity of the lateral feature between the input and output, must have a higher weight compared to *l[cor], which penalizes outputs with the sequence of [l] followed by a coronal, in order to prevent assimilation. Three possible candidates are relevant: a non-assimilated, assimilated, and epenthesis candidates. The latter, which epenthesizes a schwa between the [l] and the following coronal (10c), always loses when the noun following the [l] begins with a simple onset. Therefore, the constraint *ə]_σ, which is never violated in Moroccan Arabic, is proposed to

rule out the epenthesizing candidate in this context. $*l[cor]$ and $*ə]_{\sigma}$ are defined in (9).

(9)

- $*l[cor]$: Assign a violation mark for every lateral approximant followed by a coronal consonant in the output.
- $*ə]_{\sigma}$: Assign a violation mark for every schwa that surfaces in an open syllable.

A relevant example of the derivation of the prepositional phrase, [ldaro] "to his house", where the noun begins with a simple onset and where assimilation does not occur, is shown in (10). The proposed weights are manually specified for illustration purposes.

(10)

/l/ + /daro/	$*l[cor]$	MAX(lat)	DEP	$*ə]_{\sigma}$	\mathcal{H}	p
a. ddaro	0	-1	0	0	-5	≈ 0
b. $\text{ⵍ} \rightarrow$ ldaro	-1	0	0	0	-1	≈ 1
c. lədaro	0	0	-1	-1	-13	≈ 0

In (10), the fully faithful candidate [ldaro] wins by satisfying the highly weighted faithfulness constraint MAX(lat). [ddaro], on the other hand, violates MAX(lat) by deleting the lateral feature of the input consonant [l]. Since MAX(lat) has higher weight than $*l[cor]$, the non-assimilating candidate (10b) is more harmonic than the assimilating one (10a). Candidate (10c), which epenthesizes the schwa between the preposition [l] and coronal consonant [d], satisfies $*l[cor]$, but it violates the highly weighted constraint $*ə]_{\sigma}$, and is, thus, ruled out. Having both $*ə]_{\sigma}$ and DEP is necessary; DEP is violated by winners, as shown in (12) below, and has, therefore, a low weight value, whereas $*ə]_{\sigma}$ is never violated by an output form in Moroccan Arabic, which explains its high weight value.


Unlike the cases with a simple onset, nouns that begin with a CC sequence have a different optimal candidate. As seen in (12c), such cases involve schwa epenthesis in the output form. Let's consider the derivation of the example [lənhar] "to the day of" in (12). The analysis of this example is similar to the one with a simple onset, but it differs in that the epenthesizing candidate (12c) is optimal. Attaching the preposition [l] to [nhar] results in a sequence of three adjacent consonants. Since having such a sequence is not acceptable in Moroccan Arabic, especially in the initial position, $*CCC$ is proposed (11).

(11)

- $*CCC$: Assign a violation mark for any sequence of three adjacent consonants in the output form.

Although (12c) violates DEP, the latter has a small weight compared to *CCC, which results in candidate (12c) having a higher harmony score compared to the fully faithful candidate (12b).

(12)

/l/ + /nhar/	*l[cor]	MAX(lat)	DEP	*ə] _σ	*CCC	\mathcal{H}	p
a. nnhar	0	-1	0	0	0	-5	≈0
b. lnhar	-1	0	0	0	-1	-21	≈0
c.  lənhar	0	0	-1	0	0	-1	≈1

As can be seen, the general grammar of Moroccan Arabic allows for having a sequence of [l] followed by a coronal by assigning a high weight to MAX(lat). This sequence is not acceptable when it results in three adjacent consonant, in which case the grammar adopts schwa epenthesis as a repair strategy.

3.3 The Morpheme-specific Behavior of the Definite Article [l]

The analysis we have so far predicts that assimilation cannot occur by default to any sequence of [l] followed by a coronal consonant. In order to account for the assimilation that is exceptionally triggered in the presence of the definite article [l], a lexically-indexed version of *l[cor] is proposed. To account for the behavior of the words that begin with a CC sequence, a lexically-indexed version of DEP is also needed. While the epenthesis forms are optimal in words beginning with a CC sequence in the default context, as seen in (12), when the definite article is present, assimilation, not schwa epenthesis, occurs in such examples. This requires the use of an indexed version of DEP that prevents epenthesis specifically in the context of definite nouns:

(13)

- ***l[cor]_{DEF}**: Assign a violation mark to any instance of a lateral approximant followed by a coronal consonant that contains a phonological exponent of a morpheme specified as DEF.
- **DEP_{DEF}**: Assign a violation mark to any inserted segment that is adjacent to the morpheme specified as DEF².

²If we assume the traditional definitions of indexed constraints proposed by Pater (2000, 2009), DEP_{DEF} would raise a locality problem since it's not directly associated with an exponent of the definite article, but with a segment adjacent to it (the schwa). Therefore, DEP_{DEF} is defined in a way that specifically reflect this difference.

By assigning a high weight to $*l[cor]_{DEF}$ and DEP_{DEF} , the assimilation that occurs in definite nouns is predicted as seen in the derivation of [ddar] "the house" and [nnhar] "the day" in (14) and (15), respectively.

(14)

$/l_{DEF}/ + /dar/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	DEP 1	DEP_{DEF} 19	\mathcal{H}	p
a. [ddar]	0	0	-1	0	0	-5	≈ 1
b. [ldar]	-1	-1	0	0	0	-14	≈ 0
c. $\text{[l\ddot{a}dar]}$	0	0	0	-1	-1	-20	≈ 0

(15)

$/l_{DEF}/ + /nhar/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	DEP 1	DEP_{DEF} 19	$*CCC$ 20	\mathcal{H}	p
a. [nnhar]	0	0	-1	0	0	0	-5	≈ 1
b. [lnhar]	-1	-1	0	0	0	-1	-34	≈ 0
c. $\text{[l\ddot{a}nhar]}$	0	0	0	-1	-1	0	-20	≈ 0

In both (14) and (15), the assimilated candidates (14a) and (15a) only violate MAX(lat), resulting in a harmony score of -5 for each of them. The fully faithful candidates, on the other hand, lose by violating $*l[cor]_{DEF}$ which has a high weight. The epenthesis candidates also lose, since they violate the highly weighted constraint DEP_{DEF} . As can be seen, adding the lexically-indexed versions of $*l[cor]$ and DEP results in the morpheme-specific assimilation observed when forming definite nouns.

3.4 The Behavior of [ʒ]-initial words

We've seen that a sequence of [l] followed by a coronal is expected to resist assimilation, except when [l] is the definite article, in which case assimilation occurs. When the definite article is attached to a [ʒ]-initial noun, assimilation is applied for some words in the lexicon and not others. Based on the literature, I proposed two possible hypotheses about the predictions of assimilation patterns for [ʒ]-initial nonce words: the categorical assimilation hypothesis and the categorical non-assimilation hypothesis. The next two subsections present two different analyses under MaxEnt with LICs, and that correspond to each of the two hypotheses.

3.4.1 Categorical Assimilation Hypothesis:

As mentioned above, the categorical assimilation hypothesis assumes that [ʒ]-initial words assimilate by default similar to any other coronal-initial words, but a subset of [ʒ]-initial

words in the lexicon exceptionally resist assimilation. Building on that, assimilated [ʒ]-initial words can be accounted for parallel to (14) and (15). The indexed constraint $*\#ʒʒ$, defined in (16), is added to the set of constraints in (14) and (15). The latter is a general constraint used to prevent assimilation in [ʒ]-initial words.

(16)

- $*\#ʒʒ$: Assign a violation mark for any output that begins with a geminate [ʒ].

Since [ʒ]-initial words assimilate by default, $*\#ʒʒ$ has a small weight of 1. The derivation of the words [ʒar] "neighbor" and [ʒmil] "favor" can be seen in (17) and (18).

(17)

$/l_{DEF}/ + /ʒar/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	DEP 1	DEP_{DEF} 19	$*\#ʒʒ$ 1	\mathcal{H}	p
a. ʒʒar	0	0	-1	0	0	-1	-6	≈ 1
b. $lʒar$	-1	-1	0	0	0	0	-14	≈ 0
c. $ləʒar$	0	0	0	-1	-1	0	-20	≈ 0

(18)

$/l_{DEF}/ + /ʒmil/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	DEP 1	DEP_{DEF} 19	$*CCC$ 20	$*\#ʒʒ$ 1	\mathcal{H}	p
a. ʒʒmil	0	0	-1	0	0	0	-1	-6	≈ 1
b. $lʒmil$	-1	-1	0	0	0	-1	0	-34	≈ 0
c. $ləʒmil$	0	0	0	-1	-1	0	0	-20	≈ 0

To account for the [ʒ]-initial words that exceptionally resist assimilation, we need a lexically-indexed version of the constraint $*\#ʒʒ$:

(19)

- $*\#ʒʒ_L$: Assign a violation mark for any outputs that begins with a geminate [ʒ] and that contains a phonological exponent of a morpheme specified as L.

$*\#ʒʒ_L$ prohibits having a geminate [ʒ] at the beginning of a word, specifically in [ʒ]-initial words that do not assimilate. Assigning a high weight to $*\#ʒʒ_L$ ensures that all these exceptionally non-assimilated [ʒ]-initial words are optimal. An example of the derivation of the words [ʒil] "generation" and [ʒtimaʃ] "meeting" is shown in (20) and (21), respectively.

(20)

$/l_{DEF}/ + /ʒil_L/$	*1[cor] _{DEF} 13	*1[cor] 1	MAX(lat) 5	*#ʒʒ _L 20	DEP _{DEF} 19	\mathcal{H}	p
a. ʒʒil	0	0	-1	-1	0	-25	≈ 0
b. $\text{ɹ} \rightarrow$ lʒil	-1	-1	0	0	0	-14	≈ 1
c. ləʒil	0	0	0	0	-1	-19	≈ 0

(21)

$/l_{DEF}/ + /ʒtimaʃ_L/$	*1[cor] _{DEF} 13	*1[cor] 1	MAX(lat) 5	*#ʒʒ _L 20	DEP _{DEF} 19	*CCC 20	\mathcal{H}	p
a. ʒʒtimaʃ	0	0	-1	-1	0	0	-25	≈ 0
b. lʒtimaʃ	-1	-1	0	0	0	-1	-34	≈ 0
c. $\text{ɹ} \rightarrow$ ləʒtimaʃ	0	0	0	0	-1	0	-19	≈ 1

As seen in (20) and (21), the assimilating candidates [ʒʒil] and [ʒʒtimaʃ] violate *#ʒʒ_L, which decreases their harmony from -5 to -25. This makes the fully faithful candidate [lʒil] in (20) optimal with a higher harmony of -14 and the epenthesis candidate in (21) also optimal with a harmony of -19. As can be seen, adding a version of *#ʒʒ that is lexically-indexed to some [ʒ]-initial words makes them exceptionally resist assimilation as observed in the lexicon.

3.4.2 Categorical Non-assimilation Hypothesis:

The categorical non-assimilation hypothesis assumes that [ʒ]-initial words do not assimilate by default, but a subset of [ʒ]-initial words in the lexicon exceptionally assimilate. The default behavior of [ʒ]-initial words can be achieved by adding the general constraint *#ʒʒ to the set of constraints in (14) and (15). A high weight for *#ʒʒ predicts that assimilation will not occur for [ʒ]-initial words by the general grammar, compared to the words that begin with other coronal consonants which are expected to assimilate. The tableaux in (22) and (23) show an example of the derivation of the words [ʒil] "generation" and [ʒtimaʃ] "meeting", respectively.

(22)

$/l_{DEF}/ + /ʒil/$	*1[cor] _{DEF} 13	*1[cor] 1	MAX(lat) 5	*#ʒʒ 20	DEP _{DEF} 19	*ə] _σ 13	\mathcal{H}	p
a. ʒʒil	0	0	-1	-1	0	0	-25	≈ 0
b. $\text{ɹ} \rightarrow$ lʒil	-1	-1	0	0	0	0	-14	≈ 1
c. ləʒil	0	0	0	0	-1	-1	-32	≈ 0

(23)

$/l_{DEF}/ + /ʒtɪmaʃ/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	$*\#ʒʒ$ 20	DEP_{DEF} 19	$*CCC$ 20	\mathcal{H}	p
a. $ʒʒtɪmaʃ$	0	0	-1	-1	0	0	-25	≈ 0
b. $lʒtɪmaʃ$	-1	-1	0	0	0	-1	-34	≈ 0
c. $l\emptysetʒtɪmaʃ$	0	0	0	0	-1	0	-19	≈ 1

We've seen in (14) and (15) above that the assimilated candidates are optimal when the word begins with a non-[ʒ] coronal consonant, since those candidates incur a single violation of MAX(lat) and, therefore, have a high harmony of -5. With regards to [ʒ]-initial words, however, as seen in (22) and (23), the assimilating candidates (22a) and (23a) violate both MAX(lat) and $*\#ʒʒ$ which decreases their harmony to -25, making the non-assimilating candidate win in (22) and the epenthesis one win in (23).

The [ʒ]-initial words that exceptionally assimilate can be accounted for using two additional lexically-indexed constraints: $*l[cor]_L$, which prevents having a sequence of [l] followed by a coronal consonant, and DEP_L which prohibits epenthesis, specifically for those exceptional lexical items.

(24)

- $*l[cor]_L$: Assign a violation mark to any instance of a lateral approximant followed by a coronal consonant that contains a phonological exponent of a morpheme specified as L.
- DEP_L : Assign a violation mark to any inserted segment that is adjacent to the morpheme specified as L.

Assigning high weights to both $*l[cor]_L$ and DEP_L decreases the harmony values of the non-assimilated and epenthesis candidates and ensures that all exceptionally assimilated [ʒ]-initial words are optimal. An example of the derivation of the words [ʒar] "neighbor" and [ʒmil] "favor" is shown in (25) and (26), respectively.

(25)

$/l_{DEF}/ + /ʒar_L/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	$*\#ʒʒ$ 20	$*l[cor]_L$ 18	DEP_{DEF} 19	DEP_L 18	\mathcal{H}	p
a. $ʒʒar$	0	0	-1	-1	0	0	0	-25	≈ 1
b. $lʒar$	-1	-1	0	0	-1	0	0	-32	≈ 0
c. $l\emptysetʒar$	0	0	0	0	0	-1	-1	-37	≈ 0

(26)

$/l_{DEF}/ + /ʒmil_L/$	*l[cor] _{DEF} 13	*l[cor] 1	MAX(lat) 5	*#ʒʒ 20	*l[cor] _L 18	DEP _{DEF} 19	DEP _L 18	*CCC 20	\mathcal{H}	p
a. $\text{ʒ}ʒmil$	0	0	-1	-1	0	0	0	0	-25	≈ 1
b. $lʒmil$	-1	-1	0	0	-1	0	0	-1	-52	≈ 0
c. $ləʒmil$	0	0	0	0	0	-1	-1	0	-37	≈ 0

As shown in (25) and (26), although the assimilating candidates [ʒʒar] and [ʒʒmil] violate *#ʒʒ, they are still optimal compared to the fully faithful candidates, since the latter violate all *l[cor] constraints including *l[cor]_L which is highly weighted. The epenthesis candidates also lose by violating both DEP_{DEF} and DEP_L. We can conclude that adding versions of *l[cor] and DEP that are lexically-indexed to some [ʒ]-initial lexical items makes these lexical items exceptionally trigger assimilation.

3.5 Predictions for Nonce Words

In this paper, I follow Moore-Cantwell and Pater (2016) in their assumption about nonce words, i.e. that they would not be lexically-indexed. Therefore, the predictions for nonce words can be automatically determined, since the lexically-indexed constraints do not have an effect on nonce words. Each of the two analyses have different prediction for the behavior of novel words. In the categorical assimilation analysis, all [ʒ]-initial nonce words are predicted to assimilate, since, *#ʒʒ_L, which makes the non-assimilated existing lexical items resist assimilation, is not indexed to nonce words. The tableaux in (27) and (28) show the assimilation behavior predicted for the nonce words [ʒin] and [ʒrafa], respectively.

(27)

$/l_{DEF}/ + /ʒin/$	*l[cor] _{DEF} 13	*l[cor] 1	MAX(lat) 5	*#ʒʒ _L 20	DEP _{DEF} 19	\mathcal{H}	p
a. $\text{ʒ}ʒin$	0	0	-1	0	0	-5	≈ 0
b. $lʒin$	-1	-1	0	0	0	-14	≈ 1
c. $ləʒin$	0	0	0	0	-1	-19	≈ 0

(28)

$/l_{DEF}/ + /ʒrafa/$	*l[cor] _{DEF} 13	*l[cor] 1	MAX(lat) 5	*#ʒʒ _L 20	DEP _{DEF} 19	*CCC 20	\mathcal{H}	p
a. $\text{ʒ}ʒrafa$	0	0	-1	0	0	0	-5	≈ 0
b. $lʒrafa$	-1	-1	0	0	0	-1	-34	≈ 0
c. $ləʒrafa$	0	0	0	0	-1	0	-19	≈ 1

In the categorical non-assimilation analysis, all [ʒ]-initial nonce words are predicted not to assimilate, while $*l[cor]_L$ and DEP_L are the lexically-indexed constraint responsible for triggering assimilation in the [ʒ]-initial words that exceptionally assimilate. Since these constraints do not effect nonce words, the latter are predicted to resist assimilate as seen in the derivations in (29) and (30) for the nonce words [ʒin] and [ʒrafa], respectively.

(29)

$/l_{DEF}/ + /ʒin/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	$*\#ʒʒ$ 20	$*l[cor]_L$ 18	DEP_{DEF} 19	DEP_L 18	\mathcal{H}	p
a. ʒʒin	0	0	-1	-1	0	0	0	-25	≈ 1
b. ⵛ lʒin	-1	-1	0	0	0	0	0	-14	≈ 0
c. ləʒin	0	0	0	0	0	-1	0	-19	≈ 0

(30)

$/l_{DEF}/ + /ʒrafa/$	$*l[cor]_{DEF}$ 13	$*l[cor]$ 1	MAX(lat) 5	$*\#ʒʒ$ 20	$*l[cor]_L$ 18	DEP_{DEF} 19	DEP_L 18	$*CCC$ 20	\mathcal{H}	p
a. ʒʒrafa	0	0	-1	-1	0	0	0	0	-25	≈ 1
b. lʒrafa	-1	-1	0	0	0	0	0	-1	-34	≈ 0
c. ⵛ ləʒrafa	0	0	0	0	0	-1	0	0	-19	≈ 0

While the proposed analyses succeed in accounting for the assimilation patterns of all known lexical items as shown in the previous subsections, ss seen in Table 2, the predictions of the both analyses for the behavior of nonce words contradict the lexical statistics where assimilation was not categorical, but variable. Therefore, in the next section, I propose a learning model using MaxEnt with LICs that will predict both a fixed categorical behavior for real [ʒ]-initial words and variable behavior for nonce words.

Context	Cat. Assim.	Cat. Non-assim.	Lexicon
Vowel after [ʒ]	100%	0%	37%
Schwa after [ʒ]	100%	0%	81%
Consonant after [ʒ]	100%	0%	96%

Table 2: Comparison of Predicted Probability of Assimilation in Each Context for Nonce Words between the Categorical Analysis and the Lexicon

3.6 Conclusion

This section examined the definite article assimilation patterns in Moroccan Arabic, particularly focusing on the behavior of [ʒ]-initial words. I proposed an analysis using MaxEnt with LIC that effectively captured the default assimilation or non-assimilation by the use of a set of general constraints and accounted for exceptionality by the use of lexically-indexed constraints. The proposed analyses predicted categorical behavior for [ʒ]-initial

nonce words. In the next section, I present a learnability approach within MaxEnt with LICs that accounts for the exceptionality patterns observed in known [ʒ]-initial words and predicts variability for nonce words.

4 A Learnability Approach: MaxEnt with Unsupervised Indexation

Contrary to the predictions of the two categorical hypotheses, we have seen in section 2.4 that statistics across the lexicon suggest that assimilation is variable for [ʒ]-initial words based on the following context. Previous studies on the productivity of morphophonological alternations have shown that, when speakers generalize to nonce forms, they tend to follow lexical statistics (or frequency match) (Zuraw, 2000; Ernestus and Baayen, 2003; Hayes and Londe, 2006; Hayes, 2009; Linzen et al., 2013; Becker and Gouskova, 2016). Since the two categorical hypotheses did not predict this behavior for nonce words, in this section, I implement a MaxEnt model that will be shown to predict frequency matching behavior for nonce words. The model differs from the analyses proposed in section 3.4 in that it has less supervision. In other words, the grammar is learned by exposure to the data without providing any information about what is default and what is exceptional with respect to the behavior of [ʒ]-initial words. I will show that the proposed MaxEnt model generates weights for a set of general and lexically-indexed constraints that predict the lexical trends observed for nonce words as well as the fixed categorical behavior of the existing [ʒ]-initial words.

To examine the learnability of the variable patterns of assimilation of [ʒ]-initial words, I used a MaxEnt learning model incorporating lexically-indexed constraints. The MaxEnt implementation that was used is Harmonic Grammar in R (HGR) (Staubs, 2011), an algorithm that was created to run computations in Harmonic Grammar (Legendre et al., 1990a,b; Legendre and Smolensky, 2006; ?) using R. HGR uses batch Gradient Descent, an online optimization algorithm that is guaranteed to converge on both probabilistic and categorical distributions as long as the model has all of the information relevant to the examined pattern.

4.1 Training data

The training data consists of 104 unique items: 25 items with the preposition [l-] attached to stems beginning with all possible conditions (Tables 3), 49 items with the definite article [l-] attached to non-[ʒ]-initial stems (Table 4), and 34 items with the definite article [l-] attached to [ʒ]-initial words (Table 5). Each input has three different outputs: a fully faithful candidate, an assimilating candidate, and an epenthesis candidate. The proportion of items with each following context as well as their outcome (faithfulness, assimilation or epenthesis) matched the proportions in the lexical statistics.

Category	Subcategory	Count	Assimilated (%)	Faithful (%)	Epenthesis (%)
Preposition [l]	non-coronals + consonant	1	0%	0%	100%
	non-coronals + schwa/vowel	5	0%	100%	0%
	coronals + consonant	6	0%	0%	100%
	coronals + schwa/vowel	13	0%	100%	0%
total:		25	0%	72%	28%

Table 3: Distribution of items with the preposition [l] by initial sound and following context

Category	Subcategory	Count	Assimilated (%)	Faithful (%)	Epenthesis (%)
DA [l]	non-coronals + consonant	5	0%	0%	100%
	non-coronals + schwa/vowel	15	0%	100%	0%
	non-[ʒ]coronals + consonant	5	100%	0%	0%
	non-[ʒ]coronals + schwa/vowel	24	100%	0%	0%
total:		49	59%	30%	11%

Table 4: Distribution of items with the definite article [l] (excluding [ʒ]-initial items) by initial sound and following context

Category	Subcategory	Count	Assimilated (%)	Faithful (%)	Epenthesis (%)
DA [l]	[ʒ] coronals + consonant	10	90%	0%	10%
	[ʒ] coronals + schwa	10	70%	30%	0%
	[ʒ] coronals + vowel	14	36%	64%	0%
total:		34	62%	35%	3%

Table 5: Distribution of [ʒ]-initial items with the definite article [l] by initial sound and following context

4.2 Constraints

The following general constraints were provided to the learning model: *l[cor], MAX(lat), *#ʒʒ, DEP, *CCC, *ə]_σ, IDENT(cor), and *#CCə. The latter two were not mentioned in the analyses for both categorical assimilation and categorical non-assimilation due to their irrelevance to those analyses. IDENT(cor) prevents assimilation from happening in words that begin with a CC sequence whose initial C is a non-coronal consonant. We can see an example of how this constraint is implemented by comparing (31) with (32). Adding IDENT(cor) with a weight of 20 ensures that the assimilated candidate loses against the epenthesizing candidate, which is the expected behavior for nouns beginning with non-coronals.

(31)

$/l_{DEF}/ + /fraʃ/$	MAX(lat) 5	DEP 1	DEP _{DEF} 19	*CCC 25	\mathcal{H}	p
a. ffraʃ	-1	0	0	0	-5	≈ 1
b. $lfraʃ$	0	0	0	-1	-25	≈ 0
c. $ləfraʃ$	0	-1	-1	0	-20	≈ 0

(32)

$/l_{DEF}/ + /fraʃ/$	IDENT(COR) 20	MAX(lat) 5	DEP 1	DEP _{DEF} 19	*CCC 25	\mathcal{H}	p
a. $ffraʃ$	-1	-1	0	0	0	-25	≈ 0
b. $lfraʃ$	0	0	0	0	-1	-25	≈ 0
c. ləfraʃ	0	0	-1	-1	0	-20	≈ 1

In addition to IDENT(cor), $*\#CCə$ was used to account for the difference in the likelihood of assimilation between [ʒ]-initial words with a vowel following [ʒ] and those with a schwa following [ʒ] as observed in the lexicon and experiment. $*\#CCə$ favors assimilation, since unassimilated forms with a schwa following [ʒ] begin with a [CCə] sequence. The assimilated forms, on the other hand, begin with a geminite, which is considered one sound. We can see an example of how this constraint is implemented by comparing (33) with (34). By adding $*\#CCə$ with a weight of 14, the assimilated candidate becomes optimal against the non-assimilating one in a context where the schwa follows [ʒ]. This aligns with the general lexical trends where nouns that have a schwa after [ʒ] assimilate more than those with a vowel following [ʒ].

(33)

$/l_{DEF}/ + /ʒil/$	*l[cor] _{DEF} 13	*l[cor] 1	MAX(lat) 5	*#ʒʒ 20	DEP _{DEF} 19	\mathcal{H}	p
a. $ʒʒil$	0	0	-1	-1	0	-25	≈ 0
b. lʒil	-1	-1	0	0	0	-14	≈ 1
c. $ləʒil$	0	0	0	0	-1	-19	≈ 0

(34)

$/l_{DEF}/ + /ʒənn/$	*#CCə 14	*l[cor] _{DEF} 13	*l[cor] 1	MAX(lat) 5	*#ʒʒ 20	DEP _{DEF} 19	\mathcal{H}	p
a. ʒʒənn	0	0	0	-1	-1	0	-25	≈ 1
b. lʒənn	-1	-1	-1	0	0	0	-28	≈ 0
c. $ləʒənn$	0	0	0	0	0	-1	-19	≈ 0

The learning model was also given the following constraints that are lexically-indexed to the definite article [l]: *l[cor]_{DEF} and DEP_{DEF}. While the indexation of the definite article was manual, the indexations for the nouns in the training data were automatic. The model was given versions of *l[cor], MAX(lat) and DEP that are lexically-indexed to all 104 input forms. In other words, each input had a lexically-indexed version of each of the three constraints (e.g. *l[cor]_{3ar}, MAX(lat)_{3ar}, and DEP_{3ar} for the input /3ar/). This step is necessary to make the model determine what's default and what's exceptional in an unsupervised manner.

4.3 Results for real words

The model learned a grammar that accounts for the assimilation patterns of the sequence [l] followed by a coronal in all possible contexts. It learned the necessary weights for both the general and lexically-indexed constraints. The general constraints accounted for all cases involving the preposition [l] as well as the definite article when attached to items that begin with non-coronals or non-[3] coronals. The model did assign weight to indexed-constraint of some non-[3]-initial words that range between 1 and 5.3, but, interestingly, these weights were not necessary since the categorical behavior of these items was predicted using the general constraints only. With respect to existing [3]-initial items, the modeling results show a three way distinction based on the following context.

4.3.1 Vowel-following context

In a vowel following context, the assimilated forms had 0 weight for the indexed versions of MAX(lat) and a high weight (7.3) for the indexed versions of *l[cor] as shown in the tableau in (35) for the word [3ar] "neighbor". On the other hand, the non-assimilated forms had a high weight (8.1) for the indexed versions of MAX(lat) and a 0 weight for the indexed versions of *l[cor] as shown in the tableau in (36) for the word [3il] "generation". Both the assimilated and non-assimilated forms had the same weight (1.1) for the indexed versions of DEP.

(35)

/l _{DEF} / + /3ar/	*l[cor] _{DEF}	*l[cor]	MAX(lat)	*#33	DEP _{DEF}	DEP	*CCC	IDENT(COR)	*ə] _σ	*#CCə	*l[cor] _{3ar}	MAX(LAT) _{3AR}	DEP _{3AR}	\mathcal{H}	<i>p</i>
a. ɪ^{e} 33ar	0	0	-1	-1	0	0	0	0	0	0	0	-1	0	-11.9	≈1
b. l3ar	-1	-1	0	0	0	0	0	0	0	0	-1	0	0	-19.9	≈0
c. lə3ar	0	0	0	0	-1	-1	0	0	-1	0	0	0	-1	-27.8	≈0

(36)

$/l_{DEF}/ + /zil/$	* l [cor] _{DEF}	* l [cor]	MAX(lat)	*#33	DEP _{DEF}	DEP	*CCC	IDENT(COR)	* \emptyset _σ	*#CCə	* l [cor] _{zil}	MAX(LAT) _{zil}	DEP _{zil}	\mathcal{H}	p
a. 3zil	0	0	-1	-1	0	0	0	0	0	0	0	-1	0	-20	≈0
b. \mathfrak{E} lzil	-1	-1	0	0	0	0	0	0	0	0	-1	0	0	-12.6	≈1
c. ləzil	0	0	0	0	-1	-1	0	0	-1	0	0	0	-1	-27.8	≈0

4.3.2 Schwa-following context

In a schwa following context, the assimilated forms had 0 weight for the indexed versions of MAX(lat) and a weight of 5.8 for the indexed versions of * l [cor] as shown in the tableau in (37) for the word [zəld] ”leather”. On the other hand, the non-assimilated forms had a high weight (9.2) for the indexed versions of MAX(lat) and a 0 weight for the indexed versions of * l [cor] as shown in the tableau in (38) for the word [zəlsa] ”hearing”. Both the assimilated and non-assimilated forms had the same weight (1.2) for the indexed versions of DEP.

(37)

$/l_{DEF}/ + /zəld/$	* l [cor] _{DEF}	* l [cor]	MAX(lat)	*#33	DEP _{DEF}	DEP	*CCC	IDENT(COR)	* \emptyset _σ	*#CCə	* l [cor] _{zəld}	MAX(LAT) _{zəld}	DEP _{zəld}	\mathcal{H}	p
a. \mathfrak{E} 3zəld	0	0	-1	-1	0	0	0	0	0	0	0	-1	0	-11.9	≈1
b. ləzəld	-1	-1	0	0	0	0	0	0	0	-1	-1	0	0	-19.4	≈0
c. ləzəld	0	0	0	0	-1	-1	0	0	-1	0	0	0	-1	-27.9	≈0

(38)

$/l_{DEF}/ + /zəlsa/$	* l [cor] _{DEF}	* l [cor]	MAX(lat)	*#33	DEP _{DEF}	DEP	*CCC	IDENT(COR)	* \emptyset _σ	*#CCə	* l [cor] _{zəlsa}	MAX(LAT) _{zəlsa}	DEP _{zəlsa}	\mathcal{H}	p
a. 3zəlsa	0	0	-1	-1	0	0	0	0	0	0	0	-1	0	-21.1	≈0
b. \mathfrak{E} ləzəlsa	-1	-1	0	0	0	0	0	0	0	-1	-1	0	0	-13.5	≈1
c. ləzəlsa	0	0	0	0	-1	-1	0	0	-1	0	0	0	-1	-27.9	≈0

4.3.3 Consonant-following context

In a consonant following context, the assimilated forms had 0 weight for the indexed versions of MAX(lat) and a weight of 4.8 for the indexed versions of DEP as shown in

the tableau in (39) for the word [ʒmil] "favor". On the other hand, the non-assimilated forms had a high weight (10.8) for the indexed versions of MAX(lat) and a 0 weight for the indexed versions of DEP as shown in the tableau in (40) for the word [ʒtimaʃ] "meeting". Both the assimilated and non-assimilated forms had the same weight (1.2) for the indexed versions of *l[cor].

(39)

$/l_{DEF}/ + /ʒmil/$	*l[cor] _{DEF}	*l[cor]	MAX(lat)	*#ʒʒ	DEP _{DEF}	DEP	*CCC	IDENT(COR)	*ə] _σ	*#CCə	*l[cor] _{ʒmil}	MAX(LAT) _{ʒMIL}	DEP _{ʒMIL}	\mathcal{H}	p
a. ^{ES} ʒʒmil	0	0	-1	-1	0	0	0	0	0	0	0	-1	0	-11.9	≈1
b. lʒmil	-1	-1	0	0	0	0	-1	0	0	0	-1	0	0	-37.9	≈0
c. ləʒmil	0	0	0	0	-1	-1	0	0	0	0	0	0	-1	-20.3	≈0

(40)

$/l_{DEF}/ + /ʒtimaʃ/$	*l[cor] _{DEF}	*l[cor]	MAX(lat)	*#ʒʒ	DEP _{DEF}	DEP	*CCC	IDENT(COR)	*ə] _σ	*#CCə	*l[cor] _{ʒtimaʃ}	MAX(LAT) _{ʒTIMAʃ}	DEP _{ʒTIMAʃ}	\mathcal{H}	p
a. ʒʒtimaʃ	0	0	-1	-1	0	0	0	0	0	0	0	-1	0	-22.7	≈0
b. lʒtimaʃ	-1	-1	0	0	0	0	-1	0	0	0	-1	0	0	-37.9	≈0
c. ^{ES} ləʒtimaʃ	0	0	0	0	-1	-1	0	0	0	0	0	0	-1	-15.5	≈1

4.4 Results for nonce words

The model predictions show that both assimilated and non-assimilated [ʒ]-initial words can have some degree of exceptionality, since the lexically-indexed constraints associated with both subsets of [ʒ]-initial words are shown to have some weight depending on the context following [ʒ]. When disregarding the lexically-indexed constraints, a variable outcome is predicted for nonce words; that is, the context following [ʒ] determines the probability of assimilation. The tableaux in (41), (42), and (43) show the predictions for nonce words in the vowel, schwa and consonant contexts, respectively.

(41)

	$*[cor]_{DEF}$	$*[cor]$	MAX(lat)	$*\#33$	DEP _{DEF}	DEP	*CCC	IDENT(COR)	$*\text{ə}]_{\sigma}$	$*\#CC\text{ə}$	\mathcal{H}	p
$/l_{DEF}/ + /3in/$	12.5	0.02	5.9	5.9	6.7	8.7	24	15.4	11.2	0.9	\mathcal{H}	p
a. 33in	0	0	-1	-1	0	0	0	0	0	0	-11.9	$\approx .66$
b. l3in	-1	-1	0	0	0	0	0	0	0	0	-12.6	$\approx .34$
c. lə3in	0	0	0	0	-1	-1	0	0	-1	0	-26.7	≈ 0

(42)

	$*[cor]_{DEF}$	$*[cor]$	MAX(lat)	$*\#33$	DEP _{DEF}	DEP	*CCC	IDENT(COR)	$*\text{ə}]_{\sigma}$	$*\#CC\text{ə}$	\mathcal{H}	p
$/l_{DEF}/ + /3\text{ə}r\text{q}/$	12.5	0.02	5.9	5.9	6.7	8.7	24	15.4	11.2	0.9	\mathcal{H}	p
a. $\text{33}\text{ə}r\text{q}$	0	0	-1	-1	0	0	0	0	0	0	-11.9	$\approx .84$
b. $\text{l3}\text{ə}r\text{q}$	-1	-1	0	0	0	0	0	0	0	-1	-13.5	$\approx .16$
c. $\text{lə3}\text{ə}r\text{q}$	0	0	0	0	-1	-1	0	0	-1	0	-26.7	≈ 0

(43)

	$*[cor]_{DEF}$	$*[cor]$	MAX(lat)	$*\#33$	DEP _{DEF}	DEP	*CCC	IDENT(COR)	$*\text{ə}]_{\sigma}$	$*\#CC\text{ə}$	\mathcal{H}	p
$/l_{DEF}/ + /3r\text{a}f\text{a}/$	12.5	0.02	5.9	5.9	6.7	8.7	24	15.4	11.2	0.9	\mathcal{H}	p
a. $\text{33r}\text{a}f\text{a}$	0	0	-1	-1	0	0	0	0	0	0	-11.9	$\approx .97$
b. $\text{l3r}\text{a}f\text{a}$	-1	-1	0	0	0	0	-1	0	0	0	-36.6	≈ 0
c. $\text{lə3r}\text{a}f\text{a}$	0	0	0	0	-1	-1	0	0	0	0	-15.5	$\approx .03$

It can be seen from these tableaux that, when a consonant follows [3], assimilation is predicted 97% of the time. When a schwa follows [3], assimilation is predicted 84% of the time. When a full vowel follows [3], assimilation is predicted 66% of the time. A comparison of the predictions of the lexicon and the MaxEnt learner incorporating LICs about the behavior of [3]-initial nonce words are shown in Table 6.

4.5 Discussion

As shown in Table 6, the overall assimilation patterns predicted by the MaxEnt learning model closely match the patterns observed in the lexicon, particularly for the schwa and

Context	MaxEnt with LIC	Lexicon
Vowel after [ʒ]	66%	37%
Schwa after [ʒ]	84%	81%
Consonant after [ʒ]	97%	96%

Table 6: Comparison of Predicted Probability of Assimilation in Each Context for Nonce Words between MaxEnt with LICs and the Lexicon

consonant contexts. However, the predicted assimilation rate for the vowel context was significantly higher than the lexical statistics. Similar behavior has been observed by Hughto et al. (2019) in their investigation of exceptional and variable patterns using a MaxEnt model with lexically-scaled constraints. Hughto et al. (2019) aimed to model both variation and exceptionality in four toy languages based on Russian vowel deletion. In Russian, the vowel of a CV prefix is deleted when attached to stems beginning with a vowel or a single consonant; when the stem begins with a CC sequence, vowel deletion is variable and lexically conditioned.

One of the key findings in their study is the influence of majority patterns in the training data on the model’s predictions. Hughto et al. (2019) show that, as the percentage of triggering CC-stems in the training data increases, the model predicts a higher probability of deleting the prefix vowel before any CC-stem, following the dominant pattern more strongly. This effect is especially prominent when the majority pattern is 60%-100% of the data. As a result, the model generalized this behavior to nonce forms. On the other hand, when there is no clear majority pattern, the model’s predictions for nonce forms more closely followed the lexical statistics.

Similar to what has been shown by Hughto et al. (2019), one possible explanation for the higher assimilation rates predicted by the MaxEnt model for the vowel-following condition is the pressure from other conditions where assimilation occurs frequently or even categorically. First, there is an overall high proportion of assimilation among [ʒ]-initial words in the training data (62%). Second, all items beginning with non-[ʒ] coronals in the training data are predicted to assimilate categorically. This clearly shows that the majority patterns in the training data favors assimilation, explaining the model’s higher probability of assimilation overall and, specifically, in the vowel context.

One potential solution to this issue is to reduce the number of the items giving rise to the majority patterns as this would presumably result in a lower assimilation rate for [ʒ]-initial words with a following vowel. However, removing a large number of items beginning with non-[ʒ] coronal or eliminating all of them, for instance, would result in training data that no longer accurately represent the lexicon of Moroccan Arabic speakers. Therefore, a more effective approach is to revisit the learning models themselves and explore potential solutions to prevent such pressures from affecting the models’ behavior.

The next section presents the results of a nonce word experiment that tests the predictions of the categorical analyses and the MaxEnt learning model that were presented in the previous sections by investigating how Moroccan Arabic speakers generalize their

knowledge about the assimilation patterns of [ʒ]-initial words to nonce forms: whether they regularize (categorical assimilation/non-assimilation) or frequency match.

5 Nonce Word Experiment

To test the predictions of the categorical analyses and the MaxEnt learning model, a forced-choice acceptability judgment experiment was conducted. The experiment was designed to explore the influence of the context following [ʒ] (whether a vowel, a schwa, or a consonant follows [ʒ]) on the assimilation of nonce word. Participants were presented with assimilated and non-assimilated versions of each nonce word and were asked to choose their preferred version. The findings show that participants are sensitive to the phonological context following [ʒ], with a higher rate of assimilation when a consonant follows [ʒ], a lower assimilation rate when a vowel follows [ʒ], and an intermediate assimilation rate when a schwa follows [ʒ].

5.1 Participants

In this experiment, 32 adult Moroccan Arabic speakers were recruited. Some of them were friends and family of the author, while others were recruited through word-of-mouth. The participants were at least 18 years old and were from the cities of Fes and Rabat in Morocco. The experiment was conducted entirely online where participants were able to complete the tasks at their convenience. On average, participants spent approximately 24 minutes to complete the experiment.

5.2 Materials

The stimuli consisted of 42 words: 6 real words and 36 nonce words. Both types were chosen to represent the context following [ʒ]. Real words were equally divided into those with a consonant following [ʒ], those with a vowel following [ʒ], and those with a schwa following [ʒ]. Among these, three words began with coronals, and three with non-coronals.

As shown in Table 7, nonce words were divided into three categories based on the sound following [ʒ] and were also divided into three sets depending on what kind of sound the word begins with. There were 14 words with a vowel following the initial consonant, among which 4 begin with non-coronals, 4 begin with non-ʒ coronals and 6 begin with [ʒ]. There were 14 words with a consonant following the initial consonant classified in the same manner to those with a vowel following the initial consonant. There were 8 words with a schwa following the initial consonant, among which 2 begin with non-coronals, 2 begin with non-ʒ coronals and 4 begins with [ʒ] coronals. No explicit hypotheses were formulated regarding the word shapes selected for each condition; the chosen word

shapes were primarily selected due to their resemblance to existing Moroccan Arabic forms. For instance, the CCVC pattern is common for nouns in Moroccan Arabic, such as [ktab] ‘book’, [ħlib] ‘milk’, and [bnat] ‘girls’. While the chosen nonce words were not close enough to any real words to be noticeably similar to participants, the consonant sequences selected are attested within Moroccan Arabic phonotactics.

Patterns		non-coronals	non-ʒ coronals	[ʒ]-initial words
CV	CVC	ħan, fux	sax, tuʒ	ʒuh, ʒin, ʒas
	CVCəC	fadər, ħarən	tikəl, nadəl	ʒuʁəm, ʒirəħ, ʒaləħ
CC	CCVC	xmig, kfax	ʒjal, ʃruf	ʒfad, ʒmir, ʒbuq
	CCaCa	ħsama, xzada	ʃnara, zmada	ʒrafa, ʒmasa, ʒkala
Cə	CəCC	bərx	dənt	ʒərq, ʒəħt
	CəCCa	gəʃwa	nəħla	ʒəʃwa, ʒərqa

Table 7: The nonce words presented to participants

The words were presented in isolation without the need for a sentence frame, since assimilation is clearly observable in isolated forms. Both assimilated and non-assimilated versions of each item were recorded by the author using Praat in a quiet environment. Before being presented to participants, the recorded items were examined by a native speaker of Moroccan Arabic who is unfamiliar with the task, and who judged the pronunciations to be natural and clearly assimilated or not.

5.3 Procedure

The experiment was conducted using cognition.run, a tool created by neuroscientists for running online experiments. The experiment begins with a consent form followed by two main parts, each designed to assess participants’ preferences for assimilated versus non-assimilated versions of real and nonce words: the training trial where participants are trained on six real words (e.g. [ʃəwd] ”horse”) to get familiarized with the task, and the testing trial where participants were presented with the nonce words. In the training trial, participants were presented with audio recordings of six real words, each in both its assimilated and non-assimilated forms. These audio files were presented in a randomized order to prevent any potential order effects. The position of the assimilated and non-assimilated versions was also randomized. Participants listened to the two versions of each word and selected the one they preferred by clicking on a radio button adjacent to the audio clip. The question presented to participants was ”which definite noun version you prefer?”. The question was written in Moroccan Arabic (with Arabic script).

The testing trial of the experiment introduced the 36 nonce words, again in both assimilated and non-assimilated versions. Similar to the real words, participants listened to two versions of each nonce word and made their preference. Participants were forced to

hear both audio clips and make a selection for one of the audio clips using radio buttons before the arrow button, which takes them to the following screen, became visible. An example screen for how each word was presented is shown in Figure 1.



Figure 1: An example screen of how a certain word is presented to participants

5.4 Results

The results of the experiment followed the trends observed in the lexicon. In other words, the rates of assimilation for the definite article [l-] when followed by [ʒ]-initial nonce words were significantly influenced by the phonological context following [ʒ], i.e. whether a vowel, a schwa, or a consonant followed [ʒ]. In the vowel-following context, the assimilation rate was observed at 66%. This rate increased to 74% in the schwa-following context, suggesting a stronger tendency to assimilate. The highest assimilation rate was 86% in the consonant-following context.

Statistical analysis was conducted using a logistic regression model to further assess the influence of the following context on the likelihood of assimilation as observed in the experimental results. Using Helmert coding, The analysis revealed significant effects of the context following [ʒ] on the probability of assimilation. In the schwa-following context, the likelihood of assimilation decreases ($\beta = -0.91$, $p = 0.005$) compared to the consonant-following context. The likelihood of assimilation decreases further in the vowel-following context compared to both schwa and consonant contexts ($\beta = -0.97$, $p < 0.0001$). The model included random intercepts by participant and by item to account for variability across participants and items.

5.5 Discussion

The nonce word experiment revealed crucial insights about Moroccan Arabic speakers' generalization of the assimilation patterns. The results show that the likelihood of assimilation of [ʒ]-initial nonce words depends on the context following [ʒ]. Speakers showed

variable behavior, which aligns with the predictions of the MaxEnt learner incorporating LICs. A comparison of the predictions of the lexicon, the experiment, and the MaxEnt learner incorporating LICs about the behavior of [ʒ]-initial nonce words are shown in Table 8.

Context	MaxEnt with LIC	Experiment	Lexicon
Vowel after [ʒ]	66%	66%	37%
Schwa after [ʒ]	84%	74%	81%
Consonant after [ʒ]	97%	86%	96%

Table 8: Comparison of Predicted Probability of Assimilation in Each Context for Nonce Words between MaxEnt with LICs, the Experimental Results and the Lexicon

The assimilation rates for the consonant and schwa conditions in the experiment were, to some extent, close to the lexicon rates, i.e. a frequency matching behavior. However, the assimilation rate in the vowel context was significantly higher than the lexicon. This discrepancy may be due to certain aspect of the experimental design itself. Alternatively, it may be related to the organization of the lexicon/corpus. While the corpus used is representative of the Moroccan Arabic speaker’s knowledge of [ʒ]-initial words, it is possible that Moroccan Arabic speakers use two distinct lexicons: one lexicon contains fully integrated (inherent) Moroccan words, while the other consists of words not fully assimilated into Moroccan Arabic structure. This latter set of words often resembles MSA words and is often associated with more educated, religious and political discourse.

To accurately determine if a word is inherently Moroccan Arabic, we must examine whether it has undergone vowel reduction or deletion. When words are derived from MSA, they undergo these phonological changes: short vowels are deleted ($V \rightarrow \emptyset$), and long vowels are shortened ($VV \rightarrow V$) (Kaye, 1987; Scheer, 1997). Words that follow this pattern, as shown (44), are considered inherent Moroccan Arabic words. If we only consider such words in the corpus, the assimilation rate in the vowel context increases to 65%, closely aligning with the experimental results. Therefore, it is possible that, when predicting the behavior of nonce [ʒ]-initial words with a following vowel, participants used the lexical statistics associated with a lexicon composed exclusively of Moroccan Arabic inherent words.

(44)	MSA	Moroccan Arabic	Gloss
	kalaam	klam	‘speech’
	ʔaʕtaa	ʕta	‘he gave’
	ʒaar	ʒar	‘neighbor’

6 Conclusion

This study investigated the assimilation patterns of the Moroccan Arabic definite article [l-] focusing on the variation and exceptionality observed in [ʒ]-initial words. By

examining a comprehensive corpus and nonce word experimental data, this study challenged the previously proposed binary categorizations about the assimilation of [ʒ]-initial words and showed that the observed productive and exceptional patterns are gradient. The findings reveal that assimilation of [ʒ]-initial words is not a categorical but rather a variable phenomenon influenced by the phonological context following [ʒ]. It has also been shown that a MaxEnt model with unsupervised indexation is successful in learning the exceptionality and variation observed within the definite article assimilation patterns. This model was able to account for the observed variability in [ʒ]-initial nonce words as well as the fixed pronunciations of existing [ʒ]-initial words. This shows the potential of this model in capturing the complexity of phonological systems.

The findings from this study contribute to the ongoing discussions regarding the nature of phonological representations and the mechanisms through which phonological patterns are learned. The definite article assimilation patterns in Moroccan Arabic support the view that phonological processes can be gradient and influenced by lexical statistics. This provides an argument for the probabilistic nature of phonological knowledge and its representation. Unlike most previous studies that examine artificial language data or toy languages, this paper offered a detailed examination and application of the MaxEnt model incorporating LICs to a realistic dataset.

Another possible avenue for future research is to examine the predictions of alternative models for learning variable and exceptional patterns (Becker and Gouskova, 2016; Smolensky and Goldrick, 2016; Shih, 2018; Hughto et al., 2019) about the assimilation patterns investigated in this paper as well as similar complex patterns. Another consideration that is worth revisiting is the predictions of the MaxEnt model with LICs for the behavior of nonce words. Despite the commonly held assumption that such models predict frequency matching, we have seen that the model predicted an higher rate of assimilation for the vowel condition, compared to the lexical trends. This behavior has been acknowledged by previous studies (Moore-Cantwell and Pater, 2016; Hughto et al., 2019). However, Further research is needed to reveal the factors influencing this behavior.

References

- Al Ghadi, A. (1990). Moroccan Arabic Plurals and the Organization of the Lexicon. Faculty of Letters and Humanities, Mohammed V University.
- Becker, M. and Gouskova, M. (2016). Source-oriented generalizations as grammar inference in russian vowel deletion. Linguistic inquiry, 47(3):391–425.
- Benhallam, A. (1980). Syllable Structure and Rule Types in Arabic. PhD thesis, University of Florida.
- Blanc, H. (1969). The fronting of Semitic g and the qāl-gāl dialect split in Arabic. In Proceedings of the International Conference on Semitic Studies, pages 7–37, Jerusalem: Harrassowitz Verlag.
- Boudlal, A. (2001). The Prosody and Morphology of a Moroccan Arabic Dialect: An Optimality-Theoretic Account. VDM Verlag, Rabat. Issue: March.
- Ernestus, M. and Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. Language, 79(1):5–38.
- Freeman, A. (2016). Arabic ǰ and the class of Sun Letters: A historical and dialectological perspective. In Perspectives on Arabic Linguistics XXVII, pages 171–185. John Benjamins.
- Goldwater, S. and Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In Proceedings of the Workshop on Variation Within Optimality Theory, pages 111–120, Stockholm University.
- Harrell, R. (1962). A Short Reference Grammar of Moroccan Arabic. Georgetown University Press.
- Harrell, R. and Sobelman, H. (1966). A Dictionary of Moroccan Arabic: Arabic-English. Georgetown University Press, Washington, D.C.
- Hayes, B. (2009). Introductory Phonology. Wiley-Blackwell.
- Hayes, B. and Londe, Z. C. (2006). Stochastic Phonological Knowledge: The Case of Hungarian Vowel Harmony. Phonology, 23(1):59–104. Publisher: Cambridge University Press.
- Hayes, B. and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. Linguistic inquiry, 39(3):379–440.
- Heath, J. (1987). Ablaut and Ambiguity: Phonology of a Moroccan Arabic dialect. State University of New York Press, Albany.
- Heath, J. (1989). From code-switching to borrowing: foreign and diglossic mixing in Moroccan Arabic. Library of Arabic linguistics ; monograph no. 9. Kegan Paul International, London ;.
- Hughto, C., Lamont, A., Prickett, B., and Jarosz, G. (2019). Learning exceptionality and variation with lexically scaled maxent. Society for Computation in Linguistics, 2(1).
- Kaye, J. (1987). Government in Phonology. The Case of Moroccan Arabic. The Linguistic Review, 6(2).
- Legendre, G., Miyata, Y., and Smolensky, P. (1990a). Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: An application. In 12th Annual Conference of the Cognitive Science Society. Psychology Press.
- Legendre, G., Miyata, Y., and Smolensky, P. (1990b). Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations.

- In Proceedings of the 12th Meeting of the Cognitive Science Society.
- Legendre, G. and Smolensky, P. (2006). The harmonic mind : from neural computation to optimality-theoretic grammar. MIT Press.
- Linzen, T., Kasyanenko, S., and Gouskova, M. (2013). Lexical and phonological variation in russian prepositions. Phonology, 30(3):453–515.
- Maas, U. and Procházka, S. (2022). Nominal determination in Moroccan Arabic:. Studies in Language, 46(4):793–846. Publisher: John Benjamins Publishing Company.
- McCarthy, J. J. (1994). The phonetics and phonology of Semitic pharyngeals. In Keating, P., editor, Papers in Laboratory Phonology III: Phonological structure and phonetic form, pages 191–233. Cambridge University Press, Cambridge.
- Moore-Cantwell, C. and Pater, J. (2016). Gradient exceptionality in maximum entropy grammar with lexically specific constraints. Catalan Journal of Linguistics, 15:53–66.
- Outchakoucht, A. and Es-Samaali, H. (2021). Moroccan Dialect -Darija- Open Dataset.
- Pater, J. (2000). Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. Phonology, 17(2):237–274. Publisher: Cambridge University Press.
- Pater, J. (2009). Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. Publisher: Equinox Publishing Ltd.
- Pater, J., Staubs, R., Jesney, K., and Smith, B. C. (2012). Learning probabilities over underlying representations. In Proceedings of the twelfth meeting of the Special Interest Group on Computational Morphology and Phonology, pages 62–71.
- Ridouane, R. (2016). Leading issues in tashlhiyt phonology. Language and Linguistics Compass, 10(11):644–660.
- Scheer, T. (1997). Vowel-zero alternations and their support for a theory of consonantal interaction. page 67. Publisher: Rosenberg & Sellier.
- Shih, S.-h. (2018). On the existence of sonority-driven stress in Gujarati. Phonology, 35:327–364.
- Smolensky, P. and Goldrick, M. (2016). Gradient symbolic representations in grammar: The case of french liaison. Rutgers Optimality Archive, 1552:1–37.
- Staubs, R. (2011). Harmonic Grammar in R. Software package for studying Harmonic Grammar and Maximum Entropy Grammar in R, including features for hidden structure learning. Available at: <https://websites.umass.edu/hgr/>.
- Watson, J. (2002). The phonology and morphology of Arabic. The phonology of the world's languages. Oxford University Press, Oxford, 1. paperback ed edition.
- Zuraw, K. (2000). Patterned exceptions in phonology. PhD thesis, UCLA.