The Verb in Transitional Libyan Arabic: Morphomes, the Stem space and Principal parts

Ramli, Noura

A thesis submitted for the degree of Doctor Philosophy

Department of Language and Linguistics

University of Essex

(October 2016)

Abstract

Should we analyse Arabic morphology in terms of a morpheme-based approach or in terms of a stem-based approach? This is the question which has figured prominently in morphological debate in recent years, especially in Semitic linguistics with ablaut-rich inflectional systems. This study provides a novel synchronic account to Transitional Libyan Arabic morphology, using a stem-based approach that assesses the *morphomicity* (Maiden, 2009, p.45) of stem alternations in the verb inflectional paradigm.

This work focuses on the role of stem alternations in defining inflectional paradigmatic complexity in relation to implicative relations and inflection classes within the stem-space and principal parts morphological approaches. Following Bonami and Boyé's (2002) approach to stem alternations in French, we define an inheritance hierarchy for TLA morphomic verb stems and show how this effectively identifies a set of inflection classes in the absence of affixal allomorphy. Within Stump and Finkel (2013) principal parts model, TLA inflection class membership can be determined by principal parts as indexed stems and/or as substems. The scale of the complexity of TLA inflectional system is also measured using the Principal-Parts Analyzer (PPA) computational tool.

TLA conjugations reveal a synchronic morphomic patterning which shows sensitivity to extramorphological factors. The TLA semi-autonomous morphology is reflected by stem referencing features that provide the base for stem indexing possibilities which in turn can define TLA inflectional classes in the absence of the affix allomorphy. The results of principal parts analysis reveal that verb inflectional complexity of TLA as a Semitic language is as morphologically complex as concatenative stem based systems, posing serious empirical problems for any justifications for a unique distinctive non-concatenative morpheme-based account.

Acknowledgements

All praise is due to Allah (the Sacred and the Mighty) for everything.

I would like to acknowledge the debt I have to my supervisor Prof. Andrew Spencer and thank him for his support and for devoting his time to me, answering my questions with unflagging encouragement and patience. My gratitude also extends to Prof Jonathan Owens and Dr. Wyn Johnson for their comments and support. Among those who were able to contribute, each in their own way to this work, I would like to gratefully thank the academic and administrative staff in the department of language and linguistics. I would also like to thank my supervisory board which included Prof. Louisa Sadler, Prof. Bob Borsley and Dr. Doug Arnold. I am also grateful for my husband, my son, my parents and my family for more than words can say.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2nd SV	The second stem vowel
CA	Classical Arabic
Cond.	Conditional
C	Consonant
Cor	Coronal
DIM	Diminutives
Dist	Distillation
DU	Dual
ELA	Eastern Libyan Arabic
F	Feminine
f	Function
Fut	Future
FV	pre-formative vowel
GEN	Genitive
IC(s)	Inflection Classe(s)
Imper	Imperative
Imperf	Imperfective
Ind	Indicative
Inf	Infinitive
JLA	Al-Jabal al-Garbi Libyan Arabic
LA	Libyan Arabic
Lab	labial
LOC	Locative

M	Masculine
M	Measure
MPs	Morphosyntactic Properties
NOM	Nominative
Obj	Object
OCP	Obligatory Contour Principle
OT	Optimality Theory
Perf	Perfective
PL	Plural
PPA	Principal-Parts Analyzer
Pst. part.	Past participle
Prst.	Present
Prst. Part.	Present participle
SA	Standard Arabic
Seg	Segment
SG	singular
S.pst.	Simple past
Sbjv	Subjunctive
Subj	Subject
S	Stem slot
TAM	tense, aspect, mood
TLA	Transitional Libyan Arabic
V	Vowel

Chapter 1 Introduction

The study investigates the stem patterns of verb inflectional paradigm in TLA, mainly focusing on the nature of these stems and their distinct properties. The notion of stem employed in this study can be defined as the morphological form that is a sub-part of an inflectional form after removing the affixes, but it can also be an unaffixed inflectional form (Aronoff, 1994, Stump, 2001b). Each verb lexeme can be based on a single or multiple stems morphologically motivated through the process of semi-suppletion. The role of stem allomorphy in lexeme formation has been discussed in the description of a number of languages (e.g. Latin (Aronoff, 1994) and Sanskrit (Stump, 2001b)). However, the status of stems in the Transitional Libyan Arabic (TLA) verbal system has not figured prominently in contemporary morphological discourse, although the complexity of this language can provide a rich variety of material for morphological theory.

Traditionally, LA verbs are divided into four verbal forms: sound, double, hollow and defective (Elfitoury, 1979; Elgadi, 1986; Harrama, 1993). These different forms can be treated as inflectional classes. However, unlike languages such as Latin, which have typical inflectional classes, determined by the forms of inflectional endings (e.g. *amat* 'loves', *monet* 'advises', *regit* 'rules' and so on), in TLA, verbal series can determine the forms of verb stems (§1.2), for example, if a lexeme belongs to the sound verb form. Unlike most of the previous work on the Libyan Arabic (LA) verbal system, which has offered descriptive analyses based on the phonological system of verbs (Elfitoury, 1979; Elgadi, 1986; Harrama, 1993), this study follows a different approach using a stem-based model to account for TLA verbal stem alternations.

The aim is to consider the applicability of the notion of stem allomorphy as developed in current morphological theory to TLA verbal inflection and its implications for an exhaustive description of this system. We will assess the status of the TLA verb stem as 'morphomic' (in the sense of Aronoff 1994) in the light of data from TLA inflectional morphology.

1.1 Background

The study adheres to a stem/word-based account in line with considerable work on the Semitic verb and nominal inflectional systems, (Darden, 1992; McOmber, 1995; Ratcliffe, 1997, 2003; Benmamoun, 1999, 2003; Heath, 2003) among others, all of which suggest that the morphological complexity in Semitic languages including Arabic dialects is best handled under a stem/word based account. Therefore, the focus in this study is the nature of stem alternations (Maiden, 1992; Spencer, 2012; Bonami & Boyé, 2002) in TLA and the different factors that may contribute to the alternations within the paradigm. This work will assess the types of stems that exist in the system and how stem-typology can be handled by stem referencing features. Therefore, addressing the segmentation issue is essential.

A TLA verb lexeme can be based on one or more sets of related (identical) stems forming a 'stem-space'. Filling the stem-space can be constrained by dependency relations of varying complexity. Furthermore, the stem variations partition the lexicon into inflection classes (ICs) resulting in a distinct reclassification of the verb system. A similar idea of *partition classes* is suggested by Pirelli and Battista (2000) and a *conjugation pattern* is discussed by Hoberman and Aronoff (2003).

The main concern of this study is the verb inflectional system. However, the derivational morphology will also be considered since the stem alternation pattern of derivationally related verbs also contributes to the morphological complexity of TLA inflectional system. The 'internal organisation' of a paradigm (Zwicky, 1985, p.372;

Maiden, 1992) is not randomly structured but rather displays an 'implicative structure' (Wurzel, 1989, p.114) and dependencies across different forms avoiding redundancies within the verbal paradigm (Bonami & Boyé, 2002). Likewise, the TLA verb stem allomorphy is the locus of generalizations about the morphological complexity (i.e. ICs and implicative relations) of this system. By contrast, affixes show little or no allomorphy. Thus, there are no ICs defined by affix homophony (Chapter 4 and Chapter 5).

The aim of this work is to provide a synchronic based account for the stem alternation of TLA and the IC complexity which can shed more light on *morphology by itself* (Aronoff, 1994) along with *semi-autonomous morphology* (Maiden, 2013). Therefore, a central goal is to fill the gap in Arabic dialect morphology literature and to provide a clear picture on the central role of principal parts in accounting for the morphological complexity in the verb inflectional paradigms in TLA. This study identifies ICs using PPA software that provides a typological account for comparing the TLA system with other systems of the world. We will show that, contrary to traditional assumptions made in the literature, although TLA is a Semitic variety, it is not completely radically different from other non-Semitic inflectional systems.

1.1.1 The Variety of Arabic Under Study

The Libyan Arabic (LA) dialects are part of a language situation that can be characterized by diglossia. The spoken dialect is used in everyday conversation and the official language is the written variety that is Standard Arabic (SA) that is taught at schools and considered to be the symbol of the Arab nation and religion. LA dialects and SA have a number of different characteristics, but the two languages continue to exist side by side, each with a distinctive function.

The dialect chosen as the model for the study of LA is the dialect of Misrata that Owens (1983, 1984) classifies as part of a Transitional Libyan Arabic dialect group (TLA). Owens (1984) identifies three main types of LA which includes the Western dialect spoken mainly in the capital city, the Eastern dialect used in and around the region of Bengasi and finally TLA spoken in Misrata city and Sabha. TLA shares some features of the other dialects, but it also shows distinctive features. Therefore, 'Misrata might be considered a secondary core area by virtue of its unique traits' (Owens, 1983, p.111).

The reason for choosing the Misrata dialect as a model of TLA morphological complexity is that it is my mother tongue. The dialect is spoken in the city of Misrata, the third biggest city in Libya. One source of the data used in this study comes from folk literature, that is, poetry and proverbs, in addition to data from PhD theses concerned with Arabic morphology.

1.1.2 Framework

Traditionally, the morphological complexity of Arabic and many of its varieties has been analysed in terms of consonants and vowels (Root and Pattern approach (McCarthy, 1979, 1981). The verb stem and the affixes are treated as realizations of distinct morphemes. Many of these studies seem to focus on the phonology of the dialect and provide a descriptive analysis of the word classes under a root-based approach in which the verb inflection and derivation is accounted for through the interdigitating of consonants and vocalic patterns. In this study, we analyse TLA verbs in terms of their stems. Some TLA verbs show paradigm internal complexity where the stem alternants are morphologically conditioned (§4.3 §4.5.4 and §5.2). The dialect is well suited for such a study since the verbs show morphomic stem alternations, leading to difficulties for a root and pattern analysis. The stems show implicative relations, and a paradigm can be built on global relations between all the forms. The stem complexity will be investigated within a stem-space model and principal parts theory. Under these approaches, the stem is morphomic, that is, it is a pure form, which does not realize any morphosyntactic properties (MPs) (Aronoff, 1994).

1.2 The terminology as used in this work

Measure (form, Binyanim)

The verb measures in this study will be considered as ‘a function that maps a root into a corresponding set of templates. The result of this mapping is a stem that undergoes inflectional morphology.’ (Aronoff, 1994, p.138) (See §3.1 for further discussion). There are various terms used in the literature to refer to these verb elements. In this work, I use the term ‘measure’, a translation for the terms used in Arabic literature (wazn.SG./ʔawzaan.PL). The measures will be labelled (*M*1, *M*2 ...etc).

Verb series is the verb base form that provides the phonological structure of a verb stem by specifying the number of consonants and vowels forming the verb stem.

Stem form is the syllable structure of the stem.

Indexed stem is a stem listed in the lexicon with a distinctive/separate individual index, which in TLA is represented by any semi-suppletive stem which has a [+morphological] referencing feature.

Stem change pattern is a set of stem variants ABCD that may or may not be distinctive (indexed) stems.

Stem referencing feature is a representative referencing feature for the type of stems in a semi-autonomous morphological system. It reflects the morphological and extramorphological factors contributing to the stem selection and formation specified by binary values.

Stem alternations/stem allomorphy and multiple stems are a set of distinctive stems with [+morphological] referencing feature *vs.* ***allostems*** are a set of related stems with a [-morphological] referencing feature

Paradigm is the notion that 'corresponds to sets of forms, or to sets of abstract 'cells', which are usually defined in terms of feature bundles' motivated morphsyntactically (Stump, 1993, 2001b), and/or morphosemantically (Blevins, 2015, p.87). In other words, 'a paradigm is a set of inflected forms based on a common lexeme or stem' (McCarthy 2005, p.170). Thus, a paradigm can be based on both inflectionally- and derivationally-related forms (Blevins, 2001).

Complexity. 'The complexity of an IC system in a realized paradigm is the extent to which it inhibits motivated inferences about a given lexeme's realized paradigm from subsets of that paradigm's cells' (Stump & Finkel, 2013, p.381).

Inflection class (IC) is 'a class J of lexemes such that (i) J's members are distinguished by a common pattern of inflection and (ii) membership in J has no syntactic significance; conjugation classes and declensions classes are inflection classes. An IC system for a syntactic category C is a set S of ICs such that every lexeme belonging to C belongs to some IC in S' (Stump & Finkel, 2013, p.11). ICs are morphomic and their common inflection pattern can be based on affixation or stem alternations.

The structure of the thesis will be as outlined in the section below.

1.3 The thesis outline

Chapter 2 Roots and Patterns, Paradigms and Stems The aim of this chapter is to review the issue of the status of Arabic morphology as morpheme-based vs a stem-based. Each approach has different empirical predictions. However, the morpheme-based account for Arabic, particularly for the dialects, morphology may not be adequate. The chapter briefly discusses the literature on Arabic verb morphology as a non-concatenative and/or concatenative system. The first approach considered is the Root and Pattern or 'autosegmental morphology' approach as developed in notable works such as McCarthy (1979, 1981), McCarthy and Prince (1990a), in which the verb system is treated as having unique lexical autosegmental elements such as the consonantal and vocalic patterns, which is also a feature of other languages of the Semitic family. The alternative approach considers the stem as the base for inflectional and derivational processes, dismissing the need for assigning a unique status to Semitic systems as root-and-pattern-based. In addition, the chapter introduces a detailed account for the types of stems and stem properties that have been discussed in the literature mainly for Romance languages. This chapter also considers the implications of stem alternations for autonomous and semi-autonomous morphology. Finally, the chapter will present the two stem-based accounts: the stem-space and principal parts models, both of which serve as the framework for this study.

Chapter 3 The TLA Verb introduces a comprehensive discussion of the verb system in TLA. The chapter will discuss the verb classification based on verb measures and series, along with a brief introduction to TLA phonology and paradigmatic properties. In addition, the chapter reviews a number of previous analyses of Libyan Arabic (LA) verb morphology using different frameworks all of which provide a useful source of comparison for the account I present in this thesis. Relevant analyses will include the work on Eastern Libyan Arabic (ELA) (Owens, 1984) and al-Jabal al-Garbi (the western mountain) Dialect (JLA) (Harrama, 1993). The two different varieties show distinctive morphological features that can be compared with the data of this work. However, none of these previous studies seems to provide a systematic theoretical framework to account for the verb morphological complexity with economy of description.

Chapter 4 TLA stem complexity The purpose of this chapter is to discuss the stem complexity in the TLA verb inflectional paradigm and to consider the pattern of stem change across different measures. The stem pattern provides a new perspective on the relationship between verb measures, showing that some verbs that belong to different measures share the same stem change while related measures within the same series have distinct stem behaviour. However, factors motivating the stem alternation can vary and result in morphologically complex stems with different sets of distinct properties/features. However, any account of morphological complexity needs to address the methodological challenge of stem-affix boundaries (the 'segmentation problem'), which is crucial for stem paradigmatic patterns. In addition, the chapter reviews two different accounts of segmentation provided for different varieties of Libyan Arabic and proposes an account for TLA. A framework for representing the stem change and stem properties will be presented with an exhaustive discussion of the pattern of stem change in each verb series of all the triliteral verb measures. The chapter explains how the interaction of these synchronic conditioning effects results in semi-autonomous TLA stems.

Chapter 5 Inflection classes and implicative relations argues for a novel stem-based classification of TLA verb system based on implicative relations and paradigmatic concept of ICs. The chapter develops the stem-space account for TLA in which the morphomic stems are hierarchically defined, based on dependency relations; hence, they are lexically defined classes. The implicative relations in the system will also be assessed by principal parts that represent the network of inferences across lexemes and/or forms of each lexeme, providing an important device for determining the scale of complexity of different inflectional systems (Stump & Finkel, 2013). The results shed further light on the morphological system of TLA in relation to other non-Semitic stem-based systems.

Chapter 6 Conclusion The final chapter presents a summary of the discussions and suggestions for further future research.

Chapter 2 Roots and Patterns, Paradigms and Stems

Which framework of morphology can provide the most plausible account for the ablaut-rich morphology of Arabic? This issue of the status of Arabic morphology as morpheme-based or stem-based has figured prominently in the research literature in the last few years. These two distinct models lead to different empirical predictions and they are faced by a number of universal issues which in most cases suggest that the morpheme-based account of Arabic morphology, particularly that of the dialects, is considerably problematic. I will argue that it is important for Arabic linguistics to provide a clear picture on the importance of the theoretical notion of stem in accounting for the ablaut alternations in verb inflectional paradigms by appeal to the notion of morphology *by itself* (Aronoff, 1994). In addition, it is important to evaluate the central role of principal parts in reflecting the paradigmatic implicative relations and inflection classes in Transitional Libyan Arabic.

Traditionally, the morphology of Arabic has been treated as a non-concatenative type, having unique lexical elements including the consonantal and vocalic patterns, which are also a feature of other language families. This approach to Semitic morphology has been first invoked for lexicography in Medieval Arabic and then adapted and modified by what is widely known among most linguists as ‘Root and Pattern’ or ‘Prosodic Morphology’ (McCarthy, 1979, 1981; McCarthy and Prince, 1990a; Watson, 2002). Under this approach, exponents in morphological patterning are formalized by prosodic units of foot, syllable, mora feet among others. Therefore, the autosegmental phonology framework (McCarthy, 1981; Goldsmith, 1976; Harris, 1941) was proposed to account for derivational and inflectional morphology of both verbal and nominal morphophonological alternations in Arabic along with other

Semitic languages. This approach seemed to provide a plausible account for the morphological complexity of Arabic at that time and even earlier. However, there are theoretical hurdles raised by this morpheme-based approach.

A radically different perspective on Semitic Morphology known as the word-based/stem-based approach has been proposed. This approach promotes a uniform account of Arabic using the same universal constraints operating across different world languages; it is in many cases data driven, as is the case with this study. A small number of studies have argued that there is no need to invoke the separation of consonantal roots and vocalic elements when the morphological patterning in Arabic can be handled and effectively accounted for by the stem (McOmber, 1995; Benmamoun, 1999, 2003; Gafos, 2003; Heath, 1987, 1997, 2003; Ratcliffe, 1998, 2004; Darden, 1992; Camilleri, 2014). For example, McOmber (1995) argues that the representation of the lexical entry in Arabic can be the CCVC stem and there is no need for the consonantal root representation. Similarly, Benmamoun (1999, 2003) is in favour of a stem-based model in which the imperfective verb form has underived status and serves as the verb default form and the base for lexeme formation processes. Similar proposals are made by Radcliffe (1998, pp.23-50, 2003) who states that '…the root and pattern analysis alone is insufficient in terms of modern notions of descriptive adequacy.' Therefore, '…many of the morphological processes of Arabic must make reference to fully formed words, rather than simply to roots, templates and vowel melodies'. In other words, the Arabic lexicon is based on words rather than roots. The stem account is also proposed for Arabic dialects such as Cairene (Darden, 1992), Moroccan (Heath, 1987). The inadequacy of the root and pattern approach is also demonstrated by Optimality Theory (OT) accounts of Semitic morphology (Ussishkin, 1999, 2000; Gafos, 2003). Most of these studies have been concerned with the stem form and role of stems in word formation. However, none of the studies on Arabic has specifically addressed the issue of the morphomic status of the stem and the motivating factor(s) for stem allomorphy in Arabic verb inflectional paradigms. Furthermore, there is hardly any work investigating the role of stem allomorphy in classifying and determining principal parts and inflection classes in the Arabic system (§5.2 and §5.3).

The choice between these two different frameworks is not only essential for providing a plausible account and a better understanding of the system, but also it is crucial for modelling other theoretical domains including the system's inflectional classes, paradigm based implicative relations and also for Arabic computational linguistics. In fact, both approaches have been reflected in computational models of Arabic morphology, used as the base of 'knowledge-based methods' (Soudi et al., 2007, p.4). McCarthy's root and pattern approached has been implemented to model Arabic processing, using DATR to describe morphological dependencies in Arabic (Al-Najem, 1998). Similarly, Stump and Finkel (2013) suggested a non-concatenative (plat[1]) representation for Semitic languages for computationally capturing the principal parts analysis and inflection class complexity in the system (see Chapter 2, §4.2.1). However, other studies have developed a computational framework based on the notion of stem, such as the Lexeme-based and Stem-based approaches, motivating a fundamental change towards a uniform approach adhering to the universal rules of other systems (Cavalli-Sforza et. al., 2000).[2] Likewise, the computational analysis of principal parts can be based on stems in inflectional systems such as French in which the IC complexity is reflected by stem allomorphy (see Chapter 5, §3.3.4).

---

[1] It is a notion developed by Stump and Finkel (2013, p.40) and it is 'a format for the representation of IC systems that is essential for identifying and measuring differences in their complexity' (see Chapter 2, §4.2.1 for further details).

[2] The syllable-based morphology account is also implemented in Arabic computational studies in which the prosodic levels of feet and morae are ignored in favour of the syllable structure of a stem; thus, inflectional forms are defined by indexed syllables (Cahill & Gazdar, 1999a, b).

Given the two different approaches the aim of this chapter is to introduce some of the main assumptions of McCarthy's autosegmental morphology and show that this morphemic perspective is problematic for the theory of morphology in TLA and seek a plausible account. The alternative approach is a stem/word account which will be presented in §2.2. Then, a discussion of the notion of the stem as developed in stem-based accounts such as Anderson (1992), Aronoff (1994), Spencer (2012), Stump (2001b), Bonami and Boyé (2002) will be presented (§2.3). Section 2.4 will introduce two approaches to paradigmatic implicative relations and ICs based on the notion of stem-space and principal parts.

2.1 Non-concatenative Semitic morphology

The root-and-pattern, or non-concatenative approach to Arabic is mainly based on the observation that a great majority of the morphological processes in the language involve modifying the internal structure of a word, for example, by the process of ablaut (e.g. Arabic *labisa* 'he wore' ~ *yalbasu* 'he wears' cf. English *man* ~ *men*). By contrast, in the concatenative morphology the morphological process is applied word-externally by attaching delineated segments as (*walk* ~ *walk-s*). The account is mainly based on word formation and derivation. The other issue is how to account for the morphological forms and the formal relations between measures (derived verbs) such as *M*1 & *M*2 and *M*1 & *M*7 passive (see Appendix A for the list of SA measures). The non-concatenative morphology is the base of many morphological models of Semitic languages, including the Autosegmental Morphology (McCarthy, 1979, 1981), Root and Pattern (McCarthy & Prince, 1986) and Root and Sites (Guerssel & Lowenstamm, 1990). The difference between these versions is mainly based on the number of the proposed morphemic units and they vary in the approach followed in constructing the surface word forms.

Arab grammarians used the earliest version of this model in the medieval era, and they based their work on a morphological system that consists of two morphemes: a consonantal root which represents a broad semantic feature and a vocalic pattern for realizing the morphosyntactic feature of aspect and voice. The Semitic root is a consonantal discontinuous unpronounceable entity. In addition, it represents a sub-lexical bound morpheme without a lexical category. However, it conveys meaning and the radicals represent a single structural entity. By joining the two morpheme units, we construct the masdar/deverbal noun that is used as a default word form for deriving all the other surface forms in the language through various processes of morphophonology (affixation, vowel elision or insertion). For example, the verb *katab* has the consonantal root /KTB/ conveying the broad meaning of 'writing' when combined with the vocalic pattern /a, a, a/ the result is the verb form *kataba* 'he writes' (Boudelaa & Marslen-Wilson, 2001).

The masdar from this root is *kiitaba* based on the pattern (CVVCVCV) which can derive the noun *kitaab* 'book' by vowel deletion and insertion. As Bohas and Guillaume (1984) (cited in Boudelaa & Marslen-Wilson, 2001, pp.32-33) emphasize this traditional analysis of Arabic is not *strictly* root and pattern, since the lexeme formation proceeds from the masdar form through a tree structured derivation rather than the mapping process between root and template.[3] For instance, the root /QTL/ is assigned to the pattern CaCC to derive the masdar *qatl* 'killing' which is the base form for the lexeme formation to derive the verb perfective form *qatala* 'he killed' by modifying the stem vowel. This verb form is in turn used as the base for forming the imperfective verb form *yaqtul* 'he kills' and finally, the formation of the nominal form

[3] The view that the masdar is the base of lexeme formation is held by the Basra school while for the Kufa school the base of derivation is the perfective verb form (Boudelaa & Marslen-Wilson, 2001).

maqtal 'the killing of'. Clearly, the focus of this early version of root and pattern was defining the function and semantic features of the resulting measure itself (e.g. the measure CVCCVC consistently defines a transitive causative/intensive verb form (Ratcliffe, 1998)).[4]

Another model of the root and pattern was put forward by McCarthy (1979, 1981) using the autosegmental phonology theory as a tool to handle morphological issues in Arabic. McCarthy's original proposal has been regarded as the most influential analysis of the morphological organization of the Semitic system. Under this template morphology, the morphology of Arabic (such as underived nouns and verb stems) is accounted for by three morphemic units: a morpheme carrying the meaning is presented by the consonantal root carrying the meaning; the second morpheme is the vocalic melody representing the morphosyntactic features and an abstract morpheme known as the CV-skeleton. The vowels and consonants are represented as separate autosegments projected on different tiers each of which conveys only single morphemic information. The two types of tiers are then associated with the abstract template also known as the CV-skeleton. The role of the template is to determine the phonological surface form and to represent some morphosyntactic information. For instance, a form such as *katab* 'to write' consists of the root /KTB/ projected as a separate autogsegment while vowels *a*, *a* represent the vocalic melody; both morphemes form the CV-skeleton CVCVC. McCarthy (1981, p.388) proposes that the separate autosegmental tiers of the root consonant and vowel melodies of a word are linked into the core template by association conventions which are subject to the well-formedness conditions. These conditions were adapted from autosegmental phonology

[4] A different model of root-and-pattern was developed by Cantineau (1950) who proposed that every surface form is derived by mapping a root to a template by a set rules of association. For example, the derivation of *qatl* 'killing' and the the nominal form *maqtal* 'the killing of' requires the mapping of the root /QTL/ onto the templates CaCC and maCCaC respectively (Boudelaa & Marslen-Wilson, 2001).

constraints for prosodic features of tone and vowel harmony. For mapping morphemes onto the root and affix tiers, McCarthy (1981, p.382) uses universal association conventions some of which are very similar to the conventions formulated for tonal phonology (Goldsmith, 1976). These conventions include the following:

(1)

‘i. If there are several unassociated melodic elements and several unassociated melody-bearing elements, the former are associated one-to-one from left to right with the latter.

ii. If, after application of the first convention, there remain one unassociated melodic element and one or more unassociated melody-bearing elements, the former is associated with all of the latter.

iii. If all melodic elements are associated and if there are one or more unassociated melody-bearing elements, all of the latter are assigned the melody associated with the melody-bearing element on their immediate left if possible.’

In addition, the CV-tier employs the features [±syllabic] of which any [-syllabic] segment such as consonants and glides is projected on the C-element tier while the [+syllabic] segment reflects the V-elements. Consider the following derivation examples of the perfective active and perfective passive *M*1.

a. Perfective Active

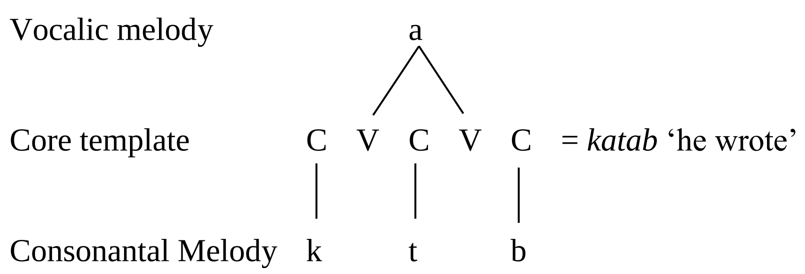


Figure 1 Autosegmental representation for the perfective active verb *katab* ‘he wrote’.

b. Perfective Passive

Vocalic melody		u		i		
Core template	C	V	C	V	C	= *kutib* ‘it was written’
Consonantal Melody	k		t		b	

Figure 2 Autosegmental representation for the perfective passive verb *kutib* ‘it was written’.

In these verb forms, the core template of the Classical Arabic (CA) root *ktb* is CVCVC to which the consonantal root elements are associated with the C-slots while the vowel *a* is initially assigned only to the leftmost vowel through the universal left-right association convention. McCarthy (1986) argues that it is not possible for the autosegmental tiers to contain the elements with the same feature unless they represent different morphemes. Identical autosegmental elements are prohibited by the Obligatory Contour Principle (OCP). Therefore, in the perfective active, there is only a single vowel *a* which is associated with the second V-position by the feature-spreading principle. An important aspect of this framework is that the same three morphemes are employed in lexeme formation for deriving forms such as the verb *kattaba* which is a *M*2 verb form with a causative meaning of form *katab*. Thus, the triliteral verb roots can have fifteen measures and the quadriliteral verbs can be classified into four verb forms. The canonical patterns of triliteral perfective forms of eight measures can be generated by the template in (2.i):

(2) The Prosodic Template (or the CV Template) for the triliteral perfective measures (McCarthy, 1981, pp.386-387):

i. $[(\left\{\begin{matrix} C \\ CV \end{matrix}\right\})$ CV ([+ seg]) CVC] ⟶

ii. V →Ø/ [CVC__CVC]

a. CVCVC	
b. CVCCVC	f. CCVCVC
c. CVVCVC	g. CCVCCVC
d. CVCVCCVC	h. CCVVCVC
e. CVCVVCVC	

This prosodic template has only the two features [syllabic] and [segmental], whose values can be represented by a consonant or a vowel depending on the measure. The scheme consists of two expansions each of which is characteristically specified for a particular verb measure. According to this prosodic template, all the generated forms have to end in a closed syllable by rule (i) while rule (ii) applies to all the patterns generated by (i) redundantly to eliminate the illicit [***CVCV**CVC] with a sequence of two light syllables. The forms in (a, b, c, d, e) can be generated by the CV([+ seg]) CVC]. By contrast, the patterns in (e, f, g) are the output of the second expansion of the curly brackets in the template, which also generates the prohibited pattern [***CVCV**CVC] subsequently blocked by rule (ii). In addition, these canonical patterns show that light syllables are not allowed to precede heavy syllables (e.g. '*CVCCVCVC') and forms that have initial consonant clusters, are the patterns which consist of at least three syllables. The resulting patterns represent the canonical forms from which any verbal stem has to be derived (McCarthy, 1981, p.386). The templates in 2.a, b, c, d, and e appear in *M*1 (CVCVC), *M*2 (CVCCVC), *M*3 (CVVCVC), *M*5 (CVCVCCVC) and *M*6 (CVCVVCVC) respectively. The template in 2.f is used in *M*7 with an n-prefix (nCVCVC), in *M*8 with a t-infix after the first consonant (CtVCVC), and in *M*9 with a geminate final root consonant (CCVC$_i$VC$_i$). The pattern in 2.g is the template of *M*10 with prefixed *st*- (stVCCVC) while the pattern in 2.h appears in *M*11 with the gemination of final root consonant (CCVVC$_i$VC$_i$) (McCarthy, 1981) (see Appendix A).

With regard to the affix material (e.g. the morpheme /t/ of *M*8, /n/ of *M*7 and /st/ of *M*10), McCarthy (1981) treats these affixes as morphemes assigned to separate autosegmental tiers and they share the same root tier distinctive features. These affixal tiers are defined by the grammar of the language, through morphological rules identifying the appropriate affixes for each measure and the correct mapping to their prosodic template position. Therefore, these affixes are linked with the core template elements by the left-to-right association conditions in (1) which ensure that the affix projection precedes the mapping of the consonantal root. For instance, *M*6 (taCVVCVC) has the affix /t/ occurring as a prefix morpheme *takaataba* 'write to each other'. The affix material is assigned first to the prosodic template and then followed by the consonantal root (Figure 3).

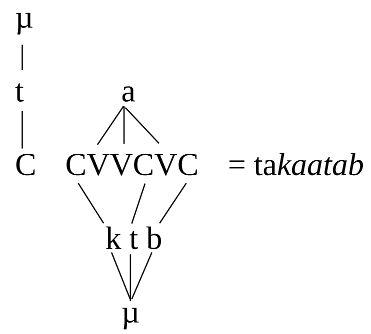


Figure 3 Autosegmental representation for *M*6 *takaatab* 'write to each other'

The output of the left-to-right association to both the affix and the root tiers results in mapping the affix material onto the leftmost slot of the template and then the root elements are projected onto the root tier, subject to the association convention in (1) which prohibits the many-to-one association. However, the affix material of verb measures can also be infixed as in *M*8 *ik(t)ataba* ‘to write’, requiring a subsequent rule of association to ensure the correct output. McCarthy handles this case by arguing that, in both *M*8 and *M*6, the morpheme /t/ is initially treated as a prefix linked with the first consonant position and subsequently the *M*8 affix is mapped to its correct tier position by a ‘flopping rule’. This rule changes the position of the affix with the feature reflexive from the first to the second slot (Figure 4):

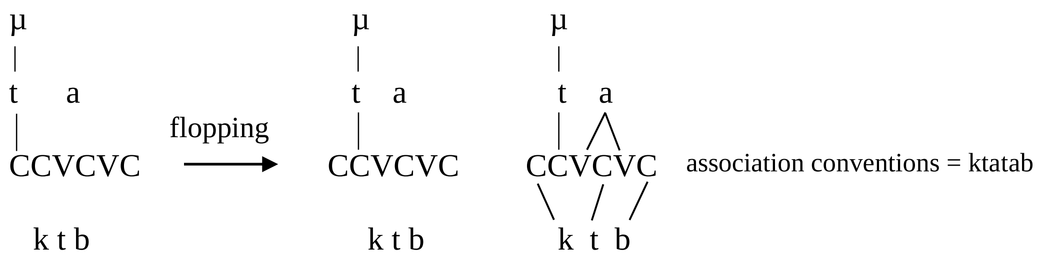


Figure 4 Autosegmental representation for *M*8 *ktatab* ‘to write’

The reflexive affix /t/ is projected first and assigned to the leftmost consonant position in the template (**C**CVCVC) and the left-to-right associations assign the root elements. Subsequently, the flopping rule shifts the affix to the second position slot which is already occupied by the root element. However, the restriction against many-to-one association presented in (i) ensures that the root element is reassigned to the other available slot resulting in the correct output.

Other cases in which the left-right mapping requires the interaction with a subsequent rule in order to produce the correct forms involve the derivation of *M*2 and M5 verbs. For example, applying the left-to-right association, the restriction on many-to-one association and the third association convention in (1) to the root C-elements in the derivation of *kattab* CaCCaC (*M*2 perfective active), yields the wrong output **katbab*. These association conventions map the root elements *ktb* from left to right, one-to-one, onto the first, second and third slots respectively (*katbaC*). The remaining empty fourth slot is associated with the root element occupying the slot on the left (**katabab*) by the third universal convention, resulting in the incorrect form**katbab* (Figure 5). McCarthy (1981, p.391) handles such cases by the 'Erasure' rule in (2) that deletes an incorrect association line between the penultimate consonant and the core template tier. Subsequent re-association of the root consonant /t/ by the general association conventions leads to the required result.

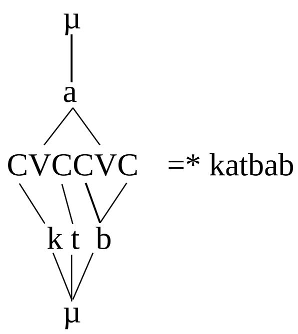
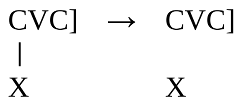


Figure 5 The application of association conventions without the Erasure rule to *M*2 perfective active.

(3) Erasure rule (*M*2 and *M*5)



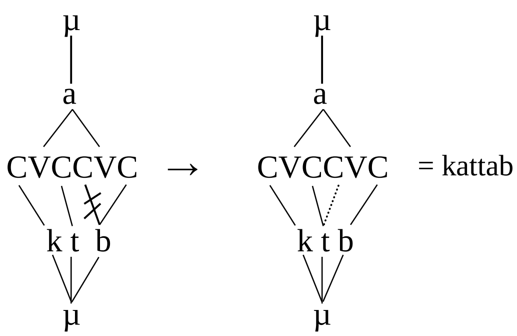


Figure 6 Autosegmental representation for *M*2 *kattab* 'cause to write'

The Erasure rule deletes the association line, linking the penultimate consonant with the root melody as in (Figure 6). After the application of the Erasure rule, the linked C-slot in penultimate position can be re-associated with the consonant /t/ in the left position by the general convention of association.

Another important aspect of this model is that inflectional categories such as aspect and voice are considered to be represented by arbitrary vocalic morpheme alternations for which McCarthy (1981, p.400) provides the non-*M*1 vocalic tier rules in Table 2.1.a whereas *M*1 verbs are associated with the ablaut classes in Table 2.1.b.

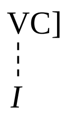
a. Non-*M*1 vocalic pattern		Examples	
		*M*2 kattab ‘cause to write’	*M*6 takaatab ‘write to each other’
Perfective active	a_2^4	kattab	takaatab
Perfective passive	u_1^3 i	kuttib	tukuutib
Imperfective passive	u a_2^4	ukattab	utakaatab
Active participle	u a_1^3 i	mukattib	mutakaatib
Passive participle	u a_2^4	mukattab	mutakaatab

b. *M*1 vocalic pattern in the second stem vowel					
class	Perf	Imperf	Examples		
1	a	i	ḍar*a*b	yaḍr*i*b	‘beat’
2	a	u	kat*a*b	yakt*u*b	‘write’
3	i	a	ʕal*i*m	yaʕl*a*m	‘know’
4	u	u	ḥas*u*n	yaḥs*u*n	‘be beautiful’

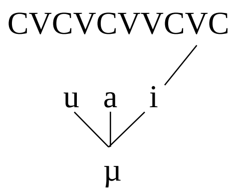
Table 2.1 the vocalic pattern in *M*1 and non-*M*1 verbs and participles.

McCarthy (1981) combines the vocalism of non-*M*1 into single sets of melodies in which the leftmost vowels replicate in the template while a stem vowel final *i* requires a special rule. This rule assigns any *i* to the rightmost template vocalic position and must be applied before the universal conventions. Table 2.1.a represents the sets of rules that are used to account for both the stem vowels in addition to the affix vocalic pattern. In the perfective active forms, the stem contains from two to four vocalic morae, all of which are the low back vowel *a* while the perfective passive can have two different vocalic patterns: the high back vowel *u* which can consist of up to three vowels and the high front vowel *i* has to be the last vowel. The same representation pattern follows for imperfective passive, the active participle and the passive participle of non-*M*1 verbs. McCarthy assumes that these vowels are separate morphemes that represent the vocalism of non-*M*1 verbs and, as pointed out earlier, the association of vowel melodies with the V-slots in the core template is generally governed by the left-right association and the spreading constraints. However, there are cases in which projecting the vocalic pattern requires pre-association to guarantee the correct vowel stem position. For instance, the association of the vowel *i* in the perfective passive and active participle to the final V-position in the template has to precede the association of the vowel *u* which occurs before *i*. Thus, overall, the approach handles the vocalism of all bases/stems (e.g. active participle of M6, *mu-takaatib* (CV-CVCVVCVC)), except active imperfective forms and *M*1, through the rules of vowel tiers in Table 2.1.a (in the given example *u*, *a*, *i*), and the pre-association of the vowel *i* with the rightmost vowel (CV-CVCVVCiC). The final procedure includes other conventions followed by the left-to-right application (*a* is associated by the spreading principle to three vocalic positions, Cu-CaCaaCiC) (p.401) (Figure 7).

(4) Vowel association



By rule (4)



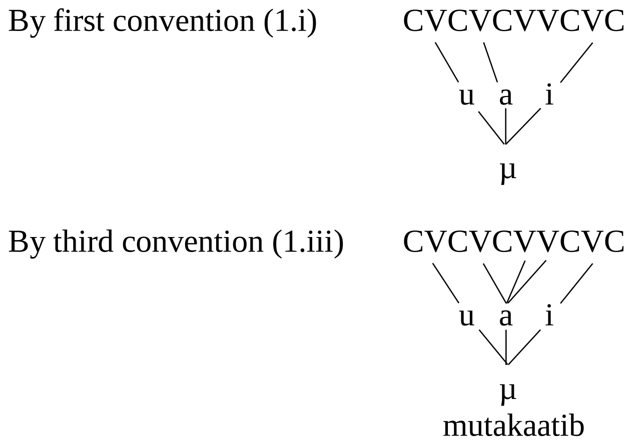


Figure 7 The derivations of the vocalism of the active participle form of *M*6 *mutakaatib*

This pattern of derivational steps for the imperfective active of non-*M*1 verbs is problematic for cases in which we have up to four variable patterns of vowel alternation composed of the prefixes and stem vowels. McCarthy briefly suggests that some of these patterns can be handled by deletion rules.

Likewise, the vocalism of *M*1 is handled differently from other measures (Table 2.1.b). The basic vocalic pattern for *M*1 active stems includes the vowel *a* for both vowels in the perfective stem and the vowel *u* for the single vowel of the imperfective stem (e.g. *katab/ktub*). However, unlike non-*M*1 verbs, some *M*1 verbs show ablaut change in the second vocalic position of the perfective stem and in both the affix vowel and stem vowel positions of the imperfective. By contrast, the first stem vowel in the perfective active is invariably *a*. The ablaut pattern in *M*1 can be lexically determined as McCarthy (1981, p.403) states that 'we can give only a lexical account of assignment of any given root to an ablaut class'. However, there can be a polarity based interpredictability through the value of [high] between perfective and imperfective sub-paradigms, (excluding class 4 (Table 2.1.b))[5] in which both aspect forms have *u* vowel) with the ‘regular semantic property of stativity’ (1981, p.403). In CA, when the vowel of the imperfective form is [+ high], the perfective counterpart would be [-high] in the second V-position. On the other hand, the presence of the [+high] vowel *i* in the perfective stem in second V-position corresponds to a low vowel *a* of the imperfective. However, McCarthy’s approach to verb vocalism has been criticised on various grounds for obscuring some important grammatical generalizations as will be discussed below.

In sum, the crucial rules of this morphemic approach to Arabic nonconcatenative morphology include the Prosodic Template, the affix specifications and their special association principles, Flopping Rule and Erasure Rule. Any other morphological alternations can be captured by independently motivated autosegmental principles and conditions.

2.1.1 An excursus on the history of the morphemic concept

The morphemic approach was used in generative grammar to account for the verbal/nominal phonological patterning in Semitic languages. The notion of morpheme has been used in many theories of morphology and usually defined as the minimal linguistic unit which has a meaning. This is the core assumption of lexical-incremental theories according to which the morpheme device can account for morphological complexity patterning in concatenative and non-concatenative language kinds (Stump, 2001b). This theory of morpheme was postulated by Bloomfield in his work (1926) *A Set of Postulates for the Science of Language.* There are many remarkable studies of the field of morphology that devote considerable attention to the notion of morpheme

[5] In Arabic, class 4 verbs with the vocalic pattern CaCuC tend to be attributive verbs (e.g. *ḥasun* ‘be beautiful’, *badun* ‘to be, become fat’) which ‘often denote permanent qualities’ of the first argument (Cuvalay-Haak, 1997, p.96).

and its status in different languages, the origin of this term and its development within Bloomfield's structuralist approach have been widely discussed. Under Bloomfieldian theory, a simple linguistic form is a morpheme and it is a meaningful phonetic form which does not resemble any phonetic-semantic component of any other form. A string of two or more constituents make a complex form which comprises of immediate constituents carrying the meaning. This means that each morpheme has two aspects, phonetic and semantic. Bloomfield (1935) identifies two types of morphemes, *free* and *bound*. Free morphemes can be a word on their own while a bound morpheme is a linguistic form which has to be dependent on other morphemes. A bound morpheme attaches to either a free or a bound morpheme in word formation; for example, *dog* ~ *dog + s*. All types of morphemes are stored in the lexicon and through grammar; the morphemes can form words and phrases.

Problems arise when we come to describe the phonological realization of morphemes. For example, the pluralization of most English nouns such as 'cats' is achieved by suffixation of the bound morpheme [z]. In addition, the choice of the plural suffix can depend on the word, for example, *dog* [z], *roses* [iz]; thus, there are three alternates of plural /s/ that are lexically determined. The phonetic modification of the morpheme [z] to [s] occurs when the word final consonant is voiceless whereas the phonetic modification of the morpheme [z] to [iz] happens when the words end in a sibilant sound, elsewhere [z] is selected. However, the application of the morpheme account seems to be problematic for the morphological processes that involve vowel variation, for example (*man* ~ *men*). According to Matthews (1993, p.71), the pluralization of this example involves two features of arrangements: the first is the modification [a] to [e] and the second is the selection (a noun is selected from a special class). Unfortunately, Bloomfield does not clarify the process that can account for this plural alternation. Therefore, a different approach is needed to account for words that pluralize neither by a morpheme nor by a process (vowel alternation). Bloomfield's account of plurals such as (*sheep* ~ *sheep*) in which there is no difference in the phonetic form of the singular and plural: /ʃip/ involves the zero element. He argues that the plural morpheme is replaced by zero, claiming that the constituent lacks form.

Many modifications to Bloomfield's account were introduced by American structuralism in response to the problems of non-concatenative processes. The work of the successors (Bloch, 1947; Hockett, 1947) in the Bloomfieldian school advocated a number of solutions to many inexplicit issues in Bloomfield's account of morphology. One of the most important works includes Harris' (1942) approach which introduced the notion of a morpheme unit in which the morpheme alternants (plural morphemes) are grouped. There is a morpheme unit which has the alternant [naif] ~ [naivz] → [singular] ~ [plural]. Harris (1942) argues that it is possible to perform the phonetic modification through a morpheme adding a specific phonetic unit, a morpheme removing that phonetic unit, or both morphemes. Therefore, the vowel alternation is considered to be a special type of morphological modification where *man* ~ *men* are not the same morphemes. Under this analysis, a zero morpheme is considered to be an alternant which, while not having any phonological content, still carries meaning. However, for Bloomfield, the zero morpheme can be a linguistic form whereas for Hockett (1947), the zero morpheme alternant can be determined by comparing it with other morphemes, but zero cannot be the only alternant for a morpheme unit.

In developing the theory of morphemes, Hockett (1947) introduced the term morph as a solution to Harris' (1942) account in which [naif] ~ [naivz] are treated as two different morphemes since the phoneme sequence [naif] is not contained in the second word. Hockett (1947) treated these cases as a portmanteau morph that has more than one morpheme, each of which contributes to the meaning of the morph.

Despite the subsequent developments of this theory, the morpheme notion remains problematic for Arabic morphological structure, as will be further discussed in the following section.

2.1.2 Issues with the root and pattern approach

Although several revisions to McCarthy's original account have been presented, for example, under the prosodic morphology hypothesis (McCarthy, 2006; McCarthy & Prince 1990a), the approach has been criticized on two grounds: first, the emphasis on the unique non-concatenation formalism of all word forms into root, vowel patterns and templates and second, the morphemic nature of the account. Unlike the earlier versions of this approach in which the Arab grammarians considered the root and pattern representation as a uniform entity, in the autosegmental model, a form is based on a composite representation consisting of three morphemes. Despite this difference, the root and pattern in both versions of the model are morpheme-based supported by different motivating views, two of which will be evaluated below.

The use of the root and pattern approach in a morpheme level morphology in Semitic languages is influenced by the degree of 'internal' modifications in the morphological system, shared by related forms many of which are enabled to have the same consonantal material. Thus, roots and patterns seem to have a strong presence in language acquisition and psycholinguistics studies (Boudelaa & Marslen-Wilson, 2001; Ravid, 2006; Shimron, 2003). Arguments for this morphemic status have distributional motivation based on the observation that roots constitute the formal and semantic core of the morphologically related words i.e. the unity of the consonantal roots and word patterns are shared by a great number of words with similar meanings. For example, in Arabic, a consonantal root can be shared by a morphological family of a great number of related forms. Thus, the word pattern maCCaC occurs in many derived forms, in most cases representing place nouns such as ‘*madxal* ‘inlet’, *maktab* ‘office’, *masbaḥ* ‘swimming pool’ (Boudelaa & Marslen-Wilson, 2001, p.33).

Similarly, Ravid (2006, p.130) presents examples from Hebrew in which the root x-š-x surfaces in many forms including: *hexešix* ‘darken (transitive and intransitive)’, *huxšax* ‘be darkened’, *xašux* ‘dark’, *xóšex* ‘a Biblical noun’, *xašexa* ‘darkness’, all linked through the same root and associated with the general meaning of darkness in this root. Moreover, Ravid (2006, p.129) argues that the consonantal skeleton is even persistent in words such as *šafan* ‘rabbit’, *subsidiya* ‘subsidy’ which are considered to be based on only a single morpheme and lack any internal structure. Yet, the CV-skeleton of such words can be mapped into complex patterns to derive new words. Many psycholinguistic findings have found evidence for the centrality of the root and template in the mental lexicon of the speaker. Nevertheless, Shimron (2003, p.24) concludes that ‘…none of this evidence precludes the possibility that other parts of Semitic words, segments that are perhaps larger than or different from roots and templates, may have their own independent status in our mental lexicon.’ In other words, the fact that ‘…Semitic language speakers are sensitive to roots and templates does not mean that such sensitivity has a role to play in every linguistic process, nor does it preclude sensitivities to other types of morphological elements’, suggesting that the different models are possible options for the speakers. Therefore, it is not clear how the autosegmental morphology as a possible synchronic account would help advance the understanding of concatenative morphology of the inflectional systems in comparison with arbitrary morphological stem allomorphy such as the system of this study (Chapter 4). Therefore, many of the supposed arguments for this morpheme account do not always hold.

The non-concatenative morpheme account of Semitic morphology is mainly based on the observation that there are co-occurrence restrictions that condition the radicals of the root and do not apply to the consonants of the affixes. Greenberg (1950) observed that only the root radicals are subject to these rules, whereas affixes are not in the scope of these restrictions. This seems to imply that the consonantal root might be functioning as an independent unit at a morpheme level representation. The restrictions disfavour the co-occurrence of identical root consonants (e.g. SNN, SSN, NSN) and/or root radicals (the first and the second or the second and the third) with the same place of articulation (PBK or TGK). Greenberg (1950) reports that although there can be no identical consonants nor homorganic ones in the first and second position of the root, many verbal roots have identical consonants in the second and the third positions. For example, *madad* ‘he stretched’ and *farar* ‘he fled’. Nevertheless, the same positions have restrictions against homorganic consonants whereas the first and third consonants have partial restrictions both homorganically and on identity. A replication of Greenberg’s study was provided by McCarthy (1994) who provided co-occurrence classes based on the place of articulation and manner features. Traditionally, Arabic co-occurrence restrictions are explained by organizing consonants into natural classes within which the co-occurrence consonants apply.

Greenberg (1950) and McCarthy (1994) present a number of main co-occurrence classes including: Labials = {/b, f, m/}, Coronal Obstruents = {/θ, ð, t, d, tˤ, dˤ, s, z, sˤ, zˤ, ʃ/}, Dorsals = {/k, g, q, χ, ʁ/}, Gutturals = {/χ, ʁ, ħ, ʕ, h, ʔ/} and Coronal Sonorants = {/l, r, n/}. They argue that consonants within the same class are subject to co-occurrence restrictions, and a consonant of a particular class can co-occur freely with a consonant from a different class. However, these co-occurrence restrictions do not include the uvular fricatives /χ/ and /ʁ/ that simultaneously group in different classes: the dorsals class and the gutturals class. In addition, there are fewer restrictions on the co-occurrence of the non-adjacent consonant root. They also note that roots containing one fricative and one stop are more common than those containing two fricatives or two stops. To account for the varying degrees of restrictions, McCarthy (1994) applies the Obligatory Contour Principle (OCP) that prohibits the co-occurrence of identical or homorganic root elements, but with two modes of application: one at the root node and the other at the place node. Thus, in the case of underlying adjacent pairs, the OCP is applied to the whole segment banning the occurrence of *mmd* or *mdd* in the underlying representation. The *mdd*-type forms can surface in double verbs in which root geminates are underlyingly represented as a single segment *md* which in turn is subject to the spreading process. The restrictions on homorganic adjacent segments are handled by applying the OCP to tiers of place features. Therefore, forms such as **ftb* are prohibited in Arabic since they violate the OCP place constraints as shown in the following illustration (Frisch et. al., 2004, p.11):

(5)

labial tier: [lab] [lab]

skeletal tier: *C C C

coronal tier [cor]

The place features used by McCarthy are [labial], [coronal], [dorsal] and [pharyngeal]. Nevertheless, although the OCP account seems to provide an ideal account, many exceptional patterns have been documented. Greenberg (1950) noted exceptions of the form /χCʕ/ that violate the restrictions on pairs of pharyngeals and uvulars. The

Buckwalter (1997) corpus lists the following exceptional patterns to homorganic co-occurrence restriction in Arabic.

C1C2	C2C3	C1C3
Labial fmC 1 *Totals* *1*	 *0*	bCf 1 bCm 9 fCm 11 *21*
Coronal stop dˤdC 1 *Totals* *1*	Ctd 4 Ctˤd 1 Cdˤd 3 *8*	dCt 2 dˤCd 2 tˤCt 3 dCtˤ 1 tCd 1 dˤCtˤ 3 tˤCd 2 dCdˤ 1 *15*
Coronal fricative sðC 2 ʃðC 4 ʃsC 1 ʃzC 1 ʃsˤC 2 ʃzˤC 2 *Totals* *12*	Cʃz 1 *1*	ʃCθ 2 ʃCð 1 ʃCs 3 ʃCsˤ 1 ʃCzˤ 1 *8*
Dorsal gkC 1 χgC 1 ʁgC 1 χqC 1 ʁqC 1 kχC 1 gχC 2 kʁC 3 *Totals* *11*	Cqg 1 Cgq 2 Cχq 1 *4*	gCk 1 ʁCq 6 χCk 1 kCχ 3 kCg 1 gCχ 2 qCg 4 χCg 5 ʁCg 1 gCq 2 χCq 5 *31*
Pharyngeal ʔχC 4 ʔħC 3 ʕhC 4 ʔhC 2 hʔC 1 *Totals* *14*	Cχʕ 2 *2*	ʔCχ 1 ʔCh 3 ʔCħ 1 χCʔ 5 χCʕ 9 ʁCʔ 1 hCʕ 8 ħCʔ 3 ʔCʕ 1 ʕCʔ 1 ʕCh 5 hCʔ 6 *44*
Coronal sonorant nrC 1 rnC 7 *Totals* *8*	Crl 2 Cnl 1 Clr 1 Cnr 7 Cln 5 Crn 9 *25*	rCl 14 nCl 22 nCr 23 lCn 11 rCn 15 *85*

Table 2.2 The exceptional patterns to homorganic co-occurrence restriction in Arabic based on six classes of same-place articulation in different consonantal position C1C2, C2C3 and C1C3 listed by the Buckwalter (1997) corpus (adapted from Alderete et al., 2013, p.56).

In Table 2.2 the rows list all the exceptions to homorganic co-occurrence restriction which is arranged in six groups with same place of articulation (labial, coronal stop, coronal fricative, dorsal, pharyngeal, coronal sonorant). The three columns present the consonantal position for the group of exceptions (C1C2, C2C3 and C1C3). The data illustrates the number of instances of exceptions and the total of those in the same-place class. For example, there are 7 roots with the pattern /rnC/ and the total number of exceptions to OCP in the coronal sonorant class in C1C2 pairings is 8. The number of the attested patterns in all classes in C1C2 positions comes to 23 from 47 roots. Overall, the total number of exceptions in C1C3 pairings outnumbers the rest of the co-occurrences in the other places (a total of 45 patterns out of 204 roots in C1C3 position and 14 patterns out of 40 roots in C2C3 position). Yet, the number of the exceptional patterns and the roots within each class are randomly distributed in the three pairings. Some classes show one pattern within a certain pairing but a variety of patterns with the other pairings. For instance, in C2C3 pairs the only co-occurrence prohibition on two pharyngeals is */Cχʕ/, but this pair is just one of several patterns in the other two pairings. Therefore, the data reveal that the OCP account does not seem to show the difference in the degree of restriction within the classes. In fact, Frisch et al. (2004) who gave a quantitative similarity based account; argue that dividing the co-occurrence into categorical group does not provide a plausible account for the exceptional consonant pairs, because OCP-Place in Arabic is a gradient constraint based on the similarity between these consonants.

Another issue with McCarthy's approach is raised by the way that identical consonants in the C2 and C3 positions in a triliteral root (e.g. *madd* 'pass sth.') are represented in the CV template. The underlying root of this example is taken to contain two radicals C1C2, /md/. The verb surface forms of this root includes *madd* ~ *madad* through morphophonological conditioning. Different versions of root and pattern consider *madd* to be derived from *madad* which in turn is produced via the underlying representation /md/ assigned to a triliteral shaped CVCVC template. Then, through intermediate processes including deletion and metathesis that are morphologically stipulated, the output *madad* is repaired to arrive at the second output *madd* (McCarthy, 1986, pp.247-248). In addition, the OCP does not justify the banning of identical root elements in C1 and C3 positions (*KLK) (Shimron, 2003).

The OCP approach also fails to account for the double roots in Arabic dialects including TLA. Watson (2002) discusses the template pattern of forms such *faakik* active participle ‘having undone’ in Cairene Arabic and in the active participle form *jaafif* ‘drying’ (cf. *jaaff* ‘dry’) of San’ani Arabic, suggesting that these examples show that double verbs have a distinct template with germinate elements in adjacent positions (in TLA *šaakik* ‘to doubt’).

Clearly, the issue with this OCP analysis is that the violations of the consonant root co-occurrence restrictions are handled by means of a whole series of intermediate arbitrary stipulations. A more economical account would therefore be preferable. Several revisions to McCarthy’s original account have been presented under the prosodic morphology hypothesis (McCarthy & Prince, 1990a, 1993). However, the approach has been criticized on several grounds including the emphasis on unique decomposition of all word forms into root, vowel patterns and templates, and the morphemic nature of the account.

Model specific problems of the root and pattern approach have been identified by Hudson (1986, pp.98-103) who provides specific problematic issues of this model based on arguments concerning ‘inexplicitness’, ‘arbitrariness’ and ‘complexity’. He argues that, in some cases, it is hard to understand McCarthy’s claims because of the inexplicitness of autosegmental grammar. For instance, the autosegmental model does not provide the 'appropriate formalization' for the left-right association. 'How do we assure that vowel features associate with [+syllabic] and consonant features with [-syllabic], McCarthy does not say, and we can only assume that some otherwise unnecessary statement of the grammar is required for this purpose.' Likewise, he considers the Erasure rule to be inexplicit since no specific context is given for the rule application. Therefore, it is possible that the application of this rule can involve any verbal forms that have the structural requirements and hence the Erasure rule would not be strictly limited for relating the perfective stem to the imperfective stem.

Another issue, which is particularly relevant to this study, is how to account for what is considered to be exceptional patterns of apophonic variation as in measure one (*M*1) verbs (§1.2 and §3.1 Hudson (1986) argues that the approach to the ablaut in *M*1 does not explain how to accomplish the exclusion of class 4 from the polarity generalizations nor does it provide a plausible autosegmental representation for the class 3 vowel melody. As Hudson (1986, p.99) explains 'in the derivation of the perfect the two vowels (lexical *i*, grammatical *a*) would compete for two positions in the template CVCVC'.[6] Although this special rule has the potential of producing the desired pattern, to guarantee the correct output, the vowel *a* had been treated as an affix which would have been projected to the correct V-position by affix-first convention, and this has not been proposed by McCarthy. The scale of this problem, however, is even more acute when considering the pattern of apophonic variation in TLA. The polarity generalization does not hold in TLA. The reason for this is that the perfective stems are represented by two sets of stems based on the quality of the second V-

[6] The vowel *a* represents perfective aspect and active voice while the high front vowel *i* is found in the ablaut class 3 and it is lexically determined (See 2.1 and Table 2.1).

position. One set of stems has [-high] while the other has [+high] (*rekab* 'he rode' *rukub-it* 'she rode' ~ *resam* 'he drew' *risim-it* 'she drew', *šeraḥ* 'he explained' ~ *šuruḥ* 'she explained'). By contrast, the corresponding vowel in imperfective forms can be [+/- high] cf. *ya-rkub* 'he rides', *yi-rsim* 'he draws', *ya-šraḥ* 'he explains' (See §3.1 and §4.3.1 for further discussion).

Another issue for autosegmental morphology is the emphasis on characterizing the morphological system as a set of independent patterns (templates) with inherent semantic features and the fact that it therefore fails to identify systematic allomorphy. Ratcliff (1998, pp.23-29) argues that autosegmental morphology is challenged by the lexeme formation of the broken plural[7] and diminutives forms. The model describes this allomorphic nominal system as sets of distinct forms with independent pattern replacement without capturing the derivational relationships of these same function forms. For instance, root and pattern morphology treats the plural stems, which are in italics in Table 2.3, as isolated template patterns CVCVVC, CVCVVCVC and CVCVVCVVC although each of them shows systematic allomorphic alternations marking the same set of functions. In each example, the stem choice is determined by the singular forms. The formation of the plural forms and diminutives is distinguished by the long vowel /aa/ in the second syllable and the diphthong /ay/ respectively. In other words, the plural/diminutive templatic patterns are composed of two parts: one part is shared with the singular forms (suffixes) while the other part shows stem internal change, marking the grammatical category.[8]

---

[7] The term broken plural refers to the plural forms that are formed by changing stem internal elements. By contrast, sound plurals are the plural forms whose formation involves a regular suffixation process (e.g. muʔlim-un 'teacher.M.SG.NOM' ~ muʔlim-uuna, 'teacher.M.PL.NOM'. In these examples, the change is external to the stem). Broken plurals have several classes each of which show distinct vowel patterns (Ratcliff, 1998).

[8] Further developments of the autosegmental morphology model suggested that such forms are formed by a prosodic input conditioned by principles of prosodic circumscription according to which the transformation from singular to plural involves morphophonologically altering the first two moras of a given form/word (McCarthy & Prince, 1990b). 'Under prosodic circumscription, […], a morphological

template			singular	plural	diminutive
CVCC	>>	CVCVVC	*kalb*un[9]	*kilaab*un	*kulayb*un
CVCCVC	>>	CVCVVCVC	*daftar*un	*dafaatir*u	*dufaytir*un
CVCCVVC	>>	CVCVVCVVC	*miftaaḥ*un	*mafaatiiḥ*u	*mufaytiiḥ*un

Table 2.3 broken plural and diminutives forms of *kalb* 'dog', *daftar* 'notebook' and *miftaaḥ* 'key'

The basic issue with this morphological analysis of Arabic is that it fails to provide an adequate grammatical description and generalizations about allomorphic alternations. In other words, a plausible account should capture inflectionally and/or derivationally based morphological oppositions without over specifying detailed morphological information. Allomorphy of this sort can be economically captured by a stem-based approach (See Chapter 4 and Chapter 5 for discussion on the TLA verb stem allomorphy).

Another issue with the autosegmental morphological analysis is that the account is based on the morpheme theory. The basic relevant morphemic view adopted by McCarthy (1981) is that the lexicon is a list of three sets of morphemes each of which must mean something. The vocalic apophony in CA verb inflection has been represented as independent morphemes that must mean something such as tense, aspect, voice. In the root and pattern approach, the characteristic features of inflectional morphology have been based on regarding the vowel patterns as the base for the morphosyntactic distinctions between verb forms. Recall that the fundamental assumption is that the inflected verb form is built on a two, three or four consonantal root which is associated with the basic meaning of the verb along with the affixes which

operation is applied to a prosodically-determined substring within the morphological category, often yielding some sort of infix.' (McCarthy, 1993, p.189).

[9] The *un-/-u* endings represent the nominative case.

are attached to the root and result in different verb forms and different inflections as in the following examples of 3M.SG forms in *M*1, *M*2 and *M*3 respectively:

Verb	Active			Passive	
	Perfective	Imperfective	Imperative	Perfective	Imperfective
katab 'to write'	katab*a*	*ya*ktubu	*u*ktub	kutib*a*	*yu*ktab*u*
kattab 'to cause to write'	kattab*a*	*yu*kattibu	kattib	kuttib*a*	*yu*kattab*u*
takaatab 'to correspond with sb.'	takaatab*a*	*ya*takaatabu	takaatab	tukutib*a*	*yu*takaatab*u*

Table 2.4 The inflectional paradigm in CA.

The verbal affixes have been described as morphemes that realize grammatical categories and which distinguish between perfective and imperfective forms. It is generally assumed that aspect-tense and subject-gender agreement are marked by suffixes in the perfective whereas in the imperfective they are marked by prefixes (Table 2.4). However, a different analysis is proposed by Bahloul (2008) who argues that aspect and tense are marked by the first vowel morpheme in the perfective verb stem while the suffix /-a/ realizes agreement morphology. By contrast, in the imperfective, aspect/tense are denoted by the prefix vowel along with the deletion of the first stem vowel.[10] Likewise, Er-Rayyan (1986, p.75) assumes that the first vowel *a* (CaCVCV) realizes aspect-tense morphology, in other words, the first vowel is considered to be a morpheme denoting the property of tense-aspect. The perfective verb vowel melody consists of three vowels each of which function independently as a morpheme, but *M*1 active and passive perfective can have different sets of vowel patterns (Table 2.5).

[10] Under a stem-based account, this view has been challenged by Benmamoun (1992, p.27) who considers the final /-a/ in *kattab-a* as a marker for subject/gender agreement and that tense/aspect morphology is not realized by phonologically overt morpheme, but rather denoted by 'an abstract morpheme' in other words, by a zero morpheme. In addition, he treats the prefix of the imperfective as a purely default form with no aspectual realization.

Verb	Gloss	Template	Vowel pattern
		Active	
kataba	‘he wrote’	CaCaCa	(a, a, a)
ġaḍiba	‘he became angry’	CaCiCa	(a, i, a)
kabura	‘he grew up	CaCuCa	(a ,u ,a)
		Passive	
kutiba	‘it was written’	CuCiCa	(u, i, a)
šuriba	‘it was drunk’	CuCiCa	(u, i, a)

Table 2.5 Vowel patterns in *M*1 active and passive perfective 3SGM form.

Bahloul (2008) adds that the first vowel marks active voice which alternates with the high back vowel *u* to achieve passivization. These examples in Table 2.5 show that the passive construction is realized by the vowel patterning *u-i* in which the vowel *u* is the only replacement for the first vowel *a* independently of the verb class (Table 2.1). Therefore, this vocalic affix alternation has been considered as an active-passive voice realization.[11] In addition, it has been claimed that the second vowel of the perfective verb forms in SA (CVC*V*CV) can be *a, u* or *i* vowels conditioned by valency of the verb. The vowel *a* usually accompanies transitive verbs while the *u* vowel is mainly used in intransitive verbs which are also stative. By contrast, the vowel *i* can represent transitive or intransitive verbs denoting ‘mental or psychological states’. The verbs with the different second vowel patterning are given in the following examples (Bahloul, 2008, p.33)[12]:

[11] Although this might be a plausible analysis for SA, the first vowel in TLA active verb forms can be *i* or *u*, any of which can be used in the passive construction.

[12] Although this might be a reasonable account for SA, this distinction does not hold in TLA in which transitive and intransitive verbs mostly have the default vowel *a*, thus, the second vowel patterning cannot be determined by transitivity:

(+ transitive) /a/	(+/- transitive) /a/	(-transitive) /a/
ḍerab-it ‘I hit’	feraḥ-it ‘I became happy’	kebar-it ‘I became old’
mesak-it ‘I held’	feham-it ‘I understood’	seman-it ‘I became fat’
šekar-it ‘I thanked’	meraḍ-it ‘I became sick’	geṣar-it ‘I became short’

(+ transitive) /a/	(+/- transitive) /i/	(-transitive) /u/
ḍaraba 'to hit'	fariḥa 'to be happy'	kasuda 'to be stagnant'
masaka 'to hold'	ġaḍiba 'to be angry'	kaṯura 'to become numerous'
šakara 'to thank'	mariḍa 'to be sick'	qaṣura 'to be short'

Table 2.6 Vocalism in *M*1 perfective verbs.

The suffix vowel realizes agreement in person, number and gender. In terms of the imperfective vowel melody, Bahloul (2008) assumes that the different vowel patterning that is internal and external to the root denotes different grammatical categories. Unlike, the perfective form, the imperfective form has no first vowel within the first and second consonant of the root, but the vowel in the prefix attached to the base has been considered to be functioning as a morpheme denoting aspect-tense category. The reason he provides for that is the prefix vowel alternation found under the passive construction as shown in the following examples:

verb	**active**		**passive**	
	perfective	imperfective	perfective	imperfective
katab 'to write'	k*a*taba	y*a*ktubu	k*u*tiba	y*u*ktabu
šarib 'to drink'	š*a*riba	y*a*šrab	š*u*rib	y*u*šrabu

Table 2.7 The vocalism in perfective active forms and the passive counteraparts in SA.

Similar to the first vowel -*a* in the perfective, the first vowel *a*- (within the prefix) in the imperfective changes to -*u*-, as illustrated in (Table 2.7). By contrast, the last vowel of the imperfective and the initial consonant of the prefix mark agreement morphology.

Clearly, in the traditional descriptions, the vowel alternations in perfective verb forms are treated as 'cumulative' grammatical categories and the same approach is applied to the imperfective. However, a fundamental problem with this account on the vocalic tier level is the claim that each vocalic element is associated with a grammatical meaning. In addition, the morpheme notion has the principal commitment of one-to-one mapping in either direction between the meaning element and its corresponding phonological form. However, this property does not hold in words with the cumulative morphs (Anderson, 1992). Consider the following morphologically assigned vocalic melody of *M*2, *M*3 and *M*4 identified by Ratcliffe (1998, p.38).

Measure	Imperf			Perf			Mood
2,3	u	a	i	a	Ø	a	Active
4	u	Ø	i	a	Ø	a	
5,6	a	a	>	a	>	a	
2,3	u	a	a	u	Ø	i	Passive
4	u	Ø	a	u	Ø	i	
5, 6	u	a	>	u	>	i	

Table 2.8 The morphological vowel melody in *M*2, *M*3, *M*4, *M*5 and *M*6 (Ratcliffe, 1998, p.38).

Ratcliffe (1998) considers the first vowel quality as a marker of voice with the low vowel *a* denoting active and the high vowel *u* is the passive marker. However, Ratcliffe does not clarify why the morphologically assigned passive vowel *u* is also found in a contradictory feature value that is the imperfective active. Likewise, traditional non-concatenative accounts characterise the vocalic melody element *a* as a single morpheme meaning 'perfective active' although the same vowel is also lexically (in Stump's (2001b) terms of lexical theory) listed as an indicator of imperfective in the passive mood in addition to the suffix vowel *a* which is used to mark indicative independently assigned by an inflectional rule. In fact, in TLA (Chapter 4), it is not possible to establish any morpheme-based principle for the arbitrary ablaut change in the verb inflectional paradigm.

In sum, the morpheme device as a pairing of a phonetic form mapped with a corresponding meaning is fundamentally flawed, leading to a problematic autosegmental analysis of the Arabic verb. As will be shown in Chapter 4 and Chapter

5, the verb inflectional system of TLA will be more challenging for this non-concatenative approach due to the arbitrary stem alternations. We conclude this section by a statement from Shimron (2003, p.11) and that is ‘a synchronic approach to Semitic languages requires systematic structural support, and this can be achieved only through a close examination of the variety of word formation processes in each and every language.’ Therefore, in this study of TLA verb inflectional morphology we follow Hoberman and Aronoff’s (2003, p.63) proposal for Maltese, and that is a Semitic language such as TLA ‘…may contain relics of root-and-pattern morphology, but its productive verbal morphology is decidedly affixal’ (§4.5 and §5.3.2). Thus, ‘…Semitic is no different from any other language family to the extent that members of a family share some peculiar trait, it is purely an accident of history ….’ (2.4.1 and §5.2).

2.2 Concatenative Semitic morphology

Although it is extensively used in Arabic inflection and derivation, the root-and-pattern, templatic form of morphological organization is not restricted to Semitic languages. The ablaut, for example, can characterize the morphological structure of Romance languages and the template system can also be found in Yawelmani. Furthermore, for many Arabic dialects including TLA, the arbitrary vocalic apophony that is paradigmatically distributed is the most distinctive and pervasive feature of inflectional changes rather than root radicals, which clearly poses a problem for a root-based approach. Therefore, a uniform paradigm based account in which a Semitic language is not considered to be a fundamentally different system is desirable.

More recently, there have been many studies of Arabic in favour of a stem-based account to verb paradigms. The main motivation for a concatenative approach is not a matter of account uniformity and simplicity, but it is based on and supported by the morphological pattern, as in the case of this study (Chapter 4). Notable stem/word-based works to Semitic languages include El-Bat (1994), Hoberman and Aronoff (2003), Ratcliffe (1997, 1998, 2003), Heath (1987, 1997, 2003, p.126) and Benmamoun (1999, 2003). The essential principle of this approach is that the derivation and/or inflection process in Arabic is based on full phonological units represented as a stem rather than dividing vowels and consonants into abstract morphemes, as it is the case with the root and pattern model. Many proponents of Semitic stem-approach emphasize the significance of the stem in the imperfective sub-paradigm as the basic morphological unit for inflection and derivation. For instance, Heath (2003) investigating the morphological process of derivation, distinguishes underived stems including the imperfective forms of *M*1 and/or singular noun from complex stems which are derived from and frequently dependent on the simple underived stems through ablaut change and affixation.

Likewise, Ratcliff (1997) and Benmamoun (1999, 2003) consider the imperfective stem to be the base for morphological processes. The importance of stems has also been confirmed by McCarthy and Prince (1990b, p.219) who further develop the original root and pattern approach by applying prosodic principles. They propose that lexical forms such as the broken plural require stems as the base of lexical formation, hence ‘the category root is …morphologically inappropriate as the basis of broken plural formation, since some derivational affixes are transferred intact.’ They give examples such as the single form *miftaaḥ* ‘key’ which alternates with the plural form *mafaatiiḥ* ‘keys’. Both of these have the affix *m*- reserved while ‘in the true-root-and-template derivational morphology…only the root consonantism carries over…’.

Similarly, Heath (1987, 1997, 2003, p.119) argues that there are no plausible principles for dividing consonants and vowels and assigning them to different representation levels forming 'triple-decker sandwiches'. The reason for this is that the ablaut templates do not realize any grammatical information in most cases. For instance, the word patterns CaCC and CiCC of the stems *kalb* 'dog' and *silm* 'peace', are purely lexical, they do not realize any morphosyntactic features. The default stem for nouns is the singular stem whereas for verbs the default is the imperfective stem. In fact, the underived stems (e.g. CCVC -*ktub*- stem 'write' in the imperfective verb forms), in many cases function as the input for the derived stems formed by the derivational operation of ablaut (e.g. CVCVC -*katab*- 'write' in the perfective verb forms). In other words, the stem alternation within the same inflectional verbal/nominal paradigm is handled by *derivational* processes which can equally apply to lexeme formation.[13] These derivational processes are conditioned by directionality. Thus, the output is a derived stem from the input which can be underived or derived, (for example, the underived stem of intransitive verb -*dxul*- 'enter' is the input for the derived stem -*daxxil*- causative verb *M*2 'put in' which also functions as the input for the derived stem *ta-dxiil* to form the verbal noun (p.116). However, Heath does not completely dispense with roots and patterns for *M*1 sound verbs. For instance, the verb 'enter' has the perfective and imperfective stems -*daxal*- and -*dxul*- respectively either of which can function as the underived input stem. The reason for that is that they are derived through mapping the ablaut with the consonants *dxl*. Yet, the same analysis cannot hold for hollow and double verbs that are inevitably stem-based (p.121).

[13] It seems that these stem-based accounts assume that the verbal paradigm can consist of forms related by inflection or by derivation with a mutual relationship between inflectional stem allomorphy and word derivation. These approaches seem to be similar to a word-based morphology (Blevins, 2006).

Benmamoun (1999, 2003) argues for a word/stem-based account challenging the empirical validity of roots in word derivation. Under his account, Arabic word formation is based on the imperfective verb as the default underived input stem with no temporal or aspectual features. He presents a unified account for verbal and nominal derivations all of which are based on the imperfective stem. Thus, under this account, a word such as *muʕallim* ‘teacher’ is formed by affixing *mu~* to the imperfective stem *yu-ʕallim* which can also act as the input for the imperative forms *ʕallim* (2003, p.110). The syllabic structure and the vocalic elements of the imperfective stem depend on the verb measure. For instance, the imperfective stem for *M*1 verbs is CCVC (e.g. *ya-drus* ‘he studies’), but CVCCVC in the other verb measures (e.g. *M*2 *yu-darris* ‘he teaches’). In the case of the non-*M*1 verbs, it is possible to establish implicative vocalic melody relations between the imperfective and the corresponding imperative (*M*2 *darris* ‘teach!’) and/or the nominal forms (*mu-darris* ‘a teacher’) since all these forms share identical stems (*darris*). This shows that the derivational relations are based on words rather than only roots. By contrast, the implicative relations based on imperfective stem and both nouns and verbs can be challenged by *M*1 verbs due to the ablaut. For instance, the 2SGM.imperf. form *ta-drus* shares the same stem with the 2SGM.imper *ʔu-drus*, but the 2SGM.perf *daras-ta* ‘study’ has a different vocalic melody. In other words, the perfective form of *M*1 is not fully implicated in the verbal derivation due to the difference in the vocalic pattern.

Furthermore, he even concludes that the arbitrary vowel change, which in our example is represented by alternation of the vowel *u* in the imperfective with the vowels *a-a* in the perfective, ‘does not argue for or against either the root to template account or word to word account.’ The reason for that as he claims is that ‘…both introduce more complexity in the system.' (2003, p.112).[14] However, clearly the complexity issue is mainly caused by overgeneralizing and not distinguishing the stem allomorphy in the inflectional paradigms from the bases of the verbal and/or nominal word formation.

The derivational dependence relations based on stems is also proposed for Arabic plural and diminutive formation by Ratcliffe (1998, p.29). He begins by discussing the redundancy of specifying separate morphemes in the form of different templates corresponding broken plural and diminutive forms such as CVCC *kalb* 'dog.SG' ~ CVCVVC *kilaab* 'dog.PL' ~ *kulaub* 'dog.DIM.SG'. He concludes that the formation of both the plural and diminutive forms is based on the nominal singular stem as a full phonological form, implying that 'morpheme-(-root) based model of Arabic lexicon argued for in McCarthy (1979) is inadequate.' (p.149). Likewise, Watson (2006, p.159) discussing the diminutive verbs in San'ani, identifies a number of cases in which the morphological unit of formation is a full phonological form such as the singular noun or an adjective. In such cases the primitives of formation have related verb meaning in addition to the unique base that cannot easily be reduced to root radicals. For instance, the diminutive verbs *tmaydar* 'to become a manager' and *twayzar* 'to become a minister' which semantically can be linked with the nominal words *mudiir* 'manager' and *waziir* 'minister' respectively. Yet, the nominal forms do not have /MDR/ or /WZR/ roots, but they are rather related to the roots /DWR/and /ʔ-z-r/ via other intermediate forms. Hence, Watson (2006) concludes that the derivation of the diminutive verbs is based on the nominal stem rather than the roots. However,

[14] Generally, the stem/word based accounts suggested for Arabic tend to be more similar to the 'abstractive model' (in Blevins' terminology (2006)) according to which the recurrent elements, which in these studies are mainly found in derivational processes, are abstracted from inputs such as derived stems (i.e. full forms). Nevertheless, this abstractive model also overlaps with the 'constructive model' (concatenative-approach) in which the primitives of all related forms is reduced to roots/stems. The overlapping between these two models is also apparent in the approaches of Ratcliffe (1997, 2003), Heath (1987, 1997) and Benmamoun (1999) in which the input is the underived stem (imperfective and/or singular noun) rather than the fully inflected form.

the account Watson introduces for diminutives in this dialect does not completely exclude the existence of roots.

From an OT point of view, the stem approach to Semitic morphology is also favoured over the root approach. Ussishkin (1999, 2000) following Bat-El (1994) presents an OT account according to which the word is the base for denominal verb formation in Hebrew, hence, the root model can be dismissed. Likewise, mainly focusing on the stem alternations within inflectional paradigms, Gafos (2003, p.49) investigates the underlying morphological unit for the two stems found in the inflectional perfective sub-paradigm of *M*1 double verbs (e.g. *madd* 'he passed sth.' cf. *madad-na* 'we passed sth.' a morphophonologically conditioned alternation). Unlike traditional approaches, he considers the stem /madd/ to be the input for the surface variant [madad] derived using OT constraints such as IO-Faithfulness, OO-Faithfulness and markedness constraints.

Clearly, the main assumptions of the stem approach to Arabic is that the basic morphological unit of lexeme formation is a fully vocalised phonological element, represented in the form of a stem or a word. One of the important implications of this approach will be that we can deploy the stem-based analysis proposed for, e.g., the Romance languages in the analysis of Semitic morphology. Although a number of studies present an enriched view of the role of the Semitic stems in lexeme formation, they focus less on the nature of the stems in inflectional morphology and the motivating factors for stem change. In addition, it is essential for any stem account to reassess the relationship between verb measures from an inflectional point of view. In this study the assumption that root and pattern modelling is unnecessary is preferred and supported by the morphological nature of stems and the distribution of the stem allomorphy in inflectional paradigms of TLA verb measures. This can also justify one of the main concerns of this study which is analysing the TLA inflectional morphology system using the stem-space account and principal parts theory, which does not only confirm the plausibility of stem accounts but also shows implicative relation patterns similar to those in languages such as French, Italian and Spanish. The following section will introduce the notion of stem and stem properties assumed generally in the literature and also used in this study for a better understanding of the verb inflection system in TLA.

2.3 The Stem approach to TLA

The principles of this stem-based study is based on Aronoff's (1994) work on inflectional paradigm stems in general and on Semitic verb measures in particular. Aronoff (1994) not only agrees with the stem analysis for Semitic languages but also takes the argument to a higher level by suggesting that the measures are stem based and they are also inflection classes determined and handled by stem allomorphy. Likewise, in this work, we consider the verb system of TLA to be based on verb measures each of which show stem alternations internal to their inflectional paradigms (§4.2 §4.5 and §5.2). These measures have their unique stem shapes and syntactic structure related by derivation processes with different grammatical function (e.g. causative, passivation among others). In fact, TLA verbs can be viewed as inflectional classes distinguished by the pattern of their inflectional stem allomorphy (§4.3, §5.2. and §5.3). The concept of stem assumed in this work is based on a fully vocalised morphological base to which inflectional affixes can be added to form inflected words. In other words, it is a device used in the inflectional verb paradigm with varying distributions and allomorphic changes depending on the verb measure/series. The following section will introduce some of the types of stems and the motivating factors conditioning the stem changes that have been widely described in the literature mainly in relation to the Romance stem system.

2.3.1 The stem

Recent years have witnessed a renewal of interest in the notion of stem allomorphy where the base form of a lexeme does not seem to have a uniform phonology in inflectional or derivational morphology. Traditionally, in the generative work, this important notion had been considered under lexical and phonological accounts, handled by '(re)adjustment rules' (Chomsky & Halle, 1968) or autosegmental association principles (McCarthy, 1979). A different approach has been proposed by Aronoff (1994) and Maiden (1992) who consider stem allomorphy as a morphological component of a lexeme; hence, the distinct stems are 'morphomic', that is, autonomously morphological phenomena. Following from Aronoff's proposal, Spencer (2012, p.98) develops the Strictly Morphomic Stem Hypothesis (SMSH) according to which: 'All stems are morphomic' in that they do not serve a role in realizing (expressing and/or altering) a phonological, morphosyntactic or a semantic property. In fact, Spencer (2013a, p.4) has argued that '…being morphomic is a necessary condition for being a stem.' This idea of *autonomous* morphology is based on the notion of 'morphome': a function that cannot be ascribed to any synchronic or diachronic 'extramorphological' factors (phonological or morphsyntactic). Under this view, differences of form (e.g. stem alternations) systematically correlate with arbitrary paradigmatic contexts that cannot be reducible to a coherent set of feature(s) (Aronoff, 1994) (§2.3.2). Yet, recent works have suggested that a morphomic stem is not always a clear-cut case of *morphology by itself,* since 'morphomicity' can show sensitivity to certain extramorphological factors due to some degree of coherence or motivation. According to Maiden (2013, pp.2-41) 'Autonomy is not an inherently absolute notion, and it is perfectly conceivable that some morphological phenomenon can be in some measure autonomous, yet, also sensitive to extramorphological conditioning which, nevertheless, falls short of fully determining its characteristics'. In the other words, 'morphomic phenomena were…. potentially present, to a greater or lesser degree, even in the presence of phonological or other kinds of 'extramorphological' conditioning'. The result is *semi-autonomous* or *partially autonomous* morphology which is at the boundary between purely morphological, phonological and/or morphosyntactic sensitivity (§2.3.4).

In this work, I propose that stems as objects of semi-autonomous morphology can be represented on a 'morphomicity' scale that represents the morphomic stem's degree of sensitivity to morphophonological, morphosyntactic and/or phonological factors (4.4 §4.5 §4.5.14.5.2, §4.5.3 and §4.5.4). It is important to emphasise that this approach does not necessarily assume that stems bear or realize these features. In this study, the stems have a representative structure associating with 'stem referencing features' with a single value 'privative', ranging from the *canonical* extreme of **morphomic** with no sensitivity to any extramorphological factors to the other canonical extreme of **non-morphomic** stem with systematic realization of morphosyntactic (e.g. inflectional stem), morphophonological or phonological features (§4.4). It is possible for a stem to embody either of these extremes. However, in many inflectional systems, stems tend to fall in between. Embodying the morphomic extreme would reflect the SMSH (Spencer, 2012). Maiden (2005) would view such a scale as a diachronic process developing 'from pure phonology to pure morphology' while for Anderson (2008), stems have phonologically conditioned environment, predicting their forms (§2.3.4). Embodying the non-morphomic extreme can be represented by stems as multiple exponence. Baerman and Corbett (2012), on the other hand, regard a stem alternation as both allomorphy and serving as part of exponence for a morphosyntactic property set (§2.3.3).

2.3.2 Morphomic stem

One of the influential recent works in morphology is that of Aronoff (1994) who explores the issue of autonomous morphology. He identifies representations and principles that belong not to syntax or to phonology, but to the field of morphology alone. Aronoff argues that stems are purely morphological entities, sound forms devoid of any semantic features. Focusing on Latin, he explores the inflectional paradigm in which some forms are built on a particular shared stem which can be characterized as morphomic. For example, the perfect participle, future participle and the supine share the same stem (the third stem) which is a sound form with no syntactic or semantic value, but has only the lexical meaning inherited from its lexemic entry. In fact, this morphomic stem is regarded as a function mapping between morphosyntax and morphophonology. In the English verb system, the morphosyntactic features for passive or perfect map onto a shared stem form (the perfect participle) by the function f_{en}. Semantically, this stem does not have any passive or perfect semantic features that are mapped into morphophonological realizations of different verbs.

Aronoff further considers the example of the morphomic stem (third stem) in Latin, which is illustrated by the forms *amāt-(um)* and *monit-(um)* (Table 2.9):[15]

lexeme	infinitive	gloss	passive perfect participle	supine	active future participle
AMŌ	amā-re	‘love’	amāt-us	amāt-um	amāt-ūr-us
MONEŌ	monē-re	‘warn’	monit-us	monit-um	monit-ūr-us

Table 2.9 The Latin *amō and moneō* inflectional forms of the perfect participle, the supine and the active future which are built on the third stem (Aronoff, 1994, pp.32-33).

[15] Likewise, in TLA as will be discussed in Chapter 3 and Chapter 4, the hollow verb stem *baan* ‘appear’ can be found in the perfective, imperfective, future and imperative.

Table 2.9 shows that both perfect participle and future participles can be used as adjectives and can be marked for gender number and case.[16] The third stem for the verb *amo* is the basis of the form *amāt-um* with the stem form *amāt.* The stem is also selected in the passive perfect participle *amāt-us* and in the active future participle *amāt-ūr-us* Similarly, the supine has the shared stem '*amat'*. The infinitive form has the stem *'ama'* preceding the suffix /re/ that is the Latin infinitive marker; hence, the traditional supine stem of *'amo'* is *'amat'* which represents the third stem. In fact, the use of the third stem as the base stem is found in the vast majority of verbs in the language including irregular verbs with suppletive stems (e.g. *premere* 'to press', *press-us* 'having been pressed' *press-ur-us* 'about to press', *ferre* 'to bear', *lat-us* 'having been borne', *lat-ur-us* 'about to bear'). Clearly, the third stem occurs in different morphosyntactic environments (e.g. active and passive), and there is no common shared property that can be attributed to this morphomic stem. Therefore, Aronoff's proposal suggests that each lexeme might associate with a set of lexically listed stems with their unique indices and the relationship between the set of stems can be handled by rules of word formation. The indexed stems are morphological entities devoid from any semantic or syntactic function. However, it is important to emphasize that the morphomic stem as a function of autonomous morphology does not necessarily exclude the possibility that these phenomena could partly correlated with paradigmatic or syntagmatic distribution (see §4.5.2 and §4.5.3 for the discussion on TLA stems).

Morphomic stems can also be found in many other languages of the world. For instance, Stump (2001b) provides detailed arguments to support the claim that Sanskrit stems are morphomic. Sanskrit has two stem kinds one of which can be divided into

[16] The perfect participle marker can be the suffix /t/ attached to the stem (e.g. '*amat'*) while in the future participle, the suffix /ur/ is added to the same stem '*amat'* before the gender-number-case masculine inflections (e.g. '[*amat-ur*]-*us'*).

two sub-types and all stems are defined morphologically by their distribution. Stump distinguishes between Strong and Weak stems, and the later can be divided into two stem variants traditionally referred to as Middle and Weakest. For instance, the possessive adjective BHAGAVANT ‘fortunate, has a two-way distinctive pattern of stem alternation that includes the Strong (bhagavant-) and Middle (Weak bhagavat) stems. Both stems can appear in the masculine and neuter paradigms, and have a morphologically motivated distribution which does not match with any morphosyntactic natural class. Within the masculine paradigm, the Strong stem fills the cells of nominative singular, dual and plural, but the accusative singular and dual. Within the neuter paradigm, the Strong stem has arbitrary distribution, occupying only the cells of the nominative/accusative plural. Likewise, the Weak (Middle) stem is morphologically conditioned appearing elsewhere in both paradigms. Given the arbitrary distribution, these stems are morphomic and they are lexically listed with their unique indices to which stem choice rules are sensitive (Stump, 2001b, pp.170-171). An example of the three-way stem distinction comes from the perfect active participle TASTHIVANS ‘having stood’ which has the Strong stem (*tasthivans*) and Weak stem variants: Middle stem (*tasthivat*) Weakest stem (*tasthus*) (2001, p.173). The choice between the Strong and Weak stems is determined paradigmatically while the alternation between Middle and Weakest stems is syntagmatically based. In the former case, the masculine paradigm has the Strong stem filling the nominative/accusative singular and dual in addition to the nominative plural cells. By contrast, in the neuter paradigm, the Strong stem only appears in a single cell that of the nominative plural. On the other hand, the Weak stem variants are determined by the phonological form of the affix, with the Middle stem appearing before a vowel-initial suffix and the Weakest stem occurring elsewhere (before a consonant-initial suffix or covert suffix).

In addition, the three paradigmatically-defined stem types correlate with with three phonologically defined stem types known traditionally as the Vrdhi, Guna and Zero grades. Stump (2001b, p.186) provides the following definition of these grades:

(6) Sanskrit grades

For any gradational nominal L, each of (a)-(c) implies the other two:

a. The Vrdhi-grade stem of L has the form Xā(R)C_0

b. The Guna a-grade stem of L has the form Xa(R)C_0

c. The Zero-grade stem of L has the form X(R̥)C_0

Ideally, the three morphological stems would map their corresponding grades with Strong, Middle and Weakest stems matching up with Vrdhi, Guna and Zero grades respectively. For instance, in the declension of PAD 'foot', the Strong and Middle stems correspond to Vrdhi and Guna respectively. However, Stump provides many cases in which the correspondence between the morphological stems and the phonological definition does match. For example, ATMAN 'soul' the Middle maps with Zero grade while Weakest stems corresponds to the Guna grade. These cases of mismatches provide further support for the morphomic status of Sanskrit stem alternation[17] (see §4.5.4 for the discussion of TLA morphomic stems). The following sections will discuss the stem sensitivity to extramorphological factors. However, arguments for an inflectional stem have been also raised for a number of different languages as will be discussed in the following section.

[17] Blevins (2003) also identifies morphomic 'first stem' in Western Germanic languages. Likewise, Baerman and Corbett (2012) identify morphomic patterns, diachronically maintained across the Chinantecan languages.

2.3.3 Stem sensitivity to morphosyntactic factors

Aronoff's (1994) discussion of the Latin morphomic stem system is based on Matthew's (1972) analysis of the supine system based on 'parasitic formation' in which the base of a morphological form is not the lexical root but the stem of another morphological form in the same paradigm. The future participle generally associated with the feature set [Future, Active] but it is based on ('parasitic on') the perfect participle generally associated with the incompatible feature set [Past, Passive] (Table 2.9).[18] This results in a property clash where the stem is associated with two distinct and incompatible MPs ((supine) stem and [ur: 'future.participle.active']). Matthews (1972) suggests using the parasitic stem whereby the future participle is derived from the perfect participle and both have the meaning of participle. The other possibility is to use the supine stem as the minimal unit from other forms can be derived. This stem choice seems to entail that, unlike the perfect and future participles, constructing from the supine does not trigger clash. The supine may be favoured over the perfect participle stem, because the former would be the neutral choice as far as voice is concerned. Nevertheless, the parasitic formation of the stem is problematic since it assumes that a form with the MPs of passive and past is constructed and then changed into another form with the MPs of active and future.[19]

---

[18] For a verb such as 'amo' 'to love', '…the Future Active amatur- seems to derive from amat- by the addition of -ur-. Or, […], amaturus comes from amatus by the change of -s to -rus. But there is no sense in which the meaning of the Future Active Participle includes that of the Past Passive Participle. Formally, amat-ur- includes amat-. But in meaning all they have in common is that both are Participles.' (Matthews, 1991, p.200). In other words, the Latin verbs are derived from an underlying stem (the passive perfect participle or the supine) which is the input for constructing the active future participle. Therefore, the active future participle stem is parasitic on the supine or the passive perfect participle. Note that Aronoff (1994) uses the label perfect participle to what Matthews (1991) refers to as the past passive participle.

[19] One possible solution is suggested by Booji (1996, p.50) who says that it is possible to '…specify that the stem in -t- has a default association with the feature values [Past, Passive], but that in the appropriate contexts-i.e. when followed by the suffix /ur/-this default is overridden.'

On the other hand, Aronoff (1994, p.35) argues that the choice of the stem is arbitrary and both participles are constructed on the third stem which has no active or passive semantic value. One piece of evidence for that comes from certain Latin verbs which exhibit a future participle, but have no corresponding past participle (e.g. *caleo* 'burn', *calitur-*, *doleo* 'suffer pain', *dolitur-*). 'In the absence of a perfect participle, it is difficult to see how these future participles can be derived, unless this stem, […] the third stem, has some independent status.' In other words, the third stem is morphomic and avoids incompatible MPs in inflected words.

Inflectional stems have been described as inflectional elements which realize or paradigmatically associate with a matching set of MPs. Baerman and Corbett (2012) develop a canonical approach to stems. In a canonical inflectional system, the morphosyntactic features have to be marked by affixes varying across the cells while the stems are lexical material that have to be the same within a paradigm. Yet, they identify a number of non-canonical instances in which stems are alternating and represent an inflectional marking object as part of multiple exponence. In fact, they distinguish two instances of deviation from the canon based on the division of feature marking between stems and affixes. For instance, there are cases in which stem alternations seem to independently realize/bear the different sets of MPs. The example inflectional system they provide comes from Modern Greek verbs which inflect for aspect, person, number and tense. Consider the following paradigm of Modern Greek 'tie':

MPs	Imperfective	Perfective
1SG NONPST	ðén-o	ðés-o
2SG NONPST	ðén-is	ðés-is
3SG NONPST	ðén-i	ðés-i

Table 2.10 The imperfective and perfective stem in Modern Greek 'tie' (Baerman & Corbett, 2012, p.57)

In this paradigm the affixes realize person, number and tense while the stem alternation appears to be marking the aspectual feature, hence *ðén* is the imperfective stem while *ðés* represents the perfective stem. However, Baerman and Corbett (2012) do not rule out the possibility of a diachronic account in which aspect was lexically coded and these two stems originally were part of two different lexemes. Other cases of canonical deviation involve varying degrees of overlapping in feature marking relationship between stems and affixes. For instance, in Nuer, a Western Nilotic language, nouns inflect for case and number as in the noun declension paradigm in (Table 2.11) in which both the stem and the affix alternations in example A seem to mark case-number features.

	A	B	C
MPs	‘bark (of dog)’	‘ear’	‘meat’
NOM SG	gua̱	ji̱th	ri̱ŋ
GEN/LOC SG	gui̱-kä	ji̱th-ka	ri̱äŋ
PL	gui̱a-ni̱	ji̱th-ni̱	ri̱i̱ŋ

Table 2.11 The nominal declension in Nuer (Baerman & Corbett, 2012, p.59).

In B, the same stem is used for the same case-number values (nominative singular, genitive/locative singular and plural) for which affixes alternate. By contrast, in C, the same case-number features seem to be realized by stem alternations only in the absence of the affixes. Nevertheless, in most cases the nominal stem allomorphy in Nuer is associated with mismatching morphosyntactic values. This observation can be used as an argument for the morphomic patterning of the encountered cases. Although in the Nuer the feature marking seems to be based on matching values between stems and affixes, these matched values do not themselves form a natural class. For instance, the singular cells in the declension of example A ‘bark’ have two stem variants one for NOM.SG and the other for GEN/LOC SG. On the other hand, other declensions do not show stem alternations as in the declension of example B. Furthermore, although the stem alternation in C shows value-split (Corbett & Mithun, 1996) based on number, it is associated with mismatching case values.

Another example of stem sensitivity to morphosyntactic factors comes from Lithuanian. According to Arkadiev (2012), the stem alternation in the verb system of this language seems to correlate with a coherent set of morphosyntactic values. Lithuanian verbal inflection is based on three stems: the Present stem, the Past stem and the Infinitive. To assess the morphological or morphosyntactic status of these verbal stems, Arkadiev (2012, p.15) presents an account based on inflectional and derivational morphology. The three stems can have paradigmatic distribution to varying degrees. For instance, both the Present stem and the Past stem associate with a coherent set of MPs, that is, present and past tense respectively. However, the Past stem does not fill all the cells with the value past tense, since the past passive participle cell is associated with the infinitive stem. In fact, all the rest of 'paradigmatic labour' is done by the infinitive stem which also occurs in the active participle of future and past habitual. Therefore, the distribution of the Present and the Past stems in the verb inflectional paradigm seems to point towards a case of inflectional stems realizing the values present and past tense respectively. Nevertheless, as Arkadiev points out, the distribution of these stems is not limited to the inflectional system: they can be used in deverbal derivation. For example, the deverbal noun in *–es(ys)* (e.g. *skambesys* 'sound', cf. *skambe* 'resound') is based on the Present stem while the deverbal noun in *–imas* is based on the Past stem (*pardavimas* 'sale' from *parduoti* 'sell').

Clearly, the inflectional status of these stems depends on whether the derivational distribution is taken into account or not. Within the inflectional system, the Present and Past stems seem to be marking the relevant morphosyntactic features while the infinitive stem is morphomic due to its arbitrary distribution. However, Lithuanian verb stems are similar to the Latin third stem in that they are both used in derivational as well as inflectional morphology. Therefore, the derivational morphology of the Lithuanian system supports the morphomic status for all the stems since the set of morphosyntactic features that might be marked by these stems cannot provide a synchronic account for the stem choice in lexeme formation (see §4.5.3 for the discussion of the morphosyntactic effects on TLA stem change). Nevertheless, stems may also be sensitive to other extramorphological factors.

2.3.4 Stem sensitivity to morphophonological factors

Phonology is another factor that can interact with morphology with respect to stem alternations. In other words, the semi-autonomous stem status can also correlate with morphophonological factors, two of which are relevant to the data of this study. The first is based on the syntagmatic distribution of stems *conditioned* by the phonological environment of affixes. The second morphosyntactic factor that interacts with the stem morphology is based on stress *conditioning ablaut change/stress mobility*. These two morphophonological effects can be illustrated by the N-pattern and U-pattern/L-pattern in Romance. Equipped with evidence from a diachronic perspective, Maiden (1992, 2005, 2009) argues for the morphomic status of stem allomorphy in different Romance languages, yet, the effect of (morpho)phonological factors on the morphological component is not completely dismissed. The N-pattern is an arbitrary label denoting a 'morphomically-driven' stem alternation that can be found across Romance. The distribution of the alternation distinguishes the 1/2/3SG and 3PL forms of the present tense in addition to the 2SG imperative from the rest of the paradigm with distinct 'root'. The stem allomorphy is characterized by vocalic alternations and differentiation based on stress placement. For instance, in Medieval French, the stress falls on the root in the 'morphologically marked' cells while the cells in the rest of the paradigm have unstressed roots.

lef	'wash'	leves	leve	lavons	lavez	levent
crief	'burst'	crieves	crieve	crevons	crevez	crievent
peis	'weigh'	peises	peise	pesons	pesez	peisent

Table 2.12 Present indicative in Medieval French (Maiden, 2009, p.52).

The original phonological environment of the N-pattern, which is systematically based on stressed roots, has remained intact across Romance. However, a stress conditioning account for the stem alternation is rejected, as Maiden (2009) points out, for many reasons. For instance, originally, the stress in Latin was phonologically predictable whereas, in Romance, the stress placement may not be determined on phonological grounds; hence, any of the last three syllables of a word can carry the stress. Therefore, the distribution of the stress in the verb paradigm is determined by the N-pattern paradigmatic distribution. The stress in the verb inflectional forms falls on the ending unless it is part of N-pattern then the stress is on the root. Another specific reason, which is of particular interest in this work, for analysing the N-pattern as conditioned by morphology independently from stress placement comes from cases in which the stress placement is clearly irrelevant to a suppletion process. Maiden (2009, p.21) gives an example case from Romansh in which the reflexes of the verb 'pull' *tirare* which originated from Latin, the third conjugation verb *trahere* 'draw' which retains a stressed root in the infinitive, are associated with the distribution of N-pattern. However, these reflexes do not have an effect on the infinitive root that indicates the independence of the suppletion from stress assignment.[20]

[20] For further detailed arguments in favour a morphomic analyses of these two patterns, the reader is referred to Maiden (2005, 2009, 2010).

The second type of alternation identified by Maiden (2009) is the L/U-patterns, both of which denote stem alternations distinguishing the present subjunctive and 1SG present indicative and additionally the 3PL present indicative (U-pattern). The L-pattern is recurrent across Romance while the U-pattern can be found only in parts of Italy and Romanian. The two patterns are characterized by two phonological effects including the 'yod-effect' and palatalization, along with velar consonant affrication. Diachronically, the phonological process of the yod-effect was originally found in Latin in which the unstressed thematic front vowels /E or I/ become yod in the 1SG. present indicative, and the present subjunctive of non-first conjugations, motivating the L-pattern. By contrast, in the 3rd and 4th conjugations, the yod effect also included the 3PL indicative, hence the U-pattern. The subsequent history of this phonological process also involved the palatalization and/or affrication of consonants before yod in addition to the modification of root-final consonants. In fact, the resultant phonological environment has considerably changed over time and has led to the introduction of new palatal/affricate consonants. Therefore, in terms of the stem alternation conditioning, although the phonological effect has persisted, it does not necessarily mean that it is the conditioning motivation for the stem change. Maiden (2009, p.69, p.148) further supports his observations against a phonological analysis by the fact that there are cases in which the phonological environment is met, but, the yod-effect fails to occur. For instance, there is no evidence for the existence of a synchronic process of phonological conditioning in Italian verb stems. Counterexamples (e.g. stan[k]i 'you tire') show that velar consonants can precede front vowels. Likewise, palatal consonants can occur before vowels that do not trigger palatalization (e.g. pa[ʎʎ]a 'straw'). Furthermore, the resulting alternations of yod and palatalization along with velar consonant affrication processes are manifested in arbitrary alternations (cf. Italian [g], [k], [j], [ʎʎ], [ɲɲ], [lg], [ɲg], [ddʒ], [ttʃ]). The arbitrariness of the stem alternation is also the nature of paradigmatic distribution of L-U patterns.

It is not a straightforward task to determine the extent to which stem *morphomicity* can interact with phonological factors and this has led to different views on the exact status assigned to stem allomorphy. One account (Spencer, 2012; Maiden 2005, 2009, 2011) favours morphologically motivated stem allomorphy which can be sensitive to phonological elements. For instance, although 'm-p stems' are syntagmatically distributed, yet they remain within the domain of a morphological ruling over the stem allomorphy and regard the phonology as a requirement of the morphology. Another account (Anderson, 1992, 2008, 2011; Carstairs-McCarthy, 1988) argues for a phonological based account of the stems, yet, admits the lexical nature of stem alternation. Therefore, for this approach, the stem allomorphy is 'phonologically predictable/conditioned' and the morphological effect implemented in partial suppletion, for example, is subsidiary. The less controversial stem change type is the purely phonologically triggered, which tends to be motivated to avoid violation of any phonological constraints. This type of stem allomorphy is considered to be 'phonologically triggered/motivated', resulting in two stem variants 'allostems' (Spencer, 2013a) of the same abstract stem.

Carstairs-McCarthy (1988, pp.17-20) was first to discuss cases which show systematic variations associated with phonological factors, but are nonetheless lexically determined. For instance, he considers the N-pattern in Italian as a function of the suprasegmental effect produced by the stress assignment; hence, he distinguishes between the 'stress-indifferent default alternant' and the 'stress alternant' (e.g. Italian *romp* ~ *rupp* of the verb *rompere* 'to break') which is 'phonologically conditioned suppletion.' Favouring a phonological approach over a morphological one was further embraced with the introduction of OT, employing constraints to account for phonologically conditioned stem allomorphy. The phonological view of stem variation is argued for by Anderson (2008, 2011, 2013, p.9) who provided a detailed discussion of the nature of stems in Surmiran of Savognin which has six classes determined by the vowels that appear in certain inflectional forms attached to the suffix. The verbs with –ar are the most productive class such as the regular verb *cantar* 'sing' which in the present indicative is based on the verb stem along with the suffixes: -Ø (1SG), -as (2SG), -a (3SG), -agn (1PL), -ez (2PL), -an (3PL) (2008, p.111). The stress is assigned to the stem/root in all the persons except for the second and third plural forms. In addition, the verb stem alternates between unstressed form [kən] and stressed form [kan]. Anderson (2008) points out that the alternation between stressed and unstressed vowel has to be morphologically significant as it distinguishes the infinitive endings of the first and second ICs from that of the fifth conjugation. Nevertheless, Anderson considers the stem alternation of the regular verbs as stem variants of one base stem.

However, this inflection system also has a set of 'alternating verbs'. The paradigm of this type displays two stems: 'unstressed Stem1' and 'stressed Stem2' (e.g. the verb *ludar* 'to praise' has Stem1 /lud/ and Stem2 /lod/, the verb *vurdar* 'to watch' has Stem1 /vurd/ and Stem2 /vard/ and the verb *fittar* 'finish' has Stem1 /fet/ and Stem2 /fit/). These two stems have different paradigmatic distribution. Stem1 occurs in the '1/2PL present indicative, imperfect, future and conditional' in addition to the 2PL imperative and the present/past participles. By contrast, Stem2 is employed in the cells of '1/3SG and the 3PL of the present indicative as well as the present subjunctive and the 2SG of the imperative'. The stem choice in the infinitive sub-paradigm is based on the origin of the verb itself. Stem1 is generally selected in the infinitive of most verbs, but if the verb is historically derived from the third conjugation of Latin, then Stem2 is used instead (e.g. *discorrer* 'speak', has the Stem1 *discor* of 1PL.PRS and the Stem2 *discurrign* of 1SG.PRS). Overall, the essential difference between the stressed and unstressed stem is the vowel quality. The unstressed syllable can have any of the three vowels: [ə], [I] or [U] while the corresponding stressed alternant can have seven/eight distinct vowels and diphthongs. Since the correspondence between the stressed and unstressed variants is not unique, it is not possible to postulate a phonological rule which derived the two stems from a unitary underlying form.

However, Anderson (2008) argues that these stem alternations have the status of 'phonologically conditioned allomorphy' status of these stem alternations. Nevertheless, his synchronic account acknowledges the fact that the phonological forms of these stems have to be specified lexically, because purely phonological rules cannot predict the stress-related ablaut change according to which different vowel qualities appear independently in each stem, that is, there is no stem vowel *interpredictability* (Bonami, 2012). Nevertheless, stem allomorphy is handled on phonological grounds by formulating OT constraints distinguishing between two stem phonological forms: one 'is suitable for having stress placement on it, and the other of which is suitable for remaining unstressed'. Although, as Maiden (2009, p.75) points out, this approach runs into difficulties accounting for suppletion, the distribution of 'augments', verbs with varying root, Anderson concludes that 'there is no warrant for invoking the further step of complete and arbitrary morphological categorization that would be implied by associating the variation with a morphome' (2013, pp.11-23). Clearly, the stem alternation such as that of the N-pattern results in types of stems that are not mere allostems, but not a clear-cut case of pure morphomes either. While Spencer (2013b) suggests treating these stems as an intermediate special type, labelled m-p stem, both Maiden (2009) and Anderson (2008) warrant the setting up of one ruling domain either a morphomic according to which each stem requires indexing or a phonologically conditioned stem.

Another type of phonologically specific stem is discussed by Aronoff (1994, p.59) who makes a distinction between abstract and concrete morphomic stems. The form of the abstract stem may vary from one verb to another as it is the case for the Latin present, perfect and the third stem, for example, unlike other classes both *'are'* and *'ere'* have a theme vowel in the present stem. 1st and 4th conjugation verbs have a theme vowel in all stems while 2nd and 3rd conjugations never have a theme vowel except in the present stem. Unlike the abstract stem, there are stems that tend to be phonologically specified and more concrete, such as the Latin [b] stem on which the imperfect of all verbs and the future of the first and second conjugation are formed. Both imperfect past tense and future of 'a' and 'e' verbs (1st and 2nd conjugations) are formed by attaching the stem component /b/ preceded by a long vowel (conjugation specific vowel $\bar{V}$) but the suffix –b- does not realize any morphosyntactic value, and hence is what Hockett (1947) calls an empty morph forming a complex stem (Aronoff, 1994, p.58).

Conjugation	Future 2SG	Imperfect 2SG	Gloss
ā	laudābīs	laudābās	'praise'
ē	mordēbis	mordēbās	'bite'
e	acuēs	acuēbās	'sharpen'

Table 2.13 Examples of phonologically specific Latin 'B' stem.

The Latin 'B' stem is 'phonologically specific', and it has an invariable constant sound form which should be distinguished from the abstract stem that has a variable sound form. However, Aronoff (1994) concludes that further research should investigate the difference between these types of stems.

2.3.5 Summary

The unresolved conversation between Maiden and Anderson and my data are at the heart of this study. Both stem predictability and unpredictability can be used as a measure for TLA stem morphomicity in addition to the extent to which extra-morphological factors are involved. A possible approach would reflect and represent the status of the stem by a binary feature system for stem 'morphomicity' (§4.5 , 4.5.1, 4.5.2, 4.5.3 and §4.5.4). Providing a plausible account for stem alternation is also essential for a better understanding of the other paradigmatic dimensions such as ICs and paradigm implicative relations. The next section will discuss two frameworks for handling the morphological complexity of stems and analysing the IC system.

2.4 Inflectional paradigms and paradigmatic dimensions

Paradigms represent the inflectional forms of a lexeme, comprising of a number of cells that realize different MPs. The size of a paradigm may vary from one language to another depending on the inflection system of a language. For example, the number of cells in a paradigm in Vietnamese is very low whereas Sanskrit is a highly inflected language and can have hundreds of cells (Stump, 2001b). Filling the paradigm cells has been the subject of much research that shows that cells in a morphological paradigm exhibit correlations. Traditionally, the notion of principal parts was used to represent predictability in Latin verbal paradigms in which four verb forms provide inferences for the whole paradigm (Aronoff, 1994). However, in recent years, this implicative structure of paradigms and morphological analyses have been based on a notion borrowed from artificial intelligence and computer science, namely, 'default inheritance'.[21] Under this view, networks consist of nodes connected by default relations that can be overridden. The members of a category (daughter node) inherit some features from the main category (mother node) by a default assumption which can be overridden by more specific information. In other words, the features of a subordinate category can be inherited from a superordinate category provided that those features are not specified. For instance, in English morphology, the form of the present participle of all verb forms (e.g. *going)* does not need to be listed as it can be inferred from the basic form (e.g. *go*) by default relations. By contrast, the inheritance hierarchy of this system needs to list the past participle form of *go* because that form overrides the form which is defined at the the default node (i.e. the form *gone* blocks **goed*) (see §2.4.1 for further discussion on default relations and the stem space). Likewise, in TLA morphology, the future form does not need to be listed since it is always predictable from the default imperfect form of any verb even for the highly irregular verbs (e.g. *yakil* 'he eats' and *byakil* 'he will eat'). By contrast, the value of the imperative form of this verb *kuul* 'eat!' has to be listed and it overrides the default values inherited by other verbs of the same class (the defective series) (e.g. *milee* 'he filled' *ya-mla* 'he fills', *b-ya-mla* 'he will fill', *amla* 'fill!') (see §3.1 §3.1.2 for TLA verb system and verb paradigm and §5.2.2 for TLA dependency relations).

The information constituting the inflectional system of a language can be represented by an inheritance tree in which the top node has the most general information and the lowest nodes have more specific information. In addition, inheritance hierarchies can list intermediate nodes which carry the values of the stem (Aronoff, 2012). The hierarchy represents an implicative network in an inflection

[21] For the use of the notion default inheritance in the theory of Network Morphology, see Corbett and Fraser (1993) who use the computer language DATR (Evans and Gazdar, 1996). For the morphological analysis of Hebrew binyanim system using KATR, the reader is referred to Finkel and Stump (2007).

system, connected by 'dependency relations' which minimize the amount of redundant information to be stored for a given lexeme (Bonami & Boyé, 2002, p.51) (§2.4.1).

Another morphological phenomenon within the paradigmatic dimension is the notion of inflection classes (ICs) which represent inflectional differences that are attributed to suffix allomorphy and/or stem allomorphy represented in organized networks. ICs are an important challenge in morphology since they do not have any identifiable morphosyntactic function, that is, they are *morphomic* (Aronoff, 1994). In addition, they are typically defined in terms of affix allomorphy in languages such as Latin and Spanish. However, Bonami and Boyé (2002) have argued that inflectional complexity can also be determined by stem sets. There are varied accounts of how paradigm complexity can be reduced by identifying the stem-space of a language and the relationships that hold among the stem allomorphs. Traditional descriptions of Arabic lack ICs but modern dialects have developed stem alternations which partition the lexicon into classes (Hoberman & Aronoff, 2003) as will be illustrated for TLA (Chapter 3 and Chapter 4) using the stem-space inflectional class account (Bonami & Boyé, 2002) (§2.4.1) and compared with the principal parts analysis (Stump & Finkel, 2013) (§2.4.2).

2.4.1 Stem space and inflection classes

In the previous sections (2.3.2, §2.3.3 and §2.3.4), I presented a number of examples which showed that in many inflectional systems, stems show allomorphic variation, hence, lexemes tend to be associated with a set of different phonological representations (i.e. stems). A morphological system can be based on a vector (stem space) with a set of different stems connected by stem (dependency) relations. The stem space has been proposed for stem alternations in different inflectional/derivational systems including Latin (Aronoff, 1994), Sanskrit declension (Stump, 2001), French conjugations (Bonami & Boyé, 2002), Spanish conjugation (Boyé & Cabredo Hofherr 2006) and Italian conjugation (Montermini & Bonami, 2011).

The evolution of this notion derives from the notion of 'Overall Distribution Schema' developed by Pirelli and Battista (2000). The idea is based on the observation that the inflectional paradigm of many inflectional systems in Romance shows stem variation which is distributed systematically across paradigms. The majority of these stem alternations are the outcome of a morphomic distribution. Montermini and Bonami (2011) set out the motivation for the stem space as below.

'*stems are indexed, and languages specify with which stem a specific cell in the paradigm should be filled. In other terms, for a given class of lexemes a language specifies a stem space, corresponding to the set of procedures that allow to select the appropriate stem for each cell in the paradigm*.' (Montermini & Bonami, 2011, pp.8-9).

The stem space model is similar to the notion of lexical storage (how words are stored in the speaker's mind in relation to each other) in that it represents the internal structure of a paradigm and how unpredictable (irregular) stems can be inferred from the predictable (regular) stems. In addition, the distribution of these stems is constrained within a paradigm as will be illustrated below for the French verb inflectional system (Bonami & Boyé, 2002).

In the French system of verb inflection, nearly all verbs exhibit the same inflectional ending, therefore, the ICs are not distinguished by the inflectional suffixes but rather by modifications internal to its stems. Given the regularity of inflectional exponents, Bonami and Boyé (2002, p.3) classify verbs in terms of their inflectional regularity as represented by their stems. Verbs such as *laver* 'wash' (Table 2.14) are regular because they share the same stem across the whole paradigm.[22] By contrast, irregular verbs can be classified by two types of inflectional form irregularity. The first type includes verbs such as *mourir* ‘die’, which are based on unexpected (alternating) stems (*mœr* ~ *mur*) while the second kind has ‘suppletive inflection forms’ (e.g. *être* ‘be’) (Table 2.14).

regular verbs		
laver ‘wash’		
	SG	PL
1	lav+Ø	lav+ ɔ̃
2	lav+Ø	lav+ e
3	lav+Ø	lav+ ə

verbs with unexpected stems		
mourir ‘die’		
	SG	PL
1	mœr+Ø	mur+ ɔ̃
2	mœr+Ø	mur+ e
3	mœr+Ø	mœr+ə

verbs with suppletive forms		
être ‘be’		
	SG	PL
1	sɥi	sɔm
2		ɛt
3		sɔ̃

Table 2.14 The stem allomorphy of regular and irregular verbs in French (in the cells of the first, second, third singular plural forms).

Inflectional irregularity is based on suppletion. However, unlike totally suppletive inflectional forms, irregular verbs with stem alternations show a systematic distribution. This is illustrated in Table 2.14 in which different verbs have different stem allomorphic patterns in the present indicative paradigm, but the stems are grouped in the same cells across the verb categories. The regular verb *laver* has one stem shared across the whole paradigm. Verbs with unexpected stems can be divided into two types: semi-regular verbs such as *mourir* ‘die’ with two different stems and irregular verbs such as *boire* ‘drink’ with three different stems (Table 2.15).

[22] Bonami and Boyé (2002, p.52) propose the general hypothesis in (i) for regular inflection
i. **Regular inflection**
‘The phonology of a regular inflectional form is a function of the phonology of some stem of the verb this form instantiates.’

Verb	Gloss	1SG	2SG	3SG	1PL	2PL	3PL
laver	wash	lav+Ø	lav+Ø	lav+Ø	lav+õ	lav+e	lav+ə
asseoir	sit	asje+Ø	asje+Ø	asje+Ø	asεj+õ	asεj+e	asεj+ə
mourir	die	mœr+Ø	mœr+Ø	mœr+Ø	mur+õ	mur+e	mœr+ə
boire	drink	bwa+Ø	bwa+Ø	bwa+Ø	byv+õ	byv+e	bwav+ə

Table 2.15 Stem alternation in French present tense paradigm (Bonami & Boyé, 2002, p.54)

Bonami and Boyé (2002) proposed that it is possible to account for inflectional irregularity using the device of a stem-space in which stem slots are linked by dependency relations. The French stem space consists of *stem slots* (zones) selected by verb paradigm slots as shown by the present tense paradigm in which the singular inflection forms share the same stem slot. The 1PL/2PL cells form a group by selecting a unique stem slot different from the 3PL cell (e.g. *asje* ~ *asεj*) (Table 2.15). Given the suppletive stem patterns, the French verb system requires up to twelve morphomic stems filling the stem-space as illustrated in Table 2.16.

stem name	**inflectional forms build on this stem**
Imperf, Prst.1.2PL	All forms of the imperfective; first and second person plural in the present
Prst.SG	All singular forms in the present
Prst.3PL	Present third person plural
Sbjv.1.2PL	First and second person plural in the subjunctive present
Sbjv.SG, 3PL	All singular forms and third person plural in the subjunctive present
Imper.SG	Imperative second person singular
Imper.PL	Imperative first and second person plural
Prst.Part.	Present participle
Fut./Cond.	All forms of the future and the conditional
S.pst.	All forms of the simple past and subjunctive imperfective
Inf	Infinitive
Pst.part.	Past participle

Table 2.16 The stem-space in French (Bonami & Boyé, 2002, p.55).

The stem space for French conjugation allows twelve stem slots with morphomic stems all of which provide inferences for the whole paradigm of a lexeme. For example, the stem listed in the slot (Imperf, Prst.1.2PL) is shared across all forms of the imperfective 1.2PL in the present across the different verbs. In addition, this stem space is sufficient to account for all the stem alternations of French verbs (except suppletive inflection forms[23]), expressing the maximal irregularity of French stem allomorphic system.[24] However, no verb has a degree of complexity that represents the maximal stem alternation pattern theoretically possible in this system. Many of the verbs have no unpredictable stems (regular verbs), and one stem is sufficient for filling the stem space, based on default strategies (Figure 8), while the irregular verbs require up to three different stems as illustrated in Table 2.17.

verb	Gloss	Imperf. Prst.1.2PL	Prst.3PL	Prst. SG	Sbjv. 12PL	Sbjv.SG Sbjv.3PL	Imper. SG	Imper.PL	Prst.part
laver	wash	A	A	A	A	A	A	A	A
devoir	owe	A	B	B	A	B	B	A	A
haïr	hate	A	A	B	A	A	B	A	A
avoir	have	A	A	A	B	B	B	B	B
être	be	A	A	B	C	C	C	C	A
valoir	cost	A	A	B	A	C	B	A	A
savoir	know	A	A	B	C	C	C	C	C
pouvoir	can	A	B	B	C	C	B	A	A
vouloir	want	A	B	B	A	C	B	A	A
faire	do	A	A	B	C	C	B	A	A

Table 2.17 Ten French verb ICs

---

[23] Examples of these forms include *être* 'be', *avoir* 'have', *faire* 'do', *aller* 'go', *dire* 'say'. Unlike verbs with alternating stems, these verbs display monosyllabic suppletive forms in the present indicative and the imperative. It is possible to expand the stem space to account for these verbs without affecting the overall structure of the model.

[24] In subsequent discussion of stem spaces and ICs, Bonami and Boyé (2002, p.62) opt to exclude four stem slots from the stem space (future/conditional, simple past, infinitive, and past participle). 'These stems are not identical to the imperfective stem *even for regular verbs*; rather, regular verbs derive these stems from the imperfective stem using a non-trivial phonological function'.

Table 2.17 lists the ten verbs that represent the possible patterns of stem alternations for filling the stem space in eight of the stem slots represented in Table 2.16. In other words, French system has ten conjugation patterns based on stem allomorphy. The first IC is represented by regular verbs such as *laver* 'wash' which has a regular stem alternation pattern (AAAAAAAA). By contrast, all the semi-regular (e.g. *devoir* 'owe' with two different stems in the alternation pattern ABBABBAA) and irregular verbs (*faire* 'do' with three different stems in the alternation pattern AABCCBAA) in this system display nine distinct ways of filling the stem space with a maximum of three stems in each alternation pattern. Nevertheless, the stem alternation is systematically distributed in the stem space and subject to a network of dependency relations that constraint the way that the stem space is filled. These relations can provide a plausible account for suppletion, showing that the unexpected stem does not occur in arbitrary cells of inflectional paradigms as explained by Bonami and Boyé (2002, p.62) below.

'*We assume dependency relations are not (in the general case) relations of default identity, but that when a dependency relation occurs between two slots x and y, there is a single phonological function f relating the two slots. The dependency relation entails that by default, the dependent slot y can be filled with f(x)- the result of applying the associated phonological function to the stem it depends on*.'

Under this view, the relations connecting between the stems are expressed by the same formalism as that of the relations between the inflected forms and their stems. Regularity follows the default expectation while the irregularity represented by unexpected stems or inflected forms are the result of deviation from the default stem-stem relations and stem-inflected form relations respectively (See §5.2.2 for application of this technique to TLA). In French, most dependency relations constraining the stem spaces are based on the identity function (stems are identical by default). The network of dependency relations in French is represented in the form of a *stem dependency tree* in which the relation between the mother stem and the daughter stem is either suppletive or identical (Figure 8, Figure 9 and Figure 10). The stem dependency tree can represent how the lexicon stores the stems of this inflectional system and avoids redundancy. Therefore, only the unexpected stems that are not identical to their mother stem need to be listed. For the rest of the stems, the grammar uses the tree geometry to derive predictable stems from the listed unpredictable ones.

For French verbal inflection, there are two types of dependency relations: simple and complex. Morin (1987) proposed that in the present tense paradigm, the stems show dependency relations in which both the Prst.3PL and Imperf, Prst.1.2PL stems are identical or suppletive. The same relation holds between the Prst.3PL and the Prst.SG stems (cited in Bonami & Boyé, 2002, p.56).

(7) Imperf, Prst.1.2PL –(identical or suppletive) → Prst.3PL –(identical or suppletive)→ Prst.SG

This simple dependency relation determines structure of the tree geometry and provides a plausible account for the regularity and semi-regularity in the French inflectional system in a descriptively economical fashion. Therefore, for regular verbs only one stem is stipulated as shown in the inheritance tree (Figure 8) in which the mother stem (Imperf, Prst.1.2PL) is inherited by default for both the Prst.3PL and Prst.SG unless other information is specified:

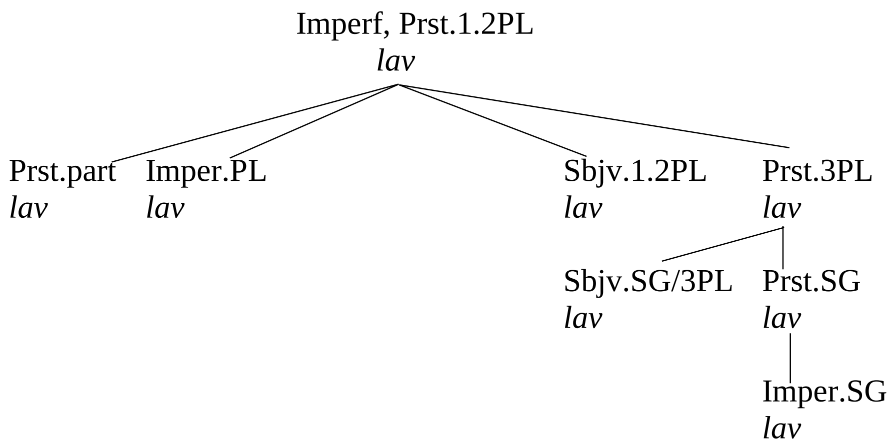


Figure 8 The stem dependency tree for the *laver* ‘wash’ (Bonami & Boyé, 2002, p.57).

The verb *laver* is a regular verb and shares the same stem across all the inflectional forms. The semi-regular verb such as *mourir* ‘die’ in (Figure 9) requires two distinct stems to be listed in the lexicon while three distinct stems are stipulated for irregular verbs (e.g. *boire* ‘drink’ in Figure 10):

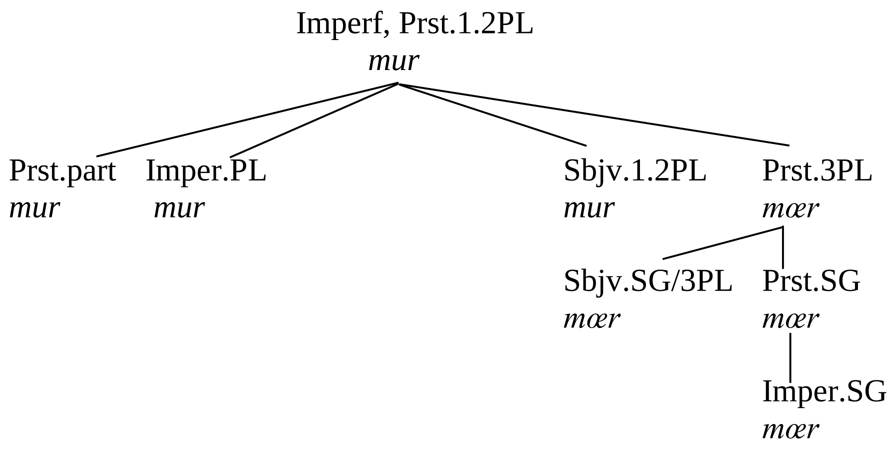


Figure 9 The stem dependency tree for the *mourir* ‘die’.

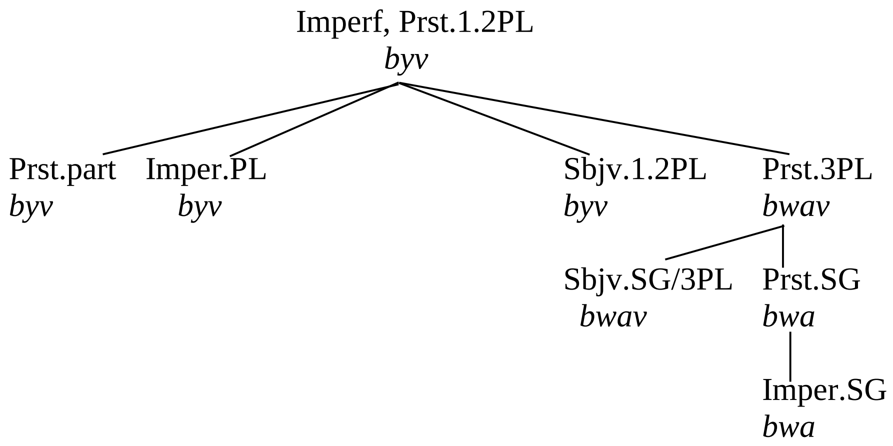


Figure 10 The stem dependency tree for the *boire* ‘drink’.

Both Figure 9 and Figure 10 show that the stems that are not identical to the mother node need to be stipulated while the rest can be predicted by using default dependency relations that may be overridden. For example, the irregular verb *boire* has the mother stem *byv* which associates with the Prst.part-stem, Imper.PL-stem and Sbjv.1.2prst-stem via the identity function. Identical stems must be used in the absence of other information specifying otherwise. By contrast, the stem slots Prst.3PL and Prst.SG override the mother default stem, stipulating two different stem *bwa* and *bwav* respectively.

However, there are stems which have complex dependency relations and have to be listed more than once. Bonami and Boyé (2002) identify fifteen verbs of which the dependency tree does not seem to reduce the phonological redundancy since each verb has two identical stems that do not seem to follow the dependency relations in (7) and hence, have to be listed in the lexicon twice. For example, the verb *savoir* ‘know’ has the stem /saʃ/ filling five different slots and the tree geometry suggested for regular and irregular verbs does not seem to provide a plausible account for the stem pattern in some verbs (Table 2.18) (Figure 11). Nevertheless, Bonami and Boyé (2002) argue that these identical stems are not randomly distributed and they can have one of the following dependency relations:

être 'be'	Imperative SG = Imperative PL = Subjunctive 1.2PL = Subjunctive SG/3PL
savoir 'know'	Present participle = Imperative SG = Imperative PL = Subjunctive 1.2PL = Subjunctive SG/3PL
faire 'do'	Subjunctive 1.2PL = Subjunctive SG/3PL

Table 2.18 Complex stem dependency (Bonami & Boyé, 2002, p.60).

Based on these distinct patterns, they propose the 'complex stem dependency' relation according to which the stem-space can have pairs of stem slots (x, y) of which the second stem is identical to the listed first stem by complex stem dependency as illustrated in following dependency tree (Figure 11):

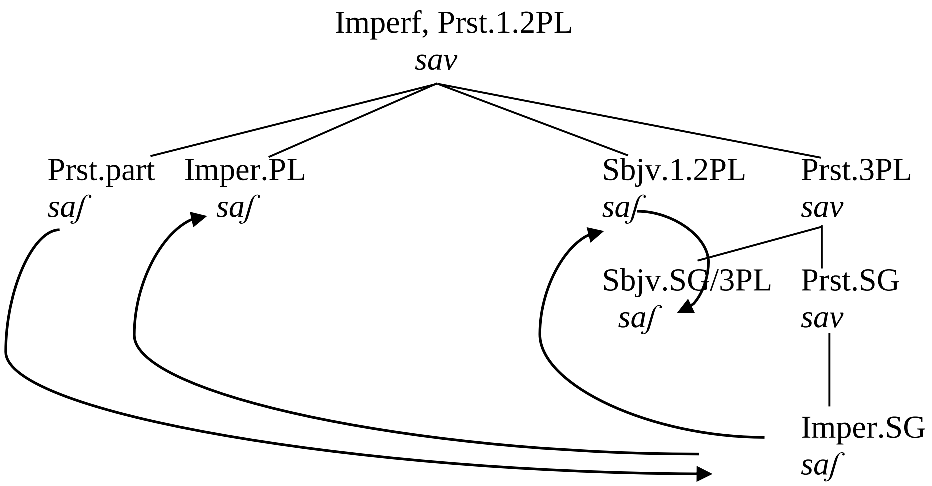


Figure 11 The stem dependency tree for the *savoir* 'know' (Bonami & Boyé 2002, p.61).

The four curved arrows in the tree geometry represent the complex stem dependency that can be found in French. For the verb *savoir*, only the present participle stem needs to be listed in the lexicon and this predicts that both the imperative and subjunctive share the same stem /saʃ/ as a result of a complex stem dependency relation on the present participle stem which is the only stem that needs to be listed in the lexicon.

Clearly, using the dependency relations modelled by the inheritance tree in filling the stem-space avoids stem redundancy in the lexicon and predicts ten ICs for the language. Although the connections in the complex stem tree can be used to represent the relations of the indirectly related stems, it is not clear why this network of complex stem relations still has to be represented in the geometry of a tree.[25]

2.4.2 Principal parts and inflection classes

Principal Parts are the minimal number of forms from which it is possible to deduce all/most of the other forms across the inflectional paradigm of a particular language system. In Latin, for instance, we can derive all ten forms of the nominal paradigm, for any of the five declension classes, by knowing the nominative and genitive singular. More recently, Stump and Finkel (2013) have formalized the notion principal parts, using them as an important tool for analysing the scale of complexity of different inflectional systems. We propose a novel stem-based classification of the TLA verb system based on implicative relations and the optimal principal parts needed to predict the remaining forms in a lexeme's paradigm using the Principal Parts Analyser (PPA), which is a computer programme developed by Stump and Finkel (2013) for analysing inflectional systems.

[25] It seems that the hierarchical structure does not hold for the verb *savoir*; yet, Bonami and Boyé (2002) still consider the dependence relation in this verb to be hierarchically based. The reason for this is possibly because the stem-space (and the principal parts approach (Chapter 5)) requires certain chosen cells to participate in the implicative relation and when faced with stem alternations such as those of the verb *savoir*, the account has to provide a solution which in this case is the complex dependency relation. This issue does not arise in other frameworks such as the entropy approach proposed by Ackerman and Malouf (2013), since in this account all cells can equally participate in the implicative relations and hence there is no need to commit to a hierarchical structure with complex dependency relations.

The complexity of IC systems can be examined through implicative relations within a lexeme’s paradigm. Traditionally, in Latin, inferences deriving the full paradigm can be achieved by principal parts that represent a subset of that paradigm. Stump and Finkel (2013, p.11) define the principal part of a lexeme L as ‘a set of cells in Ls realized paradigm from which one can reliably deduce the remaining cells in Ls realization paradigm’. Inflection systems may vary in terms of the kind of the principal parts they have: they can be regarded as realized cells (e.g. in Latin) or they can also be represented by indexed stems as in French and TLA. Both types are useful means for determining an IC membership of a lexeme. Unlike the later type of principal parts, the use of the former kind has a long history in the traditional grammar of Latin. For verb paradigms there are four implicative relations: the first person singular present indicative active, the first person singular perfect indicative active, the first perfect passive participle (supine and the present active infinitive). By memorizing these four forms, students are able to derive all the other forms of the paradigm. Examples of principal parts of five Latin verbs are illustrated in Table 2.19 (Stump & Finkel, 2013, p.12):

Conjugation	σ_1 = {1SG.pres. ind.active}	σ_2 = {1SG.perf. ind.active>}	σ_3 = {first supine}	σ_4 = {pres.active. inf}
1st	<laudō, σ_1>	<laudāvī, σ_2>	<laudātum,σ_3>	<laudāre, σ_4>
2nd	<moneō, σ_1>	<monuī, σ_2>	<monītum, σ_3>	<monēre, σ_4>
3rd	<dūcō, σ_1>	<dūxī, σ_2>	<dūctum, σ_3>	<dūcere, σ_4>
3rd (iō)	<capiō, σ_1>	<cēpī, σ_2>	<captum, σ_3>	<capere, σ_4>
4th	<audiō, σ_1>	<audīvī, σ_2>	<audītum, σ_3>	<audīre, σ_4>

Table 2.19 The principal parts of four Latin verbs.

These principal part forms participate in implicative relations to derive all the remaining cells in a lexeme’s paradigm, i.e. they are the minimum number of forms required for deriving the whole paradigm. For example, given the predictor first person-singular perfect indicative active form *laudāvī*, it is possible to deduce the person-singular pluperfect indicative active form *laudāveram*.

In language pedagogy, sets of principal parts can be characterised as unique, uniform, and optimal. The reason for the uniqueness characterization is that although lexemes can have a number of different subsets of cells that can be the predictor for the whole paradigm, only one of those subsets of cells is conventionally chosen to be the representative of the lexeme's principal parts. In Latin, for instance, the set of nominative singular and genitive singular is conventionally chosen as the principal parts for predicting the realized paradigm for AGER 'field' even though the prediction relations can also be achieved by another different set of cells such as the nominative singular and the genitive plural (Stump & Finkel, 2013, pp.14-15). In addition, the chosen principal-part set is uniform, requiring lexemes of the same syntactic categories have the same cells as their principal parts. For instance, in Latin, all nouns with different declension class membership have a uniform set of principal parts, represented by the nominative singular and the genitive singular cells. Finally, the predictor has to be optimal (adequate) which means that the principal-part set has to be represented by the minimal number of cells, but still deducing the whole paradigm and conforming with the uniformity requirement (optional). For example, the optimal set of principal parts for distinguishing the declension class membership of noun AGER consists of both the nominative singular and the genitive singular cells.

In the Stump and Finkel's (2013, p.30) approach, principal parts are the key for determining the IC membership and they can be classified into three main types presented by the schemes illustrated in Table 2.20.a, b and c. The first kind of principal parts is the traditional uniform scheme, which is static in which the set of principal parts of every IC of a lexeme of the same syntactic category can be determined by the same realized cells. In Table 2.20.a of the static scheme, W to Z represent distinct sets of MPs, each of I to VI is a different conjugation class and finally A to O represent the exponence which realizes the distinct set of MPs. In this static scheme, the set of MPs W, X and Y are the adequate set of principal parts that can deduce the whole realized paradigm in any of the six ICs. In addition, the W, X, Y cells are also minimal i.e. optimal. The second scheme for measuring the principal parts complexity appeals to dynamic principal parts, which may vary from one class to another, and hence are not necessary parallel. Conforming to the adequacy and minimality requirements, each IC has its own set of principal parts. If a lexeme's exponence for W is **a**, then we can identify L as a member of IC I. The third kind appeals to the concept of adaptive principal parts, in which all lexemes share the same first principle part, but some exponents for the MPs can be identified by the second principal part for a given IC. Under this conception, there are no further principal parts for a lexeme which has the exponent **a** because the first principal part is adequate and optimal. In contrast, a lexeme of the IC III requires two principal parts: the exponence **c** of W and the exponence **f** of X.

	ρ	σ	τ	υ
I	a	e	i	m
II	b	e	i	m
III	c	f	j	n
IV	c	g	j	n
V	d	h	k	o
VI	d	h	l	o

Table 2.20 a. Static scheme

	ρ	σ	τ	υ
I	a	e	i	m
II	b	e	i	m
III	c	f	j	n
IV	c	g	j	n
V	d	h	k	o
VI	d	h	l	o

b. Adaptive scheme

	ρ	σ	τ	υ
I	a	e	i	m
II	b	e	i	m
III	c	f	j	n
IV	c	g	j	n
V	d	h	k	o
VI	d	h	l	o

c. Dynamic scheme

In summary, the dynamic scheme necessitates the smallest optimal set of a lexeme's principal parts compared to the static and adaptive schemes. However, the dynamic scheme is more complex than the others are. The lexeme's dynamic principal parts are specifications of their associated MPs while the static and adaptive principal parts of a lexeme can be a sequence of word forms realizing sets of MPs that are the same across all the different inflectional classes (See §5.3.1, 5.3.2 and §5.3.3 for further discussion on the TLA static and dynamic principal parts).

2.4.2.1 Plats of inflection classes

The ICs of different systems vary in the degree of complexity. Stump and Finkel (2013) have developed a model for identifying and measuring ICs complexity in the format of a matrix that they call a plat (Table 2.21). The plats provide a plausible account for the implicative network among the cells of a realized paradigm.

In the matrix, the symbols *p-τ* represent different MPs while the I-VI are the ICs. The exponence of MPs are represented by **a-l** in the intersection of the column and rows.[26] If a lexeme has the exponence **a** for MP *ρ*[27] this predicts that the L belongs to IC I and that *ρ* is the optimal principal-parts set for that IC.

	ρ	σ	τ
I	a	f	j
II	b	f	j
III	c	g	k
IV	d	h	l
V	e	i	k
VI	e	i	l

Table 2.21 A plat representing a hypothetical system of ICs (Stump & Finkel, 2013, p.41).

In addition, the model is designed to avoid redundancy such that there is no need to include all MPs that share the same exponents. For example, in the plat in Table 2.22,

[26] The a-l represent the inflected word forms of the lexemes in the various ICs, which realize those MPs.
[27] The *ρ* is the morphosyntactic property (MP); the principal part is the form ~ property pair <a, ρ> (i.e. the word form **a** when it realizes property *ρ*).

it is not necessary to list the MP *υ* whose inference can be equally provided by another MP and that is *τ* which is already associated with all ICs of the given plat. Likewise, the MP *φ* does not have to be included since both the MP *τ* and MP *φ* show interpredictibility. For instance, whenever the exponent of the MP *τ* is **f** then the exponent of the MP *φ* would be **h**. According to Stump and Finkel (2013, p.42), 'where the MPs $\sigma_1,\ldots,\sigma_n$ are such that the exponence of σ_k is isomorphic to that of σ_1 for all k ($1 \leq k \leq n$), we call σ_1 the **distillation** $\sigma_1,\ldots,\sigma_n$. In the hypothetical IC system in (Table 2.22), the *τ* is the distillation of *τ, υ* and *φ*[28](see Chapter 5 §3.3.1 and §3.3.2 for the discussion on TLA distillations).

	ρ	σ	τ	υ	φ
I	a	c	f	f	h
II	a	c	g	g	i
III	a	d	f	f	h
IV	b	d	g	g	i
V	b	e	f	f	h
VI	b	e	g	g	i

Table 2.22 A plat representation for cases in which MPs are identical and isomorphic (Stump & Finkel, 2013, p.42)

2.4.2.2 The representation issue

Stump and Finkel (2013) also address the issue of data representation. They distinguish between two types of plats: hearer oriented (concrete phonology) and speaker oriented (abstract phonology) plat. In other words, plats list exponence in two ways. In the concrete (hearer) plat the exponents are given in the form of surface phonological/phonetic representations. This will mean that different allomorphic forms of the same stem (allostems) may have to be listed. In the abstract (speaker) plat the

[28] The MP *τ* would be the first distillation encountered when analysed by the PPA tool, therefore, it is a distillation of itself.

exponents are given in the form of abstract underlying representations. On this approach, allostems will be given in the form of a single representative underlying stem form. For instance, in French, under the concrete representation the verbs *aimions* /ɛm-jɔ̃/ 'we liked' and *craignions* /kʁɛɲ-ɔ̃/ 'we feared' have phonologically (slightly) different terminations. Consequently, these verbs belong to different conjugations. However, the abstract analysis reveals that these verbs belong to the same ICs constrained by same morphophonological rule (j→Ø/[palatal sonorant] _) according to which the termination /jɔ̃/ is realized as /ɔ̃/ when preceded by a palatal sonorant. In other words, both verbs have the same termination underlyingly (*aimions* |ɛm-jɔ̃| 'we liked' and *craignions* |kʁɛɲ-jɔ̃|). The alternative analysis is to assume that both verbs have the termination /ɔ̃/ while their stems end in a palatal sonorant /j/ or /ɲ/ in their concrete representation (/ɛmj-ɔ̃/, /kʁɛɲ-ɔ̃/) (pp.187-188). The corresponding abstract representations would be (*aimions* |ɛm-jɔ̃| 'we liked' and *craignions* |kʁɛɲ-jɔ̃|).

The issue of representation is essential for understanding the principal parts analyses. The two sorts of plats as we saw for French conjugation produce different results. Therefore, the choice of the plat will depend on the amount of regularity we seek for an inflectional system. Therefore, Stump and Finkel (2013) introduce sandhi rules whose function is to define a more regular plat in conformity with the speaker-oriented representation. The plat choice can also be determined by the sensitivity and the effect of the manipulations assumed, such as sandhi rules, leading to a more straightforward predictability and consequently a reduction in the number of ICs. However, if the rules of sandhi lead to a greater number of identical or similar ICs without grouping and factoring out redundancy, this can be considered as a less positive result. Thus, Stump and Finkel conclude (2013, p.353) '… there is no single correct plat for any language and what plat we choose depends on what we want the plat to represent…'

2.4.2.3 Principal parts as indexed stems

In the previous section, the principal parts of a lexeme were presented as a subset of realized word forms within its paradigm. In fact, this approach is useful if the inflectional system distinguishes IC membership by varying the exponence (the distinguisher[29]) as in Latin. However, there are languages in which all ICs inflect in the same manner, i.e. the exponents are redundant while the stems are alternating. Stump and Finkel (2013) propose that such inflectional systems should be analysed as having the principal parts based on indexed stems. For example, in French, the IC membership can be determined by stem allomorphy.

The plat structure under the stem-indexed analysis of principal parts is slightly different from the plats employed for principal parts based on cells in the realized paradigm. This type of plat is based on a stem formative that is the stem varying morphologically significant substring of a stem excluded from the theme.[30] The plats used for the French data give a list of indexed stems on the vertical axis while the distinctive stem formatives are at the intersection of a specific indexed stem with its IC membership.

The stem formatives vary depending on the representation. For example, under the concrete analysis, the stem formatives joined with the theme results in an inventory of twenty stems for each verb whereas the abstract produces nineteen stems. In other

[29] 'The substring by which a word form is distinguished from all distinct word forms in its realized paradigm is its distinguisher' (Stump & Finkel, 2013, pp.382-389).

[30] 'The theme of the realized paradigm is the invariant substring shared by all word forms in that paradigm'. For example, the past tense inflectional form *rang* has the distinguisher *ang* while the theme of the paradigm *ring* is the substring *r* (Stump & Finkel, 2013, pp.382-389). In other words, the invariable portion of an inflection form has to be the theme while the variable portion is the distinguisher.

words, the identity of the indexed stem may differ depending on the representation provided for the plats. Stump and Finkel (2013, p.52) provide abstract (speaker) and concrete (hearer) plats for French verb morphology. In the hearer-oriented plat the underlying form of the indexed stem is subject to sandhi rules.

According to Stump and Finkel (2013, p.184), traditional descriptions of French grammar have distinguished seventy-two distinct ICs based (in part) on orthographic differences. For example, orthographically the verbs: *manger* 'eat', *placer* 'place' and *aimer* 'like' belong to different ICs because, unlike *placer* and *aimer*, *manger* has an *e* in the orthography in certain contexts while *placer* is distinguished by requiring a cedilla. Despite orthographic differences, all these verbs have the same IC membership based on their phonemic transcriptions, e.g. for the 1PL.PRS forms /mãʒõ/, /plasõ/, /ɛmõ/ respectively. There are forty-nine cells in the synthetic paradigm of French verbs, the forms of which are based on the bare stem or the whole word. French verbs inflect for the MPs of indicative (present, imperfect, simple past, future), conditional (subjunctive: present and past), imperative, infinitive, present participle and past) participle (masculine and feminine). The verb paradigm also includes three persons and two numbers (singular and plural). Therefore, the total number of cells in an ordinary verbal paradigm is up to forty-nine. Each verb form is built on either the stem only or a stem attached to a suffix.

Stump and Finkel (2013, p.189) apply both analyses to the verb paradigm in French, providing speaker-oriented and hearer-oriented plats for the verb paradigm. However, the plats show that the exponents are redundant phonological information and distinctions in conjugation are expressed by stems whose paradigm distribution can be determined by their indices. The two different types of plats have different representations and different numbers of stems. The speaker-oriented (abstract) plat has 19 stem slots whereas the hearer-oriented (concrete) plat has an extra stem slot (e.g. *aimer* ‘like’ in the concrete representation has stem10 *ɛmə* and stem10a *ɛməʁ*, whereas in the abstract representation the verb only has stem10 *ɛmə*). In addition, the stems are arbitrarily distributed in the paradigm and the stem sets do not form a natural class, i.e. they are morphomic. For instance, stem6 is morphomic since it appears in the SG and 3PL forms. In addition, the indexed stems in French have dependency relations and they can be predicted by the theme of the verb and from their IC membership. Therefore, not all the stems need to be listed in the lexicon. In fact, Stump and Finkel argue that ‘a French verb’s principal parts are, in all cases, a proper subset of its inventory of indexed stems’ (p.192). Therefore, the indexed stems of a verb can be deducible from the theme along with the principle parts of a verb. For instance, the verb *aimer* lists theme /ɛm/, stem 8: /ɛmɛ/ and stem 10: /ɛmə/ in the lexicon since it is these which determine IC membership (see Appendix A for the French system abstract and concrete plats).

These two accounts of principal parts representation can also determine the number of ICs in a system, predicted by the stem theme that can vary depending on which representation is assumed. The concrete analysis of a verb paradigm results in sixty-four distinct themes which also represent the number of ICs. Under the abstract approach, sandhi rules reduce that number to 57 classes based on differences which are truly morphological. For instance, in the concrete plat, the verbs *aimer* ‘like’ and *dejeuner* ‘had’ belong to the same IC whose verb list is increased by the application of sandhi rules (a-j) to also include the verbs: *coller* ‘paste’, *lever* ‘raise’ and *ceder* ‘give up’ (Stump & Finkel, 2013, p.194) (where vowels in upper case represent alternating vowel morphophonemes, ‘C’ stands for any consonant and # a word boundary):

Sandhi rules	Examples	Gloss
a) │EC│# → /ɛC/#	│sEd│ → /sɛd/	she/he gives up
b) │EC$_1$ əC$_2$ V│ → /ɛC$_1$əC$_2$V/	│sEdəʁɔ̃│ → /sɛdəʁɔ̃/	we will give up
c) │EC$_1$əC$_2$C$_3$V│→/ɛC$_1$əC$_2$C$_3$V/	│sEdəjɔ̃│→/sɛdəjɔ̃/	we would give up
d) │E│→ /e/	│sEdɔ̃│→ /sedɔ̃/	we give up
e) │ƆC│#→/ɔC/ #	│kƆl│→/kɔl/	she/he pastes
f) │Ɔ│→/o/	│kƆlɔ̃│→/kolɔ̃/	we paste
g) │ƏC│#→ /ɛC/#	│lƏv│→ /lɛv/	she/he raises
h) │ƏC$_1$əC$_2$V│#→ /ɛC$_1$əC$_2$V/	│lƏvəʁɔ̃│→ / lɛvəʁɔ̃/	we will raise
i) │ƏC$_1$əC$_2$C$_3$V│#→ /ɛC$_1$əC$_2$C$_3$V/	│lƏvəʁjɔ̃│→ / lɛvəʁjɔ̃/	we would raise
j) │Ə│→ /ə/	│lƏvɔ̃│→ / ləvɔ̃ /	we raise

Table 2.23 Some of the sandhi rules assumed in the analysis of French ICs

The rules reduce the redundancies among the ICs and ‘allow conjugations that are superficially different to be identified at a more abstract level of phonological representation’ (Stump & Finkel, 2013, p.192). The different representation affects the results of principal parts analysis and the number of ICs. The contrasting principal parts results are the effect of the different representation. In the concrete representation, French PPA listed five static principal parts while the abstract plat is associated with six principal parts. In addition, the different accounts may require different static and dynamic principal parts. The number of static principal parts required by French verbs varies according to whether verbs have abstract or concrete representation. The seventy-two verbs used under the concrete representation require five static principal parts whereas under the abstract representation six principal parts are needed. Moreover, both types of plats share the optimal static principal parts: stem 8 (that of the first-person singular simple past), stem 10 (that of the first-person singular future) and stem 19 (that of the feminine form of the past participle). However, the abstract plat is different in also requiring another essential stem, namely, stem 17 (that of present participle) which, under the concrete analysis, is not even included in the five alternative principal parts. In fact, this distinction between the different plats is mainly due to two conjugations: *fuir* 'flee' and *rire* 'laugh'. In the hearer-oriented plat, the distinguishers of these two verbs are different in a number of stems whereas their inflection in the abstract (speaker-oriented) plat only contrasts in a single stem, stem 17 (abstract stem 17 *fuir* 'present participle form' /fɥijɑ̃/ is disyllabic but *rire* 'present participle form' /ʁjɑ̃/ is monosyllabic) (pp.202-203). Under the dynamic principal parts analysis, the two accounts do not seem to result in any stem distinction between conjugations. According to Stump and Finkel (2013), the computed dynamic principal parts for both types of plats are surprisingly alike. For example, *courir* 'run' requires two dynamic principal parts to predict the whole verb paradigm whereas *coller* 'paste' requires only one. Clearly, the plat representation can affect the results of the principal parts analysis. The following section will discuss IC complexity in relation to principal parts.

2.4.2.4 IC typology

Stump and Finkel (2013, pp.327-337) employ principal parts as a measurement tool for establishing typological comparison between different systems of ICs. They propose ten measurable correlates[31] for representing dimensions of variations in IC complexity

---

31

Measure 1. The more distillations an IC system has, the more complex it is.
Measure 2. The large the size of an IC system's optimal static principal-part sets, the more complex it is.
Measure 3. The lower the density of an IC system's optimal static principal-part set, the more complex it is.
Measure 4. The larger the size of an IC system's optimal dynamic principal-part sets, the more complex it is.
Measure 5. The smaller the average ratio of actual to possible optimal dynamic principal-part analyses for an IC system, the more complex it is.

in ten different languages computationally analysed.[32] Two of the complexity correlate are useful measures of IC transparency which is the main focus of this study and further details will be provided in Chapter 5, section 5.1. Some of the complexity measures are based on three criteria for representing dimensions of variations in IC complexity. These criteria include the following (Stump & Finkel, 2013, p.54, 57, 62):

(8)

A. How many principal parts are needed to determine a lexeme's IC membership?
B. How many dynamic principal parts are needed to determine a given cell in a lexeme's realized paradigm?
C. To what extent are particular realized cells favoured as optimal static principal part?

Stump and Finkel (2013) point out that the number of optimal principal parts required to provide inferences about a realized paradigm can vary depending on the principal part scheme assumed and the inflectional system. There are languages that have zero principal parts due to the lack of distinct conjugation patterns, as seen in Turkish verb inflection. Some paradigm systems require a small number of principal parts while others need a large number. For instance, the verb inflection system in Kwerba has ICs defined through subject agreement for first or second person plural.

Measure 6. The higher an IC system's cell predictor number (average across ICs), the more complex it is.
Measure 7. The lower an IC system's average cell predictiveness, the more complex it is.
Measure 8. The lower an IC system's average IC predictability, the more complex it is.
Measure 9. The lower an IC system's average cell predictability, the more complex it is.
Measure 10. The higher an IC system's average n-MPS entropy, the more complex it is.

Unlike the traditional approach to principal parts, the complexity measures 2 to 10 provide different dimensions for measuring IC complexity without necessarily following the traditional requirements of uniqueness, uniformity and/or optimality of principal parts (§2.4.2). These properties can limit the use of principal parts in identifying contrasting ICs within the same inflectional system and across different systems. In other words, 'there is no single measure that 'captures' an IC system's complexity'. All measures correlated with inflectional complexity and different systems can be complexity on the basis of various dimensions (Stump & Finkel, 2013, p.339).

[32] The ten languages are Comaltepec Chinantec, Fur, Icelandic, Koasati, Kwerba, Latin, Ngiti, Sanskrit and Tulu (Stump & Finkel, 2013).

The subject agreement forms define the only principal part for distinguishing the four conjugations in this system. By contrast, in the verb inflectional system in Koasati, the inventory of principal parts varies depending on the scheme. For the static analysis, the system requires two optimal principal parts while in the dynamic scheme, only one dynamic principal part is needed. Thus, distinguishing the majority of conjugations requires the listing of 2SG which only fails for conjugation $3A_{.KA}$, $3A_{k.I}$ and $3A_{.KO}$. These ICs can only be indicated by 1SG. In a dynamic principal part scheme, these ICs can be individually determined by their unique distinguishing cells. In other words, the size of principal parts required by a system for deducing the realized paradigm of a lexeme reflects the IC complexity (see Appendix A for Koasati ICs and affirmative agreement morphology). Therefore, Stump and Finkel (2013, p.55) consider, '... the complexity of an IC system as the extent to which the system inhibits motivated inferences about a lexeme's full paradigm of realized cells from subsets of its cells…'. They use the static and dynamic principal parts number as a measure for the complexity of the conjugations (see Chapter 5 §3.3.1 and §3.3.2 for the discussion on TLA IC complexity).

The second criterion distinguishes two canonical ranges of IC systems (Stump & Finkel, 2013, pp.58-60). The first is a *thin* system in which each cell of the realized paradigm of a lexeme can be distinguished by employing one and only one dynamic principal part. In the other canonical range, there is the *thick* IC system in which deducing the cells of the paradigm is done by referencing every available principal part at the same time. Different languages have different positions in the continuum between these two canonical extremes. As an example of this criterion, Stump and Finkel discuss the IC system of Fur. The system defines nineteen conjugation classes (labelled I1A, I1B, I1C, I2A, …, IVD) and requires nine distillations[33] as follows (Appendix A):

(9)

	The nine distillations in Fur verb inflection
1	{non3sbjv}
2	{non3 perf}
3	{non3prst}
4	{3SG sbjv}
5	{3SG perf}
6	{3SG prst}
7	{3PL nonhuman sbjv}
8	{3PL nonhuman perf}
9	{3PL nonhuman prst}

The ICs in this system vary in terms of the number of optimal dynamic principal parts required. For some conjugations, deducing the lexeme's realized paradigm requires the listing of one, two or three distillations as the optimal principal-part set. For example, conjugation I1A, conjugation I1B and conjugation IIID have the schematic principal-part sets {3}, {1, 3} and {1, 5, 7} respectively). All of these ICs make reference to only a single principal part at a time, thus they are classified as thin ICs. By contrast, conjugation IIIE has the principal-part set {2, 5}, but it is more similar to a thick IC since determining any given cell in the paradigm of this conjugation is achieved by having both of its principal parts referenced simultaneously.

Although inflection systems can vary in terms of the number of dynamic principal parts which they require, their average *cell predictor number* tends to be lower than the dynamic principal-part number. The cell predictor number is defined as follows: 'Where P is the realized paradigm of a lexeme belonging to IC J, J's cell predictor number is the number of dynamic principal parts required to determine a cell in P,

[33] See §2.4.2.1 for the definition of the notion *distillation*.

averaged across the distillations in P.' By contrast, the dynamic principal-part number of a lexeme represents the lexeme's 'number of dynamic principal parts on any optimal analysis' (Stump & Finkel, 2013, pp.381-382). The cell predictor number is one way for measuring IC complexity without having to be bound to the principal-part traditional requirement of uniqueness and uniformity (§2.4.2). However, in this case, the IC complexity is based on 'the role of individual cells in the networks of implicative relations' rather than the whole paradigm (p.332). Generally, less information is required to determining a single cell in contrast to the required inferences for deducing a full realized paradigm (§2.4.2). Therefore, the dynamic principal-part number can be either equal to or higher than the cell predictor number.

The calculations of the cell predictor number are subject to two factors. First, deducing the cells that participate in the optimal principal parts analysis, which is achieved by only a single principal part. Second, a realized cell can be determined by different numbers of optimal principal parts. For instance, in Fur (Appendix A), conjugation IVC has the two possible optimal principal parts: the first one includes {1, 5, 7} with predictor number 1.00 while the second one has {4, 5, 7} with a predictor number of 1.11. Examining the ten languages under these criteria showed that, although some languages show more qualities of a thick conjugation system than others do, the average cell predictor number tends to be generally low in most of these languages. For example, Kwerba, Koasati and Tulu have thin conjugation system while Icelandic is more similar to a thick system. Nevertheless, the results showed similar scores for the cell predictor number. Therefore, the complexity of an inflectional system can be determined by the cell predictor number of an IC.

The cell predictor number can also be used as a referencing tool for determining and measuring the IC complexity of a system. In other words, in a lexeme's realized paradigm, the more principal parts a cell needs in order to be deduced, the lower the possibility of deducing the paradigm as a whole.

The third criterion is based on investigating the pattern of cell choice in the optimal static principal parts analysis. In some IC systems, most of the possible optimal principal parts under the static system tend to favour a certain set of MPs. Stump and Finkel (2013) point out that Sanskrit has forty-five possible analyses of optimal static principal parts which show a generalization for the best principal part candidates. In each possible analysis, they provide three principal parts chosen from the distillations in a, b, and c respectively (pp.70-72):

(10)

a.	b.	c.
1. {1sg prst ind active}	4. {1sg prst ind middle}	13. {1sg aorist ind active}
2. {1du prst ind active}	5. {2sg prst ind middle}	14. {2sg aorist ind active}
3. {3pl prst ind active}	12.{1sg perf ind middle}	15. {1du aorist ind active}
6. {1sg imperf ind active}		
7. {3pl imper active}		

These criteria provide three dimensions for classifying variations in different principal-part systems.

2.4.2.5 IC transparency

Stump and Finkel (2013, pp.81-83) distinguish two types of canonically extreme ICs. At one canonical extreme, ‘A maximally transparent IC is an IC in whose realized paradigms each cell’s exponence determines that of every other cell; given any two cells K1, K2 in a realized paradigm belonging to an IC of this sort, there is an R-relation deducing K1 from K2.’ By contrast, ‘in a maximally opaque IC, no realized paradigm has any cell or combination of cells whose exponence determines the exponence of any other cell; in a realized paradigm belonging an IC of this sort, there are no R-relations deducing any cell from any other cell or combination of cells.'[34]

They examine the IC transparency of the Fur (see Appendix A) system to provide examples of different conditioning measures of the transparency dimension of IC complexity, since 'IC is proportional to the overall transparency of its individual ICs'. The inflectional system in Fur distinguishes nineteen conjugations based on four inflectional patterns. Stump and Finkel (2013) explain that an inflectional form in the verb paradigm of this system inflects for person (non-third, third), number (singular, plural). There is also the MP human/no human reflected in the plural of third person. All of these distinctions are realized as a subject prefix. The Fur verb can also inflect for tense/aspect/mood (TAM) including present/perfect and subjunctive realized by a suffix. In addition, the root of the verb form exhibits tone marking. A subset of the verb paradigm, which includes the plural forms of third person also has a marker for number. Therefore, the result of these inflectional properties is twelve different MPs. The overall system has nine distillations with 10, 11, 12 properties being inferred from properties 7, 8 and 9 respectively (9). The IC transparency of Fur can be determined by examining the optimal principal parts analysis under the dynamic scheme. The results show that the conjugations of this system deviate from the canonical extreme of maximal transparency to varying degrees.

Stump and Finkel (2013, p.83) provide the four criteria in (11) for distinguishing ICs within the same inflectional system. Given these criteria, IC A is more transparent than IC B:

(11)

[34] The more similar the ICs are the more principal parts are listed for a system since the similarity of ICs makes it more difficult to distinguish each IC from the other.

A. if the number of dynamic principal parts needed to distinguish lexemes belonging to A from lexemes belonging to other ICs is smaller than the number needed to distinguish lexemes belonging to B from other lexemes;

B. if the number of principal parts needed to deduce a given cell in the realized paradigms of A's member lexemes is on average, lower than the number of principal parts needed to deduce a given cell in the realized paradigms of B's members;

C. if there are more alternative optimal principal-part analyses for the realized paradigms of A's members than for those of B's members; and

D. if the realized paradigms of A's members have fewer cells whose realization cannot be predicted than the realized paradigms of B's members.

Applying the four criteria of IC transparency to Fur conjugations reveals a number of features of the Fur system. First, conjugation I1A exhibits some degree of transparency since it is possible to deduce the whole paradigm of this IC by any cells realizing present tense (Dist 6, 9 and 3), since they have unique exponents distinguishing this IC from the other conjugations. However, none of the other cells is distinctive for this conjugation. If the non-past tense cell were also unique, then the conjugation I1A would have had a higher degree of IC transparency. Unlike conjugation I1A, the rest of ICs show deviations from Fur IC transparency. Conjugation IVB requires two principal parts (Dist 4 and 7) in the optimal principal parts analysis. However, this conjugation remains similar to conjugation I1A, since in both of them only a single distillation is referenced at a time. By contrast, conjugation IIIE requires the two principal parts to participate simultaneously in deducing certain cells of the paradigm (Dist 2 and 5). In fact, conjugation IIIE also shows another pattern of deviation as it has a limited principal parts analysis, listing only a single possibility. Thus, deducing the cells of the IIIA paradigm can only be achieved by distillation 4 and no other alternatives (Appendix A).

The overall conjugation patterns in Fur can vary in terms of their transparency. According to criterion A, conjugation I1A with one principal part is more transparent than conjugation IIID with three principal parts. In addition, the application of criterion B reveals that conjugation IIIC with a cell predictor number of 1.33 has less transparency than I1A with a cell predictor number of 1.00 (Appendix A). While conjugation IIIC requires two principal parts to participate simultaneously in deducing certain cells of the paradigm (Dist 1 and 4), conjugation I1A requires only one for each cell (Dist 3). Therefore, conjugation IIIC cell predictor number is the number of dynamic principal parts required to determine a cell, averaged across the nine distillations represented in (9) (Table 2.24).

Distillations									Cell predictor number
1	2	3	4	5	6	7	8	9	
①	①④	④	④	①④	④	①④	④	④	1.33

Table 2.24 The optimal dynamic principal-part analyses for conjugation IIIC

Clearly, the concept of IC transparency can provide useful insights into the complexity and variations of ICs within the same system.

2.4.2.6 Measuring complexity

Stump and Finkel (2013, p.81) propose another objective correlation with IC complexity appealing to the IC predictability and cell predictability measures. In order to understand and evaluate the complexity of ICs of a particular system compared with other systems, they develop a metric for principal parts. Although the morphological complexity of different languages systems may not be always representable by mere numerical values, a quantitative measure remains desirable as it can provide a comparable diachronic and synchronic account of the degree of complexity within the same language and/or between different language systems. However, it is important to note that these methods can be sensitive to the nature of the data that might have to undergo certain manipulations, which can affect the results, as we will see in the case of TLA plat representation in Chapter 5. Therefore, for the purpose of this study, the quantitative method is introduced as a complement to our theoretical linguistic analysis of the morphological complexity of the TLA verb paradigms and their implicative relations. The subjective method reflects the IC complexity by two measures: predictiveness and predictability. According to Stump and Finkel (2013, pp.383-387) 'The predictiveness of a cell K in a realized paradigm is the fraction of the other cells in the paradigm that are fully determined by K.' By contrast, 'the cell predictability of a cell <*w*, σ > in a lexeme L's realized paradigm P_{L} is the ratio of (a) to (b), where (a) is the number of nonempty subsets of P_{L}'s cells whose realization uniquely determines <w, σ> and (b) is the number of all nonempty subsets of P_{L}'s cells.'[35] The IC transparency can be measured by cell predictability and/or IC predictability. The IC predictability is 'the predictability of a lexeme L's IC is the fraction of adequate (though not necessarily optimal) dynamic principal-part sets among all nonempty subsets of cells in L's realized paradigm' (Stump & Finkel, 2013, p.383). The output of the two measures is set on a fixed scale from 0:1 with a zero score reflecting complete unpredictability and/or unpredictiveness while a score of 1 denotes full predictability and/or predictiveness.[36] For measuring complexity, Stump and Finkel use the Principal Parts Analyser (PPA, Cat's Claw software). Consider the following hypothetical system in Table 2.25 (Finkel & Stump, 2007, p.42).

---

[35] The predictiveness measure will not be considered in this study.

[36] The tool is available on this link http://www.cs.uky.edu/~raphael/linguistics/analyze.html

	W	X	Y
I	a	e	i
II	b	e	i
III	c	f	j
IV	c	g	j
V	d	h	k
VI	d	h	l

Table 2.25 A hypothetical inflection class system

The hypothetical system has six ICs with three unique MPs (WXY) reflected by twelve distinct exponents. The static scheme of principal parts system requires all three MPs of each IC to be specified since there are distinctive exponents in each MP. On the other hand, for the dynamic principal parts analysis, each IC of the hypothetical system requires one optimal principal part, highlighted in Table 2.25. The principal parts analysis is reflected in the results of the IC predictability of such a system (Table 2.26). For instance, the system scores zero cell predictability in each IC represented by the lexically specified cells which are employed in the dynamic schema, as highlighted below.

Predictabilities

	W	X	Y	| Avg	Inflection-class predictability
I	0.000	0.750	0.750	| 0.500	0.571 (4 out of 7)
II	0.000	0.750	0.750	| 0.500	0.571 (4 out of 7)
III	0.750	0.000	0.750	| 0.500	0.571 (4 out of 7)
IV	0.750	0.000	0.750	| 0.500	0.571 (4 out of 7)
V	0.750	0.750	0.000	| 0.500	0.571 (4 out of 7)
VI	0.750	0.750	0.000	| 0.500	0.571 (4 out of 7)
Avg	0.500	0.500	0.500	| 0.500	0.571

Table 2.26 Cell predictability and paradigm predictability for the hypothetical IC system in (Table 2.25).[37]

[37] https://www.cs.uky.edu/~raphael/linguistics/principalPartsFull.cgi

The overall results of the cell and paradigm predictability show that the paradigm of this hypothetical system is relatively complex. From a pedagogical point of view, each IC of the hypothetical system has a unique distinctive exponence pattern which is lexically specified. Thus, the principal parts metric considers every deviation as a separate IC. In class I (Table 2.25), for example, the MP X is highly predictable (scoring 0.750) and predictive of MP Y. In contrast, although the MP W can predict the rest of the paradigm, it is completely unpredictable. Therefore, the forms of the W MP in this class are selected as a principal part. In addition, this system seems to lack a single default class since each IC has the same pattern of unpredictability and any novel lexeme in the system can be a member of any of the ICs I-VI (Table 2.26).

The learnability of this system under this approach is based on the theoretical notion of principal parts that canonically is unpredictable but predictive of other paradigm forms, rather than on the default form which canonically is predictable from other forms, but unpredictive (Brown & Evans, 2010, p.237). For the principal parts approach, paradigm complexity is *inversely proportional* to the predictiveness and predictability of a paradigm's cell (Stump & Finkel, 2013). The deduction of the whole paradigm is achieved by inferences from a minimal and adequate set of cells whereas the rest of paradigm cells, which can still predict some parts of the paradigm, may not be considered in the computation (See §5.3.3.1 and §5.3.3.2 for the discussion on TLA IC complexity).

2.5 Summary

This chapter has discussed two different approaches to Arabic verb morphology. We showed that the stem-based account is preferable to one which relies on (possibly controversial) assumptions about morphology and which assigns a special and unique status to Semitic morphology. The examination of the morphomic stem and its association with extramorphological factors have showed that the status of stems can range from autonomous to semi-autonomous. This chapter has also introduced two stem-based frameworks for the paradigmatic implicative relations and ICs.

The next chapter will introduce the verb system in TLA and discuss some of previous work on verb morphology in some LA varieties.

Chapter 3 The TLA Verb

One of the main aims of this study is to provide a systematic analysis of the verb morphology of Transitional Libyan Arabic (TLA) within a stem-based account, decomposing inflectional forms into stems and affixes realizing MPs. In TLA the affixal exponents of the verb inflectional MPs (imperfective/perfective aspect, person/number/gender agreement) show little or no allomorphy, even in irregular verbs. Thus, there are no ICs defined by affix homophony. On the other hand, the dialect is well suited for this study since the verbs show morphomic stem alternations within inflectional paradigms. We will see that the patterning of these stems within the paradigm would be difficult to capture under the root-and-pattern analysis. In addition, the inflected forms in TLA show implicative relations, hence, a verb paradigm can be based on 'global' implicative relations providing inferences of all the forms (Chapter 5). Previous analyses of LA verb morphology (Owens, 1984 on ELA, Harrama, 1993 on JLA, Elfitoury, 1976 on Tripoli LA), regardless of their theoretical approach, provide useful descriptive sources.

This chapter presents a general overview of the TLA verb system based on the traditional classification that will be reanalysed using a stem-based approach (Chapter 4). In section 3.1, the verb measures and verb series will be presented. This also provides a general description of the morphosyntactic, morphological and relevant phonological properties of TLA. In section 3.2 I will present the previous work on verb morphology in two varieties of LA including ELA by Owens (1984) and JLA by Harrama (1993). Both of these analyses assume a morphemic analysis, but they provide a very useful descriptive starting point.

3.1 TLA measures and verb series

As in many traditional descriptions of Semitic languages we can divide TLA verbs into two classes depending on whether they are formed from a triliteral (three-consonant) or a quadriliteral (four-consonant) root.[1] TLA verbs can also be categorised along a second dimension, the set of verb measures. The verb measures are traditionally defined as a systematically related set of derived verb forms, expressing distinctive grammatical features (causative reflexive, passive among others). Measure one (*M*1) is considered to be the underived verb form, and the CV templates for the other measures are derived from this by consonantal root alternations, vocalic alternation and affixation (McCarthy, 1979, 1981, 1986) (see Appendix A). I follow Hoberman and Aronoff (2003, p.63) in assuming that each verb measure 'consists of an inner stem with fixed vocalism and a fixed prosody, surrounded by fixed affixes (if they are present), and in some cases different prosodies and vocalisms in different tenses, aspects, or moods. Consonantal roots are fitted to these complex patterns.' Each measure functions as 'an inflectional class, it dictates the phonological form of the verb' (Aronoff, 1994, p.127).[2] It is generally assumed that Libyan Arabic (LA) retains eight measures, *M*1 and non-*M*1 (*M*2, *M*3, *M*4, *M*5, *M*6, *M*7, *M*8) of the ten Standard Arabic (SA) verb measures for the triliteral verbs (Table 3.1).[3] In addition, TLA verbs can belong to one of four different phonologically defined verb series, the sound verbs, double verbs, hollow verbs and the defective verbs (Table 3.2).

[1] In this study, we will only consider the triliteral verbs that are relevant to the unique stem morphomicity that can be found in TLA (§4.5 and §5.2).

[2] Note that Aronoff is using the term 'inflectional class' in a very specific sense here, which is rather different from the sense adopted in this thesis, and in most of the literature.

[3] Measure four (M4) (e.g. *ʔaxraj* 'expel') and measure nine (M9) (e.g. *ihmarra* 'become red') of SA are not attested in TLA.

Measure	example	template	Ggloss
1	leʕab	CVCVC	‘play’
2	ferraḥ	CVC$_i$C$_i$VC	‘make sb. happy’
3	šaarik	CV$_i$V$_i$CVC	‘share with sb’
5	tekallim	t(e)CV C$_i$C$_i$VC	‘talk with’
6	teʕaarik	t(e)CV$_i$V$_i$CVC	‘argue’
7	inkesar	(i)nCVCVC	‘be broken’
8	irtefaʕ	(i)CtVCVC	‘be taken’
10	staʕmil	st(a)CCVC	‘use’

Table 3.1 TLA verb measure patterns.

Measure	Regular (strong)	Irregular (weak)		
	Sound	Double	Hollow	Defective
1	CVCVC ketab ‘write’	CVC$_i$C$_i$ ḥall ‘open’	CVVC ġaab ‘be absent’	CCVV jree ‘run’
2	CVC$_i$C$_i$VC kallim ‘talk’	CVC$_i$C$_i$VC$_i$ ḥabbib ‘make sb. love’		CVC$_i$C$_i$V garra ‘teach’
3	CVVCVC naašib ‘harass’			CVVCV naada ‘invite’
5	tCVC$_i$C$_i$VC teʕallim ‘learn’	tCVC$_i$C$_i$VC$_i$ teġaššiš ‘feel upset’		tCVC$_i$C$_i$V tewaffa ‘die’
6	tCVVCVC teʕaarik ‘argue’			tCVVCV telaaga ‘meet’
7	inCVCVC inkeṣar ‘be broken’	inCVC$_i$C$_i$ ingall ‘be taken’	inCVVC inbaaʕ ‘be sold’	inCCVV indʕee ‘be infected’
8	iCtVCVC irtefaʕ ‘be taken’	iCtVC$_i$C$_i$ irtadd ‘be returned’	iCtVVC ixtaar ‘choose’	iCtCVV intsee ‘be forgotten’
10	stCVCVC staʕmil ‘use’	stCVC$_i$C$_i$ steġall ‘make use of’	stCVVC stefaad ‘benefit’	stVCCV stanja ‘appeal’

Table 3.2 The possible verb series for each measure.

*M*1 can be classified as the simple (active or basic voice) verb form from which other verb patterns can be derived, but not necessarily all. For example, the *M*1 verb (CVCVC) *šerab* ‘drink’ has a corresponding (causative) *M*2 verb (CVC$_i$C$_i$VC) *šarrib* 'make someone drink'. By contrast, the verb *rabbij* 'sing traditional songs', which belongs to *M*2, has no *M*1 counterpart **rebaj*. The main characteristic of *M*2 is doubling represented by a geminate second consonant and a *M*2 verb can belong to any verb series except the hollow series. In general, the *M*2 verbs are transitive, usually representing the intensive or causative form of *M*1. In TLA, *M*3 (CVVCVC) seems to be limited to a small number of transitive verbs (Harrama, 1993) and the semantic component of this measure represents associative action. This measure includes only sound or defective verbs. *M*5 verbs (tCV$C_i$$C_i$VC) can be derived from *M*2 verbs by attaching a prefix (t-), conveying a passive or sometimes a reflexive meaning. *M*6 (tCVVCVC) is based on *M*3 with the addition of a prefix (t-), signifying a reciprocal meaning. Both *M*7 (nCVCVC) and *M*8 (CtVCVC) can function as medio-passive (inchoative) verbs.[4] They have the same pattern as *M*1 with the addition of the prefix (n-) and the infix (-t-) respectively. The derivation of *M*10 (st(e)CVCVC) involves prefixing (st-) to the root of the verb (usually a *M*1 verb) (Elfitoury, 1976, pp.30-35).[5]

LA sound series verbs fall into two subtypes: strong verbs and assimilated verbs (Elfitoury, 1976; Harrama, 1993). The strong subtype has three distinct consonants in the root, excluding the consonants w, *y* or *ʔ*. For instance, the *M*1 verb *leʕab* 'play' and *M*2 verb *laʕʕib* 'cause to play' have the root *lʕb*; hence, they both belong to the sound series but with different verb measure templates (CVCVC and CV$C_i$$C_i$VC respectively). In traditional grammar, the SA *M*1 sound series is described as realizing

[4] The passive construction in TLA can be realized by both *M*7 and/or *M*8. Note, however, that the *M*7 paradigm of the verb *inbetah* 'lie down', for example, can have active forms alongside passive forms. In other words, the *M*7 forms do not realize a voice opposition.

[5] The present study is mainly concerned with the inflectional verb morphology and the relationship between different stem patterns that can be found across verb measures and/or series. Therefore, for further discussion on the Arabic verb measures, the reader is referred to McCarthy (1979, 1994) who provides a root and pattern based derivational analysis of these measures. A different account is proposed by Aronoff (1994) who argues that these measures are inflection classes rather than derivational ones. This raises the interesting question of how a canonical morphology based-account of inflection and derivational processes would categorize these measures.

three different patterns of vocalism, usually associated with change in meaning (see §2.1 Table 2.1). By contrast, the TLA sound series lacks this three–way variation. For instance, in SA, *M*1 sound series verbs *katab* ‘write’, *šarib* ‘drink’, and *kabur* ‘grow up’ have the templates CaCaC, CaCiC and CaCuC respectively whereas in TLA these verbs have the same pattern, CeCaC: *ketab, šerab* and *kebar*. However, it is worth noting that although the TLA sound verbs share the pattern CeCaC, they show different patterns of stem alternation in different cells of the verb paradigm (see Table 4.15, §4.3.1). The second subtype of verb has assimilated roots in which the first consonant is *y* or *w* (e.g. *M*1 *wegaf* ‘stand up’, *M*2 *yabbis* ‘cause to dry’, and *M*3 *waaṣil* ‘keep in touch’). Although the different subtypes of the sound series can be based on different consonantal roots, they all show the same paradigmatic complexity in terms of stem allomorphy within each verb measure. Therefore, the two subtypes will be treated as one verb series in the rest of this work (§4.3.1, §4.5, §5.2 and §5.3).

The second verb series in TLA is the double verbs, traditionally referred to as *geminate* verbs. This verb type is characterized by the presence of geminate consonants in (intervocalic) word-medial or word-final position. The crucial feature of this template type is that, unlike *M*1 and non-M1 sound series, it has only two distinct radicals in the consonantal root (e.g. from the consonantal root *md,* M1 ‘*madd* ‘pass sth.’, *M*2 *maddid* ‘extend’, M5 *temaddid* ‘expand’). In addition, TLA double verbs show two stems distinguished by the presence or absence of the monophthong *ee* (CCVC ~ CVCCee) before vowel initial or consonant initial suffixes respectively (§4.3.3,4.5.3). By contrast, sound verbs, with three distinct consonants in the root, do not show the monophthong-based alternations (4.3.3 and Table 4.24). In traditional grammar, the classification of the consonantal root of this series has not been consistent, treating it sometimes as a triliteral (CVC_iC_i) (e.g. for *radd* ‘return’ *rdd*) or a bililteral verb (CC) (e.g. *rd*). The autosegmental account has assumed they are biliteral, but with the same template as that of the triliteral verbs with three C slots. As we saw in Chapter 2, section 2.1 in McCarthy's analysis (1981), mapping the root of a double verb such as /rd/ to the template CVCVC derives /radad/ by the spreading rule, mapping the second consonant to both C1 and C3 positions. Under this view, the CVCVC stem form is the canonical one from which the contracted form CVCC (e.g. *radd*) is derived by a lexical rule of syncope, which assumes that C2 and C3 are derived from the same consonant in the root (recall that this device is required to avoid an OCP violation). A different account is proposed by Gafos (2003), who provides a stem-based analysis for double verbs, treating the CVC_iC_i lexically specified stems as subject to morphological conditioning. A double verb stem such as *madd* is used before a vowel-initial suffix while *madad* is selected before a consonant-initial suffix due to the restriction on the triconsonantal cluster. The stem allomorphy of TLA double verbs will be discussed in section 4.3.3.

The hollow and defective series, traditionally labelled 'weak', have the glides *w/y* in consonantal root position C2 and C3 respectively in the underlying representation. In TLA, the hollow series template is based on a long stem vowel whereas the defective series ends in a long vowel. The underlying glide in the root is realized as a long vowel in the surface form of *M*1 verbs. The traditional analysis accounts for the form of the weak series by invoking a restriction against words with intervocalic glides (*CVy/wVC). Although SA has a very few nouns with the intervocalic glide realized in the surface form (e.g. /θawrun/ 'ox'), no *M*1 verbs with the form (Cay/waC) are attested. Therefore, the weakness in the verbal root manifests itself as the appearance of a high vowel, /i/ or /u/, depending on the quality of the underlying glide, which surfaces in the deverbal nouns (madars) (e.g. SA *bayiʕ* 'selling'). However, in many Arabic dialects, there is evidence for the lexicalization of hollow verbs (see §4.3.4 and Table 4.25 for further discussion of TLA hollow verbs). Likewise, the defective series has been treated as a weak verb with two root radicals and a glide w/y which, in modern Arabic dialects, is manifested as a long vowel due to the restriction on diphthongs stem-finally (CVCVy/w ~ CVCVV). For example, the glide element in SA *nasiy-tu* 'I forgot' (template CaCiy-tu) is realized as a long vowel in Cairene Arabic and San'ani Arabic (*nasii-tu*). In TLA, the diphthong was possibly subjected to coalesce historically (see §4.3.1 for further details).

In root and pattern based approaches, the 3SGM perfective form serves as the base form (McCarthy, 1981; Guerssel & Lowenstamm, 1996) whereas stem-based models consider the 3SGM imperfective CCVC to be the default form (Ratcliffe, 1997; Benmamoun, 2003; Heath, 2003; Gafos, 2003). Although Darden presents Egyptian Arabic as a word-based system, he lists the 3SGM as the default form.

We follow the stem-based works in regarding the imperfective as the default form. However, this work does not completely dismiss the important role of the 3SGM perfective form in the network of implicative relations. With either 3SGM form, the results seem to be the same in terms of stem dependency relations (§5.2.2). Each of the verb measures can be inflected for the different MPs as will be shown in the following section.

3.1.1 Verb morphosyntactic properties

TLA verbs realize a number of morphosyntactic properties (MPs), some of them expressing grammatical meanings, others used in agreement processes. Verbs distinguish finite forms from non-finite forms (participles and masdars). The inventory of MPs for finite verb forms is summarized in Table 3.3.

MP	values
ASPECT (ASP)	perf(ective), imperf(ective)
MOOD	indic(ative), imper(ative)
TENSE (TNS)	fut(ure), non-fut(ure)
NUMBER (NUM)	SG, PL
PERSON (PER)	1, 2, 3
GENDER (GEN)	M, F

Table 3.3 TLA MPs.

The aspectual contrast is only found with [MOOD:indicative] forms. The future tense and imperative forms are only ever imperfective in aspect, and the forms of the future are derived by prefixation from the imperfective forms in a completely regular way, and therefore add nothing of interest to paradigm structure. For this reason, I will ignore the future tense forms. Finite, indicative mood, verbs agree with the subject in person (1st, 2nd, 3rd), number (SG/PL), and gender (M/F). In the imperative the verb only realizes 2nd person.

The full set of MPs is realized by affixation[6] and the same inflectional affixes for the same cells are used across different verb measures/series. TLA verb inflection involves suffixing and/or prefixing the verb stem. For instance, the inflection for the MPs in the perfective sub-paradigm is realized by suffixing the verb stem while verbs are inflected for imperfective by both prefixes and suffixes. In the imperative, agreement is realized by single suffix cumulating (2nd) person/gender. In *M*1 sound and defective series the imperfective and imperative forms also have a pre-formative vowel (FV). This is a vowel that precedes the verb stem and follows the imperfective prefix (e.g. *y*-***a***-*rkub* 'he rides', *y*-***i***-*kr*-*u* 'they hire') (see §4.1 for further details on the pre-formative vowel). TLA verbs have object pronoun suffixes in the perfective and

[6] Except for the 3SGM perfective form (Table 3.7).

imperfective paradigm.[7] Further details of the TLA affix inventory will be provided in section 4.1.3.

The active participles inflect for person and number only (Owens, 1984, p.106). They do not contribute to paradigm structure and we do not consider them here in any detail.

3.1.2 Verb morphological properties

In this study, we will be mainly concerned with the imperfective and perfective paradigms. For the most part, the pattern of stem alternation in the imperative sub-paradigm mirrors that found in the imperfective paradigm. The abstract paradigm schema consists of 24 cells as illustrated in Table 3.4 below. The schema represents all the permitted combinations of values. Each cell is generally filled by one unique inflected form, but in some cases we see syncretism, for instance, between the 1SG and the 2SGM cells in the perfective paradigm of all measures/series. This is illustrated by the following *M*1 forms: *ketab(i)t* 'I/you wrote' (sound series), *gult* 'I/you said' (hollow series), *maddeet* 'I/you passed' (double series), *nseet* 'I/you forgot' (defective series). Some irregular verbs have defective paradigms, that is, they systematically fail to fill certain cells. For example, the verb *y-ibi* 'he wants' lacks the perfective sub-paradigm, and this gap has to be filled periphrastically: *kan yi-bi* 'he had wanted'.

[7] All object suffixes except for the first person forms, can also be attached to nouns as possessive pronouns.

Mood	Indic		Imper
Asp	Perf	Imperf	
1SG			
1PL			
2SGM			
2SGF			
2PLM			
2PLF			
3SGM			
3SGF			
3PLM			
3PLF			

Table 3.4 Paradigm schema for TLA verbs.

The aspectual contrast is sometimes accompanied by stem changes.[8] For example, in *M*1 sound verbs, the vowel pattern in the perfective 1st/2nd and 3SGM is *e-a* whereas the 3SGF/PL can have *i-i* or *u-u*.[9] In the imperfective counterparts, the FV is lexically determined (§4.1). The identity of the stem vowels can be partially determined by phonological factors such as assimilation to the adjacent FV (*yu-rugd-u* 'they sleep', but *ya-ṣugṭ-u* 'they fail' *ya-simʕ-u* 'they hear') or being conditioned by the final stem consonant (*ʕ*, *ḥ* or *h*) (§4.5.1). Similarly, the other (non-sound) *M*1 verb series and some of the non-M1 verbs can also show intra-paradigmatic morphological changes in the form of stem ablaut and suffixation of the monophthongal ending *-ee*. An example of a full inflectional paradigm for each of the four *M*1 verb series and the *M*2 sound and defective series is presented in the following tables:

[8] In SA, by contrast, each measure can only have one vocalism within each sub-paradigm. For instance, *katab* 'he wrote' has the same vowel(s) for all the inflected forms in the perfective paradigm.

[9] E.g. *ketab* ~ *kitib* 'write', *resam* ~ *risim* 'draw', *ʕegal* ~ *ʕigil* 'become mature/hide', *gelaʕ* ~ *giliʕ* 'take off clothes/extract', *ʕegad* ~ *ʕigid* 'tie', *serag* ~ *sirig* 'steal', *sekab* ~ *sikib* 'pour', *seraḥ* ~ *siriḥ* 'wonder', *feham* ~ *fihim* 'understand', *felag* ~ *filig* 'burst', *fesad* ~ *fisid* 'rotten', *gesam* ~ *gisim* 'divide', *geʕad* ~ *guʕud* 'stay', *regad* ~ *rugud* 'sleep', *zerag* ~ *zurug* 'go quickly for a mission', *sekat* ~ *sukut* 'became silent', *sekan* ~ *sukun 'live'*, *ṣebaḥ* ~ *ṣubuḥ* 'be in a certain state in the morning', *feraḥ* ~ *furuḥ* 'be happy'.

verb series	sound	double	hollow	defective
MPs	ʕeraf 'know'	karr 'pull'	jaab 'bring'	gree 'read'
1SG	ni-ʕrif	n-kurr	n-jiib	na-gra
1PL	ni-ʕirf-u	n-kurr-u	n-jiib-u	na-gr-u
2SG.M	ti-ʕrif	t-kurr	t-jiib	ta-gra
2SG.F	ti-ʕirf-i	t-kurr-i	t-jiib-i	ta-gr-i
2PL.M	ti-ʕirf-u	t-kurr-u	t-jiib-u	ta-gr-u
2PL.F	ti-ʕirf-in	t-kurr-in	t-jiib-in	ta-gr-in
3SG.M	yi-ʕrif	y-kurr	y-jiib	ya-gra
3SG.F	ti-ʕrif	t-kurr	t-jiib	ta-gra
3PL.M	yi-ʕirf-u	y-kurr-u	y-jiib-u	ya-gr-u
3PL.F	yi-ʕirf-in	y-kurr-in	y-jiib-in	ya-gr-in

Table 3.5 An example of the imperfective paradigm of *M*1 sound, double, hollow and defective verbs.

verb series	sound	double	hollow	defective
MPs	ʕeraf 'know'	karr 'pull'	jaab 'bring'	gree 'read'
2SG.M	i-ʕrif	kurr	jiib	a-gra
2SG.F	i-ʕirf-i	kurr-i	jiib-i	a-gr-i
2PL.M	i-ʕirf-u	kurr-u	jiib-u	a-gr-u
2PL.F	i-ʕirf-in	kurr-in	jiib-in	a-gr-in

Table 3.6 An example of the imperative paradigm of *M*1 sound, double, hollow and defective verbs.

verb series	sound	double	hollow	defective
MPs	ʕeraf 'know'	karr 'pull'	jaab 'bring'	gree 'read'
1SG	ʕeraf-t	karree-t	jib-t	gree-t
1PL	ʕeraf-na	karree-na	jib-na	gree-na
2SG.M	ʕeraf-t	karree-t	jib-t	gree-t
2SG.F	ʕeraf-ti	karree-ti	jib-ti	gree-ti
2PL.M	ʕeraf-tu	karree-tu	jib-tu	gree-tu
2PL.F	ʕeraf-tin	karree-tin	jib-tin	gree-tin
3SG.M	ʕeraf	karr	jaab	gree
3SG.F	ʕuruf-it	karr-it	jaab-it	gr-it
3PL.M	ʕuruf-u	karr-u	jaab-u	gr-u
3PL.F	ʕuruf-in	karr-in	jaab-in	gr-in

Table 3.7 An example of the perfective paradigm of *M*1 sound, double, hollow and defective verbs.

MPs	dakkir 'remember'			bakka 'upset'[10]		
	Imperf	Imper	Perf	Imperf	Imper	Perf
1SG	n-dakkir		dakkir-t	n-bakki		bakkee-t
1PL	n-dakkr-u		dakkir-na	n-bakk-u		bakkee-na
2SG.M	t-dakkir	dakkir	dakkir-t	t-bakki	bakki	bakkee-t
2SG.F	t-dakkr-i	dakkr-i	dakkir-ti	t-bakk-i	bakk-i	bakkee-ti
2PL.M	t-dakkr-u	dakkr-u	dakkir-tu	t-bakk-u	bakk-u	bakkee-tu
2PL.F	t-dakkr-in	dakkr-in	dakkir-tin	t-bakk-in	bakk-in	bakkee-tin
3SG.M	y-dakkir		dakkir	y-bakki		bakka
3SG.F	t-dakkir		dakkr-it	t-bakki		bakk-it
3PL.M	y-dakkr-u		dakkr-u	y-bakk-u		bakk-u
3PL.F	y-dakkr-in		dakkr-in	y-bakk-in		bakk-in

Table 3.8 The inflectional paradigm of *M*2 sound series *dakkir* 'remember/remind' and defective series *bakka* 'upset'.

Starting with the *M*1 and *M*2 sound series, the perfective paradigms show that inflected forms of *M*1 show stem ablaut but this does not seem to be the characteristic of the same series in *M*2 (Table 3.7 and Table 3.8). In fact, many of the derived measures particularly of the sound and double series show (virtually) no stem ablaut in the perfective sub-paradigm, so that the vowels of the imperfective and perfective tend to be the same. However, the non-*M*1 defective series verbs show final vowel change based on ablaut and the monophthong *-ee* suffix (Table 3.8).

In the imperative, all verbs share the same base/stem with the second person form of verbs inflected for imperfective.[11] Imperative forms can be divided into two groups. One group is distinguished by the presence of the FV, possibly preceded by a glottal stop, while the second group has no FV. The first group includes *M*1 sound and defective verbs while the other measures/series (excluding *M*7 and *M*8) belong to the second group (Table 3.8).

[10] The verb *bakka* can also mean 'make sb. cry'.
[11] There are two exceptions, *klee* 'eat' and *xdee* 'take', both of which have suppletive forms *kuul* and *xuud* respectively.

MPs	Imperf	Imper
Group 1	rekab 'ride'	
2SG.M	t-a-rkub	a-rkub
2SG.F	t-a-rukb-i	a-rukb-i
2PL.M	t-a-rukb-u	a-rukb-u
2PL.F	t-a-rukb-in	a-rukb-in
Group 2	ṣaam 'fast'	
2SG.M	t-ṣiim	ṣiim
2SG.F	t-ṣiim-i	ṣiim-i
2PL.M	t-ṣiim-u	ṣiim-u
2PL.F	t-ṣiim-in	ṣiim-in

Table 3.9 Two types of imperative formation

We now turn to the phonological properties of the TLA verb system and the role of phonology in verb complexity.

3.1.3 Verb phonological properties

LA has twenty-eight consonants distinguished by place and manner of articulation (Table 3.10). Generally, the phonemes that exist in SA can also be found in TLA, though with some differences. Thus, TLA shares with SA the bilabial /b/, dental alveolar /d, dˤ, t, tˤ / and velar /k, g/ plosives, but although the uvular /q/ and glottal /ʔ/ plosives are found in the dialect, they tend to be omitted or replaced. The glottal stop may occur word-initially as in *ʔamm* 'lead prayer' *ʔamar* 'command', but it is not very common in word-medial or word-final positions, occurring only in lexical borrowings from SA (Abumdas, 1985). Harrama (1993) adds that, in most cases, this sound tends to be dropped or substituted by a glide or a vocalic sound (e.g. *seʔal* 'he asked' cf. *ya-sʔil* 'he asks', *ya-saʔl-u* 'they ask', *suʔal* 'a question').[12] In addition, the SA uvular /q/, and interdental fricatives /θ/ and /ð/ are realized as velars /g/, /t/ and /d/ respectively in

[12] The verb *seʔal* 'ask' is a *M*1 verb which is used frequently in the dialect. However, it tends to be replaced with the verb *nešad* 'ask, query' especially among the older generation who can produce the glottal stop in the perfective sub-paradigm of the verb *seʔal*, but delete it and possibly replace it with a (long) vowel *ya-saal-u* or a glide *ya-syal-u* 'they ask' in the imperfective form.

TLA. However, the sound /q/ can be found in borrowed words from SA such as /q/*anuun* ‘law’, /q/*aḍiiya* ‘case’, *ste*/q/*aal* ‘he resigned’. Among the fricative sounds, TLA can also have the voiced labiodental /v/ which is only used in loan words including /v/*ideo*, /v/*illa* and /v/*isa*. The plain /l/ can have an emphatic allophone [lˤ] in some Classical Arabic words such as A[lˤ]ah ‘God’ (cf. bi/ll/ah-i ‘by my God’ in which the plain /l/ is used when preceded by /i/) or when adjacent to an emphatic consonant (e.g. in the *M*2 defective verb /sˤ/*alla* ‘he prayed’ cf. /s/*alla* ‘basket’). TLA also has the glides /j/ and /w/ which can occur only word-initially (/w/*egaf* ‘he stood up’, /j/*ebas* ‘it dried up’, /w/*ʕee* ‘he became conscious’, /w/*alla* ‘he returned’) and word-medially (e.g. *l*/w/*ee* ‘he twisted sth.’, *la*/w:/*in* ‘he coloured sth.’, *ba*/jj/*in* ‘he clarified sth.’). In addition, TLA has the consonantal features of emphasis (pharyngealization). The primary emphatic consonants include *ṭ*, *ṣ*, and *ḍ*.

Place of articulation	Manner of articulation				
	Plosives	Fricatives	Nasals	Liquids	Glides
Bilabial	b		m		
Labio-dental		v f			
Dental-alveolar	t d tˤ dˤ				
Alveolar		z zˤ s sˤ	n	l r	
Post-alveolar		ʒ ʃ			
Palatal					
Velar	k g				j
Uvular	q	ɣ x			w
Pharyngeal		ʕ ħ			
Glottal	ʔ				
Laryngeal		h			

Table 3.10 TLA consonants

3.1.3.1 TLA vowels

Generally, there seems to be considerable disagreement between researchers of LA on the exact number of vowels in different LA varieties. The reason for this might be that phonemes have been confused with allophones especially in the presence of an emphatic sound. In addition, a short vowel in an unstressed syllable can vary in quality. The number of vowels that has been suggested ranges from 15 vowels (Griffini, 1913) (cited in Abumdas, 1985, p.41) to only eight vowels (Ahmed, 2008).

The vocalic system of verbs is based on the three vowels /a/, /i/ and /u/. Each of these vowels can be phonemically short or long. The high front short /i/ and long /ii/ can be found in words such as *bint* ‘girl’, *ʕiid* ‘Eid’, *midd!* ‘pass sth.’ and *ziid* ‘add!’. Examples of the high back vowels /u/ and /uu/ include *ḥuṭṭ*! ‘put down!’, *ḥuut* ‘fish’ and *nuuḍ* ‘wake up!’. Finally, the low vowels /a/ and /aa/ are seen in the words *gasim* ‘luck’ vs. *gaasim* ‘having shared/divided’. In addition, TLA has long mid vowels /ee/ and /oo/, as in *seef* ‘sword’, *zeet* ‘oil’ and *xoof* ‘fear’, *yoom* ‘day’ respectively. Long mid vowels form the final stem vowel of all defective verbs in *M*1 in the perfective, though they are subject to morphological conditioning (§4.3.1 and §4.5.4). Diphthongs can also be found in a few words such as *naw* ‘hot’, *ḥay* ‘alive’ and *jaw* ‘weather’. Other vowels that can be found in TLA include the short mid front vowel /e/ and the schwa, both of which occur in unstressed syllables: *resam* ‘he drew’ vs *rasim* ‘drawing’ and *gelab* ‘he turned over’ vs *gilib-it* ‘she turned over’ vs *galib* ‘turning over’. The following section will introduce a brief discussion of TLA syllable structure with particular reference to the verb domain.

3.1.3.2 Syllable structure and phonological processes in TLA

Generally, in LA varieties (Abumdas, 1985; Al-Ageli, 1995) including TLA (Elramli, 2012), syllables consist of an onset followed by a (short or long) vowel and an optional coda. Therefore, the CV(X) sequence can be considered as the basic TLA syllable structure template. A syllable can have a simple or complex onset or coda. These consonant clusters can occur in different positions including across an affix boundary (e.g. (i)*nbiiʕ* ‘I sell’ or within a word (e.g. *kamml-u* ‘finish. MPL!’). In word-final position, the consonant cluster is broken up by the epenthetic vowel /i/ (e.g. *gul(i)t* ‘I said’). The inventory of possible syllables in TLA includes the following:

Syllable	Pattern	Gloss
CV	ke.tab	he wrote
CVV	ree	he saw
CVC	jin	jinn
CVCC	radd	he returned
CCVV	lgee	he found
CVVC	gaal	he said
CCV	xab.bru	they informed
CCVC	sbiḥ	beads

Table 3.11 syllable types in TLA

his table shows examples of the different syllable types in monosyllabic or polysyllabic word forms. The CVC and CVV syllable types can occur in monosyllable or polysyllabic forms: *jee* ‘he came’ and *naa.da* ‘he called’. By contrast, the syllable type CV is not found in monosyllabic forms, possibly due to the requirement that a word should be a ‘quantitative (moraic) foot’ (Watson, 2002, p.129). In addition, TLA employs a number of syllable repair strategies including the phonological processes of syncope, vowel lengthening and epenthesis.

Syncope is a phonological process that affects high vowels when they occur unstressed in open syllables as can be seen in the following examples of *M*2 sound series:

(1)
a. kammil + u → kamml-u ‘they.M finished’
b. lawwin + in → lawwn-in ‘they.F coloured’

Epenthesis involves the optional insertion of the vowel /i/ between consonants in a two/three-consonant cluster as in the following examples of *M*1 verbs:

(2)

a. ketab + t	→	ketab(i)t	'I wrote'	cf. ketab-t-a	'I wrote it'
b. gul + t	→	gul(i)t	'I said'	cf. gul-t-(i)-l-ha	'I said to you'
c. lgee	→	l(i)gee	'he found'	cf. ma-lgaa-š	'he did not found'

The epenthetic vowel is dropped when a vowel-initial suffix is added to the word (2.a). In contrast, the addition of a consonant-initial suffix seems to trigger vowel insertion to the right of a CCC sequence (2.b) (gul[tlh]a 'I said to you'). The three-consonantal sequence includes the 1SG perfective suffix /-t/, the indirect object marker /-l-/ and the first consonant of the 3SGF object marker /-ha/. In 2.c, the epenthetic vowel /i/ is dropped when the negation particle *ma-* is attached, resulting in the resyllabification of the consonant cluster *lg* with */l/* as a coda and /g/ as an onset.

Vowel Lengthening is another syllable repair strategy that is usually triggered by an inflectional process (i.e. suffixation). For example, vowels that occur in verb-final position undergo lengthening when an object suffix is attached to the verb as in (3):

(3)

a. bana + ha	→	bana(a)-ha	'he built it (SGF.)'
b. l(i)ga + kum	→	lga(a)-kum	'he found you'
c. y(i)bi + ha	→	y(i)bi(i)-ha	'he wants it (SGF.)'
d. xalla + na	→	xalla(a)-na	'he left us'

TLA exhibits a form of vowel assimilation as shown by the following examples of *M*1 sound series:

(4)

a. tV + ftaḥ	→	ta-ftaḥ	'she opens'
b. yV + ṭlaʕ	→	ya-ṭlaʕ	'he appears'
c. nV + fraḥ	→	na-fraḥ	'I become happy'

In these examples, the FV assimilates with the stem vowel of the 3SGM imperfective form. However, there are many other cases in which these vowels show no assimilation (e.g. *ya-rkub* 'he rides' and *yu-ġsil* 'he washes'). By contrast, these vowels in the imperfective forms with a vowel-initial suffix display a different pattern of vowel assimilation. The FV and stem vowel undergoes assimilation only if the FV is high (e.g. *yu-rugd-in* 'they.F sleeps', *yi-rigb-in* 'they.F peeked' cf. *ya-rukb-in* 'they ride', *ya-libs-in* 'they.F wear') (§4.1).

Another phonological process found in this dialect is the process of regressive or progressive consonant assimilation. An example of regressive assimilation is voicing assimilation in which the consonant of the imperfective subject prefix *t-* assimilates to the first consonant of the verb after the elision of the vowel /i/ (Elramli, 2012). Examples include the following:

(5)

a. td → dd	t + diss	→	tdiss	→	ddiss	'you.M hid sth.'
b. tz → dz	t + zelbaḥ	→	tzalbaḥ	→	dzalbaḥ	'you.M fool sb.'
c. tj → dj	t + jammaʕ	→	tjammaʕ	→	djammaʕ	'you.M collect'

The examples show that the *t-* imperfective prefix becomes voiced when it precedes a voiced sound. Similarly, the prefix *t* of the perfect verb *M*5 and *M*6 can undergo total assimilation with the initial consonant of the verb as in (6):

(6)

a. ts → ss	tsallif	→	ssallif	'he borrowed'
b. tḍ → ḍḍ	tḍaarab-u	→	ḍḍaar-u	'they hit each other'
c. tz → zz	tzawwij	→	zzawwij	'he got married'
d. tš → šš	tšaawir-u	→	ššaawr-u	'they consulted each other'

Another process of assimilation is devoicing. This process affects the consonant *d* when it occurs in the final-position of a verb that has the subject suffix *–t*, *–tu* or *–ti* as in (7):

(7)
dt → tt maddid + t → maddit-t 'I extended'

Clearly, the verb domain can be subject to a number of phonological processes which can interact with morphological factors, contributing to the morphological complexity of verb stems (§4.5). The next section will consider the stress assignment in the dialect.

3.1.3.3 Stress Pattern in TLA

It is generally claimed that most LA dialects are predictable quantity-based systems (Elfitoury, 1976; Harrama, 1993; Abumdas, 1985). Stress assignment depends on the syllable weight and position. Arabic has three weights of syllable: light, heavy, and superheavy. 'Light syllables are always open, heavy syllables are open or closed, and superheavy syllables are closed or doubly closed' (Watson, 2011, p.2). Stress falls on one of the last three or four syllables following the assumed syllable-weight rules whereby a finally heavy syllable (CVV(C) or $CVVC_iC_i$[13]) systematically attracts stress (e.g. *maf.húum* 'understood', *wid.yáan* 'valleys', *dir.túu(h)* 'you.PL. did it', *juwáab* 'answer', *(i)trúdd* 'she returns'). When the final syllable is not superheavy, the stress is assigned to a heavy penultimate syllable (CVV or CVC) (e.g. *dáa.rit* 'she did', *kál.lim* 'he called', *le.ʕáb.na* 'we played', *máṣ.naʕ* 'factory', *(i)k.táa.bi* 'my book', *ketábtu* 'you.PL wrote', *fihimúuha* 'they understood her').[14] In the absence of either a final superheavy or a heavy penult, TLA stresses the antepenult which can be either short or long (e.g. *ḍú.ru.bit* she hit', *má.zir.ʕa* 'farm', *mádirsa* 'school'). Nevertheless, in some cases, the stress assignment in TLA deviates from the assumed quantity-based predictability. For instance, the rhythmic grouping of the light syllables in *fehám* 'he

[13] Geminate consonants
[14] The CVC syllable type word-finally is considered to be a light syllable, since the word-final consonant is 'extrametrical'; hence, CVC it is not considered by the stress assignment rules (McCarthy, 1979).

understood' and *naxál* 'palm-trees' seems to be weak-strong (iambs). By contrast, the light syllables in forms such as *fíhimit* 'she understood', *fíhimu* 'they.M understood', *inkúsurit*, 'it.F broke' and *fáhim* 'understanding', are possibly grouped into strong-weak pairs (trochees) since the stress is on the antepenult. In other words, TLA shows opaque stress assignment in some morphologically complex forms by stressing a light antepenultimate syllable instead of a heavy penultimate one (*yí.kit.bu* 'they.M write'). The opacity of the stress is also apparent in forms such *ge.bál.na* 'he accepted us' (cf. *gá.bil.na* 'before us') and *kí.ti.bit* 'she wrote' (cf. *ke.tá.bit*, 'I/you wrote'). The reason for the opacity is that the epenthetic vowels may not be visible for stress assignment in some Arabic dialects (Watson, 2011). In TLA, the assignment of stress in the lexicon is prior to syncope and/or epenthesis. This accounts for the stress placement in *ke.tá.bit*, 'I/you wrote' derived from underlying /ketab-t/ and *yí.kit.bu* 'they.M write' in which the second stem vowel undergoes elision (*yíkt(V)bu*).[15]

Understanding the stress pattern in verb forms is also essential for the current stem-based account, particularly for *M*1 sound verbs. Generally, the stress placement in the verb paradigm follows the pattern presented above. For instance, in the *M*1 sound 1/2 and 3SGM perfective forms, the stress is on the final syllable while the 3SGF/PL forms have the stress on the initial syllable. By contrast, all the imperfective and imperative forms of this verb series have an initial-stress placement (e.g. *ġe.rág* 'he drowned', *ġí.ri.git* 'she drowned', *yáġ.rug* 'he drowns', *yá.ġur.gin* 'they.F drowned').

Likewise, in *M*1 hollow verbs, the initial stem syllable is stressed in all the inflectional forms of both imperfective and perfective sub-paradigms (e.g. *(i)y.gúul*, 'he says', *gáal* 'he said' and *gúl.na* 'we said'). In the perfective forms of *M*1 defective

[15] For further details on the reader is referred to the relative ordering of stress, syncope and epenthesis (Brame,1974).

verbs, the monophthong or the inflectional marker vowel carries the stress, but the imperfective and imperative forms have the stress on the FV (e.g. *n(i).sée* 'he forgot', *n(i).sít* 'she forgot' cf. *yán.sa* 'he forgets', *yán.sin* 'they.F'). Finally, in *M*1 double verbs, the stress is on the penultimate syllable in the 1/2 person forms and on the first syllable in the 3SGF/PL forms in the perfective (e.g. *far.rée.na* 'we fled', *fár.rit* 'she fled'). By contrast, in the imperfective forms the stress is invariably assigned to the first syllable (i.e. stem vowel) (e.g. *(i)y.fírr* 'he flees' *(i)y.fír.ru* 'they.M flee').

In many non-*M*1 verbs, the stress is on the first stem vowel (e.g. *M*2 *ʕál.lim* 'he taught', *náa.da* 'he called', *t(e).ʕál.lim* 'he learned', *stán.jid* 'he pled'). By contrast, *M*7 and *M*8 sound series seem to have the same stress placement as that of *M*1 counterparts. Likewise, in the imperfective forms, the stress is assigned to the first (prefix) vowel (e.g. *M*1 *ke.ṣár* 'he broke' cf. *M*7 *in.ke.ṣár* 'it was broken' and *yún.ku.ṣur* 'it will be broken', *M*1 *kú.ṣu.rit* 'she broke' cf. *M*7 *in.kú.ṣu.rit* 'it.F was broken' and *tún.ku.sur* 'it.F will be broken'). In *M*2 defective series, such as in the verb *bakka* 'he upset sb.', the stress in the perfective forms is on the initial syllable in all 3rd person forms while 1/2 person forms have the stress on the monophthong *-ee-* (e.g. *bak.kit* 'she upset sb.' cf. *bak.kée.na* 'we upset sb.'). The relationship between the stress and the stem-alternation in verb paradigms will be further considered in (§4.5.2).

In the following section, we will introduce some of the previous studies on LA verb morphology, most of which are morpheme based, comparing these approaches with the current account.

3.2 Previous work on LA verb morphology

Previous analyses of LA verb morphology using different frameworks provide a useful source of comparison for the account I present in Chapters 4 and 5. Descriptions exist for ELA (Owens, 1984), JLA (Harrama, 1993) and Tripoli LA (Elfitoury, 1976). There are a good number of studies which focus on the phonological system of different LA varieties. One of the pioneering works mostly describing the variety used in Eastern Libya was done by Panetta (1943) (*L'Arabo Parlato in Bengasi*), who provided a description of the dialect as a set of guide texts to be used by Italians living in the country at that time. A more comprehensive study of the same LA variety was done by Mitchell (1952) who provided a more theoretically informed and accurate description of the phonological system, which inspired many other studies of the dialect. One of the most significant works is Owens (1984) book '*A short reference grammar on Eastern Libyan Arabic*' which provides a detailed account of the phonology, morphology and syntax of this LA variety.

The majority of the work on LA is the product of PhD studies by native speakers of different LA varieties, describing their phonological and morphological structure. Elfitoury (1976) provides 'a descriptive grammar of Libyan Arabic', focusing mainly on Tripoli LA. The phonology and phonetics of this dialect have been the focus of work by Abumdas (1985) on Zliten LA (ZLA), Al-Ageli (1995) on Tripoli LA, Elramli (2012) and Maiteg (2013) on TLA. The objective of Abumdas' thesis was to present a general overview of the phonological system. He provides comprehensive discussion of the consonantal and vocalic systems, syllable structure and the stress pattern along with a number of phonological processes. Elramli's OT account investigates assimilation processes such as the regressive assimilation of the definite article /l/ and the *M*5 and *M*6 derivational prefix /t/ both of which are subject to blocking or triggering effects of the OCP. Maiteg's (2013) Ph.D. thesis is a phonetic study of the secondary co-articulation effects of pharyngealisation. Al-Ageli's work is also constraint-based, focusing mainly on syllable structure and stress assignment in Tripoli LA. Work devoted to LA morphology has been presented in Harrama's (1993) thesis on '*Libyan Arabic morphology: al-Jabal dialect*'. He makes use of SPE-type rules to account for the morphological structure of verbs, nouns and adjectives. Unlike the majority of these studies, the present study is a synchronic surface stem-based account which emphasizes the role of *morphology by itself* in verb inflectional paradigms of TLA.

3.2.1 Owens' (1984) analysis

Perhaps one of the most influential studies of LA was done by Owens (1984) who gave a detailed description and a synchronic analysis of ELA, an urban dialect spoken in Benghazi. Owens' study introduces the basic grammatical structure of this variety 'for comparative purposes' (p.1). ELA is particularly interesting because it has a number of shared features with both the Western North African and Eastern Arabic dialects, providing an important link in the Arabic dialect continuum. For example, ELA is characterized by the lack of medial position diphthongs and *imala*, that is, the raising of a short /a/ or long /aa/ vowel in the context of /i/. This results in a long front mid monophthong, /ee/. Both characteristics are phonological features preserved by rural dialects (Owens, 2006, p.2). ELA diphthongs become monophthongs, but not after the guttural sounds *h*, *ḥ* and *ʕ* (e.g. *mawz* 'banana'> *moz*) and imala is fully reduced in the urban dialect (*nirien* 'fires' (rural) > *niran*).

In his Chapter 7, Owens (1984) provides a detailed description of the verb morphology in ELA using the root-and-pattern approach. The section on the morphology of ELA is perhaps the most detailed part of the book. Ultimately, Owens' account of the morphology uses the different syllable structures of inflectional forms within the same verb series and across the different ones as the basic framework for categorizing verbs. The basis of his account is syllabification and epenthesis. He argues that there are two vowel classes in the system, including V (non-elidable vowels) and v (elidable high vowels). Each category has a different distribution. Owens (1984, pp.20-21) argues that the distribution of the v category is governed by syllable structure and he postulates the syllable structure rule (SSR) in (8). Examples are provided in Table 3.12.

(8)

SSR1 C_C #

SSR2 a. V:
b. (CCV C_2 _) C_3 _ C_4 C_5
c. C
d. #

SSR3 #_ CC

SSRB # ^{e}(CC_C)_C #
C

before SSR	after SSR	gloss
SSR1 šift#	= šif(i)t	'I saw' (cf. TLA šufit)
SSR2.a. $kaa_1t_2b_3l_4k_5an$	= $kaat_2$ (i) b_3 (i) l_4k_5an	'he has written for you'
SSR2.b. $minti_1s_2b_3l_4h_5um$	= $minti_1s_2$(i) b_3 (i) l_4h_5um	'related to them by marriage'
SSR2.c. $gas_1s_2m_3l_4k_5am$	= gas_1s_2 (i)m_3(i) l_4k_5am	'he divided for you'
SSR2.d. $\#_1t_2m_3s_4k_5i$	= t_2(i)m_3(i)s_4k_5i	'you grabbed'
SSR3 #dbuuš	= #(i)dbuuš	'loads'

Table 3.12 Examples of the v class insertion after C_2 (Owens, 1984, pp.20-21).

These rules summarize the phonological contexts in which the v category is inserted. According to SSR1 and SSR3, given a series of three members CC# or #CC the vowel is inserted after C_2 starting from the right (Table 3.12).[16] The second phonological context for the vowel insertion according to SSR2 is after C_2. The rules provide the distribution of the v category in a sequence of five consonants. In this case, there are two instance of epenthetic vowels. The first vowel occurs between the second and third

[16] The # represents a pause which is also included as part of the phonological context for the SSR.

consonants (C_2_$C_3C_4C_5$) while the second vowel is inserted between the third and fourth members of the sequence (C_2_C_3_C_4C_5) (Table 3.12). The application of the v insertion in this context has to begin from the left in the identified sequence since proceeding from the right derives incorrect forms. It is also worth noting that the first part of SSR2 is optional and covers four contexts which include V:, CCV, C and # all followed by C_2 (8). In these four optional parts V:C_2, (SSR2.a) CC_2 (SSR2.c) and #C_2 (SSR2.d) are equivalent contexts added to account for different words not necessarily of the same category. However, the $CCVC_2$ structure in SSR2.b is specifically chosen to handle vowel insertion in active participle forms of *M*7 (Table 3.12) or M8 in ELA. In addition, given that the first part of SSR2 is optional, the second part (C_3_C_4C_5) can also account for the vowel insertion in context of a three-consonant sequence (9.c):

(9)

a. yik[tb#]	yikt*i*b#	'he writes'	
b. yi[ktb]u	yik*i*tbu	'they write'	
c. kitab[tlk]am	kitabt*i*lkam	'I wrote for you'	(ketabtilk*um* TLA)
d. [#db]uuš	#*i*dbuuš	'loads'	

In (9), where there is a sequence of three consonants (tb#, ktb, tlk and #db), the high vowel of the v class is inserted between the first two consonants of a three-member sequence, which in (9) are represented by *tb* (SSR1) *kt* (SSR2.d), *tl* (SSR2.c) and *#d* (SSR3), starting from right to left. It is worth noting that the deletion of the first stem vowel in yi[kt*i*b] is the result of the phonological restriction whereby, in an open syllable (CV), a short high vowel undergoes elision because of the ban on a sequence of two short open syllables CVCV (Owens, 1984, p.12). Clearly, the context of insertion in a sequence of three member has to be handled by SSR1 SSR2 and SSR3.

These SSR rules fail to produce the correct syllabification for a set of *M*1 sound verbs such as *ismʕit* 'I heard' which the given SSR would syllabify as **simʕit*. Therefore, Owens (1984, p.21) also postulates the SSRB that has a special boundary marker #^{e} marking a specific class of forms including: *ismiʕit* 'I heard', *išribna* 'we drank' and *ilbis* 'he dressed'.[17] The context for the insertion according to SSRB is shown in the following table.

Verb form	SSR
#^{e} smʕt#	
#^{e} smiʕit#	by SSR B
#^{e} ismiʕit	by B, non-optional part

Table 3.13 The application of the SSRB to the form ELA *ismiʕit* 'she heard'.

Table 3.13 shows the application of the SSRB to classes of perfective sound verbs. Although these v vowel category rules may account for some cases of high vowel /i/ alternation (deletion and insertion), they appear to have become somewhat unwieldy as there are a number of exceptions where the insertion of the v class vowel is not applied in the expected environment. For instance, there is no vowel insertion into a C_iC_i-sequence. However, an epenthetic vowel is introduced when the first C_i is either the future maker *b*- or an inflectional prefix (Owens, 1984, p.22):

(10)

b(i)*b*aat	'he will stay over'
t(i)*t*taka	'she leans over'
t(i)*t*manna	'she wishes'
b(i)*b*an	'he will appear'
n(i)*n*tisa	'I will be forgotten'

Similarly, vowel insertion does not apply before the negation and interrogative particle *š* (e.g. *ma-diritš* 'I did not do') and some quadriliteral verbs show a CCC-sequence

[17] Cf. TLA forms *semaʕ-t, šerab-na* and *lebas* respectively.

where the insertion of the v class vowel does not apply (e.g. *itarj(i)m* ‘he translates’ but *itarjm-a* ‘he translates it’ cf. TLA *itarijm-a*). In addition, there is no v insertion in the sequence VC_1CC where C_1 is *w* or *y* (e.g. *ḥawšna* ‘our house’) or in connected speech where the phonological context of v class is followed by a vowel-initial form (e.g. *šifit* vs. *šiftil-bint* ‘I saw the girl’) (Owens, 1984, pp.22-23).

Assuming that # is part of the phonological context, behaving just as a consonant does not seem to provide a plausible account of the v-class distribution. The attested inflectional forms can be accounted for just by assuming that ELA is a ‘coda’ language (that is, with right-to-left syllabification). For instance, in the #CCV form *#id.buuš*, the syllabification algorithm gives us #C.CV, as in *#d.buuš*. The initial consonant, therefore, has to be treated as the coda of the first syllable resulting in *idbuuš*.

Having established the framework based on syllabification and syllable structure rules, Owens introduces the verb morphology of ELA based on one simple verb and seven derived ones. Interestingly, this account of verb measures and inflectional paradigms is very similar to a word-based account. He classifies verbs into classes and sub-classes based on the syllable structure of a selected set of inflectional forms. Thus, he distinguishes two verb forms: a completive form (perfective) with a past meaning and an incompletive form that describes an uncompleted (imperfective) action or a state the meaning of which is usually habitual or generic. He describes each of these forms as a ‘morphological verb’ (1984, p.104) to distinguish them from the active participle. In addition, these forms can be further distinguished by affixation that refers to subject markers realizing person, number and sometimes gender (§4.1):

	Completive		Incompletive	
	SG	PL	SG	PL
1	kitab-it	kitab-na	ni-ktib	ni-kitb-u
2 M	kitab-it	kitab-tu	ti-ktib	ti-kitb-u
F	kitab-ti	kitab-tan	ti-kitb-i	ti-kitb-an
3 M	kitab	iktib-o	yi-ktib	yi-kitb-u
F	iktib-at	iktib-an	ti-ktib	yi-kitb-an

Table 3.14 ELA verb paradigm of *kitab* 'to write' (Owens, 1984, p.104).

The completive forms have the subject markers as suffixes whereas the subject marker in the incompletive verbs is realized as a prefix. Furthermore, in both forms, number appears always as a suffix that sometimes cumulates gender, as seen in Table 3.14. Owens (1984) claims that the vowel alternation in each cell can be phonologically determined (§3.2.1).[18]

Similar to the traditional analysis of Arabic verb measures, Owens (1984) classifies verbs into basic *M*1 and derived measures from the basic verb form by affixing a consonant, a vowel or both. However, unlike the traditional root and pattern approach, Owens' classification of the basic verbs (M1) groups them under six main categories (similar to the notion of ICs) based on the number of root radicals in addition to the syllable structure of inflectional forms within the same verb paradigm. In these six categories, he includes the sound, hollow, double verb series, vowel-initial (assimilated verbs) or final verbs (defective series verb), single consonantal root verbs, and the imperative form.[19]

The sound verb series of the basic verb form (*M*1) is usually described as a three-consonantal root verb with different vowel qualities, particularly when inflected for imperfective. However, Owens (1984, p.109) argues that, in ELA, this first category

[18] It is worth noting that the *n*- prefix of the first person singular and the *-u* suffix of the first person plural of the imperfective verb are characteristic of North African dialects (Versteegh, 2014).
[19] Confusingly, it seems that Owens (1984) views the imperative form as a separate verb class rather than as part of the inflectional paradigm.

can be further categorized based on the syllable structure of the 3SGM and 3SGF completive (perfective) forms and the 3SGM and 3PLM in the incompletive (imperfective). Hence, the ELA sound verb series is divided into five syllable-structure verb classes, consisting of three completive forms and two incompletive forms as illustrated in the following examples:

(11)

a. Completive classes

1. CVCVC	kitab	'he wrote'	iktibat	'she wrote'
2. CCVC	ilbis	'he dressed'	ilbisat	'she dressed'
3. CCC	išrib	'he drank'	širbat	'she drank'

b. Incompletive classes

4. CCC	yi-ktib	'he writes'	yi-kitbu	'they write'
5. VCCVC	ya-šrab	'he drinks'	ya-šrubaw	'they drink'

Class 1 has the syllable structure CVCVC, which is historically derived from the form CaCaCa. The affixation of a vowel-initial suffix to this form creates the forbidden sequence of two open syllables (Owens, 1984, p.109):

(12)

kitabat a. iktibat 'she wrote'
vCCVCVC
b. kitbit
CvCCVC

The first vowel in the form *kitabat* (CV.CV.CVC) is in an open syllable and it undergoes elision. The result of this process is the word-initial cluster *kt* (i.e. *ktibat*). Therefore, in Owens' account, an epenthetic high vowel /i/ is inserted word-initially to give *iktibat.*[20] By contrast, the incompletive forms of this class can belong to either Class 4 or Class 5 (11.b). The one and only form found in Class 2 is *ilbis* 'he dressed' ~ *ilbis-at* 'she dressed' which also belongs to Class 3 since the 3SGF form has the

[20] Note that the insertion rule SSRB does not apply in TLA cf. *kitib-it*.

(CVCCVC) syllable structure (e.g. *libsat* 'she wore') that is subject to SSRB (Table 3.13). The incompletive form of all Class 3 verbs belongs to Class 5 (VCCVC).[21] Given the syllable structure of each verb class in both completive and incompletive forms, Owens (1984, p.109) concludes that sound verbs can belong to three main conjugations:

Conjugation	Completive	Incompletive	examples
1	vCCvC (<CCC)	VCCVC	išrib 'to drink'
2	CVCVC	VCCVC	dibaḥ 'to slaughter'
3	CVCVC	v-CCvC (<CCC)	kitab 'to write'

Table 3.15 The three conjugations of sound verbs in ELA (Owens, 1984, p.109).

The *M*1 weak verbs are also classified into classes and sub-classes. According to Owens (1984), ELA weak verbs consist of verbs with a one or two consonant root. The two-consonant root verbs can be further divided into three classes: long vowel-initial incompletive forms (Class A) (Table 3.16), forms with long medial vowel (Class B) (hollow series), and finally, forms ending in a vowel (Class C) (defective series). Each class can be further subdivided. For instance, Class A consists of a limited number of verbs, but Owens subdivides the class into two subgroups:

Group	Completive	Incompletive	examples
1	CVCVV	VVCC	xada 'take', kila 'eat'
2	VVCVC	VV CVC	aamar 'order'

Table 3.16 Class A of weak verbs in ELA (Owens, 1984, p.112).

[21] It seems that Owens (1984) treats the initial vowel as part of the stem (e.g. *gital* ~ *y-ugtil* 'to kill') (§4.1.1).

	Completive		Incompletive	
	SG	PL	SG	PL
1	kileet	kileena	naakil	naakilu
3 M	kila	kilo	yaakil	yaaklu
Imperative				
2 M	kul		kulu	
F	kuli		kulan	

Table 3.17 The paradigm of the verb *kila* 'eat' in ELA (Owens, 1984, p.112).

The verb *kilee* 'eat' is classified as an irregular verb because of the stem pattern in the imperative forms, which shows a suppletive stem instead of the default stem (§5.2.1). By contrast, group two verbs have an initial glottal stop (*M*4 of CA) which, in TLA, can only be found with a few verbs borrowed from CA and follows CA verb inflectional paradigm.

Class B includes the hollow verbs that show ablaut. The 3SGM form of each class always has a long low vowel *aa* whereas the 1/2 person forms have either *i* or *u* as the stem vowel. Owens (1984) claims that the alternation between *i* and *u* vowels can be phonologically determined since the forms with a high back vowel occur with pharyngealized consonants stem initially and *i* elsewhere. The incompletive form can have *ii*, *uu*, *aa* as the stem vowel (Table 3.18):

	a. baan 'appear'	b. faat 'pass by'	c.ṭaaḥ 'fall'
1SG	n-baan	n-fuut	n-ṭiiḥ
1PL	n-baan-u	n-fuut-u	n-ṭiiḥ-u
2SG.M	t-baan	t-fuut	t-ṭiiḥ
2SG.F	t-baan-i	t-fuut-i	t-ṭiiḥ-i
2PL.M	t-baan-u	t-fuut-u	t-ṭiiḥ-u
2PL.F	t-baan-in	t-fuut-in	t-ṭiiḥ-in
3SG.M	y-baan	y-fuut	y-ṭiiḥ
3SG.F	t-baan	t-fuut	y-ṭiiḥ
3PL.M	y-baan-u	y-fuut-u	y-ṭiiḥ-u
3PL.F	y-baan-in	y-fuut-in	y-ṭiiḥ-in

Table 3.18 The incompletive sub-paradigm in ELA hollow verbs.

However, it is not rare to find counter-examples where the phonological conditioning environment is not met. The back vowel *u* can occur in a non-velar environment while the front vowel *i* can appear next to pharyngealized consonants.[22]

For Class C, weak verbs (defective verbs), Owens (1984, p.114) establishes three groups based on the quality of the final vowel in the completive and incompletive forms. The first group has a final vowel *i* in the completive, but ends in the vowel *a* in the incompletive (e.g. *ilg**i**it* 'I met', *ilgi* 'he met', *yalg**a*** 'he meets'). The second class has the stem final vowel *e* in the completive and *i* in the incompletive (e.g. *rum**e**et* 'I threw', *ruma* 'he threw' *yirm**i*** 'he throws'). The third group has *e* as the final stem vowel and *a* in the incompletive (*bid**e**et* 'I started', *bida* 'he stated' *yabd**a*** 'starts'). In addition, the verb final vowel may undergo lengthening before a consonant-initial suffix. However, Owens does not provide any theoretical explanation for that, nor does he rule out the possibility of a vowel shorting process.[23] For weak verbs, Owens also identifies two verb categories which have only one consonantal root, these include the i-completive verb (e.g. *j**i**it* 'I came', *ja* 'he came' and *iji* 'he comes' cf. TLA *jee*) and the e-completive verb (*r**e**et* 'I saw', *ra* 'he saw' and *ira* 'he sees' cf. TLA *ree*).[24]

The final verb category of *M*1 is the double verb series. According to Owens (1984), verbs with gemination do not have class variations, hence there is only one form ending in a final geminate sequence. However, in TLA, two categories of this verb class can be identified based on the verb vocalism, according to which the

[22] E.g. from TLA *faat ~ fuut* 'pass' *maat ~ muut* 'die' *daar ~ duur* 'turn around' *daar ~ diir* 'do', *saar ~ ṣiir* 'happen', *saar ~ siir* 'walk on track', *daaʕ ~ diiʕ* 'announce', *ḍaaʕ ~ ḍiiʕ* 'lose', *ṭaaḥ ~ ṭiiḥ* 'fall', *ṭaab ~ ṭiib* 'be cooked, accept', *ṭaaf ~ ṭuuf* 'circumambulate Mecca'.

[23] In TLA, the 3SGM form ends in the monophthong *ee* which might suggest that it is a case of vowel shorting rather than lengthening since the MPs of this form are realized by Stump's (2001, p.53) Identity Function Default.

[24] The verbs *haat!* 'give.IMPER', *haak!* 'take.IMPER' and *yibi* 'he wants' also exist in TLA. The paradigm of the first two verbs is defective, as these verbs only occur in the imperative. The last verb is only found in the imperfective while the perfective sub-paradigm is filled by periphrasis *kan yibi* 'he wanted'. However, the perfective stem of this verb does occur in the negated form: *ma-ba-š* 'he did not want'.

perfective forms have the low back vowel *a* while the imperfective can be either a high front *i* or a high back *u*, (e.g. *radd* ~ *rudd* 'to return' *rann* ~ *rinn* 'to ring') (§4.3.3).[25]

In contrast to his analysis of *M*1 verbs, Owens' analysis of non-*M*1 verbs basically highlights the available measures and their phonological shape, but without proposing a sub-classification within any measure.

Clearly, Owens' approach to ELA verb morphology and his detailed description of the verb inflectional morphology based on the syllable structure of inflected words is perhaps a move towards a word-based account, suggestive of a novel framework for Arabic verb classification and possible ICs. One issue raised by Owens' account is overclassifying verbs based on different factors, leading to a slightly confusing account. The ELA *M*1 verb is based on six verb categories in addition to their sub-classes. The first class is sound verbs which are categorized as a three-consonantal root class and further divided by the different types of syllable structure in selected forms of the imperfective and perfective paradigms. On the other hand, the weak verb classification is done based on the position of the stem vowel while each of the double verbs and imperative verbs (verb with defective paradigms) are listed as separate classes. The alternative stem-space account would handle the ELA verb system by stem indexing, capturing the paradigmatic similarities and differences in a more economical fashion.

The following section will present another study on verb morphology that makes extensive use of phonological rules to account for the verb inflectional system in another LA variety, even if the analysis does not hold for a number of verbs.

[25] Further examples of the vocalic pattern in double verbs include: *ġamm ~ġumm* 'cover up', *xaṭṭ ~ xuṭṭ* 'draw a line', *xaff ~xiff* 'weigh less', *raff ~ riff* 'blink', *haff ~hiff* 'move quickly', *gašš ~ gišš* 'brush', *kašš ~ kušš* 'snap at sb.' and *xašš ~ xušš* 'narrow'.

3.2.2 Harrama's (1993) analysis

Most studies of the verb morphology of LA dialects have been widely based on a phonological account, arguing that any form internal alternations in different measures can be explained by phonological rules, mainly assimilation. Harrama's (1993) study deals with the morphological structure of JLA. Although the focus of his work is the morphological component of this dialect, he provides an extensive discussion of the phonological system along with major phonological processes to account for the derivation and inflection of triliteral and quadriliteral verbs. The verb paradigm is presented under a root-based morphology in which the root of a verb is analysed as consisting of consonants 'interdigitated' with vowel patterns representing MPs, sometimes accompanied by affixes. Any verb form internal alternations are handled by phonological rules/processes. This synchronic analysis uses diachronic referencing in which CA is the underlying representation of JLA.

Similar to ELA and TLA, Harrama (1993) points out that JLA has eight verb measures each of which has the 3SGM perfective form as the default base form. Some measures can have different verb series, sound, hollow, double and/or defective. The JLA inflectional paradigm is based on the perfective, imperfective and imperative in which the verb form internal vocalic pattern is regarded as the realization of MPs. The vocalic pattern is subject to various phonological processes. The most interesting feature of this dialect is the ablaut change. The inflectional forms of *M*1 sound verbs involve the optional raising of the stem vowel *a* to *i* or *u* in the 3SG.F/PL forms (Table 3.18). However, Harrama makes three important observations about the nature of this vocalic pattern. First, the ablaut in these forms is optional. Second, it is conditioned by the sociolinguistic factor of age and tends to be employed by older generations. Finally, Harrama hedges his phonological account for determining the vowel quality since for any phonological rule for this ablaut, 'exceptions do exist' (1993, p.92) (Table 3.19)[26]:

MPs	fataḥ 'he opened'	ṭalag 'he let go'
1SG	fataḥ-t	ṭalag-t
1PL	fataḥ-na	ṭalag-na
2SG.M	fataḥ-t	ṭalag-t
2SG.F	fataḥ-ti	ṭalag-ti
2PL.M	fataḥ-tu	ṭalag-tu
3SG.M	fataḥ	ṭalag
3SG.F	fataḥ-it~ fitaḥit	ṭalag-it ~ ṭulag-it
3PL.M	fitaḥ-u	ṭalag-u ~ ṭulag-u

Table 3.19 The perfective sub-paradigm of JLA sound verbs *fataḥ* 'he opened' and *ṭalag* 'he let go' (Harrama, 1993, p.93).[27]

Similarly, *M*7 and *M*8 verbs show vowel alternation when appended by vowel-initial suffixes, the stem vowel undergoes vowel raising in the 3SGF form (e.g. *nḍarab* 'to be hit' ~ *nḍurabit*). Harrama argues that the stem vowel alternation can be determined by the environment, hence, the stem vowel *a* is raised to *u* when the first consonant is an emphatic sound (*ṣ, ṭ* and *ḍ*). The other environment for alternation is when the first radical is *r*, *g* or *x* and the second is possibly an emphatic, or in cases where the first consonant is *w*. To account for these alternations, Harrama (1993, p.95) proposes the following rule:

[26] Similar vocalic alternations exist in TLA, however, unlike, JLA, the alternation is morphologically motivated and it is not optional.

[27] Unlike TLA and ELA, the second third person plural forms do not distinguish gender, and the masculine form is used for these plural forms.

(13)

a → u {ṣ, ṭ, ḍ, w} ___ CaC + V(C)

{g, x, r} ___ emphatic

↘ i elsewhere

The presence of the emphatic consonant can determine the vowel quality, which tends to be a high back vowel *u* in JLA and in TLA (§4.5.1). However, as will be discussed later (§4.3.1 and §4.5.4), the high back vowel can also occur with non-emphatic sounds. Likewise, the high front vowel is a possible stem vowel when the first consonant is *r*, *x*, *g*.[28] Unlike *M*1, no ablaut change occurs in JLA non-*M*1 sound verbs. However, the conjugation of the 3SG.F/PL forms in these measures involves metathesis between the second stem vowel and the second radical (e.g. *M*10 *staʕmil* 'he used' ~ *staʕimlin* 'they.F used') (Harrama, 1993, p.98).

As with the perfective paradigm, Harrama (1993) provides a phonological account for the inflection for the imperfective aspect. When a *M*1 sound verb is inflected for the imperfective, it can have one of the following patterns: -CCiC, -CCaC or –CCuC, as can be seen in the *ya-fraḥ* 'he feels happy', *yi-xdim* 'he serves', *yu-drub* 'he hits'. Harrama claims that the stem vowel *i*, *a* or *u* is determined by phonological rules. The stem vowel in the underlying representation is /i/ alternating between /u/, /i/ and, in some cases, /a/, depending on the surrounding consonants. The stem vowel is /u/ when one of the radicals is an emphatic sound (e.g. *ṣ*, *ṭ* or *ḍ*) and C_2 or C_3 is not a

[28] TLA examples of vocalic alternations include *rekab* ~ *rukub* 'ride', *rejaʕ* ~ *rujuʕ* 'return', *šerab* ~ *šurub* 'drink', *šerak* ~ *šuruk* 'tear up' compare with *regab* ~ *rigib* 'peek', *rejam* ~ *rijim*, 'throw devil with stones', *šerad* ~ *širid* 'be miles away', *šerak* ~ *širik* 'invoke other Gods along with Allah'.

guttural consonant (e.g. *ḥ*, *ʕ*, *ʔ*, *x* or *ġ*). For example, *yu-ḥfuḍ* 'he memorizes' and *yu-ṣbur* 'he is patient', but not in verbs such as *yaṣġir* 'he/it becomes small' and *ya-ḍġaṭ* 'he compressed' and interestingly Harrama also lists *yu-ṭbux* 'he cooks' as an example of this conditioning environment (cf. TLA *ya-ḥfiḍ* 'he memorizes' vs. *yu-rfuḍ* 'he refuses', *ya-ṣġir* 'he/it becomes', cf. *ya-ḍġuṭ* 'he compresses').

Another case for a high back stem vowel discussed by Harrama is one in which C_2 is *k*, *g* or *r* and the following sound is a voiced non-guttural sound. Examples include *yu-skun* 'he dwells', *yu-rgud* 'he sleeps' and *yu-xruṭ* 'he peels' (but not in *ya-šraḥ* 'he explains'). Although these verbs also have a high back vowel in TLA, there are a number of cases in which the same environment has a high front vowel (e.g. *yi-srig* 'he steals', *yi-škim* 'he pokes', *yi-ʕgid,* 'he ties', *yi-mrid* 'he crawls' and *yi-friz* 'he spots'). In addition, Harrama's phonological account for the 3SGF/PL forms of the imperfective is counter-intuitive. He suggests that the stem vowel of these forms can either be a high front or back vowel which first undergoes elision since it is a high unstressed vowel in an open syllable, but after the application of syncope, an epenthetic vowel is inserted and subject to a high and back vocalism alternation. In other words, the input of this rule is the same as the output.

Non-*M*1 imperfective forms are handled by the phonological process of syncope. The final stem vowel surfaces as the vowel *i* which may undergo elision, depending on the environment. For example, in the *M*3 verb form *tsaafr-u* 'they travel' (cf. *y-saafir* 'he travels') the vowel *i* is a high, unstressed vowel in an open syllable; hence, it undergoes elision.

The phonological account is extended to other verbs series. For example, Harrama points out that when the double verb series of the *M*1, *M*7, *M*8 and *M*10 are extended by the third person suffixes of the shape -V(C) or zero, the vocalic pattern of the stem does not change. In contrast, the addition of a consonant-initial suffix triggers the insertion of a long mid front vowel between the last stem consonant and the suffix (e.g. *ḥall* 'he opened' ~ *ḥalleet* and *nkarr* 'it/he was pulled' ~ *nkarreet*).

Unlike double verbs, the conjugations of *M*1, *M*7, *M*8 and *M*10 of hollow verbs involve vowel alternation in the perfective paradigm, as seen in *ġaab* **~** *ġubt* 'be absent' and *jaab* ~ *jibt* 'bring sth.' In forms with consonant-initial suffixes, the stem vowel is shortened and raised to *i* or *u* whereas the vowel of forms with vowel-initial suffixes, remains unchanged. Harrama (1993, p.101) accounts for the vowel alternation by considering the underlying representation of the verb which has the glides *y* or *w* that undergo deletion due to their intervocalic position. The result of the deletion of the glides *y* or *w* is the raised vowels /i/ and /u/ respectively. The vowel of non-*M*1 hollow verbs undergoes shortening in the 1/2 person forms (e.g. *M*8 *xtaar* ~ *xtart* 'choose'). Harrama argues that the vowel shortening occurs when a consonant-initial suffix is attached, but the shortening is blocked when a vowel-initial suffix or a zero suffix is attached. However, as will be explained in Chapter 4, the underlying glide representation does not always correspond to the stem vowel. In addition, the vowel-shortening rule does not apply in all cases. For instance, when a consonant-initial marking object is attached to the 3SGM form, the vowel remains long although the environment of vowel shortening rule is met (e.g. *xtaar-na* 'he chose us' cf. *xtar-na* 'we chose').

The JLA defective verb of all measures has the same inflectional pattern in the perfective. The 3SGM form has the final vowel *a* in all measures. This vowel is realized as a monophthong *ee* when a consonant-initial suffix is appended (e.g. *bana* 'he built' *baneet* 'I built' *ban-it* 'she built', *ʕadda* 'he went away' *ʕaddee-t* 'I went away', *ʕadd-it* 'she went away'). According to Harrama (1993), the presence of the stem vowel *ee* in defective verbs is due to the process of coalescence which applies to the underlying form of the verb (*banya*, *ʕaddja*) where the *ay* is realized as *ee*. On the other hand, the attachment of the vowel-initial suffix triggers the deletion of the glide *y* or *w* from the underlying form as they occur between two vowels. However, the dialect also has verb forms such as *lawa* ‘he twisted’ and *šawa* ‘he grilled’ in which the process of coalescence does not apply. Although the phonological account can provide a plausible account for some patterns of conjugation in JLA, it leaves many others unexplained.

3.3 Summary

This chapter has introduced a general description of the TLA verb system and the relationship between the different verb measures. Previous work on verb morphology of different LA varieties acknowledges the distinctive inflectional patterns and provides a useful source of comparison for the morphological paradigmatic complexity, most of which is challenging for purely phonologically based analyses. Some of these complexities are also a feature of TLA inflectional system, and I shall argue that these are most economically handled by means of an analysis in terms of stem-space and principal parts (§5.2).

Chapter 4 TLA stem complexity

In this chapter, we will discuss stem complexity in the TLA verb inflectional paradigm and consider the pattern of stem change across different measures. The stem pattern provides a new perspective on the relationship between verb measures, showing that some verbs that belong to different measures share the same stem change pattern while related measures within the same series have distinct stem behaviour. However, factors motivating the stem alternation can vary and result in morphologically complex stems with different sets of features. However, any account of morphological complexity requires addressing the methodological challenge of stem-affix boundaries, which is crucial for understanding the stem paradigmatic pattern.

In section 4.1 I review different accounts of stem-affix segmentation provided for two LA varieties (§4.1.1 and §4.1.2) and propose an account for TLA (§4.1.3). A framework for representing the stem change and stem properties will be presented in section 4.2 and section 4.4 . A discussion of the pattern of stem change in each verb series of the triliteral verb measures will be provided in section 4.3 . In section 4.5 , we will discuss the interaction of synchronic conditioning effects resulting in semi-autonomous TLA stems.

4.1 Segmentation

In a canonical approach, the morphs of a particular system are linearly organized, with each morph occupying a fixed position class and realizing a specific feature. However, the morphotactics of many languages exhibit a range of deviations from this ideal system. These include ‘fused/cumulated exponence’ in which a morph realizes more than one feature, ‘extended exponence’ in which two or more morphs realize the same feature and 'zero exponence' (absence of overt exponents for expressing some feature) (Crysmann & Bonami, 2012, p.1), as well as empty morphs, morphs with no semantic and syntactic function (Matthews, 1972). Generally, the identification of the position class of any complex word in any language may not be straightforward.[1]

With the aim of 'identifying the stem', Spencer (2012, p.88) provides a number of principles for establishing a segmentation algorithmically to determine how to assign the stem and affixes to position classes in a morphologically complex word. Before the application of these principles, we need to identify the conjugation of the morphologically complex word. The reason for this initial step is to identify the organization order/concatenation of similar affixes with respect to the stem in particular and the affixation of ICs in general. Firstly, the procedures of segmentation require treating all syncretism as a single cell. The next step involves determining the recurrent component pieces that realize the MPs identified by the 'templatic' organization of the system and subject to the linear order requirement. Faced with ambiguous segmentation of the remaining morphs/pieces, maximize the morphomic stem by the 'Stem maximization principle' (SMP) according to which if a recurrent word partial fails to realize unambiguously a single morphosyntactic property set assume that partial is part of a stem (pp.98-99).

A different approach to the economy of description is presented by Laporcaro (2012), arguing for an analysis that maximizes the endings (affixes). In his analysis of Sardinian, a Romance language closely related to Italian, he claims that the theme vowel has to be segmented as part of an inflectional ending rather than as part of the

[1] Determining the segmentation is crucial in the inferential-realization models in which morphologically complex words are decomposed into sub-components (morphs). However, in the word and paradigm models, the specific segmental composition of exponents and stems may not be a central issue since the MPs are properties of the whole word. Therefore, the stem is the sub-part of the word which can predict the maximum number of paradigm forms with the minimum number of functions (Montermini & Bonami, 2011).

stem as implied by the SMP. Therefore, he concludes that IC systems and/or the irregular verb allomorphy is most economically encoded in the endings. Laporcaro's exposition contains a number of misunderstandings relating to inferential-realizational morphology and stem-based analyses. In particular, it makes wrong assumptions about the way that stems are defined and deployed in inferential-realizational models.

Laporcaro (2012, p.24) considers the three conjugation classes of Sardinian regular verb, shown in Table 4.1, comparing the maximizing stems approach with his maximizing endings approach.

Maximizing stems

	1SG	2SG	3SG	1PL	2PL	3PL
1st class	ként-o	kénta-s	kénta-t	kantá-mos	kantá-des	kénta-n
2nd class	krésk-o	kréske-s	kréske-t	kreskí-mos	kreskí-des	kréske-n
3rd class	pált-o	pálti-s	pálti-t	paltí-mos	paltí-des	pálti-n

Maximizing endings

	1SG	2SG	3SG	1PL	2PL	3PL
1st class	ként-o	ként-as	ként-at	kant-amos	kant-ades	ként-an
2nd class	krésk-o	krésk-es	krésk-et	kresk-ímos	kresk-ídes	krésk-en
3rd class	pált-o	pált-is	pált-it	palt-ímos	palt-ídes	pált-in

Table 4.1 The present indicative verb paradigm for *kantáre* 'sing', *kreskér* 'grow' and *paltíre* 'leave'.

In the maximizing stems approach, the theme vowel is considered to be part of the stem in 1/2 SG and 3PL.[2] In 1st/3rd Class forms, the stems of the 2/3SG and the PL forms are syncretic (apart from stress). In the 1/2PL cells of the 2nd Class, however, a different stem is selected from that of the rest of the subparadigm (*kréske* ~ *kreskí*). Laporcaro argues that this shows that a maximizing endings approach is going to be more economical than a maximizing stems approach.

[2] In 1SG, the /-o/ does not recur elsewhere in the paradigm nor does it associate with particular feature(s), so the forms /kánto, krésko, pálto/ should really be analysed as morphomic stems, but we will follow Laporcaro's segmentation for purposes of exposition.

According to Laporcaro, in the maximizing stems approach, the stem *kreskí* will be derived from the form *kréske* by stress shift and final vowel substitution e ~ i. By contrast, the result of maximizing the affix is that all classes show stem syncretism in all the present indicative cells, indeed, the stem for all these cells is simply the verb's root form. Laporcaro (2012, p.25) claims that the maximizing endings approach is more economical since it avoids 'positing any ad hoc readjustment rule' for regular and irregular stem allomorphy. Consider for example the stem allomorphy of the verb *bier* 'see' in Table 4.2.

	1SG	2SG	3SG	1PL	2PL	3PL
Maximizing stems	bí-o	bíe-s	bíe-t	bidí-mos	bidí-des	bíe-n
Maximizing endings	bí-o	bí-es	bí-et	bid-ímos	bid-ídes	bí-en

Table 4.2 The present indicative verb paradigm for *bíer* 'see'

This verb has two stem variants: one with an i-ending (*bí-o*) and the other with a d-ending (*bid-ímos*). According to Laporcaro, this type of verb should be treated as a verb with an i-ending which is subject to the morphonological process of *d*-insertion to avoid hiatus (**bi-ímos* → *bid-ímos*). He next claims that the maximizing stems approach would have to assume that the stem alternants are derived by means of the phonological readjustment rules shown in (1) (Laporcaro's example (23)):

(1)

bíe + mos → (stress shift) bié+mos → (vowel substitution) bií+mos → (*d*-insertion) bidí+mos

This would lose the morphotactic generalization that the d-stem alternant is selected precisely when the following affix begins with /i/.

However, this is simply a misunderstanding of the way that stem formation and stem selection rules work in models like Paradigm Function Morphology (Stump, 2001b). Suppose for the sake of argument that the optimal grammar of Sardinian does indeed need to reflect directly the selection of the d-alternant and the i-initial suffix. In a model such as the Paradigm Function Morphology approach stem selection can be determined either purely in terms of morphological and morpholexical (morphomic) properties or (partially or fully) in terms of surrounding phonological context. That is, we can distinguish paradigmatic stem selection from syntagmatic stem selection. For the 2nd Class verbs (Table 4.1), one possible analysis (depending on the details of the rest of the verb system) would be to propose a Stem Formation Rule which defined an indexed allostem of the theme-vowel extended stem forms, along the lines 'if Xe is the form of Stem-n for a 2nd Class verb, then there is an allostem Xí'.[3] A paradigmatic rule of Stem Selection would then select the Xí (allo)stem in the appropriate morphosyntactic contexts. For verbs of the *bíer* subclass, we would simply define allostem alternants, |bi-|, |bid-| whose selection is determined syntagmatically in Stump's (2001) terms (§2.3.2). In other words, the Stem Selection Rule for these forms is sensitive to phonological form and selects the |bid-| alternant before /i/ and the |bi-| alternant elsewhere. To the extent that we need to relate the two allostem alternants explicitly we can set up a Stem Formation Rule which states that for any Stem-n form |X| in the relevant lexical class, where Stem-n denotes the index of the |bi(d)| stem, there is an allostem |Xd|. Such an approach is neither more nor less economical than the morphophonological solution provided by Laporcaro (though one could argue that it has the benefit that it does not commit us to a derivational model of phonology).

[3] Alternatively, we may simply define two distinct indexed stems, Xe and Xí, depending on how the rest of the paradigm patterns.

Laporcaro also raises the issue of what is the most economical way of distinguishing ICs. For Laporcaro, ‘economy’ is partly defined in terms of the number of times information distinguishing IC is encoded. This is a somewhat unclear notion, but we will assume that it presupposes a feature-counting metric. Laporcaro (2012, p.29) considers the infinitive forms in the regular 1st, 2nd, 3rd Class verbs shown in Table 4.3, segmented according to the assumptions of the two analyses:

	1st class	2nd class	3rd class
Maximizing stems	kantá-re	kréske-r	paltí-re
Maximizing endings	kant-áre	krésk-er	palt-íre

Table 4.3 The infinitive forms in the regular verb *kantáre* ‘sing’, *kreskér* ‘grow’ and *paltíre* ‘leave’

He shows that two infinitive endings need to be defined in Sardinian, *-re* for 1st/3rd Class, and *-r* for 2nd Class. He argues that, under the maximizing stems analysis, ‘IC information would have to be encoded twice, on both stems <…> and endings.’ He then claims that the maximizing endings analysis provides a more economical account since ‘IC contrasts are represented only once, on the ending, which is affixed to the lexical root, rather than to a stem containing a final vowel’ (p.29).

What this seems to mean is the following. On the maximizing stems analysis, we assume two infinitive endings, *-re/-er*, selected by realization rules sensitive to the inflectional class features of the verb lexeme. The Stem Selection Rules for the infinitive select a stem form, call it Stem2, which is defined as Xa, Xe, Xi for the three ICs (Table 4.1). The realization rules for Sardinian clearly need to mention the exponent of the infinitive property (say, [VFORM:infinitive]) and clearly need to distinguish the (presumably default) exponent *-re* from the 2nd Class exponent *-er*. To that extent we can say that the lexical class property ‘2nd Class’ is encoded in the grammar at least once. Depending on the exact details of Sardinian grammar, the Stem Selection Rules for the infinitive forms may or may not need to mention explicitly a morphosyntactic property of the form [VFORM:infinitive], because it could well be that the selection of this particular stem is the default for the verb system as a whole. In that case it would be very difficult to argue that the maximizing stems analysis led to an 'uneconomical' grammar. Let us suppose, therefore, that the Stem Selection Rules are obliged to refer to the property [VFORM:infinitive]. In Laporcaro's terms, this would be relatively 'uneconomical', because that property would be mentioned twice in the rule system. Intuitively, Laporcaro's point is that a grammar which appeals to multiple or extended exponence is in some sense more complex than one which can define exponence as a one-to-one mapping between properties and exponents.

By contrast, on the maximizing endings analysis advocated by Laporcaro there are three infinitive endings, *-are/-er/-ire*, each of which uniquely defines the IC of the verb root to which it attaches. As a result, a 'simple statement of exponence' results in 'relatively transparent affixes' (Laporcaro, 2012, p.29). However, it is difficult to see exactly what this implies for the grammar of Sardinian as a whole. First, being a member of an inflectional class is primarily a property of lexemes (and also of stems, as Aronoff, 1994, points out in detail). The only way that the infinitive ending itself could actually determine IC would be if the infinitive endings could attach freely to the root of any lexeme and hence define its IC membership. However, this is clearly not the case. It is the IC of the verb that determines its infinitive endings, not the other way around. In other words, given the assumptions of an inferential-realizational model, Laporcaro's claim here is simply incoherent. [4]

[4] This section benefited greatly from a number of comments from Prof. Spencer. A. I am grateful to Prof Brown, D. for raising this point.

There are a number of other misunderstandings in Laporcaro's exposition but what we have said here should be sufficient to demonstrate that his analysis of Sardinian is not relevant to the question of maximizing stems.

The aim of the following section is to review the segmentation analysis provided for different varieties of LA segmentation comparing these accounts with the first description provided for TLA based on two main principles: stem maximization and strictly morphomic stem hypothesis (Spencer, 2012). The segmentation that I propose for TLA (Chapter 4) favours an analysis that maximizes the stem since this approach can provide a plausible account for this inflectional system based on 'simple statement of exponence' (§4.1.3, and §4.1.1 on Owens §4.1.14.1.1). However, I also consider the option of maximizing the ending by treating monophthong and the FV as part of the inflectional affix (see §4.1.3) (see §4.1.2 on Harrama's view). I show that the latter view does not seem to be an economic account for the TLA system (§4.1.3 and §5.3.1).

4.1.1 Owens' (1984) segmentation

Owens' (1984, p.105 and p.156) segmentation analysis of ELA is a root and pattern-based account in which the linear order of the affixes follows the sequence in 0 which provides the morphotactic requirement (Spencer, 2012) for ELA. The first and second slots are specified for the negation particle ma- and person (imperfective) prefixes respectively. Starting from the verb root, preceding from left to right, the suffixes of the perfective and person occupy the fourth slot in the sequence. The number/gender properties are realized in the fifth slot and by accumulative affixes and the fifth position is followed by the indirect object marker (in Owens' analysis), then by the object suffix and finally the negation clitic.

(2) Negation(ma)-person-**root**-person-number/gender-l-object pronominal-negation(š)

However, the verb root can also be preceded by the imperfective marker while the imperative form in ELA only shows number/gender agreement realized as a suffix.[5] Owens (1984) provides a unified account of segmentation for all verb forms and verb series, emphasising the redundancy of the exponents/affixes in determining the inflection class of ELA (Table 4.4).

	Perfective		Imperfective		Imperative	
	SG	PL	SG	PL	SG	PL
1	-t	-na/ -na̲	n-	n-…-u/-a̲w		
2 M	-t	-tu	t-	t-…-u/-a̲w	-	-u/-a̲w
F	-ti	-tan	t-…-i/-ay	t-…-an/-a̲n	-i/-ay	-an/-a̲n
3 M	-	-o	y-	y-…-u/-a̲w		
F	-at/ -a̲t[6]	-an/ -a̲n	t-	y-…-an/-a̲n		

Table 4.4 Owens' (1984, p.104) segmentation of affixes for ELA verbs.

The segmentation above eliminates affix alternations across all verb measures/series mainly due to the segmentation of the FV, so that it is not part of the prefix. According to Owens (1984), the occurrence of the FV is phonologically conditioned, appearing before forms with an initial underlying consonant cluster, and hence can be found in *M*1 sound and defective verbs series. By contrast, in *M*1 hollow/double and non-*M*1verbs, prefixes in the imperfective are preceded or followed by an epenthetic vowel that is subject to various phonological processes.[7] Nevetheless, Owens (1984) classifies

[5] Similarly, in SA, the imperative is derived from the imperfective by deleting the person prefix but inserting a glottal stop and a vowel as a syllable-structure repair rule due to the ban on consonant clusters (Benmamoun, 1999).
[6] According to Owens (1984), the vowel of the 3SGF subject suffix /-a/ in ELA is sometimes conditioned by the prosody of the preceding syllable.
[7] For instance, the vowel can be realized as a glottal stop or result in lengthening of the stem initial consonant when C_1 is a nasal or a voiceless consonant (e.g. *n:dunn* 'I think'). In addition, in *M*5 verbs such as *y(i)-tikallim* 'you talk', the vowel following the prefix is an epenthetic vowel which never occurs in the imperative forms (Owens, 1984, p.26). Owens (2013, p.465), points out that in Moroccan Arabic the FV is always a schwa while Uzbekistan Arabic, Baghdadi and Cairene have /a/ for 1SG form and /i/ in 2/3SG forms.

the FV into two types based on the verb class. The first type is a v-class vowel that occurs in Class 4 verbs (Table 4.4). The second type is a V-class vowel found in Class 5 verbs (§3.2.1. (11)).[8] The variants in the suffixes *an* ~ *a̱n* are determined by vowel assimilation while the suffix alternation in *i* ~ *ay* and *u* ~ *aw* are morphophonologically conditioned determined by ELA verb class. Thus, the suffixes *i/u* are used in Class 4 while *ay/aw* suffixes associate with Class 5.

	Perfective		Imperfective		Imperative	
	SG	PL	SG	PL	SG	PL
1	kitab-i[9]t	kitab-na	n[10]*i*-ktib	n*i*-kitb-u		
2 M	kitab-it	kitab-tu	t*i*-ktib	t*i*-kitb-u	*i*ktib	*i*kitb-u
F	kitab-ti	kitab-tan	t*i*-kitb-i	t*i*-kitb-an	*i*kitb-i	*i*kitb-an
3 M	kitab	iktib-o	y*i*-ktib	y*i*-kitb-u		
F	iktib-at	iktib-an	t*i*-ktib	y*i*-kitb-an		

Table 4.5 Owens (1984, p.105) segmentation for ELA *M*1 sound series *kitab* 'write' (Class 4).

Although neither type of FV is not included as part of the prefix, Owens' classification of the FV in terms of his classification of vowel classes is not quite clear. He considers the FV of Class 4 in the sound verb series as an epenthetic v-class vowel which should not receive stress. On the other hand, the FV of Class 5 is treated as a V-class and it is always stressed. Nevertheless, in both cases the FV is actually stressed. In terms of the quality, the FV is realized as a high long vowel when it precedes a sound verb series which has y/w as the initial stem consonant (e.g. *wegaf* 'he stood up', *y(u:)guf* 'he stands up', *yebas* 'it dried up', *y(i:)bis* 'it dies up'). Likewise, the epenthetic vowel

[8] Recall from (§3.2.1) that a v-class vowel is a short vowel which can undergo deletion or vowel harmony and lengthening while a V-class vowel is non-elidable vowels.

[9] This is an epenthetic v-class vowel.

[10] Abu-Haidar (1985, p.193) points out that the use of the *n*- prefix as 1SG marker and the -*u* suffix as the 1PL marker in the imperfective is a characteristic of North African dialects. ELA seems to be a linking point between the North and Eastern dialect continuum by sharing some features of North African dialects and others of Eastern dialects.

which precedes the prefix *y-* in the hollow series is subject to regressive assimilation, hence the long vowel /i:/ in (e.g. *i:-guul* ‘he says’).

In addition, Owens (1984) treats the monophthong *ee,* which is a case of imala, in the defective series as part of the stem rather than the suffix. However, he suggests that the *ee*-realization can be a synchronic and morphophonologically conditioned process in the defective series in which the stem final vowel undergoes lengthening when followed by a consonant-initial suffix. The same inclination pattern in the form of a monophthong *ee* is found with the double verb series. However, unlike the defective series, the monophthong in this case is blocked with C-initial object suffixes, which raises the question of the segmentation of *ee* as part of the stem or part of affix (§4.1.3).

With regard to object suffixes, Owens (1984) provides the affix inventory in Table 4.3 for ELT. Unlike, subject suffixes, the vowel-initial object suffixes can co-occur with a stem final vowel in the defective series and the final vowel of a subject suffix (e.g. *bina-a* ‘he built it’, *yarjook* ‘they are waiting for you’ cf. *yarjo* ‘they are waiting’). The addition of the vowel-initial object suffix to defective verbs, leads to a segmentation issue (§4.1.3).

	SG		PL
1	-ni		-na/-na̱[11]
2 M	-a̱k	~ -k	-ka̱m
F	-ik		-kan
3 M	-a ~ -a̱ ~ V:/h[12]		-hum[13]
F	-ha ~ -ha̱		-hin

Table 4.6 The object suffixes in ELA (Owens, 1984, p.92).

11 The /a/ and /a̱/ vowels are low front and back vowels respectively. Therefore, the two suffix variants are a case of an allomorphic alternation determined by vowel harmony.
12 In ELA, the vowel of the vowel-initial subject suffixes is conditioned by front or back vowel harmony, while the vowel of the initial-vowel object suffix following 1SG and 2SG suffix *-t* in the perfective is always a front vowel.
13 This suffix can occur as a ‘self-standing pronoun’ (Owens, 1984, p.93).

4.1.2 Harrama's (1993) segmentation

In his study of JLA (§3.2.2), Harrama's (1993) segmentation is based on a morpheme-based account in which suffixes and prefixes are concatenatively attached to the verb stem and the suffixes in perfective alternate with zero morphemes in the corresponding imperfective forms. In addition, he suggests that the suffix /ʔi-/ is treated as an exponent of the imperative mood, which alternates with zero morphemes depending on the verb series.

	Perfective		Imperfective		Imperative	
	SG	PL	SG	PL	SG	PL
1	-t	-na	ni-…Ø	ni-…-u		
2 M	-t	-tu	ti-…Ø	ti-…-u	ʔi- ~ Ø… Ø	ʔi- ~ Ø…-u
F	-ti	/	ti-…-i	/	ʔi- ~ Ø…-i	/
3 M	Ø	-u	yi-…Ø	yi-…-u		
F	-it	/	ti-…Ø	/		

Table 4.7 Harrama (1993, p.91, p.106) segmentation of affixes for JLA verbs.

Unlike Owens' analysis, the affixal inventory provided by Harrama proposes that the FV is part of the prefix in the imperfective and imperative sub-paradigms, which can be subject to a number of phonological processes, including vowel harmony and syncope, according to which a high-unstressed vowel in an open syllable undergoes elision.[14] Therefore, the JLA segmentation analysis of the hollow and double verbs is in line with sound and defective series, by suggesting that the FV in the former set of series has undergone syncope.[15]

There are two issues with this proposed analysis. First, the syncope process suggested is problematic since the FV is retained in the PL.Imperf forms of the sound series (e.g. *tu-gutl-u* 'you kill' cf. (*i*)*t-maddid-u* 'you extended sth.') where although

[14] The front vowel /i/ is the default FV. Similarly, Watson (2002) treats the FV as part of the prefix for both San'ani and Cairene Arabic. Unlike Owens (1984), Harrama (1993) agrees with Elfitoury (1976) on the segmentation of the imperative sub-paradigm in which inflection prefixes have that shape *ʔi ~ Ø*.
[15] In this study, we treat the former case as vowel insertion.

the condition for the elision process is met, the FV is still appended to the prefix. Secondly, some FVs might assimilate with the imperfective stem vowel, yet counterexamples are not rare (e.g. *y-u-gtil* 'he kills' and *ya-rkub* 'he rides'), showing that the FV in *M*1 sound verbs is a morphological element (§3.2.2).

In addition, what is interesting about this segmentation is that both Harrama (1993, p.71) and Owens (1984) agree that the phonological extension *ee* is not part of the exponent suffix in the double and defective series. However, unlike Owens, Harrama assumes an underlying representation based on SA for the defective series ending in a glide that undergoes syncope when a vowel-initial suffix is appended.[16] With consonant-initial suffixes, on the other hand, the glide undergoes a process of coalescence with the second stem vowel; hence, the monophthong *ee* is part of the stem. We might agree that the monophthong *ee* is part of the stem, yet, suggesting that the synchronic process is to treat this long vowel as a glide can be problematic. The reason for rejecting this analysis is because *ee* does not surface as a glide even in participles (e.g. *nisee* 'he forgot' ~ *naasii* 'forgetting').

Likewise, the double series segmentation is based on a SA abstract form in which the phonological extension *ee* originates as the second stem vowel. This vowel undergoes various phonological processes in the surface representation, including metathesis with C3, vowel raising and lengthening in which the suffix consonantal stem is attached to the stem. By contrast, when the subject suffix is vowel-initial, then V2 of the stem is subject to a process of metathesis with C2 and assimilation with V1.[17]

However, Harrama's segmentation of the double and defective verbs may be problematic, for the two reasons. The first problem is considering SA forms to be the

---

[16] For Harrama (1993), an example of a JLA defective series ending in a glide can be *maš-ay* with CVCVGlide syllable structure rather than a vowel final stem *maša* 'walk'.

[17] 1. stem + C(V) radad+t *raddat* ~ *raddeet* 'I returned'
2. stem + V(C) radad+u *raadd* ~ *raddu* 'they returned'

input of JLA. Although it is logically possible that a child learning LA might be able to figure and use an underlying form based on the SA form, this assumption remains poorly motivated, given that there are few, if any, relevant studies of LA acquisition. Secondly, the monophthong *ee* in both series is treated as the second stem vowel affected by phonological adjustments. Although this might be a reasonable analysis for defective verbs, the monophthong in the double series is best treated as a phonological extension of the stem.[18]

4.1.3 The segmentation of TLA

The segmentation proposed for TLA verb forms deals with 'the segmentation problem' (Spencer, 2012, p.92) and the challenge of determining the position classes when they deviate from the canonical morphotactics (Stump, 1993; Montermini & Bonami 2011; Corbett & Baerman, 2006). In this study, we assume that the exponents are unified across measures and verb series. However, segmentation issues might be raised by the FV in the sound and defective series. The segmentation problems can also be caused by the monophthong *ee* in double and defective verbs, and by the defective verb final long vowel *aa* when object pronominals are affixed to the defective series.

[18] In her analysis of defective verbs in Cairene Arabic and San'ani Arabic, Watson (2002, p.144) treats the mid-long *ee* vowel as part of the verb stem. In Cairene Arabic, both vowels in the defective verb CVCV are identical in terms of quality. In the case of the i-final defective verbs, in Cairene Arabic, the final vowel undergoes vowel lengthening before a consonant-initial suffix (e.g. *nisi* 'he forgot', *nisii-t* 'You.M forgot'). By contrast, in San'ani Arabic, the final stem vowel in the 3SGM unsuffixed form is long; hence, it remains long in the rest of perfective sub-paradigm (e.g. *nisii* 'he forgot', *nisii-t* 'You.M forgot'). The a-final defective verbs in both Cairene and San'ani Arabic end with vowel *a* in the 3SGM form (e.g. *rama* 'he threw sth.'). However, the Cairene Arabic stem has *ee* before consonant-initial suffixes while San'ani has *ay* in the same phonological context. According to Watson (2002, p.145), historically, (as in Harrama's 1993 analysis), the root of defective verbs was represented as $CCC_{glide(y/w)}$ which in many modern Arabic dialects is replaced by a long vowel before consonant-initial suffixes as a result of a diachronic phonological process of coalescence of diphthongs. On analogy with defectives, double verbs before a C-initial suffix obligatorily have the long-mid vowel *ee* which is treated as 'an allomorphic adjunct'. Watson (2002) argues that both double *M*1 and sound/defective has the same base, which is a double trochee. In both Cairene and San'ani Arabic, pre-vocalically, the double verb stem in the perfective sub-paradigm has the shape CVCC. In San'ani Arabic, by contrast, double verbs have two stem variants CVCC+V ~ CVCVC+C. However, modern dialects, TLA and Cairene Arabic, employ a different strategy: on analogy with the defectives, a low long vowel /ee/ is added.

In section 3.1.1, we showed that TLA is a fusional system, in which portmanteaux morphs are found in adjacent slots. Verbs agree with the subject NP and object NP in person number and gender. TLA morphotactics for linear order requires the negation affix/clitic *ma-* to precede the subject affix on the left edge of the stem and the perfect person number/gender subject affixes precede object suffixes (Table 4.8):

1	2	3	**4**	5	6	7	8
Neg	Fut marker	Subj marker	**stem**	Subj marker	Ind obj marker	Obj marker	Neg
ma		y	**(i)** kitb	u	l	na	š

Table 4.8 The position classes of portmanteau morphs in TLA in *mayikitbulnaš* 'They do not write to us' (following the formatting in Crysmann & Bonami, 2012).

The 'templatic organization' in Table 4.8 is shared by all verb measures, thus, following Owens (1984), we provide the first segmentation account for TLA and suggest that all verbs inflect in the same manner, that is, there is unified exponence for all verbs (Table 4.9).

	Subject						Object[19]	
	Perfective		Imperfective		Imperative			
	SG	PL	SG	PL	SG	PL	SG	PL
1	-(i)t	-na[20]	n-	n-…-u			-ni	-na
2 M	-t	-tu	t[21]-	t-…-u	Ø	-u	-ik	-kum
F	-ti	-tin	t-…-i	t-…-in	-i	-in	-ik	-kin
3 M	Ø	-u	y-	y-…-u			-a(a)(h)[22]	-hum
F	-it ~ aat-	-in	t-	y-…-in			-ha	-hin[23]

Table 4.9 The segmentation of affixes for TLA verbs.

The affixes show a fusional pattern with regard to position class. The suffix /-t/ realizes the MPs of 1/2SGM in the perfective sub-paradigm. In the imperfective, the slot 3 prefix /t-/ realizes aspect and number. The two /t/ affixes are an example of homophony. Another observation about the TLA affix inventory is that the subject and object markers of 1PL are homophones, but occur in different slots: 5 and 7 respectively (Table 4.8).

The unified account of exponents might seem to suggest a simple segmentation procedure, but the FV and the monophthong can raise a ‘segmentation problem’. There are two possible FV segmentation analyses. First, one might assume that the vowel is part of the prefix as an exponent of MPs, as has been suggested for Cairene and San’ani Arabic in which the vowel is part of a fusional exponent expressing imperfective aspect/voice (Watson, 2002). Similarly, Harrama (1993) assigns the FV to the

[19] When object suffixes are added to subject suffix the V or C of the suffix undergoes lengthening (e.g. *kitib-u* ‘they.M wrote’, cf. *kitib-uu-hin* ‘they wrote them.F’ and *kitib-in* ‘they.F wrote’ cf. *kitibinn-a* ‘they wrote it.M’). According to Mitchell (1960, p.386) the reason for this adjustment process is to avoid and minimize the sequencing of ‘short open syllables’. Owens (1980) adds that the vowel change in the 3SGF suffix from high vowel /i/ to a low vowel /a/ when the 3SGM object is added /-aat-a/ seems to suggest that this suffix vowel originates as /-at/ across the whole paradigm. In TLA, this suffix is realized as a long low vowel /-aat/ when followed by the object suffixes /a/ and /Vk/.

[20] The 1PL subject and 1PL object suffixes show affix homophony (Spencer, 1991, pp.211-212).

[21] The prefix (t-) is subject to the phonological process of assimilation.

[22] Owens (1984, p.93) points out that, with regard to the 3SGM.obj marker, ‘in connected speech the lengthened vowel form without the *-h* is the more usual.’

[23] The 3PL.M.obj cannot combine with 2PL.F suffix **tin+hum* and instead the 2PLM subj and 3PLM.obj suffix combination (*tu-hum*) is used with feminine gender. However, the 2PLF and 3PLF suffixes (*tin-hin*) is a possible morph combination. This can be an example of an unmotivated affix syncretism, showing that the subject-object markers represent an inflectional paradigm.

subject/prefix position class, not as an exponent, but rather as a phonologically conditioned prefix allomorph which is subject to elision and vowel harmony. However, in TLA, the FV is lexically determined, as it can be a low back vowel /a/, high front /i/, or a high back vowel /u/. Therefore, treating the FV as part of the prefix would result in three lexical sets of exponents for the sound series (*a*, *i*, *u*) and two sets for the defectives (*i*, *a*). If we accept this analysis, then, similar to the Romance theme vowel (in Spanish for example), the FV would have the function of a theme vowel distinguishing three inflectional categories just within these two verb series. By contrast, the rest of the verb series and measures would form separate classes with a different set of exponents. However, this segmentation would violate the ‘economy of exponence’ (Spencer, 2012) in TLA in which *M*1 double and defective series and all non-*M*1 verbs have no FV and share the same exponents (Table 3.5 and Table 3.8). Therefore, a different analysis is required.

The second possible account is to consider the FV as a phonological extension of the stem; hence, it is not part of an affix. This analysis has been adopted for Maltese (Camilleri, 2013) because the FV is conditioned by the weight (in moras) of the stem, as seen by the syncretism across the perfective and imperfective sub-paradigms. In defective verbs, the FV would require this segmentation analysis, since syncretic stems can be found in a subset of this verb series (e.g. *y-a-nsa* ‘he forgets’ cf. *nsa-a* ‘he forgets us’) (Table 4.11). Therefore, in this case, the FV can be segmented as a phonological extension outranked by paradigmatic syncretism.

Although the phonological extension segmentation can apply to TLA, we suggest that the FV in sound series is part of a morphologically complex stem by the SMP (Spencer, 2012). The FV only occurs with *M*1 sound and defective series in the imperfective while in *M*5, *M*6 and *M*8 verbs, the prefixes are followed by the vowel *i* which we treat as an epenthetic vowel as it does not occur in the imperative. In *M*1, on the other hand, the FV is necessary in the imperative form of the relevant verb series even though MPs in the imperative are only expressed by suffixes and no prefixes are required. However, as we will see in §5.2.2, the stem of imperative is always parasitic on the imperfective in all measures including *M*1 hollow and double. Given that the FV is obligatory in the imperative form, then it is possible that the stem boundaries of the verb series in question should include the FV and the glottal stop is a consonant prothesis as a syllable structure repair process (Watson, 2002).

In addition, the FV in TLA can vary in terms of quality, as mentioned above, it can be *i*, *u*, or *a* in sound verbs and *i or a* in defective verbs.[24] Although the FV of the defective series is always in harmony with the stem vowel, it never assimilates to a suffix vowel (e.g. *y-a-nsa* ‘he forgets’ cf. *a-ns-i* ‘you 2SGF forget it!’ *a-ns-u* ‘you 2PLF forget!’). Likewise, in some cases of sound verbs, it can assimilate to the first or second stem vowel; yet, there are many examples where this is not the case. Therefore, the FV is lexically determined as compared with the stem vowel of the 3SG imperfective form (Table 4.10):

FV	2ndSV	Example	Gloss	FV	2ndSV	Example	Gloss
i	i	yi-ktib	‘he writes’	u	u	yu-rgud	‘he sleeps’
u	u	yu-rgud	‘he sleeps’	a		ya-rkub	‘he rides’
	i	yu-ġsil	‘he washes’	i	i	yi-ktib	‘he writes’
a	a	ya-smaʕ	‘he hears’	u		yu-ġsil	‘he washes’
	u	ya-rkub	‘he rides’	a		ya-ṭḥin	‘he grinds’
	i	ya-ṭḥin	‘he grinds’	a	a	ya-smaʕ	‘he hears’

Table 4.10 The vocalic pattern of the sound series FV in contrast with second stem vowel (2nd SV).[25]

[24] Unlike TLA, in Nigerian Arabic, the FV is always realized as a low vowel /a/ in the 1SG form, but alternates between the high vowel /i/ and /u/ in the 3SGM/PL forms. These different values of the FV can be found with the 3SGF and 1PL forms and second person form depending on the dialect. For example, it is invariably a high front /i/ in the Eastern dialect and always a low vowel /a/ in the southwest dialects (Owens, 1993, p.106).

[25] The sound verb of *M*7 has the prefix *Vn-* in the perfective and *CVn-* in the imperfective (Owens, 1984, p.122). In the perfective of this measure, the prefix vowel is always /i/, but in the imperfective, the prefix vowel can be /i/ or /u/.

If the stem vowel is a low *a*, then the FV will be a low *a*.[26] Likewise, if the FV is a high front vowel *i*, then the stem vowel will be *i*.[27] However, if the FV is a high vowel, then the second stem vowel cannot be low. By contrast, a high stem vowel <u>can</u> occur with low FV. The segmentation of the FV in sound series as part of the stem is favoured since it does not affect any systemic syncretism relation within the inflectional paradigm. Therefore, applying the SMP does not affect the dependency relations between sub-paradigms, but rather it results in a *complete* segmentation. In other words, maximizing the stem with the FV can be the economic approach, as it does not induce a different indexing for the stem (§5.2).

Similarly, a problem might be encountered with the segmentation of the monophthong *ee* in the *M*1 double verbs and in *M*1/non-*M*1 defective verbs (Table 4.11). The monophthong *ee* has a syntagmatic distribution across the perfective sub-paradigm when the verbs are marked for subject. In other words, the presence of this vowel seems to be associated with the phonological context of the subject suffix: the monophthong is found before a consonant-initial suffix in both double (*M*1 *gallee-na* 'we took') and defective verbs (*M*1 *nsee-na* 'we forgot). In the defective verbs, the monophthong *ee* is also present when no suffix is added (*M*1 *nsee* 'he forgot'). Otherwise, this vowel is not observed (Table 4.11).[28] Unlike double verbs, in defective

*M*7 CVn	*M*1 (FV)	pass.imperf	act.imperf	Gloss
i	i	y-in-kitib	yi-ktib	write
i	a	y-in-dibaḥ	ya-dbaḥ	slaughter
u	u	y-un-gutil	yu-gtil	kill
u	a	y-un-ṣubaġ	ya-ṣbaġ	dye

[26] There is a historical explanation for some of the arbitrariness. The *a* FV goes back to a stem CCaC, which apparently is still maintained when there is guttural at C3. This is a rather important development, showing that TLA is different both from ELA and from Tripoli LA (Owens, personal communication).
[27] In modern dialects, the FV *i* is preserved from Proto-Semitic and this phenomenon is known as *taltalah* (Versteegh, 2014, p.187).
[28] In Cairene Arabic, the double verb alternates between CVCVC and CVCC before Vowel-initial suffix and Consonant-initial suffix respectively (Watson, 2002).

verbs, the monophthong of the 3SG unsuffixed form alternates with *i* or *a* in the imperfective sub-paradigm of all measures (*M*1. *nseet* ~ *na-nsa* 'forget 1SG', *M*2 *bakkeet* ~ *nbakki* 'upset') and with the vowel *a* in all defectives of all measures with object markers (Table 4.11).

		Imperfective		Perfective				
		subject		subject			object	
Verb	Gloss	1PL	3SGM	1Pl.Perf	3SGM	3SGF	1PL	3SGM
gall	take	n-gill-u	y-gill	gallee-na	gall	gall-it	gall-na	gall-a
radd	return	n-rudd-u	y-rudd	raddee-na	radd	radd-it	radd-na	radd-a
nsee	forget	n-a-ns-u	y-a-nsa	nsee-na	nsee	ns-it	nsa-a-na	nsa-a
nfee	exile	n-i-nf-u	y-i-nfi	nfee-na	nfee	nf-it	nfa-a-na	nfa-a
bakka	upset	n-bakk-u	y-bakki	bakkee-na	bakka	bakk-it	bakka-a-na	bakka-a

Table 4.11 The distribution of the monophthong *ee* in TLA.

Similar to the FV, the monophthong *ee* can have two possible segmentation analyses. It can be part of the stem or the affix. For double verbs, treating the monophthong as a phonological extension (as an adjunct) of the stem is a possible analysis that preserves stem syncretism within the perfective sub-paradigm. Nevertheless, the object markers (including the object suffix -*na*) do not induce a phonological extension when attached to the unsuffixed 3SGM form (*gall-na* 'he took us' cf. *gallee-na* 'we took') although the phonological context for the extension is met.[29] Therefore, for TLA, we opt for segmenting the vowel *ee* as part of the stem.[30] Likewise, in the defective series, the vowel *ee* can be treated as a stem final vowel.

Further segmentation issues might arise when object suffixes are appended to the stem in the defective series. Consider the verb *nfee* 'exile', for example (Table 4.11). Before a consonant-initial suffix (e.g.1PL.obj -*na*), the monophthong *ee*

[29] It is also worth noting that that the monophthong is not deleted in the following environment *maddee-t-a* 'I passed it SG.M'.
[30] The arbitrary unmotivated stem choice in double verbs (e.g. *gall* ~ *gallee*) shows that the subject/object paradigm is inflectional (§4.3.3 and §4.5.4).

alternates with the long vowel *aa* (e.g. *nfa-a-na*). However, the same pattern is found with object vowel-initial suffixes *-a* and *-Vk* (e.g. *nfa-a* and *nfa-ak*), leading to a segmentation ambiguity.

The question now is should the ending *aa* be part of the stem, part of the affix or divided between the stem and affix? Or is it the case that we have two different stems; one with a long vowel before a consonant-initial suffix and the other with a short vowel and in harmony with the vowel-initial suffix. We proposed that the ending *aa*, before consonant-initial suffixes, is best analysed as (CCa-a) with the second vowel being a phonological element due to the morphophonological process of vowel lengthening. Therefore, the vowel *-a* is neither part of the stem nor of the affix before consonant-initial suffixes. In other words, before consonant/vowel-initial object suffixes, there is only one stem-shape, (CCa). There are two reasons for this segmentation. First, the object marker *-a* is always clearly associated with the exponents 3SGM/dative, and when this marker is added to the 3SGM form, the result is (CCaa); hence, the segmentation (CCa-a). The second reason is that we are able to obtain the stem/partial syncretism with the perfective sub-paradigm when with obj-markers. In other words, the SMP should not outrank the segmentation of pieces on the right edge of the defective stem, which result from vowel lengthening process.

Clearly, the decomposition issues encountered in TLA appear to be clustered mainly around the stem-affix boundaries rather than affix-affix boundaries. Preserving the stem boundaries has priority over the affix. Therefore, TLA segmentation algorithm has to consider stem-segmentation first ('stem-insertion first') in contrast to affix-segmentation (Crysmann & Bonami, 2012). In addition, it has to take into account the SMP (Spencer, 2012), but also maintain the stem extension segmentation to achieve a 'simple statement of exponence' and avoid complicating the assignment of inflection classes because of increasing the number of indexed stems.

(2) Initial TLA segmentation algorithm given the full paradigm

a. Identify systematic syncretism (Spencer, 2012)
b. Identify a potential syncretic stem (across the whole paradigm)
c. Identify potential affix boundaries (affix-segmentation)
d. Identify affixes realizing a MP (across ICs) and cluster them
e. Cluster systematic partial affixal syncretism (right-to-left)
f. Maximize the stem (SMP) iff (f) does not violate (b)
   (e.g. for FV (y)-{u(gtil)}).
g. Segment a subcomponent near stem boundaries as a phonological extension iff (b) is violated
   (e.g. for defective [nsa]-{a}-(n.a) 'he forgot us')
   (e.g. for defectives (y)-{a}-[nsa] 'he forgets')

Having identified TLA stem-affix boundaries in verb inflectional paradigms, the next section will discuss the stem-shape and stem allomorphy in all measures of TLA.

4.2 Representing stem change

TLA verbal paradigms display a systematic stem variation that involves changes to the 'stem-shape' by altering the syllable structure, ablaut and addition of the 'monophthong' extension. These stem changes can be motivated by morphological factors (Maiden, 2009, 2011; Bonami & Boyé, 2002) and extramorphological factors including phonological (Anderson, 2011), morphophonological (Spencer, 2013b; Maiden, 2011) and morphosyntactic (Baerman & Corbett, 2012). In this section, and following (Boyé & Cabredo-Hofherr, 2006), the stem change includes all patterns of stem-shape variations. The stem variation pattern will be represented by the metaparadigm in which every stem that marks a stem change has a different letter. It is worth noting that this representation is different from the 'stem-space' in which only indexed stems that are morphologically motivated are represented (see Chapter 5, §5.2 and §5.2.2). The reason for using this framework for classifying TLA stems is that it can determine the factors that control the change and provide a deeper understanding of the theoretical status of the stem.[31]

To establish the characteristics of this framework, we consider the paradigm of the regular verb types in the language and that is *M*1 sound series. For instance, the inflectional paradigm of the verb *ġesal* 'wash' shows the different stems across imperfective, imperative and perfective (Table 4.12). The pattern of the change can also vary depending on whether the verb is marked by subject affixes or object suffixes as show in the meta-paradigm of *ġesal* (Table 4.13).

MPs	Subject			3SGM.Obj	
	Imperfective	Imperative	Perfective	Imperfective	Perfective
1SG	n-uġsil		ġesal-t	y-uġsil-ni	ġesal-ni
1PL	n-uġusl-u		ġesal-na	y-uġsil-na	ġesal-na
2SGM	t-uġsil	uġsil	ġesal-t	y-uġusl-ik	ġusul-ik
2SGF	t-uġusl-i	uġusl-i	ġesal-ti	y-uġusl-ik	ġusul-ik
2PLM	t-uġusl-u	uġusl-u	ġesal-tu	y-uġsil-kum	ġesal-kum
2PLF	t-uġusl-in	uġusl-in	ġesal-tin	y-uġsil-kin	ġesal-kin
3SGM	y-uġsil		ġesal	y-uġusl-a	ġusul-a
3SGF	t-uġsil		ġusul-it	y-uġsil-ha	ġesal-ha
3PLM	y-uġusl-u		ġusul-u	y-uġsil-hum	ġesal-hum
3PLF	y-uġusl-in		ġusul-in	y-uġsil-hin	ġesal-hin

Table 4.12 The paradigm of *M*1 sound series *ġesal* 'wash'.

[31] Camilleri (2014, p.39) developed a similar approach for Maltese stem alternations as 'a-theoretic reference to a pattern of distinctly-indexed stems'. By contrast, in this study, the stem change does not always necessary associate with a distinct stem index, but it is essential to show all the interactions of the different stem conditioning factors. Stem indexing in TLA will be discussed in Chapter Five.

MPs	Subject			3SGM.Obj	
	Imperfective	Imperative	Perfective	Imperfective	Perfective
1SG	A		C	A	C
1PL	B		C	A	C
2SGM	A	A	C	B	D
2SGF	B	B	C	B	D
2PLM	B	B	C	A	C
2PLF	B	B	C	A	C
3SGM	A		C	B	D
3SGF	A		D	A	C
3PLM	B		D	A	C
3PLF	B		D	A	C

Table 4.13 The stem metaparadigm of the verb *ġesal* 'wash'.

The paradigms in Table 4.12 and Table 4.13 show that within the imperfective sub-paradigm with subject suffixes, the stems are in allomorphic relations and these same stems are used in the imperative.[32] Likewise, within the perfective sub-paradigm with subject markers, the 1/2 person forms share the same stem while the 3rd person forms show two stem variants. With regard to inflectional forms with object markers, stem change can be observed when an object marker is attached to any of the unsuffixed forms such as the 1SG, 2SGM, 3SG imperfective forms and the 3SGM perfective form. However, no stem change is triggered when an object suffix is appended to the verb, but preceded by a subject marker. In addition, the listing of all possible stem changes is redundant due to the systematic predictability across paradigm cells (Montermini & Bonami, 2011; Bonami & Boyé, 2008; Boyé & Cabredo-Hofherr, 2006). Therefore, to avoid the redundancy, we will conventionally choose the forms in Table 4.14 for each sub-paradigm as representatives of a set of cells with the same stem.[33] The stem

[32] The imperative and the future forms are always parasitic on the imperfective forms. Therefore, the same stems are selected in cells with the corresponding person, number, gender MPs.
[33] These selected cells will also be essential for the discussion of the optimal principal parts and the implicative relation among different measures introduced in (chapter 5).

selected in these cells will be the basis for the pattern of stem change in each verb measure/series.[34]

Subject					3SGM.Obj	
Imperfective		Perfective			Perfective	
3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM

Table 4.14 The representative forms for TLA stem change.

For the imperfective sub-paradigm, the first representative cell is the 3SGM that represents every unsuffixed form and any form with a consonant-initial object suffix. The second selected cell is that of the 3PLM form which includes all forms with a vowel-initial subject or object suffix. Across the perfective sub-paradigm, the 1PL form shares the same stem with the 1SG and 2 person forms. The reason for listing this form is that it is essential for illustrating the contrast between the stem patterns when a *M*1 double verb is marked with the 1PL.obj suffix.[35] Although the 3SGF/PL forms share the same stem in all verbs, the 3SGM can have a variable pattern depending on the verb series and measure. Therefore, the cells of both 3SG forms will represented. Finally, the representative paradigm will also list the stem changes of the unsuffixed perfective form 3SGM form when the 3SGM and 1PL object markers are attached to the form. The reason for representing the 3SGM.obj sub-paradigm, is that it shows distinctive pattern of stem change distinguishing different sub-patterns within the same verb series (§4.3.1).

[34] Note that the sub-paradigm represented by the 3SGM.obj form will not be represented with intransitive verbs, *M*7/8 verbs (e.g. *inkesar* ‘it was broken’) or M6 verbs (e.g. *telaaga* ‘they met each other’).
[35] See footnote 19.

4.3 TLA stem change

4.3.1 Sound series

The set of lexemes that belong to the *M*1 sound series (CVCVC) are the regular set of verbs. However, this verb series displays a variety of distinct stems which are chosen in particular cells within their inflectional paradigm. The data in Table 4.15 present an example of the stem changes in ten transitive and intransitive sound verbs.

verb	Gloss	Subject					3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
ʕegal	hide	y-iʕgil	y-iʕigl-u	ʕegal-na	ʕegal	ʕigil-it		
ʕegal	mature	y-aʕgil	y-aʕigl-u[36]	ʕegal-na	ʕegal	ʕigil-it		
semaʕ	listen	y-asmaʕ	y-asimʕ-u	semaʕ-na	semaʕ	simiʕ-it	semaʕ-na	simiʕ-a
ḍerab	hit	y-uḍrub	y-uḍurb-u	ḍerab-na	ḍerab	ḍurub-it	ḍerab-na	ḍurub-a
getal	kill	y-ugtil	y-ugutl-u	getal-na	getal	gutul-it	getal-na	gutul-a
šerab	drink	y-ašrub	y-ašurb-u	šerab-na	šerab	šurub-it		šurub-a
ṣehad	burn	y-aṣhid	y-aṣuhd-u	ṣehad-na	ṣehad	ṣuhud-it	ṣehad-na	ṣuhud-a
refaʕ	carry	y-arfaʕ	y-arufʕ-u	refaʕ-na	refaʕ	rufuʕ-it	refaʕ-na	rufuʕ-a
werat	inherit	y-uurit	y-uurt-u	werat-na	werat	wirit-it	werat-na	wirit-a
weṣaf	describe	y-uuṣuf	y-uuṣf	weṣaf-na	weṣaf	wuṣuf-it	weṣaf-na	wuṣuf-a
*M*1 Stem change		A	B	C	C	D	C	D

Table 4.15 The stem change in *M*1 sound series.

The paradigm of the TLA regular verb is based on four abstract stems that vary from one set of cells to another within the same paradigm (Table 4.15). In fact, these stems differ in two respects: they exhibit different stem-shapes and vowel alternation. For instance, both stems C and D are based on stem-shape (CVCVC) showing that, unlike many varieties of LA, a sequence of two open syllables is allowed in TLA in the sound

[36] cf. *ya-ʕurg-u* 'they sweat' in which the stem vowel is a high back vowel while FV is the low vowel *a*

series.[37] Therefore, when a C-initial suffix is added to the stem of the 1SG/2SGM perfective forms, the TLA requires a phonological readjustment rule of (CVCVC) of epenthetic vowel (CVCVC-v-C) to retain the same stem-shape across the 1/2 and 3SGM cells; hence, stem C. Within the imperfective, the stem-shape is different from that of the perfective; hence, stems A and B have the stem-shapes (VCCVC) and (VCVCC) respectively. According to Owens (1980), these stems are subject to phonotactic constraints by which the syllable structures *(C-VCV_1CV_2C) and *(C-VCV_1CV_2C-V) appear to be disallowed. Therefore, the second vowel is deleted before a V-initial suffix (stem B) and the first vowel is deleted elsewhere (stem A). It seems that a sequence of two short open syllables is allowed in stem D, but not in stem A. In fact, Watson (2002) points out that each of the perfective and imperfective paradigms is based on a different template. Therefore, we propose that each stem in the imperfective and perfective sub-paradigms has a distinct stem-shape since there is no evidence for a synchronic process for deriving one stem from another (Table 4.16):

	Stem-shape
Stem A	VCCVC
Stem B	VCVCC
Stem C	CVCVC
Stem D	CVCVC

Table 4.16 The stem-shape in *M*1 sound series.

In addition, the stem-shape of the sound series is further distinguished by two types: simple and complex stems, dividing the aspect paradigms: stem A and B are complex while stems C and D are simple. The imperfective stems are complex since they have

[37] According to Mitchell (1960), in a sequence of two open syllables, one of the vowels has to undergo elision; hence, the structure *CVCVCV is forbidden in ELA. Likewise, in Tripoli LA, according to Al-Ageli (1995), short unstressed vowels in open syllables are syncopated due to the restriction on open syllables. Owens (1980, p.111) describes the second vowel of stem D as an ‘epenthetic vowel peculiar to Misurata’.

an extra component that the FV added to the stem. However, unlike the phonologically specific concrete stems in other languages, such as the b-stem in Latin (Aronoff, 1994), the dental stem in Western Germanic (Blevins, 2003) and the *r* stems in Spanish (Spencer, 2012), the FV is a lexical component, i.e. the quality of the vowel is lexically conditioned (cf. the monophthong /ee/ in other series). Therefore, stems A and B are phonologically specific vowel-initial complex stems with two bases (CCVC and CVCC). In addition, the stem change can be distinguished in terms of ablaut. Unlike stem C, each of stem A, B and D shows an ablaut change pattern which is lexically conditioned apart from two cases (assimilated verbs e.g. *wegaf* 'sand up' and verbs ending in *h*, *ḥ* or *ʕ*) (§4.5.2). In stem A, the vowel can be either high front/back vowel or a low back vowel whereas the stem vowel(s) of stem B and D can only be a high vowel (Table 4.15). A similar pattern of stem change can be observed in *M*7 and *M*8.

In section 3.1 , we showed that the relationship between *M*1 and *M*7/*M*8 is traditionally assumed to be derivational, based on the grammatical function of (semi)passive. However, the correspondence between these measures does not always match in meaning or it is completely absent. Unlike *M*1, the pattern of stem change in these measures does not involve altering the stem-shape, and it is only distinguished in terms of ablaut (Table 4.17):

verb	Gloss	Subject			
		Imperfective		Perfective	
		3SGM	3PLM	3SGM	3SGF
insemaʕ	be heard	y-in-simaʕ	y-in-simiʕ-u	in-semaʕ	in-simiʕ-it
inḍerab	be hit	y-un-ḍurub	y-un-ḍurub-u	in-ḍerab	in-ḍurub-it
ingetal	be killed	y-un-gutil	y-un-gutul-u	in-getal	in-gutul-it
inšerab	be drunk	y-in-šurub	y-in-šurub-u	in-šerab	in-šurub-it
inḥebas	be imprisoned	y-in-ḥibis	y-in-ḥibis-u	in-ḥebas	in-ḥibis-it
inkesar	be broken	y-un-kusur	y-in-kusur-u	in-kesar	in-kusur-it
*M*7 Stem change		A	B	C	B
irtefaʕ	be carried	y-i-r<t>ifaʕ	y-i-r<t>ifiʕ-u	i-r<t>efaʕ	i-r<t>ifiʕ-it
ilteḥam	be soldered	y-i-l<t>iḥim	y-i-l<t>iḥim-u	i-l<t>eḥam	i-l<t>iḥim-it
intefax	be bloated	y-u-n<t>ufax	y-u-n<t>ufux-u	i-n<t>efax	i-n<t>ufux-it
*M*8 Stem change		A	B	C	B

Table 4.17 The stem change in the sound series of *M*7 and *M*8.[38]

Both *M*7 and *M*8 show three stem changes. In fact, the stems in *M*7 can be associated with *M*1 perfective stems since they all can share the same stem-shape (CVCVC) and in some cases the same/similar ablaut change. Therefore, *M*7 stem A has the same stem-shape (CVCVC) as that of *M*1 stem D (CVCVC), but the ablaut of these stems is not always the same. By contrast, stem C of both measures (e.g. *getal* cf. *in-getal*) and both *M*7 stem B and *M*1 stem D (*gutul-it* cf. *in-gutul-it*) share the same stem-shape and ablaut. A same relationship exists between the stems of *M*1 and *M*8 (Table 4.15 and Table 4.17).

In terms of stem complexity, the stem pattern in the imperfective sub-paradigm, in contrast to *M*1, is based on basic stems with no phonological extension. Clearly, the main characteristics of the stem change in these measures is ablaut while the stem shape

[38] The 1 and 2 person perfective forms of these measures share the same stem as that of the 3SGM. For example, *inḥebas-na* 'be prisoned', *intefax-na* 'be bloated/be praised/gain weight', *inkeṣar*-na 'be broken, be humiliated' and *inkemaṭ*-na 'suffocate'. These person forms of these measures are also used in folk music and poems.

is shared across the whole paradigm. However, not all sound series measures distinguish stems by ablaut.

Unlike *M*1 and *M*7, the pattern of stem change in the sound series of *M*2, *M*3, *M*6 and *M*10 only distinguishes stem variants by related stem-shape alternation as shown in Table 4.18. In section 3.1 , we showed that *M*5 and *M*6 can be derived from *M*2 and *M*3 respectively by adding the affix *-t*. Therefore, these measures are effectively showing the same stem change. Although these measures have different syllable structure in their stems, they follow the same pattern of stem change. Therefore, it is possible that the relationship between this unified pattern and the measures is a morphomic relationship (Bonami & Boyé, 2008).

verb	Gloss	Subject					3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
ʕallim	teach	y-ʕallim	y-ʕallm-u	ʕallim-na	ʕallim	ʕallm-it	ʕallim-na	ʕallm-a
ʕaamil	deal	y-ʕaamil	y-ʕaaml-u	ʕaamil-na	ʕaamil	ʕaaml-it	ʕaamil-na	ʕaaml-a
teʕallim	learn	y-teʕallim	y-teʕallm-u	teʕallim-na	teʕallim	teʕallm-it	teʕallim-na	teʕallm-a
teʕaamil	deal	y-teʕaamil	y-teʕaaml-u	teʕaamil-na	teʕaamil	teʕaaml-it	teʕaamil-na	teʕaaml-a
staʕmil	use	y-staʕmil	y-staʕml-u	staʕmil-na	staʕmil	staʕml-it	staʕmil-na	staʕml-a
stanjid	plead	y-stanjid	y-stanjd-u	stanjid-na	stanjid	stanjd-it	stanjid-na	stanjd-a
Stem change		A	B	A	A	B	A	B

Table 4.18 The stem change in the sound series of *M*2, *M*3, *M*5, *M*6 and *M*10 respectively.

The examples in Table 4.18 show that the stems in this pattern are distinguished from each other only in one respect, namely the presence or the absence of the final stem vowel. Thus, stem A is chosen in the unsuffixed forms or before a consonant-initial suffix while stem B occurs before a vowel-initial suffix.[39] In fact, this stem pattern is the most *regular*[40] or the less complex stem pattern with the minimum stem change exhibited by the most common measure (*M*2) in contrast with the most irregular or most complex stem pattern with the maximum stem changes displayed by most common measure *M*1 within the regular verbs (sound series).[41] In the *M*1 sound series, there seem to be restrictions on the process of syncope since the short unstressed vowel in the 3SGF has to be retained (*simiʕ-it* 'she heard'). By contrast, in these measures, the dialect seems to not only allow the process of syncope of unstressed short vowels in open syllables, but also it behaves as a differential syncopating type of dialect, by only allowing the elision of a short high front vowel.[42]

Clearly, the stem change pattern within the sound series can vary across measures with two stem variants to others with up to four different stems. The factors inducing the stem change in TLA will be discussed in detail in section 4.5 . In the next subsection, we will consider the complexity of stem changes and the nature of each pattern in the defective series.

---

[39] The same stem pattern is selected when the 2SG object marker *-ik* is attached to the stem as in *kallm-ik* 'he called you' in contrast with the 1SG form *kallim-(i)t* 'I called you' in which the vowel preceding the suffix is epenthetic. This could be taken to show that the 2SG object marker is indeed a vowel-initial suffix with no affix allomorphy (*k~ik*) as it has been suggested for ELA (Table 4.6).

[40] Bonami and Boyé (2002, p.52) define regular inflection as 'the phonology of a regular inflectional form is a function of the phonology of some stem of the verb this form instantiates'. According to this definition, regular verbs select the same stems across the whole paradigm while irregular verbs have multiple stems. However, for TLA, the characteristic of regular verbs is to have multiple stems.

[41] This observation might also have more general implications for the canonical paradigm in which stems should be the same (Corbett, 2009), but clearly in TLA the prototypical verb (sound series *M*1) has a stem pattern in which stems are not the same (ABCD).

[42] Dialects have been classified into two types based on the quality of the vowel which undergoes syncope in open syllables. The first type is differential dialects (e.g. Egyptian Arabic (Kenstowicz, 1980)) in which only the high vowels /i/ and /u/ undergo elision whereas in non-differential dialects (e.g. Tripoli LA, Al-Ageli, 1995) the short vowel is deleted irrespective of the vowel quality. This dialect distinction is attributed to Cantineau (1939).

4.3.2 Defective series

In section 3.1 , we said that verbs of the defective series are traditionally classified as weak, because their phonological form consists of just two consonants and they end in a vowel. Similar to the sound (§4.3.1) and double series (§4.3.3), the stem-shape of this series can also be divided into basic and complex. However, unlike any of the other series, the stem-shape can be complex in two ways by having a phonological extension at both edges of the stem. Thus, the complex stem has the FV stem-initially in the imperfective and has the monophthong *ee* (imala)[43] stem-finally in the perfective. In addition, the stem change in this series is the result of the final vowel ablaut and presence or absence of the (final) stem vowel (Table 4.19):

verb	Gloss	Subject					3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
n(i)see l(i)gee	forget found	y-a-nsa y-a-lga	y-a-ns-u y-a-lg-u	nsee-na lgee-na	nsee lgee	ns-it lg-it	nsa-a-na lga-a-na	nsa-a lga-a
Stem change 1		A	B	C	C	B	A	A
n(i)fee k(i)ree	exile hire	y-i-nfi y-i-kri	y-i-fn-u y-i-kr-u	nfee-na kree-na	nfee kree	nf-it kr-it	nfa-a-na kra-a-na	nfa-a kra-a
Stem change 2		A	B	C	C	B	D	D

Table 4.19 The stem change in *M*1 defective series.

The paradigm of *M*1 is based on a set of three stem-shapes: (CCV), (CC), and (CCee). While stem-shapes (CC) and (CCV) are found across the sub-paradigms, the stem CCee can only associate with the perfective sub-paradigm, representing distinct cells with clashing MPs (Table 3.7). The stem alternation classifies *M*1 of this series into two sub-sets of stem change based on the vocalic pattern of the imperfective and the systematic stem syncretism. Thus, *M*1 defective verbs differ in terms of the ablaut in the

[43] In TLA defective verbs, the monophthong is possibly a diachronic process of *imala* (the raising of the stem final vowel *a* to *ee*). In JLA, the defective verbs have /a/ stem final vowel. By contrast, the *ee* in TLA double verbs is not a case of *imala* (J. Owens, personal communication, February 5, 2015). In this work, I will refer to the two cases as a monophthong.

imperfective sub-paradigm in which the stem vowel of the 3SGM can be either *a* or *i*. The a-vowel final verbs can have three stem changes while the i-vowel final verbs have four. The paradigmatic behaviour of the verb former group is characterised by syncretic stems which are found across the aspect sub-paradigms. Thus, stem A fills the cell of the 3SGM imperfective, and it is also used for the 3SGM.obj sub-paradigm when the exponent is a vowel initial object marker. By contrast, the i-vowel final verbs have more stem variants. This stem complexity is due to the absence of the stem syncretism; thus, it has stem D in contrast with stem A of the a-vowel final verbs through ablaut. Despite the ablaut difference between these two verb sub-sets, they both have the same sub-pattern of stem change, that is, ABCCB when the exponents are subject markers. In addition, they both have a monophthongal induced stem change across 1/2 and 3SGM forms and similar paradigmatic behaviour can be observed across the other measures of this series.

Nevertheless, not all measures with this verb series can have the dual stem complexity of *M*1 defective verbs. In addition, different measures can share the same stem pattern despite their contrastive/different sets of stem-shape. For instance, the stem variation within *M*7 and *M*8 correlates considerably with that of *M*1. Both *M*7 and *M*8 have four stem-shapes: CiCV, CiC, and CiCee with the same paradigmatic distribution as that of *M*1: CC is used before a V-initial suffix and CiCee is used before a C-initial subject marker. By contrast, the stem-shape CCV is only found in the imperfective and it has the same ablaut change as *M*1 verbs. The following paradigm illustrates the stem change pattern in *M*7 and *M*8:

verb	Gloss	Subject				
		Imperfective		Perfective		
		3SGM	3PLM	1PL	3SGM	3SGF
indiʕee inkiwee intisee	be infected cauterised be forgotten	y-in-diʕi y-in-kiwi y-i-n(t)isa	y-in-diʕ-u y-in-kiw-u y-i-n(t)is-u	in-diʕee-na in-kiwee-na in(t)isee-na	in-diʕee in-kiwee i-n(t)isee	in-diʕ-it in-kiw-it i-n(t)is-it
Stem change		A	B	C	C	B

Table 4.20 The stem change in the defective series of *M*7 and *M*8 respectively.

Clearly, the pattern of stem change in these measures is parasitic on that of *M*1, although the stem-shape of *M*8 can be differentiated from *M*1 in having an affixal consonant within the stem.

Unlike *M*7 and *M*8, the rest of the measures of this series display different but related stem change pattern based on ablaut and the monophthong *ee*. Consider the paradigm of *M*2, *M*3, *M*5, *M*6, and *M*10 (Table 4.21):

	Gloss	Subject					3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
bakka naada	upset invite	y-bakki y-naadi	y-bakk-u y-naad-u	bakkee-na naadee-na	bakka naada	bakk-it naad-it	bakka-a-na naada-a-na	bakka-a naada-a
Stem change		A	B	C	D	B	D	D
temanna telaaga stanja	wish meet purify	y-temanna y-telaaga y-stanja	y-temann-u telaag-u stanj	temannee-na telaagee-na stanjee-	temanna telaaga stanja	temanna-it telaag-it stanj-it	temanna-a-	temanna-a
Stem change		A	B	C	A	B	A	A

Table 4.21 The stem change in the defective series of *M*2, *M*3, *M*5, *M*6 and *M*10 respectively.

The stem-shape of non-*M*1 verbs can vary depending on the measure; however, it is still similar to the stem-shape of *M*1, in that the change is induced by ablaut and the monophthong. In addition, the stem change can divide these measures into two subsets based on ablaut and stem syncretism. For example, *M*2 and *M*3 can have three stem-shapes with the same syntagmatic distribution as that of *M*1 (CVCCV ~ CVCC ~

CVCCee for *M*2 and CVVCV ~ CVVCee ~ CVVCV for *M*3). These measures belong to the i-vowel final verbs which, again in line with *M*1, lack stem syncretism across the subject and object sub-paradigms.

By contrast, each of *M*5, *M*6 and *M*10, which have different verb-bases but same stem-shape alternations. These measures belonging to the a-vowel final verbs that show the stem syncretism pattern. Thus, the 3SGM imperfective stem A is also used to fill in the cells of 3SGM perfective stem when the attached suffix is a C-initial object marker. Nevertheless, the crucial difference between all of these measures and *M*1 is that of the structure of stem-shape and in turn the stem change pattern of the 3SGM perfective form. We saw that in *M*1 (Table 4.19) the stem of this form is syncretic with that of 1/2 person forms in the perfective since both have the stem-shape CCee. By contrast, in the measures in question, the stem of the 3SGM forms does not have the monophthong -*ee*, but rather it is characterised by a low final vowel *a*. Therefore, in the a-final vowel verbs, the 3SGM perfective form is syncretic with the 3SGM imperfective form while in the a-final vowel verbs has a different stem; hence, stem D.

4.3.3 Double series

In section 3.1 , we saw that the main characteristics of this series is based on the presence of two (*M*1) or three (two Cs stem-medially and one C stem-finally) (*M*2 and *M*5) geminate consonants depending on the measure. The stem in *M*1 verb of this series is similar to the defective series in having the long vowel *ee* extension, but it differs from both the defective and sound series in lacking the FV. In addition, the stem-shape can vary in two aspects: the presence/absence of the monophthong *ee* and ablaut. Consider the following paradigm of *M*1.

verb	Gloss	Subject					3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
gall	take	y-gill	y-gill-u	gallee-na	gall	gall-it	gall-na	gall-a
radd	return	y-rudd	y-rudd-u	raddee-na	radd	radd-it	radd-na	radd-a
stem change		A	A	B	C	C	C	C

Table 4.22 The stem change in *M*1 double series.

This series of *M*1 displays three stem changes, the ablaut resulting in a coherent stem change distinguishing perfective vs imperfective sub-paradigms. Stem A in the imperfective cells can have the vocalic pattern of either *i* or *u* while the perfective stem vowel is always the low back vowel *a*.[44] It is also worth noting that the ablaut cannot be deduced by the consonantal features since either high vowel can co-occur with consonants of different place of articulation.[45] In other words, the ablaut change is lexically conditioned. In addition, within the perfective sub-paradigm, the stem change pattern is the result of the presence of the monophthong which, although it seems to be phonologically conditioned, occurs only before a C-initial subject marker (e.g. *gallee-na*), and no monophthong realization is induced before a C-initial object marker (*gall-*na).[46] In fact, the monophthong results in a disyllabic stem with the heavy syllable carrying the stress (e.g. *gallée-na*). Thus, in terms of the stem-shape, the basic stem CVC_iC_i occurs in all cells, but stem C is represented by the complex stem CVC_iC_iee. Traditionally, the double series was represented by the same template as the sound series. According to Watson (2002, pp.146-148), double verbs in San'ani Arabic have

[44] In Maltese, there is a bi-implicative relation across this series, which helps to deduce the ablaut change: when the perfective of the 3SGM form has the low /a/, the imperfective ablaut change is /o/. By contrast, when the perfective form has the vowel /e/, then the ablaut in the imperfective form will be either /e/ or /i/ (Camilleri, 2014, p.83).
[45] Examples of stem A include *rinn* 'ring', *riff* 'move quickly', *rigg* 'become thin', *rušš* 'sprinkle' *fišš* 'blow off', *dimm* 'condemn', *gišš* 'mop', *kišš* 'feel frightened', *kušš* 'show dislike of people', *šikk* 'doubt sb.', *šukk* 'stab sb.'
[46] TLA also allows the affixation of a consonant-initial suffix to a nominal word form with a final geminate, resulting in a trimoraic syllable (e.g. *jadd*, *jadd.na* 'our grandfather').

a disyllabic template in which the overt second syllable was only triggered by consonant-initial suffixes (e.g. *habb-a* ‘he loved’, *habab-na* ‘we loved’), the verb is realized as a monosyllabic stem produced by a resyllabification ‘Double Rule’. In TLA, on the other hand, double verbs only show a monosyllabic base irrespective of the morphophonological environment.[47] Therefore, as was pointed out by Watson for Cairene, we argue that, synchronically, TLA double verbs have only the basic stem-shape CVC_iC_i with a signal vocalic pattern that, in template terms, would possibly mean that the double rule has become lexicalized. One evidence for the lexicalization, as argued by Watson for San’ani and Cairene Arabic, is that TLA has a few verbs of which the shape stem in the surface representation is CVC_iVC_i (e.g. *jafif* ‘dry up’). Therefore, it appears that the phonological process which resulted in CVCC with the allomorphic variation of the same stem CVCVC+C vs CVCC+V may no longer be applicable in TLA.

Across the non-*M*1 counterparts, there are three patterns of stem change which can result in grouping the different measures of this series into the same classes of stem change: *M*1 with *M*10 (Table 4.23 and Table 4.24), *M*2 with *M*5 (Table 4.24). While *M*10 shares the same stem change pattern with *M*1, *M*7 displays different paradigmatic complexity. In other words, the pattern of stem change can show a morphomic stem-change relation by associating same pattern with measures of different stem-shapes.[48] In addition, the stem change of this series can display mutual-patterns characterised by absence of ablaut in the imperfective sub-paradigm or the monophthong in the

[47] Broselow (1976) argues that double verbs in Cairene are mapped to a triliteral verbal template in the underlying representation. This account is based on the observation that the active and passive participles of the double series share the same template as that of the triliteral verbs; hence, the second C of the CC is associated with two templates in adjacent positions.

[48] Bonami and Boyé (2008) discuss four ways of having morphomic relations which include morphomic stems, morphomic classes.

perfective sub-paradigm.[49] Consider the paradigms of non-M1 double series (Table 4.23 and Table 4.24).

verb	Gloss	Subject				
		Imperfective		Perfective		
		3SGM	3PLM	1PL	3SGM	3SGF
ingall irtadd	be taken be rejected	y-in-gall y-i-rtadd	y-in-gall-u y-rtadd-u	in-gallee-na i-rtaddee-na	in-gall i-rtadd	in-gall-it i-rtadd-it
stem change		A	A	B	A	A

Table 4.23 The stem change in the double series of *M*7 and *M*8 respectively.

verb	Gloss	Subject					3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
ḥaddid	specify	y-ḥaddid	y-ḥaddid-u	ḥaddid-na	ḥaddid	ḥaddid-it	ḥaddid-na	ḥaddid-a
tesammim	be poisoned	y-tesammim	y-tesamim-u	tesammim-na	tesammim	tesamim-it		
stem change		A	A	A	A	A	A	A
steġall	exploit	y-steġill	y-steġill-u	isteġallee-na	steġall	steġall-it	steġall-na	steġall-a
stemarr	continue	y-stemirr	y-stemirr-u	istemarree-na	stemarr	stemarr-it		
stem change		A	A	B	C	C	C	C

Table 4.24 The stem change in double series of *M*2, *M*5 and *M*10 respectively.

Both *M*7 and *M*10 are the same as *M*1 in terms of stem-shape, since each of them have two stem-shapes basic and complex: (st)-CVC_iC_i ~ (st)-CVC_iC_i-ee. By contrast, *M*2 and *M*5 paradigms have only a basic stem of the shape (t-)$CVC_iC_iVC_i$.[50] These measures are also contrastive in terms of their pattern of stem change. For instance, in the stem change of *M*7, the imperfective and perfective sub-paradigms have syncretic stems by sharing the same stem A which, in turn, is also syncretic with stem C of *M*1 (Table 4.22). In *M*7 and *M*8, the stem of the 3rd person perfective forms is used for the imperfective forms, hence these measures have the stem change pattern AABAA which is a mutual stem change pattern, also found in *M*2 and *M*5.

[49] This case is true when the suffixes attached are subject markers (*gallee-na* 'we took').
[50] The stem change in *M*2, *M*5 and *M*10 may possibly be subject to phonological constraints of degemination or consonant deletion in connected speech since it is difficult to produce three double consonants in a sequence without a vowel. However, in *M*10 the ablaut is always *i* with no high vowel *i* or *u* alternation.

Unlike all the other measures of this series, *M*2 and *M*5 show no stem change across the sub-paradigms. In other words, *M*2 and *M*5 display a unified pattern of stem change. However, *M*7 and *M*8 stem change is different from the pattern in *M*2 and *M*5 in that the 1/2 person forms have a different stem characterized by the presence of the monophthong *ee*. With this sub-pattern, *M*7 and *M*8 is more similar to *M*1 and *M*10.

Likewise, the pattern of stem change in *M*7 is the same as *M*2 and *M*5 in that the imperfective cells show no ablaut and have the same stem as the perfective cells. Thus, the TLA double series verbs is clearly characterised by sharing the different (sub) patterns of stem change within a single paradigm.[51]

Another important observation is that the pattern of stem change might be considered as having a morphomic relation with the verb measure/stem-shape. For instance, *M*1, *M* 7/8 and *M*10 share the same set of stem-shapes, but only *M*1 and *M*10 have the stem change AABCC.CC. By contrast, the stem change of *M*7/8 is AABAA which is also associated with *M*7 of the defective (Table 4.20). In other words, the different measures of the double series can have the same set of stem-shapes, but associate with different patterns of stem change. Furthermore, the same stem change can be associated with different verb series.

---

[51] The stem change of *M*7 and *M*8 may possibly be an example of stem heteroclisis in line with the heteroclisis denoted by declension exponents found in Latin because of deponency. According to Baerman (2007, pp.14-15) deponency, which expresses ‘morphological mismatches’, can lead to paradigmatic gap/defectiveness which in turn is filled by heteroclisis, a borrowed exponent/form of a different class. For example, the declension of the Latin *balneum* ‘bath’ is a mixture of the second and first declension noun in the singular and the plural respectively. Another example of heteroclisis provide by (Baerman, 2007) comes from Gothic in which two inflection classes are: strong and weak. While the strong class uses ablaut to inflect for the preterite, in the weak verb preterite paradigm, the inflection is realized through the addition of a dental suffix (-t, -d or -s). By contrast, Gothic has some verbs known as the preterite-present because of their deponent paradigms which are based on only preterite forms. The present tense of these verbs is formed according to the ablaut pattern found in the preterite of the strong verb. By contrast, the preterite paradigm of the preterite-present verbs inflects as that of the weak verbs.

4.3.4 Hollow series

The set of verb lexemes that belong to the *M*1 hollow series display two/three stem change patterns. Unlike sound series, the stem-shape of this series is based on a monosyllabic syllable structure which appears to be subject to phonological constraints but only within the subject agreement sub-paradigms. An example of the stem variation in four transitive/intransitive hollow verbs with different vocalic patterning is presented in (Table 4.25).

Unlike M1 sound series (§4.3.1), the stem change in the paradigm of *M*1 hollow series is based on maximally three stem changes and minimally two. Each of stem pattern differs in two respects: they exhibit different stem-shapes and vowel alternation. Both stems A and C have the stem-shape CVVC that undergoes a syllable structure change, resulting in stem-shape CVC of stem B.

verb	Gloss	Subject					Subj. 3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
baaʕ	sell	y-biiʕ	y-biiʕ-u	biʕ-na	baaʕ	baaʕ-it	baaʕ-na	baaʕ-a
faat	pass by	y-fuut	y-fuut-u	fut-na	faat	faat-it	faat-na	faat-a
ṣaad	fish	y-ṣiiḍ	y-ṣiiḍ -u	ṣuḍ-na	ṣaaḍ	ṣaaḍ -it	ṣaaḍ-na	ṣaaḍ-a
Stem change 1		A	A	B	C	C	C	C
baat	stay over	y-baat	y-baat-u	bit-na	baat	baat-it[52]		
xaaf	be frightened	y-xaaf	y-xaaf-u	xuf-na[53]	xaaf	xaaf		
Stem change 2		A	A	B	A	A		

Table 4.25 The stem change in M1 hollow series.

While the imperfective sub-paradigm shows stem A across all cells, the pattern within the perfective shows uniformity across perfective third person cells compared with the 1/2 person forms represented by stem B. In fact, the shape of stem B and C might

[52] The verb *naam* 'sleep' and *baan* 'appear' have the same stem pattern as the verb *baat*.
[53] In ELA, the same verb is inflected as *xif-na* (Owens, 1984).

appear to be subject to phonotactic constraints. While stem C is used before a V-initial subject suffix and has a monosyllabic structure, stem B is chosen before a C-initial subject suffix due to a syllable structure repair of unstressed vowel shortening. However, this repair strategy of ‘Closed Syllable Shortening’ (CSS) to avoid ‘stray Erasure’ does not hold before a C-initial object suffix and the stem-shape CVVC is retained (Watson, 2002, p.66).[54] In fact, the CSS in TLA is never obligatory between the stem and negation particle *š* or the prepositional complement *l* particle (e.g. *gaal-(i)-l-ha* ‘he said to her’).

In addition, there are four main vocalic patterns which divide hollow verbs into three-stem change patterns all of which have the long low back vowel *aa* in stems C/A in the third person forms. The verbs *baaʕ* ‘sell’ and *faat* ‘pass by’ represent the first pattern in which the frontness or backness of the high long stem vowel in stem A provides inferences for the stem B vowel that only differs in length. On the other hand, the stem vowel implicative relation does not hold in the second pattern, which is represented by the verb *ṣaad* ‘fish’. The vowel quality of stem B is determined lexically. The vocalic pattern of the final pattern is attested in a few frequent verbs (e.g. *baan* ‘appear’, *ʕaaf* ‘disgust by certain food’) in which the vocalic pattern of stem C (*aa*) in stem change 1 is also used within the imperfective sub-paradigm; hence, the resulting pattern of stem change has only two variants. Thus, the shape of the three hollow stems can be represented as in the following table:

	Stem-shape
Stem A	CVVC
Stem B	CVC
Stem C	CVVC

Table 4.26 The stem-shape in M1 hollow series.

[54] Similarly, in Cairene Arabic, hollow verbs are subject to CSS when subject or object suffixes are attached. By contrast, in San’ani Arabic, hollow verbs undergo vowel shortening before a subject suffix only. With object suffixes, which have a *h*- or *n*-initial consonant, San’ani Arabic keeps the CVVC stem-shape, but introduces an epenthetic vowel between the suffix and stem (Watson, 2002).

Unlike the M1 sound series (§4.3.1), the stem consonantal *ḥ* as a final C in M1 hollow series does not seem to induce a phonological restriction on the stem vowel; hence, the consonant *ḥ* can be preceded by a front/back high vowel (e.g. *saaḥ* ~ *siiḥ* ~ *siḥ* 'melt' vs. *raaḥ* ~ *riiḥ* ~ *ruḥ* 'be lost'). It is also worth noting that the consonantal features do not seem to affect the vocalic quality of the stems vowels (e.g. *daar* ~ *diir* ~ *dir* 'do' vs *daar* ~ *duur* ~ *dur* 'turn around' and *daag* ~ *duug* ~ *dug* 'taste' vs. *daag* ~ *Ø* ~ *diig* 'be narrowed down'). In fact, the traditional explanation of the vocalic pattern in hollow verbs assumes that the stem vowels are derived from the underlying root glides (i.e. /i/ from /j/ and /u/ from /w/) (see §3.1). In TLA, the quality of the root glide does not always match/correlate with the quality of the stem A or stem B vowel (Table 4.27).

root	gloss	stem A	stem B
rwḥ	be lost	riiḥ	ruḥ
ṭwḥ	fall	ṭiiḥ	ṭuḥ
ṣwm	fast	ṣiim	ṣum
ṣyḥ	shout	ṣiiḥ	ṣuḥ
ṣyd	fish	ṣiid	ṣud
ṭyr	fly	ṭiir	ṭur
ġyr	be jealous	ġiir	ġur
ḍyf	add	ḍiif	ḍuf
ḍyʕ	be lost	ḍiiʕ	ḍuʕ
xyn	betray	xuun	xun
syg	drive	suug	sug
gys	try on	giis	gis
gym	pray at night	guum	gum
gym	take away sth.	giim	gim
gyd	pull with robe	guud	gud
ġyb	be absent	ġiib	ġub

Table 4.27 The glide in the root of the hollow verbs compared the stem vowel in TLA.

Broselow (1976), discussing hollow verbs in CA, argues for the lexicalization of these vocalic patterns based on the observation that the stem of the derived form of the hollow verb such as the active participle always has the shape CVyVC with a glide *-y* regardless of the front-back vowel alternation and the underlying glide. In the examples provided in (Table 4.25), the relevant deverbal nouns are *bayaʕ, fayit, sayid, bayit and xayif* respectively.

Further evidence for lexicalisation comes from the stem change of the *M*7, *M*8 and *M*10 hollow series. The relationship between *M*1 and *M*7/*M*8 in terms of the stem-shape and ablaut seems to be strongly correlated particularly in the use of stem C (third person perfective stem). Consider the following fragment of the perfective and imperfective sub-paradigm of the measures in question:

verb	Gloss	Subject			
		Imperfective		Perfective	
		3SGM	3PLM	3SGM	3SGF
baaʕ ṣaad	be sold fish	y-in-baaʕ-u y-in-ṣaaḍ	y-in-baaʕ-u y-in-ṣaaḍ-u	in-baaʕ in-ṣaaḍ	in-baaʕ-it in-ṣaaḍ-it
laam[55]	blame	y-i-l(t)aam	y-i-l(t)aam-u	i-l(t)aam	i-l(t)aam-it
stem change		A	A	A	A

Table 4.28 The stem change in hollow series of *M*7 and *M*8 respectively.

These measures display only one stem variant all based on stem C of *M*1 hollow series. In other words, regardless of the root glide, the same vocalic pattern is shared across aspect sub-paradigms within third person forms. However, in verbs such as *jaab* 'bring' it is possible to have corresponding 1PL form *nin-jaab-u* 'we are brought' which has the same stem vocalic pattern /aa/ as the third person forms. In addition, the 1 and 2 perfective forms tend to have the same stem as that of M1 (e.g. *jib-na* 'we brought' and *in-jib-na* 'be brought/be born', *lum-na* 'criticize' and *iltum-na* 'be blamed'. Thus, the stem-shape of this measure can be CVVC or CVC.

[55] Other examples of hollow M8 do not have a corresponding *M*1 (*štaag* 'tmiss', *xtaar* 'choose', *ḥtaaz* 'need', *ḥtaal* 'deceive' and *ḥtaar* 'be confused' (Harrama, 1993, p.83).

The stem pattern of M10 is similar to that of *M*1 as it shows three stems. Nevertheless, both stem B and C are based on the same vowel quality while stem A shows vowel quality change (Table 4.29).

verb	Gloss	Subject					3SGM.Obj	
		Imperfective		Perfective			Perfective	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
stefaad	benefit	y-stefiid	y-stefiid-u	stefad-na	stefaad	stefaad-it	stefaad-na	stefaad-a
stem change		A	A	B	C	C	C	C

Table 4.29 The stem change in *M*10 hollow series.

The stems B and C are the result of the phonological process of syncope; hence, both stems have the same vocalic quality. Although stem A shows vowel change, all verbs of this measure appear to have the same stem vowel quality and no clear co-relation with the quality of the middle root C (e.g. *steʕiid* ~ *steʕad* ~ *steʕaad* ‘recover’).

4.3.5 Summary

This section has introduced a description of the stem change pattern across the measure provided and their corresponding series. Clearly, the overall pattern of stem change can be based on altering the stem-shape and/or ablaut. The relationship between the measures and the pattern of stem change seems to have a morphomic nature, since different measures can display the same stem pattern. On the other hand, the different stem patterns can associate with the same measure series. However, the stem change can be motivated by different factors according to which not every stem variant can form a distinct stem with a lexical indexing. The following section will discuss the morphomicity of TLA stems.

4.4 Representing TLA semi-autonomous morphomic stems

The stem change in TLA can be affected by different factors. In some cases, there seems to be only one motivating condition while in other stem patterns the effect is more complex. The factors might act as ‘stem referencing features’ which can describe the morphological nature of TLA verb stems. Based on the different stem types discussed in the literature, including morphomic (m-stem) (Aronoff, 1994), morphophonological (m-p.stem) (Spencer, 2013; Maiden, 2009), morphosyntactic (m-s.stem) (Baerman & Corbett, 2012) and phonological (p.stem) Anderson (1992, 2008, 2011) (§2.3) there seems to be a scale of different stem properties which can determine the stem complexity. TLA stems can be motivated by the interaction of these morphological and extramorphological factors. Understanding the nature of the autonomous and/or semi-autonomous TLA stem change, we propose the framework in (Table 4.30).

Stem type	morphological and extramorphological stem factors			
	Morphological	morphophonological	morphosyntactic	phonological
m-stem	+	-	-	-
m.p-stem	+	+	-	-
m.s-stem	+	-	+	-
p-stem	-	-	-	+

Table 4.30 The stem referencing features.

TLA stem can be defined by four properties. The plus and minus signs respectively indicate the presence or absence of sensitivity to particular factor(s) that correlate(s) and/or condition(s) the stem alternation. Therefore, the [–morphological] stem means that there is a level of stem form/choice predictability provided by extramorphological factor(s). At the level of arbitrariness, any stem that has a [+morphological] referencing feature means it cannot be *fully* predicted by any other factor; hence, it is lexically determined and morphologically induced.[56] However, the arbitrary stem-shape, ablaut, blocking of a possible monophthongal realization which acquires the stem the [+morphological] referencing feature can co-occur with some/all extramorphological factors provided in (Table 4.30).

The reason for listing these factors in this order is to represent the idea that the canonical suppletive stem (Corbett, 2007) and/or the maximally morphomic stem (Spencer, 2012, 2013a) may possibly be the canonical nature of TLA stems as a Semitic language. Nonetheless, TLA stem alternations can show deviation from this ideal. The tendency of this deviation seems to be due to morphophonological factors, then possibly a morphosyntactic factors, and finally at the end of this scale of non-canonicity come phonological factors. In fact, TLA stem can have referencing features at different points of the scale and a representative way of pointing to these points of the scale is best shown by binary marking (with a +/- sign). However, it is important to emphasise that TLA stems do not display all the +/- possibilities of stem conditioning factors. The discussion will only be based on the properties that stipulate stem change in TLA verbs.

4.5 TLA Stem referencing features

In section 4.3 , we showed that stem change is a common feature of TLA inflectional paradigm. The statue of TLA stems in different measures can be defined by different referencing features. The stem alternation seems to be systematically associated with a set of particular properties including phonological (§4.5.1), morphophonological (§4.5.2), morphosyntactic (§4.5.3), and/or morphological (§4.5.4) properties. The

[56] A (maximally) morphomic stem in autonomous morphology should have a [+morphological] referencing feature with no sensitivity to any other factors. By contrast, a semi-autonomous morphology includes morphomic stems that are morphologically induced, [+morphological], but can show sensitivity to extramorphological factors. It is important to note that the + sign does not donate the realization of particular extramorphological factors, but rather it reflects the sensitivity and correlation of the morphomic stem with these factors.

interaction of these synchronic conditioning effects results in semi-autonomous TLA stems.

4.5.1 Phonological effects on the stem change

The paradigmatic pattern of TLA stem alternation can be determined by phonologically conditioned behaviour which, nevertheless, interacts with arbitrary morphological stipulation in some measures. However, the effect is that not all stems need to be listed in the lexicon since some of these changes can be due to phonological adjustments required by syllabification constraints. For instance, the pattern of stem change in the sound series in each of *M*2, *M*3, *M*5 and *M*6 (Table 4.18) can also be due to phonotactic constraints of syllabification, resulting in variation of the same stem. Thus, in both *M*2 *ʕallim* 'teach' and *M*5 *teʕallim* 'learn', the final stem vowel occurs post gemination, and undergoes vowel syncope when a vowel-initial suffix is attached to the stem. According to Watson (2007, p.339), the process of a 'postgeminate high vowel deletion' is a distinctive characteristic of Arabic VC dialects.[57]

In some cases, as Owens (1980) points out for TLA, an unstressed short vowel becomes a high vowel when it occurs in a closed syllable word-finally. For instance, in stem A of sound series *M*2 (e.g. TLA *gássim* cf. ELA *gássam* 'he divided') (Table 4.18), the stress is on the first syllable and the final vowel is raised. However, he also

[57] Kiparsky (2003) presents a semisyllable account of syllables in Arabic, which divides the dialects into three types each of which distinguished and linked by sharing certain syllabification patterns. Under this account, Arabic has three different types of dialects based on the syllabification of the morphologically derived CCC clusters. In the first, second and third types, the CCC cluster is syllabified as CVCC (VC dialects), CCC (C dialects) and CCVC (CV dialects) respectively. For instance, the epenthesis in CV dialect occurs to the right of the CCC cluster (CCvC) (e.g. TLA /gult-la/ [gultila] 'I said to him') whereas in the VC dialects, the epenthesis is inserted to the left (CCC=CvCC) (e.g. /gilt-la/ [gilitla] 'I said to him'). In addition, in VC dialects, the imperfective inflectional forms with a V-initial suffix has the syllable structure CVCC-V(C) as in: TLA [jikitbu] 'they write'. By contrast, C dialects simply drop the vowel ([jiktbu]), and CV dialects retain the vowel as in CCiC ([jiktibu]) in the corresponding cases. Therefore, TLA can pattern as a CV or VC dialect whereas Tripoli LA is a VC dialect as classified by Watson (2007). Although this semisyllable approach can account for a wide set of Arabic syllable-related dialects, TLA (along with other dialects (Watson, 2007)) shows syllabification phenomena of different dialect types.

notes that, in the forms to which a consonant suffix is attached (TLA stem C) (e.g. *gassím-na* 'we divided'), the stress is on the penultimate syllable and yet the vowel quality is not affected. The second stem vowel in stem A (*y-gassim* (3SGM.imperf) ~ *y-gassm-u* (3PLM.imperf)) seems to behave as a structural vowel which can be stressed, but can undergo elision. In fact, this account seems to contradict Owens' description of the syllable structure according to which vowels can be either 'structural V' which can carry stress or unstressed elidable 'epenthetic v' which is unstressed and elidable (§3.2.1). According to Watson (2007), in CV Arabic dialects[58], the epenthetic vowel is always treated as a regular vowel in terms of undergoing lexical processes including vowel shortening and stress to which VC dialects epenthesis is not visible. In TLA, the second vowel in form *gassim-na* 'we divided' carries stress, but in the form *gul-(i)-t* 'I said', the vowel *i* never receives stress.

Clearly, the vowel of the former form is likely to be a stem vowel while in the latter the vowel can be considered as an epenthetic vowel which is always unstressed and can be elidable. The stem change pattern in sound series of the measures in question can be influenced by phonological factors based on phonotactic constraints resulting in phonologically conditioned change of the same stem. Therefore, the stem pattern in non-*M*1 (except *M*7 and *M*8) sound series effectively constitutes a phonological environment for stem selection.

Likewise, we might tentatively suggest that phonological factors can also contribute to the morphophonological apophonic effect in some verb series. In sound series *M*1, for example, verbs show ablaut across both sub-paradigms resulting in stem A, stem B and stem D (Table 4.15). Some vocalic patterns of these stems are variable while others seem to be correlated constantly with certain sounds. The distinctive

[58] See footnote 57

vocalization of stem A and stem B is of a particular interest since in certain cases the vowel alternations seem to be corresponding to a network of specific sounds including (*h, ḥ , ʕ*, and emphatic sounds). Otherwise, the vocalic alternation is lexically stipulated, resulting in fully independent stems (with different indexing). Consider the ablaut in the imperfective sub-paradigm of *M*1 sound series in (Table 4.31):

Stem A	Stem B	gloss	apophony
rʕim	riʕm	desire	i~i
ḍrub	ḍurb	hit	u~u
dbaḥ	dibḥ	slaughter	i~a
gtil	gutl	kill	i~u
smaʕ	simʕ	hear	a~i

Table 4.31 The ablaut changes in stem A and stem B *M*1 sound series.

The vowel of stem A is alternating (i.e. it can be either *i*, *u* or *a*). However, the vocalic pattern is not completely arbitrary since the ablaut change that leads to the stem-shape CCaC is phonologically conditioned by C3. In this series, when the verb stem A ends in the laryngeal sound *h* or the pharyngeal sounds *ḥ* or *ʕ*, then the stem vowel will be the low vowel *a*. It is important to note that the selection of the stem CCaC is restricted in two ways. First, the conditioning consonants have to be a stem final consonant. Second, the conditioning only affects the stem A vowel. In other words, the effects of these stem final consonants on neighbouring vowels is blocked elsewhere. There are many examples in which the consonant *h* occurs as C1/C2 but in which the preceding/following vowel is a high front or back vowel and lacks the lowering effect (e.g. *ʕeraf ~ n-iʕrif/n-aʕruf* ‘know’, *ʕezaf ~ n-iʕzif* ‘play a musical instrument’, *ʕeraḍ~ n-uʕruḍ* ‘display’ *ʕegas ~ n-iʕgis* ‘reverse’, *reʕaš ~ n-urʕuš* ‘shiver’, *reʕaf ~ n-irʕif* 'bleed through noise').[59] In addition, unlike stem A, the corresponding stem D can only have a high vowel as a second stem vowel (*ʕuruf*, *ʕizif*, *ʕuruḍ*, *ruʕuš*, *riʕif*). In many Semitic languages, according to Watson (2002), the laryngeals, pharyngeals and the uvular fricatives belong to the natural class of gutturals. For instance, Arabic, Maltese and Ethiopian Semitic pattern together in lowering a neighbouring vowel such as the FV which is realized as low /a/ when the stem initial consonant is a guttural. However, unlike TLA, the lowering effect of these sounds in these languages is not restricted by their stem final position.

A case of phonologically induced ablaut is observed in stem D of the *M*1 sound series in which the stem-shape can be either CuCuC or CiCiC. As it will be argued in section 4.5.2 and section 4.5.4, it may not possible to predict the stem choice in this case; hence, the stem selection is morphologically stipulated. However, in one case, the ablaut in stem D seems to be determined by root radical dependency when one of the consonants is a pharyngealized phoneme. Thus, if the verb has the consonants *ṣ, ṭ, ḍ*, or *ẓ,* stem D tends to have stem-shape CuCuC. In other words, emphatic consonants can block CiCiC in stem D, and vowel backness is induced (e.g. *ṣegaṭ* ~ *ṣuguṭ* 'fail', *ṭelab* ~ *ṭulub* 'request', *ṭelaʕ~ṭuluʕ* 'come out'). However, it is interesting to note that

[59] According to Owens (1984, p.111), in some verb classes in ELA (§3.2.1), the shape of the stem appears to be determined by the verb root radicals. Thus, if the verb has the guttural /h, ḥ or ʕ/ as C_2 or C_3, it will belong to class 3 with the perfective stem-shape vCCvC and imperfective stem-shape CCVC (class 5). In other words, the shape of the stem has dependency on the phonetic feature of the C forming the root. However, he also provides the verb *išrib* 'he drank' as a class 3 verb. He emphasizes that if the verb has the stem-shape CVCVC (class 1) in perfective and VCCVC in imperfective (class 5), the verbs of this class will tend to have the consonants in question as C_2 and C_3. He adds that there are only three verbs with C_2 or C_3 as /h, ḥ or ʕ/ that can belong to both class1 and class 4 (imperfective stem shape v-CCvC) simultaneously based on aspect. These verbs include yu-*gʕud* 'he stays', *yi-lḥim* 'he welds' (TLA *ya-lḥim*), *ti-ṣbiḥ* 'be in the state of' (TLA *ta-ṣbaḥ*) (i.e. these have the same aspect stem-shapes (aspect classes) as *kitab* 'write' verb). It seems that the overall tendency is that verbs with the relevant consonants will restrict the imperfective stem-shape to that of class5 while the perfective can have the stem-shape of class 3 or class 1. Clearly, the relationship between the stem-shape and phonetic features of the radicals seems to be arbitrary rather than a systemic synchronic process for determining the verb class. The knowledge of the consonant does not always determine the inflection class of the verb. We might suggest that these verbs show a morphomic stem (based on arbitrary choice of syllable structure/stem-shape) within the perfective sub-paradigm.

the emphatic assimilation does not seem to always affect the FV nor the low vowel conditioning in the imperfective forms (e.g. *y-u-ṭlub* 'he requests' cf. *y-a-ṣguṭ* 'he fails'[60], and *ya-ṭlaʕ* 'he comes out').

Likewise, in the double series, these emphatic sounds tend to be associated with the *u*-ablaut (e.g. *ḍann* ~ *ḍunn* 'think', *ṣarr* ~ *ṣurr* 'pack up', *raṣṣ* ~ *ruṣṣ* 'press'). Borg (1997) points out that, diachronically, the emphatic/guttural environment is associated with /aa/ which is synchronically realized as /uu/, but /ii/ elsewhere. In TLA, however, this observation does not seem to hold in the defective and hollow series. For instance, in the defective series the vowel of stem A can never be a high back vowel, it is either a high front vowel /i/ or a low vowel /a/ (*gḍee* ~ *yi-gḍi* 'do a task', *rḍee* ~ *ya-rḍa* 'accept', *ṭlee* ~ *yi-ṭli* 'paint') (Table 4.19). Similarly, there are hollow verbs (Table 4.25), in which stem A (e.g. *ṣiid* 'fish', *ṣiim* 'fast', *ṭiib* 'be cooked, accept sth.' *ḍiif* 'add') and the corresponding stem C (e.g. *ṣaad, ṣaam, ṭaab, ḍaaf*) have the front long /ii/ vowel and a back low vowel /aa/ respectively. Moreover, the *M*1 sound series with plain consonants can also choose the stem-shape CuCuC.[61] According to Owens (1984, p.37), in ELA verbal stem change (in imperfective) based on the consonantal phonological effect might be associated with a consonantal scale determined by the tendency for co-occurring with back prosody vowels.[62]

---

[60] (cf. ZLA & Tripoli LA *yu-ṣguṭ* 'he fails').

[61] Borg (1997) argues that Cypriot Arabic shows retention of the historical /a/ which does not raise to /e/ when in the neighbourhood of a historically emphatic environment. Likewise, Heath (1989, p.75) notes that in Moroccan Arabic, emphatics can have a backing and lowering effect on adjacent vowels (e.g. in the loanword French mɔkɛt 'moquette' vowels are realized as MokeT).

[62] The numbering of the scale from 1 to 5 represents the level of the association between five sets of consonants and the back prosody. Level 1 of the scale has the consonant sets that are most likely to associate with the back prosody ascending to the level 5 consonant set the members of which are less likely to occur with a back vowel. Within level 1, there are two sets which are equally likely compared with the sets from 2 to 5. However, the co-occurrence of a back vowel and the consonant(s) in 1.a set is even more probable than with those of the 1.b set.

(3) 1.a. (ḍ, ṭ, ṣ), b. (x, ġ), 2. (r, l), 3. (g, b, m, f), 4. (w), 5. (n, b, f, d, t, z, s, y, k, ʕ, h, ḥ, š)

The first set of consonants can be divided in two main groups. The first three consonants (1.a) are positioned high in the scale as they always occur with a back vocalic pattern. The first subset, which consists of the emphatic sounds, is not only associated with back vowels, but also spreads the backness feature to adjacent syllables. By contrast, the second subset (1.b) does not necessarily affect the prosody of the rest of the syllables. Owens points out that the vocalic alternation tends to have morphological distribution since it occurs in 'pairs of related morphological forms' (1984, p.38). Nevertheless, the vocalic pattern adjacent to the consonantal set 2 to 5 cannot be predicted (§4.5.4).[63]

Clearly, the paradigmatic behaviour of the verb series discussed above show that, in some verbs, the stem change can be due to phonological stipulation as we saw in the case of vowel syncope. By contrast, the phonological effect associated with the stem change seems to be morphologically conditioned in other verbs (ablaut change to a back vowel). In other words, stem A of *M*2, *M*3, *M*5, *M*6 and stem D of *M*1 in sound series can have a [+phonological] sensitivity (Table 4.32).

stem change	morphological and extramorphological stem factors			
	Morphological	morphophonological	morphosyntactic	phonological
Stem A	-	+	-	+
Stem D	+	+	+	+

Table 4.32 The Stem referencing features in stem D in *M*1 sound series and stem A of *M*1, *M*2, *M*3, *M*5 and *M*6 sound series.

[63] It is worth noting that the back prosody in Owens' (1984) analysis is represented by the back vowel /u/ and low back vowel /a/. In fact, in ELA low vowels tend to correlate with back prosody more than high vowels (e.g. ELA *yagbal* cf. TLA *yagbil* 'he accepts sth.').

In these cases, stem A of the verbs in question has a [+phonological] property due to the phonological stipulation of the stem-shape. Likewise, stem D can also be [+phonological] since that the knowledge of the phonological emphatic feature of the stem consonant can predict the stem vocalization. In addition, the minus sign indicates that the factor may not affect the stem change. Nevertheless, there are cases in which all factors can contribute to the stem change as will be further discussed in the following sections. Next, we will consider the paradigmatic effects stipulated by interaction of phonological and morphological rules.

4.5.2 Morphophonological effects on the stem change

Every stem change in every measure/series can have a syntagmatic distribution determined by the phonological shape of the affix. However, this does not always guarantee a fully predictable stem. For instance, the vowel-initial exponent can determine *allomorphic* stem alternation (Spencer, 2012, 2013a) characterised by stem vowel elision or blocking of the monophthongal realization (e.g. *M*1 double series, *raddee-t* 'I returned' vs. *radd-it* 'she returned'). In some verb series, the stem change, akin to ablaut, can have reference to a syntagmatic distribution; however, the stem form remains arbitrary (e.g. *rebak* ~ *ribik* 'confuse sb.', *rekab* ~ *rukub* 'ride', *regab* ~ *rigib* 'peek', *gerab* ~ *gurub* 'touch/approach'). In addition, the morphophonological mapping between the stem with the monophthong *ee* and the consonant-initial exponents can be overridden (*rakkee-na* 'we hit badly' vs *rakk-na* 'he hit us badly'. Therefore, the distributional correspondence is complex and can be a statement of *'morphology by itself'*.

In the *M*1 sound series, morphophonological interactions can affect the stem-shape and as a result, the stem alternation can have reference to a syntagmatic distribution according to which a specific stem is chosen by the morphophonological environment effect induced by a V/C-initial prefix/suffix as a subject or object exponent (Stump, 2001b) (Table 4.16). In the imperfective paradigm, stem A and B have different stem-shapes (e.g. *y-a-smaʕ* ~ *y-a-simʕ-u* 'hear') (Table 4.15). While Stem A occurs before consonantal allomorphs (or in unsuffixed forms) created by the subject/object exponents, stem B is chosen in a pre-vocalic position. In other words, the context created by the exponent leads to a formal stem variation resulting in 'allostems' based on the stem-shape. However, the stems remain unpredictable. Likewise, stem D in the perfective sub-paradigm is only selected in a pre-vocalic position whether it is before a subject or an object suffix (e.g. *gurub-it* 'she touched', *gurub-a* 'he touched him') while stem C occurs elsewhere (*gerab, gerab-na, gerab-ti* 'he, we, you.SGF touched sth/sb.'). The following table represents the stem selection in *M*1 sound series when sensitive to extramorphological factors.

	Subj. or Obj. suffix
Stem B	V-initial
Stem D	
Stem A	C-initial
Stem C	

Table 4.33 The syntagmatic distribution of stems in *M*1 sound series.

Another morphophonological dimension is represented by the stress assignment which appears to be sensitive to some patterns of stem change in this series. In the imperfective stems, the stress is on the initial syllable formed by the FV. By contrast, in the perfective sub-paradigm, stem change D correlates with initial stress placement whereas stem B has penultimate stress (*gútul-it, getál, y-úgtil* 'kill') (Chapter 3, §1.3.3). Although the stem change coincides with the stress shift as an element of morphophonological alternation, the stem choice and form is independently achieved i.e. 'the alternations are irremediably as 'unnatural' products of stress' (Maiden, 2005, p.106). For instance, it is not possible for the stress to trigger the stem change since the same stem D is also selected in the *M*7 (Table 4.17); nonetheless, the stress is on the initial syllable, that is, the prefix *in-* rather than the stem.[64] In addition, recall from Chapter 3, section 1.3.3 that in TLA, a heavy syllable attracts stress. In verbal paradigms, stress can be shifted to the vowel subject suffix when a V-initial marker such as the object 3SGM suffix *-a* is added (e.g. *kúruh-it* ~ *kuruh-áat-a*/*-ik* 'she hated him/you').

Another example of a morphophonological stem-shape conditioning in the *M*1 sound series is related to the set of verbs which has the glide *w/y* as the initial consonant C1, traditionally known as assimilated verbs, displaying an allomorphic stem variation. Unlike the perfective forms in which the glide is retained stem-initially, in the imperfective and the imperative forms, the C1 glide is always absent (Table 4.15). In fact, when the glide *w/y* is preceded by the FV, it undergoes assimilation, being manifested as a high back vowel *uu* (*werat* ~ *yuurit* 'inherit') or a high front vowel *ii (yebas* ~ *yiibis* 'dry up') respectively.[65] The assimilation leads to vowel lengthening; hence, altering the typical stem-shape of this series in the imperfective sub-paradigm: CV-CCVC »[66] C-VVCVC (e.g. *wezan* ~ *yu-uzin* ~ *yu-uzn-u* 'weigh', *werat* ~ *yu-u-rit, yu-urt-u* 'inherit'). Therefore, it seems that if a complex stem produced by morphology results in unattested sequences in the stem-shape, a phonological process is triggered. In other words, similar to CA (Chekayri & Schee 2003), TLA seems to have a

[64]A similar view is argued by Maiden (2009) for the stress and N-pattern in Italian. However, a different view is suggested by Anderson (2008) for Rumansh stem alternation (§2.3.4). It is also apparent from the data in Table 4.9 and Table 4.13 that TLA stem patterning is similar to the Surmiran verb stems (Anderson, 2008), whose distribution is determined by stress pattern (§2.3.4). However, in the TLA case, the alternate stem forms cannot be easily related in purely phonological terms (§4.5.4).
[65] Watson (2002, p.19) points out that although all Arabic dialects have reserved these glides, in some dialects they are realized as vowels when preceding a consonant word-initially, following a consonant word-finally or between consonants.
[66] The symbol ≫ means that the change from the expected stem-shape to the resultant form is morphophonologically conditioned.

phonotactic constraint which disallows the sequences (*uwC) and (*iyC) in the imperfective forms of the *M*1 sound series.[67] However, the sequences (wuC), (wiC) and (yiC) are permissible, occurring in stem D (*wizin-it* 'weight', *wuʕud-it* 'promise', *yibis-it* 'dry', *yiʔis-it* 'fed up'). It is also possible to have the sequence (awV) as a result of adding the negation (particle) to stem B/D (e.g. *ma-wezan-iš* 'he did not weight' *ma-wizin-it-iš* cf. the same sequence word-finally *nwee* ~ *yi-nwi* 'intend', *rwee* ~ *yi-rwi* 'water', *xwee* ~ *yax-wa* 'miss'). Clearly, the syntagmatic distribution of A, B, C and D stems in the *M*1 sound series indicates that these stems can have a [+morphophonological] referencing feature representing their behaviour as m-p stems (Table 4.34):

stem change	morphological and extramorphological stem factors			
	Morphological	morphophonological	morphosyntactic	phonological
Stem A	+	+	-	+
Stem B	+	+	-	+
Stem C	-	+	+	-
Stem D	+	+	+	+

Table 4.34 The stem referencing features in *M*1 sound series.

The table shows that every stem in *M*1 of this series has a [+morphophonological] feature, because of their morphophonologically conditioned stem-shapes.

Similarly, the defective series of all measures and the *M*1 double series show stem changes with sensitivity to morphophonological factors. This conditioning is manifested in the presence or absence of the monophthongal-induced vowel change stem-finally in the verbs in question. In section 4.3.1, we showed that measures of the defective series have different stem-shapes. They all have the monophthong /ee/ before

[67] The same morphophonological effect is found in the imperfective forms of *M*1 defective series *wʕee* ~ *yuuʕa* 'to become unconscious'. In other types of words, the sequences (uwC) and (iyC) seem to be allowed (e.g. *buwaaṣ* 'screaming', *ġuwaaṣa* 'submarine', *ṭagiya* 'hat', *masraḥiya* 'a play', *ṣiyad* 'fisherman').

consonant-initial subject suffixes (e.g. *M*1 *nseet* 'I forgot', *M*2 *bakkeet* 'I upset sb.'), but in *M*1, the unsuffixed 3SGM form is based on the stem with the monophthong *ee* (e.g. *nsee* 'forget'). In TLA, the stem-shape CCee alternates with stem-shape CC (*nsee-t* 'I forgot' *ns-it* 'she forgot') (Table 4.19 and Table 4.21). These stems are syntagmatically distributed in the perfective sub-paradigm. Thus, the stem-shape with a consonantal ending CC in any measure (i.e. without the VV ending represented by monophthong value *ee*) is selected before a vowel-initial subject suffix. The same morphophonological correlation is found within the imperfective sub-paradigm (*ya-nsa* 'he forgets', *ya-ns-u* 'they forget'). However, the stem final vowel of this series in any measure is retained in the object paradigm in all phonological contexts (e.g. *nsa-a* 'he forgot him'). It seems that the syntagmatic distribution is only partially induced.

stem change	morphological and extramorphological stem factors			
	Morphological	morphophonological	morphosyntactic	phonological
	defective series (a-vowel final verbs)			
stem A	+	-	-	-
stem B	-	+	-	+[68]
stem C	-	+	+	-
	defective series (i-vowel final verbs)			
Stem A	+	+	-	-
Stem B	-	+	-	+
Stem C	-	+	+	-
Stem D	+	-	-	-

Table 4.35 The stem referencing features in *M*1 defective series.

Stem A in the i-vowel final verbs has the [+morphophonological] feature since it correlates with the phonologically conditioned context of the exponent while stem A of the a-vowel final verbs lacks this conditioning feature. In both sets, stem B has a

[68] A [+morphophonological] and [+phonological] means that the stem is predictable and does not require listing. By contrast, a [+morphological] and [+morphophonological] means the stem is unpredictable by the features and need to be listed.

[+morphophonological] feature. Stem D is only in the i-vowel final verbs and has a [-morphophonological] representation with no reference to the context of exponents.

Likewise, the stems in the *M*1 double and hollow series have in part morphophonologically conditioned stems (Table 4.22 and Table 4.25). The stem change of these verbs has the pattern ABC in which A and C have the same stem-shapes (CVCC in double series and CVVC in hollow series), but with ablaut change. Stem B, on the other hand, is different in both cases. For instance, in the double series stem B (CVCCee, *raddee-t* 'I return sth.') is a complex stem due to the presence of the monophthong *ee* while stem B of the hollow series has the shape CVC (e.g. *biʕ-t* 'I sold') with a short vowel. Stem B, in both verb series, is selected before a consonantal subject suffix. Nevertheless, this stem is not selected before a consonantal object; thus, the double verb selects the non-complex stem-shape CVCC (*radd-na* 'he returned us' cf. *raddee-na* 'we returned') and the hollow verb uses the CVVC stem-shape with the long vowel (*faat-na* 'he passed by us', *fut-na'* 'we passed by sth/sb'). The paradigmatic behaviour of these verbs seems to imply that stem B can be defined morphophonologically in some phonological contexts but not others; therefore, the stem choice remains morphomic.

stem change	morphological and extramorphological stem factors			
	Morphological	morphophonological	morphosyntactic	phonological
	double series			
Stem A	+	-	+	+
Stem B	+	+	+	-
Stem C	+	-	+	-
	hollow series			
stem A	+	-	-	-
stem B	+	+	+	+
stem C	+	-	+	-

Table 4.36 The stem referencing features in *M*1 double and hollow series.

Stem B in both series has the [+morphophonological] referencing feature since it has a syntagmatic distribution. However, the sensitivity to the phonological shape of exponent is only possible with subject suffixes; hence, the referencing feature [+morphological].

Clearly, the stem change pattern discussed in this section can involve m-p stems the choice of which is determined by morphophonological terms. Nonetheless, these stems remain morphomic since the sensitivity to the morphophonological factors is not consistent with all types of conditioning exponents. In the next section, we will show that in addition to the semi-syntagmatic distribution, some stems correlate with paradigmatic cells which represent a natural class of MPs.

4.5.3 Morphosyntactic effects on the stem change

This section will discuss the paradigmatic behaviour of the stem change in which the resultant stem can be sensitive to the morphosyntactic value of the form. In §3.1.1, we introduced the different MPs realized by prefixes and suffixes in the imperfective, but by only suffixes in the perfective. However, in some verbs, the stem selection can show sensitivity to morphosyntactic factors. A set of resultant stem changes can be attributed to paradigmatic cells forming a natural class. In other words, the stem change can be characterized by a feature split that represents the grouping of the different stems paradigmatically.

The stem sensitivity to the morphosyntactic factor is expressed by the [+morphosyntactic] stem feature. The presence of this referencing feature means that within the change pattern, stems can be unified by reference to a particular MP value in comparison with the whole pattern. However, it does not necessarily mean that the stem is realizing a feature value in line with the exponent. Therefore, the fact that the stem change can be defined in reference to the paradigmatic context does not exclude stem morphomicity (Maiden, 2009). Evidence for this can be based on the stem change in derived measures which when compared with *M*1 show an MP clash, i.e. associate with cells whose MPs do not form a natural class. For instance, in the *M*1 sound series there are four instances of stem change ABCD (Table 4.15). In addition, each of the two stems (AB) and (CD) is chosen in a different sub-paradigm, possibly representing a case of Aspect-split (Corbett, 2015). Thus, both stem A and B occur in the non-perfective sub-paradigms (imperfective y-*u-gtil* 'he kills' ~ *y-u-gutl* 'they kill' and the corresponding imperative forms *ugtil* ~ *ugutil-u*) while stem C and D can only associate with the perfective. In other words, the stems represent a grouping of a natural class only of an aspect value [69] (Table 4.37).

	forms built on this stem (subject markers)	Split
Stem A	Imperf (3SGM. 1SG, 2SGM), Imper(2SGM)	Non-perf
Stem B	Imperf (1.2.3PL, 2SGF), Imper(2SGF/PL)	
Stem C	perf (1.2SG/PL, 3SGM)	Perf
Stem D	perf (3SGF/PL)	

Table 4.37 The paradigmatic distribution of sound series *M*1.

Although across the whole paradigm, the split in aspect feature can characterise the pattern of stem change, the stem selection within the perfective sub-paradigm associated with cells of contradictory person and number features that is particularly acute in the paradigmatic distribution of stems C and D. The lack of this conditioning factor in these stems can be represented by the [-morphosyntactic] feature. By contrast, morphosyntactic effects can be more clearly induced in stems C and D which selected

[69] Baerman and Corbett (2012, p.60) present a similar case of stem alternation sensitive to morphosyntactic factors from Koyi Rai in which the stem alternation can be defined in terms of person values determined by a kind of hierarchy of person and number realization.

in the perfective sub-paradigm in this measure; hence, the [+morphosyntactic] referencing feature (Table 4.34).

Likewise, a paradigmatically distributed account can be proposed for the stem change pattern in the hollow and double series. The stems in these verbs correlate with a bundle of the same feature values with the affixes i.e. can be found in paradigm cells that form a natural class in terms of MPs. In Table 4.22, Table 4.25 and Table 4.26, we showed that both *M*1 double and hollow series (stem change 1) have the pattern of stem change which includes three different stems ABC of which stem C seems to realize the third person and perfective aspect group of features.

stem change	forms built on this stem (subject markers)	Split	double series	hollow series
Stem A	Imperf, Imper	Non-perf	rudd	ṣiid
Stem B	perf (1/2SG/PL)	1/2Perf	raddee	ṣud
Stem C	perf (3SG/PL)	3Perf	radd	ṣaad

Table 4.38 The paradigmatic stem distribution in double *radd* 'return' and hollow verbs *ṣaad* 'fish' (stem change 1).

The above table shows that stem C in both series occurs in the third person form cells in the perfective. Therefore, stem C has a [+morphosyntactic] feature (Table 4.36). Although the pattern of change in stems A and (B and C) shows an aspect split, each of stems A and B associates with a contradictory set of features in terms of person and number. Indeed, in hollow series with stem change pattern 2 (ABA) (Table 4.25), stem A occurs in both imperfective and perfective sub-paradigms; hence, no aspect split.

This section showed that the stem change in TLA can have a paradigmatic distribution. The following section will discuss the morphomicity of different stems in different series and show that despite the stem's sensitivity to phonological, morphophonological or morphosyntactic factors, the stems remain fundamentally morphomic.

4.5.4 Morphomic effects on the stem change

The previous sections showed that TLA stems have sensitivity to extramorphological factors. However, neither the syntagmatic nor paradigmatic distribution can completely exclude the morphomic status of the stem. The semi-autonomous morphology in TLA is shown by different pieces of evidence including ablaut, (semi)-morphophonological conditioning and derivational morphology.

An example for semi-autonomous morphology in TLA can be reflected by the morphological complexity internal to the paradigms of *M*1 sound series which has the morphologically-triggered stem change through the ablaut in stems A B and D (Table 4.15 and Table 4.31). We saw that the ablaut targets the stem-vowel in the imperfective sub-paradigm and the 3SGF/PL perfective cells. Within the imperfective, the ablaut of stem A is lexically driven (Table 4.10 and Table 4.15) while in the case of stem B, the stem vowel appears to have a semi-autonomous status. It can assimilate with the lexically determined high FV, but not when the FV is low vowel *a*. Therefore, when the FV is the high front vowel *i* or back vowel *u*, then stem B will be CiCC or CuCC respectively. The fact that the stem vowel in imperfective forms with the FV *a* is unpredictable seems to support a lexical conditioning account for the ablaut in stem B rather than a phonological one. In addition, the choice of the FV is lexically determined, as we have established (§4.1.3), and segmented as part of the stem, giving another dimension for the morphological status of this stem. In other words, stem B needs to be listed as an m-p, morphologically induced stem, but it also has the reference feature of [+morphophonological].

Likewise, ablaut targets stem D of *M*1 sound series (CiCiC or CuCuC) (Table 4.15). Although this stem type can bear the [+phonological], [+morphophonological] and [+morphosyntactic] referencing features, it can still be morphomic. The ablaut change in this stem is lexically-specified and morphologically triggered.

3SGM	3SGF	gloss
regab	rigib-it	peek
gerab	gurub-it	approach
rebak	ribik-it	confuse
rekab	rukub-it	go up/board
gelab	gilib-it	turn over
gebal	gubul-it	accept

Table 4.39 The ablaut of stem D in *M*1 sound series.

In addition, the ablaut process is further lexically conditioned since it occurs only with ablaut-motivating suffixes: subjects and objects.[70] Therefore, there are cases in which a phonologically similar context does not trigger the stem vowel change, weakening the phonological motivation for the stem change. For instance, the forms 1/2.SG have a consonant-initial suffix, and the stems are always preceded by an epenthetic vowel /i/, resulting in the same phonological identity of the ablaut triggering suffixes. However, this phonological environment does not trigger ablaut. In the following table, the verbs of the first column have the 3SGF subject suffix which associates with ablaut and in the second and third column a very similar phonological context does not induce the change.

verb	gloss	Perfective		
		3SG.F	1SG/2.SG.M	3SG.M/Negation
gebal	'accept'	gubul-it	gebal-(i)-t	ma-gebal-(i)-š
gesam	'draw'	gisim-it	gesam-(i)-t	ma-gesam-(i)-š
gelab	'turn over'	gilib-it	gelab-(i)-t	ma-gelab-(i)-š

Table 4.40 Stem C and stem D of *M*1 sound series in a prevocalic position.

[70] Similarly, in Dakota, ablaut is lexically conditioned by particular ablaut-triggers (Kirchner, 2009).

The above data show that the ablaut stem change in the sound series in the presence of a triggering suffix is idiosyncratic rather than a case of 'phonologically conditioned allomorphy' (Anderson, 2008).

If we allow ourselves to cross the boundaries of inflection, we can obtain further evidence for the morphological autonomy of stem D ablaut. There are nominal forms which are identical to the 3SGM perfective forms in terms of stem-shape and they show possessor agreement by adding possessive pronominal suffixes which are identical to the Obj suffixes. However, unlike the verbal paradigms, the vowel-initial suffix of the possessor suffix does not result in stem suppletion as shown in Table 4.41:

Gloss	Perfective			noun	Genitive	
	3SGM	3SGF	3SGM. obj.3/2SGM		2/3SG	3PLM
kill	getal	gutul-it	gutul-a/ik	gátel	gatl-a/ik	gatel-hum
gift	rezag	rizig-it	rizig-a/ik	rízig	rizg-a/ik	rizig-hum
study	deras	diris-it	diris-a/ik	dáres	dars-a/ik	dares-hum
burn	ḥerag	ḥurug-it	ḥurug-a/ik	ḥáreg	ḥarg-a/ik	ḥareg-hum
search	beḥat	biḥit-it	biḥit-a/ik	báḥet	baḥt-a/ik	baḥet-hum
draw	resam	risim-it	risim-a/ik	rásem	rasm-a/ik	rasem-hum
hit	ḍerab	ḍurub-it	ḍurub-a/ik	ḍáreb	ḍarb-a/ik	ḍareb-hum
mature	ʕegal	ʕigil-it	/	ʕágel	ʕagl-a/ik	ʕagel-hum

Table 4.41 The verbal *M*1 sound series stem D vs nominal stems.

Table 4.41 shows that although the object and possessive suffixes are homophonous, they induce different changes on the verbal and nominal stems respectively. In the verbal paradigm, the vowel-initial 2/3SGM object suffixes /a, ik/ correlates with ablaut in stem D whereas the 2/3SG possessive suffixes triggers allomorphic variation of a single stem. Therefore, the stem alternation in the *M*1 sound series associates with a syntagmatic distribution. Nonetheless, they are morphologically specified.[71]

[71] Maiden (2013) presents similar arguments for the velar palatal alterations in Italian.

With regard to the stress shift in stem D (Chapter 3, §1.3.3), we showed that the stress is assigned on the penultimate syllable in the 3SGF perfective form and on the final syllable in the 3SGM and 1SG perfective forms. One might claim that the stem alternations are syntagmatically conditioned by the phonological environment of the stress pattern, implying that the multiple stems in this series are conditioned by a morphophonological process.[72] However, as Maiden (2005) argues for the Romance verb N-pattern, there are several objections to a stress-based account to stem change. Consider the stem alternation in different verb forms with different stress positions in (Table 4.42):

	3SG.F/PERF D	3SG.M/PERF	1SG/PERF C
affirmative	dúrub-it	deráb C	deráb-(i)t
Obj	durub-áat-a	dúrub-a D	deráb-t-a
negation	ma-durub-at-áa-š	ma-durub-áa-š	ma-derab-t-áa-š

Table 4.42 The stress pattern in stem D and stem B in *M*1 sound series *ḍerab* ‘hit’.

The stress in 3SGF perfective form (stem D) is on the short antepenultimate syllable /ḍu/ whereas the 1SG perfective form has the stress on the penultimate syllable although both forms share the same syllable structure. In addition, adding the negation clitic or the 3SG.OBJ suffix to stem C (*ma-ḍerab-t-áa-š* ‘I did not hit him’) or stem D (*ḍurub-áa-ta* ‘she did not hit him’) respectively, results in vowel lengthening which attracts the stress, without affecting the stem choice in the encountered cells. It is also worth noting that the vowels in the morphomic pattern of the verb alternate in an arbitrary fashion, since any of the three vowels: /a, u, and i/ can occur in stressed or unstressed position (Table 4.41). The stress pattern in stem D is possibly a diachronic reflex of the stress pattern in an earlier form of LA closer to CA in which the stress is

[72] Anderson (2008) argues for a stress-based account for the stem alternations in Swiss Rumantsch (§2.3.4).

possibly phonologically motivated. Therefore, stem D stress pattern can be a feature of the morphomic conditioning while the stress is possibly incidental to the ablaut process.

Another example of the synchronic morphomic patterning in TLA comes from the *M*1 double and hollow series (Table 4.22 and Table 4.25). In each of these verbs, stem B shows 'phonological incoherence' compared with the other stem changes as a result of the *ee* extension in the double series, and ablaut and vowel shortening in hollow series.

In addition, stem B shows 'functional incoherence' by occupying cells with different number and gender features. In (§4.5.2), we have showed that the effect of the morphophonological factor only applies with subject suffixes while object suffixes do not induce the same stem change; hence, the morphological effect is still intact. In fact, crossing the inflectional boundaries can also provide further support for the morphomic status of the monophthongal realization in stem B in double verbs. In nominal forms such as *radd* 'a reply' (cf. *radd* 'he replied'), the declension of the nominal form is all based on the stem *radd* with no *ee* extension even when the phonological context for the *ee* ending is met (*raddna* 'our reply').

The morphological status of stem B in hollow verbs is reflected in the ablaut (e.g. *saag* ~ *sug* 'drive', *gaas* ~ *gis* 'try on sth', *daar* ~ *dir* 'do sth', *daar* ~ *dur* 'go round sth.' *gaam* ~ *gim* 'take sth away' *gaam* ~ *gum* 'pray at night') (Table 4.27). Therefore, stem B of these verb series can have a [+morphological] referencing feature (Table 4.36). Likewise, the morphomic status of *M*1 stem C in double and hollow verbs (Table 4.22 and Table 4.25) can be supported by derivational morphology which can rule out the purely morphosyntactic account (§4.5.3). In these verb forms, stem C of all the transitive hollow verbs may not always be analysed as a multiple exponent since along with the derivation suffix, it can also specify passive (imperfective) in 1/2SG/PL

(Table 4.23 and Table 4.28). In addition, there are a number of hollow verbs in which stem C in third person perfective cells is also selected in the imperfective and the imperative cells (Table 4.25).

Similarly, stem A in the *M*1 defective series has a [+morphological] referencing feature (Table 4.35) since the ablaut is lexically determined; hence, it can be either *i* or *a* (e.g. *yi-nfi* 'exile', *ya-nsa* 'forget') (Table 4.19). Although the final stem vowel in the imperfective sub-paradigm is always in harmony with the FV, the final stem vowel itself has a lexical status. The defective stem A (stem change 1) and D (stem change 2) with a final low vowel *a* is morphologically selected in 3SGM/3SGM.Object form cell in the perfective sub-paradigm (e.g. *nsa-ak* 'he forgets you' and *nfa-ak* 'he exiles you'). Interestingly, the addition of the (V)k object suffix to 3SGF perfective form (*ns-it* and *nf-it* 'she forgot/exiled') leads to a subject suffix vowel lengthening but also requires the selection of the stem with an *a*-ending (e.g. *nsa-at-ik*, *nfa-at-ik* 'she forgot/exiled you').[73]

Clearly, the stem complex behaviours, and the interaction of morphological factors with extramorphological factors in selecting/determining the stem does not always rule out the morphomic stem status. Therefore, TLA stems fundamentally fall within semi-autonomous morphology.

---

[73] In the present study, we presented the stem alternation when a bound object suffix is attached to the 3SGM perfective form only, since this is the only unsuffixed base form. In all measures, any subject suffixed stem, did not show any further stem alternations when an object is added. However, in *M*1 defective verbs, the 3SGF and 3PLM forms, which have the vowel subject suffixes *-it* and *-u* respectively (e.g. *nf-it* and *nf-u* 'exile'), showed that when object suffixes are added to the inflected form, the vowel suffixes undergo vowel lengthening (*nf**a**-at-ik* 'she exiled youSGM., *nf-u-uk* 'They exiled youSGM' cf. *yi-nf**i*** 'he exiles' and *nfa-ak* 'he exiled youSGM'). However, in the 3SGF perfective form the suffix vowel is lengthened and lowered, but not in the 3PLF perfective form (*nfa-at-ik* cf. *nf-in-ik* but not **nfa-an-ik*). This observation is true of all *M*1 defective transitive verbs. In fact, this observation adds further morphological complexity to the *M*1 defective series which will not be further considered in this study but which could be pursued in future work.

4.6 Summary

This chapter has discussed the segmentation issue and proposed the segmentation algorithm for TLA based on the SMP in addition to reserving syncretism. The chapter also introduced the stem change patterns in TLA verb measures/series and examined the morphomicity statues of the resultant stems using the scale based on the stem referencing features. It was shown that the phonologically induced stem change resulted in stems that may be not completely unpredictable whether at the level of syllable structure or vocalization. In addition, we discussed the morphophonologically induced stem alternation and showed that every stem change in most measure and series can have a syntagmatic distribution formed by the morphophonological shape of the affix. Nevertheless, the phonological effect of exponent's shape varies from triggering stem change to no conditioning effect. Therefore, some stems remain morphomic. We also showed that the stem change in TLA can be sensitive to their paradigmatic distribution by showing aspect split. However, these stems are selected in cells with contradictory person/number features. The morphomic status of TLA stem is supported by ablaut and derivational morphology.

Overall, we have found that some stems represent semi-autonomous morphology. In some cases, only one property is required to explain the stem change. In other cases, a set of different properties can define the stems which can be morphologically induced with an arbitrary index (Chapter 5). These findings will help us establish that stem indexing in TLA has to be based on semi-autonomous morphological change, hence, any stem with a [+morphological] feature will require an individual index regardless of the degree of morphomicity.

Chapter 5 Inflection classes and implicative relations

5.1 Introduction

In Chapter 2 (§2.4.1 and §2.4.2), I introduced two approaches (the stem space and principal parts) for modelling the network of implicative relations among different cells of a lexeme's paradigm, which in turn organize complex inflectional systems. Both models also provide useful tools for understanding the role of stem alternation and formation in determining conjugation classes.

The advantage of the two approaches is that they aim for economy of description for inflection class complexity (§1.2). Inflection classes (ICs) do not have any identifiable morphosyntactic function, they are *morphomic* (Aronoff, 1994). This raises the question of how ICs can be defined cross-linguistically. In traditional descriptions, ICs are defined in terms of affixal homonymy/allomorphy. However, under the stem-space approach, Bonami and Boyé (2002) have argued that inflectional complexity in French can be defined by stem sets connected by 'dependency relations' defined by intermediate nodes of the inheritance hierarchies to minimize the amount of redundant information to be stored for a given lexeme (§2.4.1). The stem space is based on the notion of modelling regularity and defining regular relations between cells of a paradigm as networks in a stem-space connected by default relations while irregular relations are expressed by complex dependency relations. In other words, a regular lexeme requires the listing of only one stems while irregular lexemes can vary (§2.4.1 and §5.2). Similar to French, in TLA, MPs are expressed by affixation. However, the affixal inflectional exponents in all verbs including the irregular ones do not show allomorphy and instead TLA shows stem alternations distinguishing distinctive cells of the imperfective and perfective verb paradigms, which can be characterized as morphomic. In addition, the TLA verbal paradigms exhibit implicative relations; hence, some inflectional forms need not be listed since they are always predictable from the default form of any verb even for the highly irregular verbs (§5.2).

Likewise, reducing inflectional complexity to the principal parts schemas has the advantage of economy: this model identifies a set of cells which can provide inferences about the lexeme's whole paradigm. The lexeme with irregular forms requires fewer principal parts for predicting the lexeme's whole paradigm since one exceptional form allows the deduction of the rest of the lexeme's cells. By contrast, the more regular lexemes are the higher the principal parts. Measuring inflectional complexity in TLA requires addressing the representation issue (Chapter 2 §4.2.2 2.4.2.2 and §5.3.1) in addition to determining the plat input formatting that most adequately reflects the TLA inflectional system (§5.3.2 and §5.3.3). Thus, four possible typological classifications for TLA system will be examined.

The aim of this chapter is twofold. Firstly, we will propose a stem-based account for the TLA verb inflectional system, showing that the stem-space in this dialect can eliminate the amount of phonological information that needs to be stored in the lexicon. Following Bonami and Boyé's (2002, p.51) approach to French, we also show that the morphomic stem alternations in TLA can be hierarchically defined based on 'dependency relations'. This will provide further support for the model of semi-autonomous morphology proposed by Maiden (2009, 2013) discussed in the previous chapter. The morphological analysis of TLA stem allomorphy based on the stem space model is an 'abstractive' and 'thematic' approach (Blevins, 2006, p.533) and the framework will enable us to provide a breakdown of TLA conjugation classes based on the shared pattern of implicative relations between the stems. Secondly, we will provide a principal parts analysis of the TLA verb inflectional system. We will make use of the computational Principal Parts Analyser (PPA) developed by Stump and Finkel (2013) for generating three types of principal parts and providing statistical predictiveness measures within the verb paradigm.

The structure of the chapter will be as follows. Section 5.2 is devoted to discussing the abstractive approach to the morphologically induced stem allomorphy in the TLA verb inflectional system. I will use the stem space tool (§2.3 and §2.4.1) to represent the allomorphic stem distribution in all verb measures. We will also discuss the stem dependency relations and explain how the stem allomorphy of verbal paradigms can partition the lexicon into ICs based on the pattern of the stem change as distributed in the stem space, providing a new classification of the TLA verb system (§5.2.2). Following (Aronoff, 1994; Stump, 2001; Bonami & Boyé, 2002; Boyé & Cabredo Hofherr 2006; Montermini & Bonami, 2011; Pirelli and Battista 2000), I propose a stem model in which only morphologically stems are listed. For identifying the TLA stem space, I will first consider the functions linking different stems within a verb series and/or measure (§5.2.2). The relationship between the stems of one particular group of verbs (the defective and double series) shows systematic and predictable relations which allow the reconstruction of the whole paradigm of that verb once we know the form of one stem (§5.2.1). This group of verbs therefore requires us to postulate one slot in the stem space whereas verbs with alternating stems (sound and hollow series) will require a distinct slot for each stem with a [+morphological] referencing feature (§5.2). The distribution of the indexed stems is subject to dependency relations (§5.2.2) and the stem conjugation pattern results in six verb ICs (conjugation classes) (§5.2.2).

Section 5.3 is devoted to the application of the principal parts theory introduced in section 2.4.2 to the TLA inflectional system. The main aim is to investigate the complexity of the different ICs within the TLA system itself. I will mainly focus on TLA IC transparency- ‘the extent to which the realized paradigms of IC’s member lexemes deviate from the canonical ideal of maximal transparency.’ (Stump & Finkel, p.114). The IC transparency will be assessed by the four transparency criteria (Chapter 2, §4.2.5) and measured by IC predictability and cell predictability (complexity measure 8 and complexity measure 9) (Chapter 2 §4.2.4 and §4.2.5). As pointed out by Stump and Finkel (2013, p.83), the IC transparency measures ‘afford a finer-grained differentiation among ICs within a single system.’ For that reason, the discussion will not seek to develop a detailed typological comparison of TLA system with other inflectional systems. I will also introduce the different possible input formatting for TLA plat and evaluate the different representations that can adequately account for TLA complexity. In this section, I will assess the implicative structure of the paradigm of different measures/series by identifying a cell (dynamic principal-part scheme) or a collection of cells (adaptive or static principal-part schemes) that can provide inferences about the whole paradigm. The extent to which inferences can be motivated for a particular verb’s paradigm can be used to reflect the degree of the complexity of verb’s IC. Principal parts modelling requires us to construct a plat for TLA, and this raises crucial representation issues. One of those issues is how to identify a plat type that can provide a plausible account for the implicative structure and ICs (see Chapter 2, §4.2.2 for the discussion on the concrete and abstract plats and §5.3.1, §5.3.2 and §5.3.3 for TLA plat representation). A subsequent issue is how to determine the formatting of inputs within each plat and the choice of the representative items (see Appendix B for the TLA exponent-based plat and stem-based plat).

These issues require us to evaluate four different approaches to modelling the TLA principal parts and IC complexity. Three of these analyses will be based on three plats with different representations for the input while in the fourth approach, the stem space chart developed in section 5.2 will be used as a plat for TLA principal parts analysis (§5.3.2). The analysis, in which conjugation classes are based on the outer layer of the inflected forms, applies the segmentation procedures suggested by Stump and Finkel (2013) in addition to maximizing the ending to the TLA verbs. The result is distinctive patterns of TLA exponence rather than the unified account of exponents across all verbs as presented in (§4.1.3). The variation in affixes distinguishes different conjugation classes (e.g. *ketab-t* 'I/you write' and *madd-eet* 'I/you passed') as represented in the plat based on terminations of the inflected word (§5.3.1). However, this plat can be problematic and can obscure the measurement of the IC complexity based on stem allomorphy. In addition, maximizing the affixes raises the stem-affix boundary issues discussed in (§4.1.3, §4.1.3 and §5.3.1). The segmentation defended in this work, which is descriptively the most economical, reveals that nearly all the verbs of this system share the same inflectional affixes. Therefore, it is possible to consider a TLA lexeme's principal parts to be based on the stems (§5.3.3) and the TLA plat input formatting is represented by the stem formatives while the affixes are redundant and not included in the representation. Although this approach allows us to distinguish TLA ICs based on stem allomorphy, manipulating the input formatting of the stems results in increasing the number of distinct stem forms. Stump and Finkel (2015) suggested a plat type for Semitic languages such as Hebrew, in which the input for an inflected word is represented as discontinuous components. This plat represents the inflected word vocalic pattern and all the affixes including derivational ones, while the radicals of the consonantal root are, in effect, variables (expressed by hyphens). We will compare this plat type to the stem-based plat. The crucial issue with this plat is that it assumes the problematic root and pattern approach (Chapter 5, §3.3.1).

Finally, we also propose a novel approach to the principal parts analysis of TLA based on employing the stem space chart as a plat the input of which is a representation of the contrast between the stem conjugation patterns of six ICs (§5.3.2). This type of plat provides the greatest amount of information for drawing inferences over a stem-based system. However, it will be shown that the principal parts analysis based on the stem conjugation pattern does not necessarily match the schemes provided by a plat with stem indices.

When considering the complexity of TLA IC system, the relative transparency of TLA verb measure/series paradigms will be assessed by the four transparency criteria introduced in Chapter 2 (§4.2.4), repeated here for convenience (Stump & Finkel, 2013, p.83).

> A. if the number of dynamic principal parts needed to distinguish lexemes belonging to A from lexemes belonging to other ICs is smaller than the number needed to distinguish lexemes belonging to B from other lexemes;
>
> B. if the number of principal parts needed to deduce a given cell in the realized paradigms of A's member lexemes is on average, lower than the number of principal parts needed to deduce a given cell in the realized paradigms of B's members;
>
> C. if there are more alternative optimal principal-part analyses for the realized paradigms of A's members than for those of B's members; and
>
> D. if the realized paradigms of A's members have fewer cells whose realization cannot be predicted than the realized paradigms of B's members.

IC transparency correlates with IC predictability and cell predictability measures of IC complexity[1]:

---

[1] Other related measures of IC complexity include the following:
Measure 4. The larger the size of an IC system's optimal dynamic principal-part sets, the more complex it is. (related to IC transparency criteria A)

Measure 8. The lower an IC system's average IC predictability, the more complex it is.
Measure 9. The lower an IC system's average cell predictability, the more complex it is.

The advantage of these particular measures is that they enhance the use of principal parts by excluding the three traditional requirements of uniqueness '(only one of a lexeme's principal-part sets is privileged to with the label 'principal part')', uniformity (lexemes belonging to the same syntactic category have the same cells in their realized paradigms as principal parts)' and optimality '(each set of principal parts is as small as it can be, given the requirement of uniformity)' (§2.4.2) (Stump & Finkel, 2013, p.325). These characteristics can obscure the role of dynamic principal parts in identifying contrasting conjugations within the same inflectional system. Contrasting individual ICs can be achieved by the IC predictability measure which calculates 'the ratio of actual dynamic principal-part sets' to 'the candidate principal-part sets' the members of which are restricted to some arbitrary number (*m*=4). In other words, IC predictability is 'the ratio a conjugation's IC predictability is the percentage of actual dynamic principal-part sets among possible sets of a specified maximum size *m*' (Stump & Finkel, 2013, p.334). The reason for the relativization is that a large subset of lexeme's cells realizing distinct distillations is inevitably an adequate principal parts set. Likewise, the cell predictability measure is an important correlate of IC transparency and complexity. 'The average cell predictability of an IC J is the average cell predictability of the cells in a realized paradigm belonging to J; an IC system's

Measure 5. The smaller the average ratio of actual to possible optimal dynamic principal-part analyses for an IC system, the more complex it is. (related to IC transparency criteria C)
Measure 6. The higher an IC system's cell predictor number (average across ICs), the more complex it is. (related to IC transparency criteria B)
These measures abandon the traditional requirement of uniqueness and uniformity of principal parts. Each IC can choose different cell(s) of their paradigms for the dynamic principal part set in addition to having varying number of possible alternative analyses. The degree of inferences that can be provided by the selected cell can be reflected by the IC's cell predictor number.

average cell predictability is the average of its ICs' average cell predictabilities' (Stump & Finkel, 2013, p.334). The cell predictability measures the predictability of the realization of a cell within an IC on the basis of all of the other available realizations (within the same IC) which can determine the value of the realisation in question (see Chapter 2, §4.2.6 for further details on these two measures). For TLA, the complexity will be measured using the PPA tool applied to two plat types: the exponent-based plat and the stem-formative plat, providing contrasting results for the IC complexity of TLA as a root-based compared with a stem-based inflectional system.

In sum, the discussion of TLA and principal parts will consider four possible classifications for TLA verb inflection: the first type of principal parts analysis will be based on a hypothetical approach that will consider the TLA conjugation pattern distinguished by the exponents of the inflection word form. This hypothetical frame will be justified by a concrete phonological representation of the affixes along with alternative segmentation (§5.3.1). Section 5.3.2 presents the second kind of TLA principal parts modelled by the stem space framework and the stem change distribution. The last two kinds of analyses, developed according to the plat formatting suggested by Stump and Finkel (2013) for the Arabic of Bukhara (Miller, 2014) and French respectively, will be presented in section 5.3.3. The former takes a root and pattern perspective while the latter is a stem-motivated account (§5.3.3.2 and §5.3.3.1). For these two models, we will also discuss inflection-class transparency based on the four typological criteria and supported by two measures: inflectional-class predictability and cell predictability to provide precise figures for the IC transparency.

5.2 TLA stem-space

In Chapter 4, we explained that TLA verbs show stem changes due to the interaction of different motivating factors. The morphologically induced stems are characterised by ablaut change and the monophthong (-*ee*) distribution. The pattern of stem change varies across the different verb series and measures. In many non-*M*1 verbs, stem allomorphy is morphophonologically conditioned, i.e. they are allostems. By contrast, *M*1 verbs of different verb series and some non-*M*1 defective and hollow verbs are associated with a set of morphomic stems.

By basing our analysis on stems and their relations we develop a novel way of classifying verbs in Arabic, distinct from that based on the root-and-pattern interplay of root consonantism and vocalism. The verb regularity under the stem space account is based on the predictability of stem allomorphy while irregularity is associated with a greater degree of unpredictable allomorphy, requiring a larger number of arbitrary stem indices (Montermini & Boyé 2012).

In TLA, complexity and irregularity can be modelled by a unified stem change pattern and organizing unpredictable stem allomorphy by implicative relations. Following Aronoff (1994), Bonami and Boyé (2002), Pirelli and Battista (2000), we therefore propose that in TLA, the stem allomorphy across the different verb measures/series patterns together, partitioning the lexicon into ICs. We will use the framework suggested for French (Bonami & Boyé, 2002), Italian (Montermini & Boyé, 2012) and Spanish (Boyé & Cabredo-Hofherr, 2006) among others to define the TLA stem space. In Bonami and Boyé's work, the stem space is built on abstracting an allomorphic stem pattern which involves stem suppletion. However, in this study, the TLA stem space will be based on any stem associated with stem change bearing the [+morphological] referencing feature. This can give rise to a distinct stem with ablaut change, or to unconditioned monophthong elision. Thus, these types of unpredictable stems can be described as semi-suppletive stems with semi-autonomous morphology that individually motivates unique stem slots in the construction of the stem space.[2] However, in TLA the exact number of indexed stems is not always unambiguously determined, particularly when the partial-suppletion involves a stem final vowel, such as the stem change pattern in the defective and double verb series. Therefore, it is essential for the stem space modelling to determine the stem boundaries.

5.2.1 The decomposition procedure

In this word and paradigm approach, the decomposition of complex forms into subcomponents is achieved by deducing the maximum number of forms with the minimum number of functions. Therefore, the segmentation problem seems to be less relevant for the stem space analysis. However, for the purpose of motivating slots in the stem space, Boyé and Cabredo-Hofherr (2006) suggest a number of criteria for decomposition based on suppletion. According to these criteria, an inflected verb form can belong to two decomposition types. The first type consists of two subparts (stem + ending) while the second type is a single unanalysable listed form. Although the stem itself is indecomposable, it can determine the stem-affix boundaries. Thus, to identify the verb form segmentation, Boyé and Cabredo-Hofherr (2006, p.2) propose two hypotheses, (1) and (2):

(1) Hypothesis-1: 'The ending is the invariant suffix that appears with suppletive stems.'

(2) Hypothesis-2: 'The verbal forms which are idiosyncratic with respect to their endings are excluded from the comparisons establishing the invariant suffix'

a. am + o (analysable stem)

[2] By 'semi-suppletive stem' we mean a stem X which is systematically related to stem Y syntactically and semantically, but formally distinct from Y through a morphomic process of vowel modification.

b. caig + o (stem-suppletion, analysable)
c. sé (form-suppletion, unanalysable)

The task still remains of deciding which of the two hypotheses takes precedence in certain cases. If we apply Hypothesis-1 on its own to the form *amo*, for instance, we will find that there is no part of the word form that can be segmented so as to provide an invariant inflection suffix. The only candidate for such an affix is the (traditional 1SG marker) *-o*, but this affix is lacking in the suppletive form *sé*. However, by Hypothesis-2 we can exclude the *sé* form, intuitively, because it is too idiosyncratic (it is a whole word suppletive form). This would then allow us to treat the *-o* endings of *amo*, *caigo* as inflections and would justify the segmentations *am-o*, *caig-o*.[3]

Applying the Boyé/Cabredo Hofherr segmentation procedures to TLA is not a straightforward task. The base of decomposition under this analysis is mainly determined by suppletion, and this can vary depending on the verb measure/series. However, we can examine the application of Hypothesis-1 using the two idiosyncratic verbs *ree* 'see' and *jee* 'came'. In these verbs, the FV in the imperfective sub-paradigm has to be assigned to the stem since the invariant affix of all verbs is the *y*-segment (Table 5.1).[4]

By Hypothesis-1 the *ee*-monophthong would be part of the stem, given that it is an invariant suffix across all verbs (Table 5.1). The application of this hypothesis only to defective series verbs treats the inflectional form as an analysable (stem + ending) type with different patterns of stem-allomorphy (Table 5.2). However, it is important to

[3] The application of Hypothesis-2 provides the opposite results of the strictest interpretation of the SMP (Stem Maximisation Principle), according to which the *o*-ending would have to be segmented as part of the stem because that ending does not occur in forms such as *sé*.
[4] By contrast, the partial suppletion in the verbs *klee* 'eat'~ *ya*-kil and *xdee* 'take' ~*ya-xid* shows that the FV is combined with the *y*- element to form an invariant affix.

emphasize that, in TLA, the essential criterion for determining the decomposition is based on the invariant suffix that co-occurs with different measures and series.

verb	Gloss	Subject		
		Imperfective	Perfective	
		3SGM	1PL	3SGM
ṣehad	burn	y+*a*ṣhid	ṣehad+na	ṣehad+Ø
nsee	forget	y+*a*nsa	ns*ee*+na	ns*ee*+Ø
nfee	exile	y+*i*nfi	nf*ee*+na	nf*ee*+Ø
gall	take	y+gill	gall*ee*+na	gall+Ø
baaʕ	sell	y+biiʕ	biʕ+na	baaʕ+Ø
jee	come	y+ji	j*ee*+na	j*ee* + Ø
ree	see	y+ri	r*ee*+na	r*ee* + Ø
klee	eat	y+*a*kil	kl*ee*+na	
xdee	take	y+*axid*		

Table 5.1 The decomposition of the FV and the monophthong *ee* under the stem space model.

verb	Gloss	Subject Perfective 1PL
jee	come	j+*ee*[5]+na
ree	see	r+*ee* +na
nsee	forget	ns+*ee*+na
nfee	exile	nf+*ee*+na
gall	take	gall+*ee*+na

Table 5.2 The decomposition of the monophthong form as analysable (stem-ending) type.

This decomposition gives rise to a systematic distribution of allomorphs, which can be represented by a stem space. According to Boyé and Cabredo-Hofherr (2006, p.5), ‘the

[5] By contrast to the other defective series of *M*1 and non-*M*1 verbs, the verbs *klee* ‘eat’ and *xdee* ‘take’ show stem formation variability which does not affect the FV in the imperfective (e.g. *ya-kil* ‘he eats’, *ya-xid* ‘he takes’). However, the effect of the variability is present on the termination (*ee*-monophthong) of the imperfective and the imperative (e.g. *kuul* ‘eat!’) sub-paradigms that show stem suppletion with a hollow series stem-shape (CVVC).

maximal stem space gives the maximally possible number of stem-suppletions'. That is, the distinction between different stems has to associate with a stem-suppletion pattern which can in turn determine the number of stems and the size of the stem space. Thus, the stem space in TLA is a vector which represents the maximum abstraction of the theoretically possible stems on which each verb measure and series are built. This means that the maximal stem space in TLA is derived from the maximum number of stems with ablaut and/or unconditioned monophthong elision displayed by the pattern of stem change which can be obtained from different verb measure/series.

Consider the stem alternations in Table 5.3, which illustrates just those morphologically induced stem changes which occur across all the verb series/measure. Each unpredictable stem is represented by a different letter.[6]

verb	Gloss	*M*	Imperfective		Perfective		
			3SGM	3PLM	1PL	3SGM	3SGF
sound series							
getal	kill	*M*1	ugtil A	ugutl B	getal C	getal C	gutul D
ʕallim	teach	*M*2	ʕallim A	ʕallm A	ʕallim A	ʕallim A	ʕallm A
ʕaamil	deal	*M*3	ʕaamil A	ʕaaml A	ʕaamil A	ʕaamil A	ʕaaml A
teʕallim	learn	*M*5	teʕallim A	teʕallm A	teʕallim A	teʕallim A	teʕallm A
teʕaamil	deal	*M*6	teʕaamil A	teʕaaml A	teʕaamil A	teʕaamil A	teʕaaml A
insemaʕ	be heard	*M*7	insimaʕ A	insimiʕ B	insemaʕ C	insemaʕ C	insimiʕ B
irtefaʕ	be carried	*M*8	irtifaʕ A	irtifiʕ B	irtefaʕ C	irtefaʕ C	irtifiʕ B
staʕmil	use	*M*10	staʕmil A	staʕml A	staʕmil A	staʕmil A	staʕml A
defective series							
nsee	forget	*M*1	ansa A_i	ans A_j	nsee B_i	nsee B_j	ns B_k
bakka	upset	*M*2	bakki A	bakk A	bakkee A	bakka A	bakk A
naada	invite	*M*3	naadi A	naad A	naadee A	naada A	naad A
temanna	wish	*M*5	temanna A	temann A	temannee A	temanna A	temann A

[6] Note that, unlike the representation approach presented in Chapter 4 in which every stem change regardless of the motivating factor is represented by a different letter, in this chapter the letters only reflect the morphologically motivated stem.

telaaga	meet	*M*6	telaaga A	telaag A	telaagee A	telaaga A	telaag A
inkiwee	cauterised	*M*7	inkiwi A	inkiw A	inkiwee A	inkiwee A	inkiw A
intisee	forgotten	*M*8	in(t)isa A	in(t)is A	in(t)isee A	in(t)isee A	in(t)is A
stanja	purify	*M*10	stanja A	stanj A	stanjee A	stanja A	stanj A
				double series			
gall	take	*M*1	gill A	gill A	gallee B	gall B	gall B
ḥaddid	specify	*M*2	ḥaddid A	ḥaddid A	ḥaddid A	ḥaddid A	ḥaddid A
tesammim	poisoned	*M*5	tesammim A	tesammim A	tesammim A	tesammim A	tesammim A
ingall	taken	*M*7	ingall A	ingall A	ingallee A	ingall A	ingall A
irtadd	be returned	*M*8	irtadd A	irtadd A	irtaddee A	irtadd A	irtadd A
steġall	exploit	*M*10	steġIll A	steġIll A	steġallee B	steġall B	steġall B
				hollow series			
baaʕ	sell	*M*1	biiʕ A	biiʕ A	biʕ B	baaʕ C	baaʕ C
baat	stay over	*M*1	baat A	baat A	bit B	baat A	baat A
injaab	be brought	*M*7	injaab A	injaab A	injib B	injaab A	injaab A
iltaam	blame	*M*8	il(t)aam A	il(t)aam A	il(t)um B	il(t)aam A	il(t)aam A
stefaad	benefit	*M*10	stefiid A	stefiid A	stefad B	stefaad C	stefaad C

Table 5.3 The morphological stems in the representative forms of the subject sub-paradigm across all measures/series.

The data in Table 5.3 show that the *M*1 sound series, which is the most frequent verb form, has multiple stems in the perfective and imperfective aspect sub-paradigms while non-*M*1 verbs have less/no unpredictable stem variation (see Table 5.8 and the listing in (3) below for further details on the defective and double series stem alternations). Should each morphological stem of each series (including *M*1 defective verbs *klee* ‘eat’ or *xdee* ‘take’) be included in abstracting the stem space, then the result will have to include eight different stems cells obtained from all the series. However, within each series, there can be no more than four indexed stems. These eight slots are associated with a morphologically conditioned stem in at least one of the verb series in *M*1.

Nevertheless, using the verb paradigm of *M*1 defective verbs *klee* and *xdee* alongside other measures as the main criteria for generating the maximal stem space results in increasing the size of the stem space. Thus, in this case the maximal possible isolated morphomic stems found in the system is shown in the partition illustrated in Table 5.4.

Subject						3SGM.Obj	
Imperfective		Imperative	Perfective			Perfective	
3SGM	3PLM		1PL	3SGM	3SGF	1PL	3SGM
S1	S2	S^i	S3	S4	S5	S^j	S^k

Table 5.4 The initial TLA stem space.

The initial stem space in Table 5.4 represents the morphologically motivated instances of stem change that can be found in all verb measures/series of the system, including the verbs *klee* 'eat' and *xdee* 'take'. It has eight stem slots and this size is mainly motived by the partial suppletion in S^i slot that can be found only in the stem inventory of the two verbs *klee* and *xdee*. It is important to note that the reason for considering the verbs *klee* and *xdee* in the initial formation of stem space is based on the account presented by Boyé and Cabredo-Hofherr (2006), according to whom the stem space is derived by analysable stem-suppletion while unanalysable forms have to be listed and excluded from the organization of a stem space. Both *klee* and *xdee* have analysable stem-suppletion forms: the forms in the imperative are analysable and have 'idiosyncratic stem-suppletion', thus, they should have a slot designated in the stem space. However, enlarging the stem space based on the stem pattern of only two verbs would depend on 'the regularities one wants to be able to express.' (Montermini & Boyé, 2012, p.72). Given that this is not a systematic suppletion-pattern in the language, we opt for excluding these verbs from determining the stem space organization.

The upshot is that the maximal stem inventory required in the system will be reduced to *potentially* include seven slots for seven stems with the [+morphological] feature.

Subject					3SGM.Obj	
Imperfective		Perfective			Perfective	
3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
S1	S2	S3	S4	S5	S^j	S^k

Table 5.5 TLA stem space excluding the verbs *klee* ‘eat’ and *xdee* ‘take’.

The seven stem slots are filled by stems that in principle can be partial suppletive. Each of these stems correspond to a [+morphological] stem change in one or more verb series of *M*1 which represent the maximal possible stem-suppletion pattern that can be found in the system. For instance, S1 can be identified as an indexed stem by the stem change pattern in any series of *M*1 while S3 is represented by *M*1 double and/or hollow series. By contrast, S2 and S5 can be motivated from the *M*1 sound series. S^j is listed due to the morphologically induced monophthong elision in the *M*1 double series whereas S^k has a [+morphological] stem feature in the defective series (i-vowel final verbs). In sum, seven stem slots in the stem space are sufficient to provide the ‘overall distribution schema’.[7] It is worth noting that the allomorphic variation displayed by the verb lexemes varies with the verb series/measure rather than being a property of individual regular lexemes[8] (Table 5.6).

[7] The term is used by Pirelli and Battista (2000) for representing the notion of stem space.
[8] A regular lexeme is represented by a single stem while a less regular lexeme can be based on more than one suppletive-stem (Bonami & Boyé, 2006).

series	verb	Gloss	Subject					3SGM.Obj	
			Imperf		Perf			Perf	
			3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
sound	ṣehad	burn	aṣhid	aṣuhd	ṣehad	ṣehad	ṣuhud	ṣehad	ṣuhud
hollow	šaaf	see	šuuf	šuuf	šuf	šaaf	šaaf	šaaf	šaaf
defective	nfee	exile	infi	ifn	nfee	nfee	nf	nfa	nfa
double	gall	take	gill	gill	gallee	gall	gall	gall	gall
defective	garra	teach	garri	garr	garree	garra	garr	garra	garra
stem space			S1	S2	S3	S4	S5	S^j	S^k

Table 5.6 Examples of the [+morphological] stems (in greyed cells) which motivate TLA stem space.

Table 5.6 shows that the verb inflectional complexity in TLA can be characterised by a stem space that consists of a set of stems each of which is connected to a slot in the inflectional paradigm. In the stem slots in the imperfective paradigm, the inflectional forms can be built on the same or distinct stems, depending on the verb series/measure. By contrast, the 1/2SG/PL perfective inflectional forms share the same stem in the stem space slot in all verbs. The 3SG/PL perfective inflectional forms may select the same or distinct stem slots: the 3SGM slot and 3SGF/PL slot (Table 5.3 and Table 5.6). In fact, the patterns of stem change in most non-*M*1 sound verbs partition the whole paradigm into a set of cells which tend to be filled by the same stem. Since different verbs have different patterns of stem change, consequently not all verbs can have the same indexed stems in the same stem slots. In other words, the stem space slots/indexed stems in different verbs may correspond to the same stem base (e.g. *M*2, *M*3, *M*5, *M*6 sound series) or to four different stem bases (*M*1 sound series and defective (*i*-vowel final verbs)) (Table 5.6). Nevertheless, it is important to note that it may not be necessary to postulate a separate slot in the stem space for the 3SGM form for the reason that although it shows a partial suppletive change in *M*2, *M*3, *M*5, *M*6 defective verbs and *M*1 hollow series, it can be constructed from other already listed stems.

Nevertheless, '…storing more than one is harmless' since the overall structure of the model represents the principal parts of lexemes and increasing the size of the stem space ultimately 'depends on the regularities one wants to be able to express.' (Montermini & Boyé 2012, p.72).[9]

If we assume a single stem for the 1PL and 3SGM forms of the *M*1 sound series and the majority of other verb series, then the S3 slot for all verb measures and series in general may lead to analytical problems when it comes to determining the conjugation pattern. The reason for that is that the stem zone coincides with two different stems in certain verb measure/series. The term 'stem zone' refers to a cell/slot in the stem space specified for suppletions. The stem zone coincides with a cell or cells of the paradigm. For instance, the hollow series has BC or BA stem change (Table 5.3) and the defective series of *M*2/*M*3 have CD stem change. In other words, the inflectional forms that are built on these indexed stems can vary depending on the verb series. Therefore, in some cases it is important to specify the verb series in the stem distribution (Table 5.7).

[9] The listing of the affixless 3SGM form as an indexed stem means that it has to be treated as an unanalysable form-suppletion in non-defective verbs (e.g. *maat* 'he died'). Boyé and Cabredo-Hofherr (2006) provide a similar analysis for the 2SG imperative stems such as *pon* 'put', *haz* 'make' which are treated as listed form-suppletion rather than an analysable stem-suppletion (*haz* + Ø).

Indexed stem	forms built on these stems in different measures/series
S1	a. imperfective/imperative (*M*1 sound [before Ø.subj or Ø/consonant-initial.obj suffixes], *M*1 & non-*M*1 hollow, *M*1 & *M*10 double, *M*1 & non-*M*1defective [before Ø.subj. or any obj. suffixes]). b. imperfective/imperative/perfective (non-*M*1[10]sound, *M*2, *M*5, double)
S2	a. imperfective/imperative (*M*1 sound [before vowel-initial subj/obj suffixes], *M*1 & non-*M*1defective [before vowel-initial subj. suffixes]) b perfective (non-*M*1defective [before vowel-initial subj. suffixes])
S3	a. 1/2 perfective (non-*M*1 defective, *M*1 & *M*10 double), b. 1/2 perfective & 3SGM (*M*1 sound and defective)
S4	a. 3 perfective (*M*1& *M*10 hollow, *M*1&*M*10 double) b. 3SGM perfective (non-*M*1defective)
S5	perfective 3SGF/3PL (*M*1 sound, *M*1 defective)
S^j	perfective 3SGM (*M*1 double [before consonant-initial obj. suffix])
S^k	perfective 3SGM (*M*1 & non-*M*1defective [before vowel-initial obj suffix])

Table 5.7 The distribution of TLA indexed stems in the stem space.

Table 5.7 illustrates that filling the slots of the TLA stem space has to be based on a systematic stem distribution capable of deriving a network of implicative relations (principal parts). However, no verb can achieve the maximal theoretical degree of stem pattern complexity displayed in Table 5.6 and Table 5.7. In fact, in the majority of the verb measures, including the sound series (*M*2, *M*3, *M*5, *M*6 and *M*10) and double series (*M*2, *M*5 and *M*10), only one indexed stem is required to reconstruct the whole paradigm.

In addition, as Montermini and Boyé (2012, pp.72-73) pointed out for Italian, the relationship between the stems and/or stems and forms can be derived by a function. The output of the function can be the same as the input (identity function). However, the function 'may alter the phonology of the input, this alternation may, but need not be phonologically motivated.' In TLA, the distribution of the indexed stems in the *M*1 sound and hollow series (§4.3.1 and §4.3.4) cannot be defined by a set of default

[10] *M*7 and *M*8 are not included.

relations and expressed by functions as in (3), due to the ablaut exhibited by the indexed stems S1, S2, S3, S4 and S5 (Table 5.6). These stems are morphologically induced each has a lexically determined vocalism and they cannot be defined in terms of stem to stem relations, reducing the five indexed stems to one single representative stem (see §4.3 for further discussion on TLA stem change). By contrast, the S^j and S^k stems are based on the vowel final alternation in addition to the semi-autonomous morphological characteristics displayed by the unconditioned vowel elision pattern for which it is possible to establish a set of default relations ('stem to stem relations'), connecting the stems in defective series, as illustrated in Table 5.8 (cf Montermini & Bonami, 2011, p.11).

3SGM.imperf-3PLM.imperf	1PL.perf-3SGM.perf	1PL.perf-3SGF.perf	1PL.perf-3SGM.obj	3SGM.imperf-1PL.perf	3SGM.imperf-3SGM.obj
S1-S2 XV-X	S3-S4 Xee-Xee Xee-Xa	S3-S5 Xee-X	S3-S^k Xee-Xa	S1-S3 VXV-Xee XV-Xee	S1-S^k VXV-Xa
ansa-ans infi-inf bakki-bakk	nsee-nsee nfee-nfee bakkee-bakka	nsee-ns nfee-nf bakkee-bakk	nsee-nsa nfee-nfa bakkee-bakka	ansa-nsee infi-nfee bakki-bakkee	ansa-nsa infi-nfa bakki-bakka

Table 5.8 Stem-to-stem relations in *M*1 and non-*M*1 defective verbs.

The formalism adopted in (Table 5.8) reflects the following:

(3)

a- S1 is related to S2 by truncating the final stem vowel;
b- S3 is related to S4 by identity in *M*1 or replacing the final vowel with [a] in non-*M*1;
c- S3 is related to S5 by truncating the final stem vowel;
d- S3 is related to S^k by replacing the final stem vowel [ee] with [a];
e- S1 is related to S3 by truncating the FV and replacing the final vowel with [ee];
f- S1 is related to S^k by truncating FV and replacing the final vowel with [a] or maintaining the final vowel [a].[11]

[11] The similar mapping defines the double series stem alternation in *M*1 and *M*10

stem relation in double series	S1-S2 X-X	S3-(S5/S^j) Xee-X
*M*1	gill-gill	gallee-gall
*M*10	steġill-steġill	steġallee- steġall

These relations identify a class of verbs characterized by final stem vowel alternations all of which can be expressed by a function. In the defective series, from S1, it is possible to predict the whole paradigm while other surface stems cannot determine the ablaut in S1. Similarly, as Montermini and Bonami (2011) point out, Italian has cases in which a single stem zone is associated with more than one surface stem each of which is not clearly identified as an indexed stem by lack of a suppletive case.[12] In addition, there are verbs such as *lava* 'to wash', which show final vowel stem alternations in different stem zones of suppletion, but the stems themselves are not the reason for providing the different stem slots, and so they need not be indexed. In other words, each of these stems can be derived by a set of functions expressing the morphologically conditioned alternation of the thematic vowel.

Likewise, in TLA, stem zones S1 to S5 are reflected by the morphologically conditioned ablaut or (partial-suppletion) in the M1 sound and hollow series. By contrast, the stem zones S^j and S^k are justified by [+morphological] stems in the defective series motivated by final vowel ablaut, and in double verbs reflected by the absence of the *ee*-monophthong which is morphophonologically conditioned otherwise. If we opt to treat the final stem vowel in the defective verbs in line with the account proposed for the thematic vowel in Romance languages Spanish, Italian and French (Bonami & Boyé, 2002; Boyé & Cabredo-Hofherr, 2006), then the defective (of all measures) and double (M1) series can be derived from the listed stem S1 and/or S3. The relations between S1 and/or S3 and other stems will correspond to final stem-

---

a- S1 is related to S2 by identity;
b- S3 is related to S5/S^j by truncating the final stem vowel.

[12] Likewise, Spanish has four different surface stems for the inflected forms in the stem slot S6 which is identified by a single suppletive case in 'Preterite, Imperfective 1&2 Subjunctive and Future Subjunctive'. In other words, suppletion is the essential criterion for licensing a stem zone in the stem space while vowel alternations at the edge of the stem do not necessarily require a separate stem slot (Boyé & Cabredo-Hofherr, 2006, p.8)

vowel variation or truncation. Similarly, in the double series, the stems can be linked to each other by the identity function or a function specifying the addition of the monophthong. The result of this analysis is that the TLA verb monophthong is similar to the thematic vowel of Romance languages in conditioning regular stem change within that paradigm of regular verbs. In other words, M1 and non-M1 defective and M1 double verbs can be classified as regular verbs that require one indexed stem. This indexed stem the undergoes stem allomorphy located at the site of the monophthong or stem final vowel. Consequently, the TLA stem space would be reduced to five slots while S^j and S^k motivated by the verbs in question[13] may be treated as surface stems morphologically induced by the final stem vowel variation and the default relations between these stem functions. Consider Table 5.9, which shows stem alternations of the defective series in different measures.

verb	Gloss	Subject					3SGM.Obj	
		Imperf		Perf			Perf	
		3SGM	3PLM	1PL	3SGM	3SGF	1PL	3SGM
Underlying stems		S1		S3				
nsee	forget	ansa	ans	nsee	nsee	ns	nsa	nsa
nfee	exile	infi	ifn	nfee	nfee	nf	nfa	nfa
bakka	upset	bakki	bakk	bakkee	bakka	bakk	bakka	bakka
naada	invite	naadi	naad	naadee	naada	naad	naada	naada
tewalla	take over	tewalla	tewall	tewallee	tewalla	tewall	tewella	tewella
surface stems		X_i	X_j	Y_i	Y_j	Y_k	Y_l	Y_l

Table 5.9 The stem allomorphy of M1 and non-M1 defective series (the XY letters represent the surface allomorphic stem variations of stem A in (Table 5.3)).

Recall from Chapter 4, §4.3.1, that in the M1 defective series, the imperfective 3SGM stem in the subj-sub-paradigm can only occur with unsuffixed forms while the

[13] It is worth noting that as in the defective series, in the sound series, S^k is associated with independently motivated ablaut. However, it shows a stem-to-stem relation to S4 and can be expressed by the identity function (e.g. S^k *ʕuruf-a* ‘he knew him’~ S4 *ʕuruf-it* ‘she knew’).

imperfective 3PLM stem occurs before a V-initial suffix. However, the stem of the imperfective 3SGM form (or any unsuffixed form) is also used before V/C-initial object suffixes. Likewise, the perfective 1PL/3SGM stem can have the *ee*-monophthong in M1, while in non-M1 the corresponding stem has the final vowel *a* which is truncated before V-initial subject suffixes in all measures. By contrast, the monophthong of the 1PL/3SGM stem in M1 alternates with the vowel *a* in the 3SGM.obj-sub-paradigm.

Under the stem space approach, the stem pattern of defectives can be handled by two sub-functions. The first one generates the stem (F_s) while the second realizes the inflected form by combining the stems and the relevant suffixes (F_{+A}). In other words, the relationship between the stem and the forms is expressed by the same formalism as that for the relationship between the stems (Montermini & Bonami, 2011). Therefore, through the combination of the sub-functions, the inflectional forms of the defective series (Table 5.9) are morphologically derived from two underlying stems S1 and S3 as shown in (4) (adapted from Boyé & Cabredo-Hofherr, 2006, p.7).

(4)

Underlying Stem ⇒ Surface Stem ⇒ Inflected Form
a. F_S: Underlying Stem ⇒ Surface Stem
b. F_{+A}: Surface Stem ⇒ Surface Stem + A = Inflected Form
c. F_{+A} (FS(Underlying Stem)) = Inflected Form

For the S1 stem, there are two surface stems X_i and X_j respectively for the imperfective 3SGM and 3PLM. In (5) and (6), we give the functions FS1X_i and FS1X_j, realizing the surface stems from the underlying one.

(5)

F_{S1X_i}: S1 ⇒ Surface Stem for imperfective 3SGM: $S1X_i$
S1 ⟶ $S1X_i$
identity

F_{S1X_i} (ansa) = ansa
F_{S1X_i} (infi) = infi
F_{S1X_i} (bakki) = bakki

(6)

F_{S1X_j}: S1 ⇒ Surface Stem for imperfective 3PLM: $S1X_j$
S1 ⟶ $S1X_j$
truncate final V

F_{S1X_j} (ansa) = ans
F_{S1X_j} (infi) = inf
F_{S1X_j} (bakki) = bakk

The two surface stems are distinguished only by the presence or absence of the final vowel. Likewise, the three different surface stems for the inflected forms in the ablaut change zone (S3) can be derived by the functions $FS3Y_i$ and $FS3Y_j$ in (7) and (8):

(7)

F_{S3Y_j}: S3 ⇒ Surface Stem for perfective 3SGM: $S3Y_j$
S3 ⟶ $S3Y_j$
identity

a. F_{S3Y_j} (nsee) = nsee
F_{S3Y_j} (nfee) = nfee

b. F_{S3Y_j} (bakka) = bakka
F_{S3Y_j} (naada) = naada
F_{S3Y_j} (tewalla) = tewalla

(8)

F_{S3Y_k}: S3 ⇒ Surface Stem for perfective 3SGF: $S3Y_k$
S3 ⟶ $S3Y_k$
truncate final V

a. F_{S3Y_k} (nsee) = ns
F_{S3Y_k} (nfee) = nf

b. F_{S3Yk} (bakka) = bakk
F_{S3Yk} (naada) = naad
F_{S3Yk} (tewalla) = tewall

(9)
F_{S3Yl}: S3 ⇒ Surface Stem for perfective 3SGM.Obj.3SGM: $S3Y_l$
F_{S3Yl}: S3 ⇒ $S3Y_l$ = S3 with a modification of final V: ee ⇒ a, a ⇒ a

a. F_{S3Yl} (nsee) = nsa
F_{S3Yl} (nfee) = nfa

b. F_{S3Yl} (bakka) = bakka
F_{S3Yl} (naada) = naada
F_{S3Yl} (tewalla) = tewalla

This morphological model is based on a realizational analysis according to which functions applied to the stem realize the inflected forms. These functions do not assign features to the stem, but only express a set of features realized by the inflected forms. Therefore, the inflected forms based on S1 and/or S3 can be derived by functions F_N that realize these inflected forms by combining the two sub-functions as represented by the formalism for the constructing the perfective forms of the verb *nfee* 'exile' in (10):

(10)

a. F_1 (S3) ⇒ perfective 3SGM e.g. F1 (nfee) = *nfee*
b. F_2 (S3) ⇒ perfective 3SGF e.g. F2 (nfee) = *nf*
c. F_3 (S3) ⇒ perfective 3SGM.obj.3SGM e.g. F3 (nfee) = *nfa*

where

F_1 (S3) = $F_{+Ø}$ (F_{S3Yj} (S3)) = (S3) + Ø
nfee ⟶ nfee + Ø
keep the stem, add + Ø

F_2 (S3) = F_{+it} (F_{S3Yk} (S3)) = (S3- ee) + it
nfee ⟶ nf+ it
truncate ee, add + it

F_3 (S3) = F_{+a} (F_{S3Yl} (S3)) = (S3 (ee ⇒ a)) + a
nfee ⟶ nfa + a

modify final V (ee => a), add + a [14]

Clearly, the distinction between surface stems in the defective series of any measure shows that the distinguishing criteria identifying these classes of verbs are the morphologically motivated modification/or truncation of stem final vowels. Thus, the relationship between S1 and S3 can be related by assuming a function that deletes the FV and modifies the stem vowel of S1 to the monophthong *ee* in M1 and the low vowel *a* in non-M1. In other words, the stem alternation in this series can be connected by systematic default relations which help reconstruct the whole paradigm of the verb through one and/or two listed stem(s), reducing the stem space in Table 5.5 to five zones (Table 5.10).

Subject				
Imperfective		Perfective		
3SGM	3PLM	1PL	3SGM	3SGF
S1	S2	S3	S4	S5

Table 5.10 TLA stem space (final version)

It can be clearly seen that based on the stem space framework, the paradigm can be partitioned into five zones for the [+morphological] stems associated with ablaut. In addition, this approach promotes limiting the stem variation suggested in the zone S^j

[14] The functions F_n for the inflection forms based on the two sub-functions of S3 might possibly be formalised as follows:

a. F1 (S1) ⇒ imperfective 3SGM e.g. F1 (infi) = *yinfi*
b. F2 (S1) ⇒ imperfective 3PLM e.g. F2 (infi) = *yinfu*

where F_1 (S1) = F_{+y} (F_{S1Xi} (S1)) = y + (S1)
infi ⟶ y + infi
keep the stem, add + y

F_2 (S1) = $F_{+y\text{-}/\text{-}u}$ (F_{S1Xj}(S1)) = y + (S1- final V) + u
infi ⟶ y + inf + u
truncate final V, add + y, add + u

and S^k by considering them as surface stems derived by functions from an underlying stem.

The following section will discuss the local stem relations that structure the stem space and which defines the different ICs based on the stem alternations.

5.2.2 Inflection classes and stem relations

The work on stem spaces in Romance languages (Bonami & Boyé, 2002; Montermini & Bonami, 2011) distinguishes two types of allomorphy; regular and irregular. In regular allomorphy the stems (which need not be identical) are linked by regular relations derived from only one indexed stem. The irregular type of stem change necessarily involves lexicalised suppletive stems with distinct forms that cannot be determined from another listed stem. Based on these patterns of stem allomorphy it is thus possible to distinguish regular and irregular inflection classes.

Likewise, the stem allomorphy in TLA can distinguish different ICs, which leads to the reclassification of the verb measures and series into different sets of verb classes. The combination of these stem slots represents the conjugation pattern in TLA which can have six different classes across all verb forms, four of which are patterns of the M1 verb. The criteria for different conjugations is based on the pattern of stem change across all the stem space slots. If a verb has the stem alteration pattern AA and another verb has AB then these two verbs will be assigned to two different conjugations (Table 5.11).

verb	Gloss	verb series.M	Imperfective		Perfective		
			3SGM	3PLM	1PL	3SGM	3SGF
getal	kill	sound *M*1	ugtil A	ugutl B	getal C	getal C	gutul D
ʕallim	teach	sound *M*2	ʕallim A	ʕallm A	ʕallim A	ʕallim A	ʕallm A
insemaʕ	be heard	sound *M*7	insimaʕ A	insimiʕ B	insemaʕ C	insemaʕ C	insimiʕ B
gall	take	double *M*1	gill A	gill A	gallee B	gall B	gall B
baaʕ	sell	hollow *M*1	biiʕ A	biiʕ A	biʕ B	baaʕ C	baaʕ C
baat	stay over	hollow *M*1	baat A	baat A	bit B	baat A	baat A
Stem space			S1	S2	S3	S4	S5

Table 5.11 The conjugation pattern in TLA based on stem alternations.

Table 5.11 shows that the pattern of stem allomorphy gives rise to six different ICs, shown schematically in Table 5.12. Class I is the dominant IC in the system and it includes the lexemes in which the *allostems* (Spencer, 2012) are derived from a single underlying stem. For instance, the sound series {*M*2, *M*3, *M*5, *M*6, *M*10} has regular stem allomorphic relations. Likewise, the defective series of any measure has 'RELATED STEMS' (Boyé & Cabredo-Hofherr, 2006) allomorphy based on the default relation which express variations located at the site of the final stem vowel. By contrast, the other classes, including the regular verb series specify suppletive stems, resulting in two to four isolated stems. For example, the M1 sound series has the highest number of unrelated stems. In fact, unlike Romance languages, in TLA, the class with stem alternations is represented by what is traditionally considered the regular verb series of the default (simple/basic) verb measure.

Class	Imperfective		Perfective			verb measures and series
	3SGM	3PLM	1PL	3SGM	3SGF	
SS	S1	S2	S3	S4	S5	
I	A	A	A	A	A	sound {*M*2, *M*3, *M*5, *M*6, *M*10} double {*M*2, *M*5, *M*7, *M*8} defective {*M*1, *M*2, *M*3, *M*5, *M*6, *M*7, *M*8, *M*10}
II	A	A	B	A	A	hollow {*M*1, *M*7, *M*8}
III	A	A	B	B	B	double {*M*1, *M*10}
IV	A	B	C	C	B	sound {*M*7, *M*8}
V	A	A	B	C	C	hollow {*M*1, *M*10}
VI	A	B	C	C	D	sound {*M*1}

Table 5.12 The TLA ICs under the stem space framework.[15]

Unlike class I, the pattern of stems in class II-V cannot always be defined in terms of one another due to morphologically conditioned ablaut change which results in one or a set of isolated stems. However, Bonami and Boyé (2002) propose that the distribution of unrelated stems inside the stem space is not entirely arbitrary and it can show some degree of organisation. Consider the following stem relations in Class II-VI.

verb class	stem relations
Class II	{S3},{S1, S2, S4, S5}
Class III	{S1, S2}, {S3, S4, S5}
Class IV	{S1},{S2, S5},{S3, S4}
Class V	{S3},{S1, S2},{S4, S5}
Class VI	{S1},{S2},{S5}{S3, S4}

Table 5.13 The set of related stems in Classes II-VI.

The verb categories can be based on two to four unrelated sets of related (including identical) stems. Therefore, filling the stem space is constrained by dependency relations of varying complexity. The default stem is S1 (the stems of the imperfective sub-paradigm) which predicts most of the stem space. Therefore, similar to French, the

[15] Specifying a separate slot for 3SGM form does not affect the number of inflection classes in the system. However, it can show further variations or distinctions among the different categories.

TLA stem space can be constrained by the dependency relations (11) that hold among the stem sets in a verb paradigm.

(11) **The dependency relation in TLA verb paradigms**

{Non-perfective [3SGM $^{\text{IDENTICAL OR PARTIAL-SUPPLETION}}$ → 3PLM]} $^{\text{IDENTICAL OR PARTIAL-SUPPLETION}}$ → {Perfective [3SGF/PL$^{\text{IDENTICAL OR PARTIAL-SUPPLETION}}$ → 3SGM $^{\text{IDENTICAL OR PARTIAL-SUPPLETION}}$ →1.2SG/PL]}.

This generalization works for most of the verb measures/series. TLA has a dependency relation between the two stems used in the perfective sub-paradigm. The 3SGF/PL form is either identical with or unrelated to the 3SGM stem; and the 1/2SG/PL is either partially-suppletive or identical with the 3SGM stem. However, in M1 sound, hollow (with stem change pattern 1) and double verbs, all the stems in the perfective sub-paradigm are partially suppletive with the default stems in the imperfective sub-paradigm. By contrast, in the M1 and non-M1 defective verbs, the dependency between the stems tends to be based on identity relations (Table 5.9). These dependency relations define an inheritance hierarchy. For example, M1 double and hollow series have the dependency relations illustrated in (Figure 12 and Figure 13 respectively).

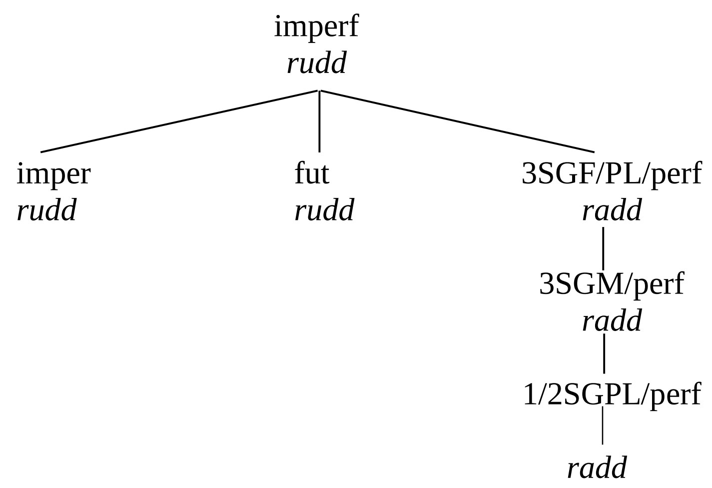


Figure 12 The stem dependency tree for the double verb *radd* ‘return’

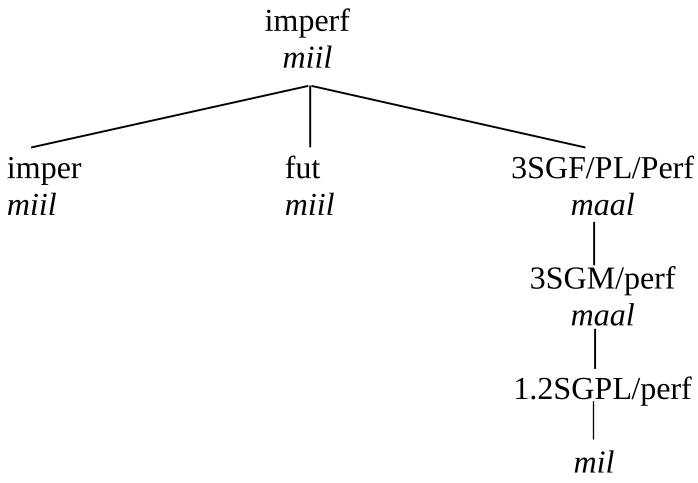


Figure 13 The stem dependency tree for the hollow verb *maal* ‘lean’

The TLA dependency account relies on the observation that there are relations of identity and/or ablaut change between the stems. The dependency relations within the perfective are as follows: in sound verbs, the 3SGF/PL/perf is a semi-suppletive form of the stem zone 3SGM which has the identical dependency relation with the stem in the 2SG/PL/perf slot. By contrast, in hollow verbs, the stem in the 3SGM/perf slot is identical to that in the 3FSG/PL/perf slot and suppletive with the 12SG/PL/perf stem zone. In other words, three distinct stems need to be stipulated for hollow verbs as shown in the inheritance tree (Figure 13).

These patterns of dependency relations have important implications for inflectional regularity and ICs in TLA. In French (Bonami & Boyé, 2006), a regular lexeme is generally associated with a single stem (e.g. *lav*) while semi-regular verbs show suppletion. Unlike French, in TLA uniformity of the stem choice and the ICs based on the same stem is not characteristic of the regular sound verbs, which stipulate distinct stems across the sub-paradigms, while some non-M1 verbs use only a single stem for all the inflectional forms. Therefore, regularity in this dialect may not be based on stem uniformity but rather it seems that the canonical stem has to vary in the verb paradigm of TLA. The fact that it is more regular to have multiple stems than to have just one stem is reminiscent of the situation in English: regular verbs have a present stem and a past (dental) stem: (*walk* ~ *walked*), but some irregular verbs have only one stem (*put* ~ *put*).

In summary a stem-based account for TLA verbs shows that inflectional complexity can be defined by stem sets provide by the stem space and connected by 'dependency relations' between nodes in the inheritance hierarchies to minimize the amount of redundant information to be stored for a given lexeme. In addition, the TLA stem space consists of stem slots selected by five paradigm slots based on unrelated indexed stems, thus, the stem space is constrained by dependency relations of varying complexity. Clearly, using the dependency relations modelled by the inheritance tree in filling the stem space of verbs avoids stem redundancy in the lexicon and predicts six ICs for TLA.

The following section will analyse the inflectional complexity and the ICs of TLA verbs under the principal parts framework as developed by Stump and Finkel (2013).

5.3 TLA principal parts

Stump and Finkel's (2013) work on morphological typology has formalized the notion of principal parts as a computational tool for measuring the complexity of inflectional systems (§2.4.2). In this section, we will introduce the four approaches for representing TLA principal parts system based on different types of plat representation in addition to the novel approach of using the stem-space chart as a plat. We will also discuss TLA IC transparency, using the PPA computer programme to measure IC complexity.

Recall from chapter 3 that in TLA, the inflectional exponents which express the morphosyntactic features in all verb measures including the irregular ones, do not show systematic/regular affix allomorphy. Instead, similar to French (Bonami & Boyé, 2002), TLA shows stem alternations distinguishing distinctive cells of the verb inflectional paradigm. The internal variation among the stems can be used as the main determinant of ICs determined by implicative relations within a lexeme's paradigm. However, the proposal that TLA affixes are highly redundant inflectional exponents is dependent on the segmentation analysis assumed and defended in this study (§4.1.3). However, if we had proposed a different decomposition for the inflectional forms, specifically by maximizing the affix, the analysis would make the outer layer of the inflected forms the source of inferences for the implicative relations in the paradigm and IC membership.

We can identify three possible analyses for TLA principal parts and IC complexity each of which is distinguished by the presence/absence of various simplifying 'manipulations' when annotating the plats. The first analysis will be based on hypothetical TLA plats based on affix allomorphy defined by different ways of segmenting the FV and the monophthong as part of the affixes (§5.3.1). Another variant of the TLA plat will involve 'multiple stem segments'/'multiple stem contexts'. In this case, the principal parts can be based on the indexed stems for which the plat input format represents stem formatives. For this type of plat, we determine the segmentation of the word form by specifying the 'theme' and 'distinguishers' (§5.3.3.2). The final possible analysis is based on what Stump and Finkel's proposal for Semitic languages, the plats are based on multiple stem segments consisting of three substems (similar to the root-and-pattern based account) (§5.3.3.1).

5.3.1 TLA Principal parts based on outer layer of verb form

In Chapter 4 we showed that the synthetic inflectional verb paradigm can be based on either the bare stem or the whole word (a stem and the exponents). TLA verbs inflect for the MPs of aspect, person and number, and an ordinary verbal paradigm can have twenty-four cells. We also showed that in most cases, the inflection of a verb lexeme is represented by the modification of the stem. The exponents might exhibit some degree of modification either alone or in combination with their stems. This raises segmentation issues. Therefore, for constructing the TLA plat it is essential to specify the stem modification and/or the affix decomposition/segmentation.

Consider, for example, the inflectional forms in the paradigm of *M*1 double series *radd* 'return', sound series *gebal* 'accept' and defective series *rmee* 'throw' in which each inflectional form consists of two main substrings: the theme and the distinguisher (Table 5.14 and Table 5.15).[16] The inflectional forms in Table 5.14 show that the substrings that do not show any phonological variation across all the cells serve as a theme (*r* for *radd, g* for *gebal, rm* for *rmee*)[17]. The distinguisher of each inflectional form is the (morpho)phonologically variable substring, resulting from the fusion of the inflectional affix with the stem and it has the role of distinguishing all the forms from each other (Table 5.15).

---

[16] See §2.4.2.3 and footnote [29] and [30] in Chapter 2.

[17] Note that identifying the theme is determined by the whole inflectional paradigm of each verb measure/series. Therefore, the theme of the double series *radd*, for example, is identified as *r* since the substring (distinguisher) *add* in the perfective paradigm alternates with *udd* in the imperfective paradigm (*r**add*** ~ *y-r**udd***). The results of the PPA tool also supports this account for the verb *gall* ~ *y-gill* 'take' (Table 5.41).

	1SG	1PL	2SGM	2SGF	2PLM	2PLF	3SGM	3SGF	3PLM	3PLF
perf	radd***eet***	radd***eena***	radd***eet***	radd***eeti***	radd***eetu***	radd***eetin***	radd	radd***it***	radd***u***	radd***in***
imperf	***na***g***bil***	***na***g***ublu***	***ta***g***bil***	***ta***g***ubli***	***ta***g***ublu***	***ta***g***ublin***	***ya***g***bil***	***ta***g***bil***	***yu***g***ublu***	***yu***g***ublin***
imp			***i***rm***i***	***i***rm***i***	***i***rm***u***	***i***rm***in***				

Table 5.14 The conjugation of *M*1 double series *radd* ‘return’ in the perfective, sound series *gebal* ‘accept’ in the imperfective and defective series *rmee* ‘throw’ in the imperative (distinguishers in bold italics).

<table>
<tr><td colspan="3">Enlarged stem</td><td rowspan="2">Affix</td></tr>
<tr><td colspan="2">stem</td><td rowspan="2">ee</td></tr>
<tr><td rowspan="2">r</td><td>add</td><td>t</td></tr>
<tr><td colspan="3">addeet</td></tr>
<tr><td>theme</td><td colspan="3">distinguisher</td></tr>
</table>

<table>
<tr><td rowspan="2">Affix</td><td colspan="3">Enlarged Stem</td><td rowspan="2">Affix</td></tr>
<tr><td rowspan="2">a</td><td colspan="2">stem</td></tr>
<tr><td>t</td><td rowspan="2">g</td><td>ubl</td><td>in</td></tr>
<tr><td colspan="2">ta</td><td colspan="2">ublin</td></tr>
<tr><td colspan="2">distinguisher</td><td>theme</td><td colspan="2">distinguisher</td></tr>
</table>

<table>
<tr><td colspan="3">Enlarged Stem</td><td rowspan="2">Affix</td></tr>
<tr><td rowspan="3">i</td><td colspan="2">stem</td></tr>
<tr><td rowspan="2">rm</td><td>i</td><td>u</td></tr>
<tr><td colspan="2">u</td></tr>
<tr><td>distinguisher</td><td>theme</td><td colspan="2">distinguisher</td></tr>
</table>

Table 5.15 The Theme-distinguisher analysis of *M*1 double series *raddeet* ‘return’ (1SG.perf), sound series *tagublin* ‘accept’ (2PLF.imperf) and defective series *irmu* ‘throw!’ (2PLM.imp).

According to Stump and Finkel (2013), in the simplest cases of form decomposition, the themes and the distinguishers can be identified as the stem and the inflectional affix respectively. However, this corresponding relation cannot be maintained across all forms as in some cases, a stem segment has to be classified as a distinguisher that can be informative of an IC membership. For instance, the theme of the inflectional forms in the *M*2 double series (e.g. 2PLF form *haddidtin* ‘you threatened’ with the theme *haddid* in the sub-paradigm of *haddid*) coincides with the stem, and the inflectional affixes are represented by the distinguisher (in this example, the suffix -*tin* is the distinguisher). By contrast, the distinguisher of the 2PLF form *raddeetin* is -*addeet* although the relevant exponent is *-tin*. The theme of the paradigm *radd* is *r* (cf. *radd* 'he returned' ~ *rudd* 'he returns') when combined with the distinguisher in the 2PLF.perf form, the result is the enlarged stem *raddee.* Similarly, the theme of the 3PLF.imperf. form *tagublin* (*M*1 sound series) is *g* which combines with the distinguisher *-ubl* to form the stem while the inflectional affixes are represented by the distinguisher *t-* and *-in*.[18]

In addition, the stem extensions that include the FV and the monophthong are treated as part of the enlarged stem. Nevertheless, as pointed out in §4.1.3, a possible segmentation analysis is to treat these extensions as part of the inflectional affixes. The consequence of this analysis for TLA system under the principal parts account would be twofold. First, it would be possible for IC membership to be identified by the terminations of the morphological word. Second, the principal parts from which we can derive inferences about this IC system would have to be regarded as a realized cell (a whole form, cf. Latin). Furthermore, assuming a plat analysis based on the exponents and the realized cells could also raise the representation issue of the concrete vs. abstract plats according to which the identification of verb *terminations* across all the different ICs has to depend on whether one assumes a phonemic or morphological representation. In fact, this type of analysis for TLA would treat the segmentation issues and the representation issues in parallel. Consider the following phonemic representation of the exponents of 2SG/PLM inflectional form across different measures and verb series in cases in which the FV and monophthong is segmented as part of the affix.

[18] In the inflectional form *kitibaata* 'she wrote it.3SGM' (cf. *kitib-it* 'she wrote') the segment *k* is the theme of the paradigm *ketab* 'he wrote' while the affixes are the substrings *it-a*. However, the distinguishers are *-aat-a* where the *i-* of the affix *-it* undergoes vowel lengthening and assimilation.

M	series		gloss	IC	Imperfective		Imperative		perfective	
					SG	PL	SG	PL	SG	PL
					2*M*					
*M*1	sound	getal	kill	1	tu-	tu-…-u	u-	u-…-u	-t	-tu
		ṣebaġ	dye	2	ta-	ta-…-u	a-	a-…-u	-t	-tu
		ketab	write	3	ti-	ti-…-u	i-	i-…-u	-t	-tu
	hollow	gaal	say	4	it-	it-…-u	Ø	Ø…-u	-t	-tu
	double	radd	return	5	it-	it-…-u	Ø	Ø…-u	-eet	-eetu
	defective	ns-ee	forget	6	ta-...-a	ta-…-u	a-…-a	a-…-u	-eet	-eetu
		mš-ee	walk	7	ti-...-i	ti-…-u	i-…-i	i-…-u	-eet	-eetu
Non-*M*1	defective	rabb-a	parent	8	it-…-i	it-…-u	-i	Ø…-u	-eet	-eetu
		stanj-a	purify	9	ti-…-a	ti-…-u	-a	Ø…-u	-eet	-eetu
	any[19]	kammil	finish	4	it-	it-…-u	Ø	Ø…-u	-t	-tu
*M*7	sound	ingetal	be killed	1	tu-	tu-…-u			-t	-tu
		insemaʕ	be heard	3	ti-	ti-....-u			-t	-tu

Table 5.16 The phonemic representation of the exponents of the 2SG/PLM inflectional form across different verb measures/series.

The examples in Table 5.16 show that the consequence of segmenting the FV and the monophthong as part of the affix and the concrete (full phonetic) representation is that we define a varying set of exponents that classify the verb system into seven ICs. For instance, in the perfective sub-paradigm, the termination of the 2SGM is realized differently across the different verb series/measures. The *M*1 sound verb form /ketab-t/ 'you wrote' is the 2SGM perfective verb form ending with /-t/ that is a termination realizing the property set {2SGM.Perf}. However, when we maximize affixes, the double and defective verbs would have the 2SGM forms such as *radd-eet* 'you returned', *ns-eet* 'you forgot' respectively, with the exponent /-eet/ rather than /-t/ in the perfective paradigm. The sound verb 2SGM imperfective forms forms *tu-ġsil* 'you wash', *ta-lʕib* 'you play', *ta-dbaḥ* 'you slaughter' and *ti-ktib* 'you write' could all be considred to have a prefix ta-, if we take the FV as part of the exponent realizing the property set {2SGM. impf.}. By contrast, the corresponding distinguisher in hollow, double and non-*M*1 verb forms with the concrete representation is realized as /it-/ (e.g.

[19] Except the defective series.

it-guul 'you say', *it-rudd* 'you return'). In other words, a phonemic representation of an affix-based segmentation means that the distinguisher realizing the property set {2SG.M. impf.} can vary, thus giving rise to different ICs.

Consequently, the concrete analysis (hearer-oriented plat) of the endings would predict different ICs, depending on the terminations (Table 5.16). The concrete plats assuming maximized affix segmentation have the following plat input formatting:

MPs			Perfect	Imperfective	Imperative
SG		1	-t	na~i~u-	
	M	2	-t	ta~i~u-	a~i~u-
		3	Ø	ya~i~u-	
	F	2	-ti	ta~i~u-...-i	a~i~u-...-i
		3	-it	ta~i~u-	
PL		1	-na	na~i~u-…-u	
	M	2	-tu	ta~i~u-…-u	a~i~u-...-u
		3	-u	ya~i~u-…-u	
	F	2	-tin	ta~i~u-…-in	a~i~u-...-in
		3	-in	ya~i~u-…-in	

Table 5.17 The concrete plat of the terminations in *M*1 and *M*7 sound series.

MPs			Perfect	Imperfective	Imperative
SG		1	-eet	in-	
	M	2	-eet	it-	Ø
		3	Ø	iy-	
	F	2	-eeti	it-…-i	-i
		3	-it	it-	
PL		1	-eena	in-…-u	
	M	2	-eetu	it-…-u	-u
		3	-u	iy-…-u	
	F	2	-eetin	it-…-in	-in
		3	-in	iy-…-in	

Table 5.18 The concrete plat of the terminations in *M*1 double series.

MPs			Perfect	Imperfective	Imperative
SG		1	-eet	na~i-…-a~i	
	M	2	-eet	ta~i-…-a~i	a~i-
		3	-ee	ya~i-…-a~i	
	F	2	-eeti	ta~i-….-i	a~i-...-i
		3	-it	ta~i-…-a~i	
PL		1	-eena	na~i- ….-u	
	M	2	-eetu	ta~i-….-u	a~i-...-u
		3	-u	ya~i-…-u	
	F	2	-eetin	ta~i-….-in	a~i-...-in
		3	-in	ya~i-….-in	

Table 5.19 The concrete plat of the terminations in *M*1 defective series.

MPs			Perfect	Imperfective	Imperative
SG		1	-t	in-	
	M	2	-t	it-	Ø
		3	Ø	iy-	
	F	2	-ti	it-…-i	-i
		3	-it	it-	
PL		1	-na	in-….-u	
	M	2	-tu	it-…-u	-u
		3	-u	iy-…-u	
	F	2	-tin	it-…-in	-in
		3	-in	iy-…-in	

Table 5.20 The concrete plat of the terminations in *M*1 hollow verbs and non-*M*1 verbs.

It is important to emphasize that this enriched representation of the affixes is the consequence of two simultaneous analyses: the segmentation based on maximizing the affix and the concrete phonetic analysis of affixation. Therefore, this representation of exponents enhances the points of contrast between verb lexemes, resulting in at least nine different ICs based on the outer layer of the verb. The termination of each IC has to correspond to a MP, but it varies depending on the ablaut of the FV (Table 5.17), the presence/absence and the alternation of the monophthong (Table 5.18 and Table 5.19), and epenthetic vowel (Table 5.20). The termination contrast in each IC is reflected by a different letter in the following concrete plat representations.

IC	ρ	σ	τ	υ
I	a	a	i	q
II	b	b	j	q
III	c	c	k	q
IV	d	d	l	q
V	d	d	l	r
VI	e	e	m	s
VII	f	f	n	s
VIII	g	g	o	t
IX	h	h	p	t

IC	Example	ρ	υ
I	getal	a	q
II	ṣebaġ	b	q
III	ketab	c	q
IV	gaal	d	q
V	radd	d	r
VI	ns	e	s
VII	mš	f	s
VIII	rabb	g	t
IX	stanj	h	t

Table 5.21 The plat of TLA ICs based on verb exponents (affix-based segmentation) (ρ ={imperf}, σ ={fut}, τ ={imper}, υ= {perf}).

In this matrix, the *ρ*-*υ* represents the different MPs while the I-VIII are the ICs. The exponents of MPs are represented by **a/t** in the intersection of the column and rows. In addition, the model requires us to avoid redundancy by excluding the MP *τ* ={fut} which is identical to the MP *ρ* = {imperf}. Likewise, the MP *τ* ={imper} is isomorphic to the MP *ρ* ={imperf}, since any inferences provided by the former MP can equally be obtained from the latter MP. Therefore, *ρ* is the distillation of both *τ* and *σ*.

In addition, the plat provides implicative relations among the ICs. Thus, if a verb lexeme L has the exponent **a** for MP *ρ* ={imperf} we can infer that the verb belongs to IC_I and the MP *ρ* will be the optimal principal part for that IC. By contrast, if a lexeme has the exponent **d** for MP *ρ* = {imperf}, then it is essential to check the MP *υ* ={perf} to uniquely identify each IC. The reason for this is that the MP *ρ* distinguishes three sets of ICs: {I to III} with the exponent **a**, **b** and **c** respectively (prefix with a FV) and {VI and VII} with the exponent **e** and **f** (prefix with the FV and the monophthong alternates with *a* and *i*). The final set is the ICs {VIII and IX} which are identified with the exponents **g** and **h** (monophthong alternates with *a* and *i*) respectively. However, both ICs {IV and V} have the exponent **d** (a prefix and an epenthetic vowel). Since the ICs IV and V can be not distinguished by the MP ρ alone, the MP υ has to be listed as the second principal part that can determine each IC.

Clearly, an affix-based segmentation causes the exponents to vary. Nonetheless, the MP set cannot always unequivocally identify each IC membership on its own. Although segmenting the FV and the monophthong as part of the affix, together with the concrete analysis of the exponents is one possible analysis, it can obscure the essential implicative relations based on the morphological patterns of the inflectional system for the reasons presented and defended discussed in §4.1.3. Therefore, it is important to factor out the morphophonological operations to reveal the real parallelism and points of contrast between the ICs. The abstract analysis of the inflectional endings, in addition to maximizing the stem, will result in unified affixes for all verb measures/series and treat the exponents as redundant phonological information. The concrete plats (Table 5.17, Table 5.18, Table 5.19 and Table 5.20) show various effects due to a number of (morphophonological) operations that have to be factored out under the abstract representation (11).

(11)

a. $\emptyset \rightarrow [i] / \left\{ \begin{array}{l} \_C.C_1VC_2C_2\# \\ \_C.C_1VVC_2\# \end{array} \right\}$ [20]

b. $[ee] \rightarrow \emptyset / \_ V$ (cf. maddeet-a ‘he passed it.3SG.M’ (monophthong is retained))

c. $\emptyset \rightarrow [ee] / C_1VC_2C_2 \_ C\#$ (this does not account for obj-suffixes with an initial C)

d. $V \rightarrow [ee] / C_1C_2\_C\#$[21] (this is not a synchronic process)

e. $\emptyset \rightarrow V / \left\{ \begin{array}{l} C\_.C_1VC_2VC_3\# \\ C\_.C_1C_2V\# \end{array} \right\}$

[20] This rule also includes the non-*M*1 verbs.
[21] This rule also applies to non-*M*1 defective series.

The abstract account and the SMP assume that all verbs have the same/syncretic exponents. Under this approach, the affix for realizing the property sets {1PL.Perf} is /-na/ and the termination /-ee/ should be considered as a morphologically-conditioned empty morph in the double series or a stem final vowel in the defective series, but in all cases it has to be segmented with the stem to achieve a simple statement of exponence. One would assume that the distinguisher joining these stems is /na/ which, by the rule in (c), has to be realized as /-eena/. However, it is not possible for these morphophonological rules to account for all instances of the monophthong, as they do not explain the absence of the monophthong before C-initial object suffixes. Similarly, factoring out the procedures in (e) in addition to stem maximization in *M*1 sound and defective verbs, and postulating underling representations |y-ugtil|, |y-ansa|, |y-guul| and |y-rudd| (cf. /yu-gtil/, /ya-nsa/, /iy-guul/ and /iy-rudd/ respectively), reveals that the property set {3SG.M. imperf.} is realized by [y] across all the ICs. Therefore, the speaker-oriented plat characterised by SMP and the abstract representation of the distinguishers in all measures can have the following realizations:

			Perf	Imperf	Imper
SG		1	-t	n-	
	M	2	-t	t-	-
		3	-	y-	
	F	2	-ti	t-...-i	-i
		3	-it	t-	
PL		1	-na	n-...u	
	M	2	-tu	t-….-u	-u
		3	-u	y-…-u	
	F	2	-tin	t-...-in	-in
		3	-in	y-...-in	

Table 5.22 The speaker-oriented plat of TLA verb terminations.

Clearly, basing the segmentation on the 'economy of exponence' in addition to the slight contrast between the concrete and abstract plats of the distinguishers confirms the redundancy of the exponents in the system, hence the principal parts analysis based on a set of realized cells (as is the case for Latin) can be dismissed. As was shown in §5.2 , the inflection system of TLA is based on multiple stems and a plausible principal-parts analysis has to determine the IC membership by plats of indexed stems. However, there are three types of plat representation that can be applied. Two of these are the product of the stem-based approach: a plat based on indexed stems vs a plat based on theme-stem formatives. The third type is a sub-stem (exponent) plat that in essence resembles a root and pattern based approach. The following section will present the different systems of plat representation and the implications for the principal parts analysis and ICs in TLA.

5.3.2 TLA principal parts based on indexed stems and ICs based on stem-space

In section 5.2 we identified five indexed stems which can be associated with a range of different MPs, but which do not necessarily form a natural class. Across the set of verb lexemes considered, the indexed stems from one to five cover the twenty-four MPs. However, the patterns of stem alternation and stem syncretism vary across conjugations, resulting in six ICs. In addition, these stems are associated with the same set of MPs in each IC, thus reducing the twenty four MPs to five distillations: that is ρ = {3SGM.imperf}, σ ={3PLM.imperf} τ ={1PL.perf}, υ ={3SGM.perf} and ϕ ={3SGF.perf} represented by the indexed stems as illustrated in Table 5.11.

It is possible to apply the three schemes of principal parts introduced in section 2.4.2 to the indexed stems without subjecting the stem to plat related manipulations or using the computer programme devised for principal parts analysis. In this case, the system of conjugations can be based on the patterns of stem change upon which principal parts can be employed to distinguish ICs. This analysis is in essence similar to Bonami and Boyé's (2002, 2006) approach to ICs, which is determined by the distribution of the stem allomorphy across the stem space. It is also possible to analyse the indexed stems from a different dimension, by following the traditional principal parts framework and investigating the complexity of ICs based on the number of dynamic principal parts required to distinguish each IC.

Under the static conception of the principal parts, the same set of stem slots with their corresponding pattern of stem change will have to determine the membership of every IC. Based on the stem allomorphy, the set of the stem slots S3 and S5 is the minimal and adequate subset of the stem space for uniquely determining each IC membership. In other words, the stem change pattern represented in S3 and S5 is the optimal principal-part set under the static scheme. For instance, IC_I is distinguished by stem pattern A~A across S3 and S5. By contrast, in any other combination set of stem slots, the pattern of stem change will not be able to satisfy the condition of adequacy and/or minimality. For example, the stems S1 and S2 do not form an adequate optimal set of principal parts since they fail to uniquely identify any IC. Likewise, the stem set S3, S4 and S5 does not serve as the optimal principal parts set because although it is adequate, it is not minimal: the proper subset {S3, S5} constitutes an adequate principal-parts set (Table 5.23).

IC	ρ	σ	τ	υ	φ
I	A	A	A	A	A
II	A	A	B	A	A
III	A	A	B	B	B
IV	A	B	C	C	B
V	A	A	B	C	C
VI	A	B	C	C	D
stem space	S1	S2	S3	S4	S5

Table 5.23 The static principal parts based on the TLA stem space.

Unlike the static scheme, the adaptive principal-parts set consists of the root property set alongside any other indexed stem which is able to define the optimal principal-parts set for a particular IC. In TLA, all verb lexemes can have the stem change pattern in S5 as their first principal part in the adaptive optimal sequence. If the verb has stem alternation D in S5, then the verb will belong to IC_{VI} and no other principal parts are needed. On the other hand, if the verb lexeme has the stem alteration A or B, then the stem slot S3 will serve as the second principal part. Thus, if the verb has the stem alternation A in S5 and stem A in S3, then the verb belongs to IC_{I} (Table 5.24).

IC	ρ	σ	τ	υ	φ
I	A	A	A	A	A
II	A	A	B	A	A
III	A	A	B	B	B
IV	A	B	C	C	B
V	A	A	B	C	C
VI	A	B	C	C	D
stem space	S1	S2	S3	S4	S5

Table 5.24 The adaptive principal parts based on the TLA stem space.

In the third scheme, the principal parts are dynamic, and they do not have to be the same for all lexemes or necessarily be shared between ICs. Therefore, the six conjugations in TLA can vary in terms of the number of slots and the corresponding stem patterns required for their minimal and adequate set of dynamic principal parts. For instance, a verb lexeme belonging to IC_I requires the stem change pattern of only one stem slot (S3) to deduce the verb's entire paradigm. However, for determining the IC_{II} membership, it is not sufficient to say that IC_{II} has the conjugation pattern B in stem slot S3, because of the ambiguity of this conjugation pattern which is also found in C_{III} and C_V, hence, the pattern of S4 is also selected. Therefore, IC_{III}, the adequate dynamic principal part set has to consist of two stem slots. By contrast, IC_{VI} requires only one dynamic principal part, S5 (Table 5.25).

IC	ρ	σ	τ	υ	φ
I	A	A	A	A	A
II	A	A	B	A	A
III	A	A	B	B	B
IV	A	B	C	C	B
V	A	A	B	C	C
VI	A	B	C	C	D
stem space	S1	S2	S3	S4	S5

Table 5.25 The dynamic principal parts based on the TLA stem space.

The principal parts schemes presented above were based on what Bonami and Boyé refer to as the conjugation pattern which is used to determine the IC in the system based on the stem alternation across each IC within the stem space.[22] Under each scheme, most IC principal parts sets employ the stem conjugation pattern of two stem slots as the minimal set. However, it is also possible to employ the conjugation pattern of one stem slot under the dynamic system.

Nevertheless, the conjugation pattern based plats seem to be obscuring the stem dependency relations and IC complexity in TLA. Although the six ICs show different

[22] This is basically similar to the exponent-based plats in which the different letters in the cell interaction is used as the base to formalise the principal parts system. In other words, stem conjugation pattern is used in the same way as an exponence change.

conjugation patterns across the stem slots of the stem space, they are not considerably different in terms of IC complexity based on the number of dynamic principal parts required. For instance, IC_I requires only one stem to fully predict the whole paradigm, in other words, IC_I is transparent in the sense that each stem slot determines that of every other stem slot. By contrast, the predictability of IC_{III} and IC_{IV} whole paradigms requires maximally two different stems. The IC_{VI} is less transparent employing five different stems. The implicative relations and IC predictability based on the number of required indexed stems in each IC reveals different sets of principal parts in the three schemes. Consider the following plats based on the indexed stems that have to be listed in order to deduce the stem pattern in the whole paradigm of each lexeme/IC.

IC	S1	S2	S3	S4	S5
I	1	1	1	1	1
II	1	1	2	1	1
III	1	1	2	2	2
IV	1	2	3	3	2
V	1	1	2	3	3
VI	1	2	3	3	4

Table 5.26 The static principal parts.

IC	S1	S2	S3	S4	S5
I	1	1	1	1	1
II	1	1	2	1	1
III	1	1	2	2	2
IV	1	2	3	3	2
V	1	1	2	3	3
VI	1	2	3	3	4

Table 5.27 The adaptive principal parts.

IC	S1	S2	S3	S4	S5
I	1	1	1	1	1
II	1	1	2	1	1
III	1	1	2	2	2
IV	1	2	1	1	2
V	1	1	2	3	3
VI	1	2	3	3	4

Table 5.28 The dynamic principal parts.

The minimal adequate subset from which it is possible to deduce the stems of the whole paradigm is based on the number of indexed stems required in each IC. This approach not only changes the members of the principal parts set in each scheme compared with the set of principal parts based on the stem alternation pattern, but can also alter the IC system transparency. For instance, in *M*1 sound verbs, the minimum number of stems that are required to deduce the stems of the rest of the paradigm are listed in S1, S2, S3 and S5. Given that the sound series of *M*1 shows the maximum number of indexed stems in the system, it is possible to use four stem slots to represent the static scheme of principal parts. Under this scheme, the same set of indexed stems will have to determine the membership of every IC. In other words, the set of indexed stems {S1, S2, S3, S5} is the minimal and adequate subset of the stem space for uniquely determining each IC membership. Therefore, each IC requires the listing of four principal parts. However, under the adaptive and dynamic systems, the plats of differentiate four principal parts comprised of different sets of stem slots. In fact, in both schemes, IC_{I} has only one member in the principal-parts set represented by the indexed stems of stem slot {S1} which is conventionally chosen to predict the whole paradigm. By contrast, other ICs require up to four members in the principal-parts set.

This section has introduced the principal parts analyses of TLA, which can be described along two dimensions. The first approach discussed ICs based on the pattern of stem change as distributed in the stem space. The other dimension considered IC variations based on the number of indexed stems needed by each IC to infer a lexeme's whole paradigm. The essential difference between the two types of plats is reflected in the implicative relations and the number of dynamic principal parts required by each IC. In the former, for instance, IC_{VI} can be determined by just one dynamic principal part while IC_{III} requires two; hence, IC_{VI} is more transparent than IC_{III}. In the indexed stem plat analysis, on the other hand, IC_{VI} is the most opaque class, with four principal parts. Different principal-parts analyses can also result from the effects of the computer programme due to the sensitivity to the plat-specific manipulations, as will be discussed in the following sections.

5.3.3 Two types of TLA Plats

This section will discuss the principal-parts analysis based on two types of input plats: the first is what I will refer to as the exponent-based plat (substems-based plat) (§5.3.3.1) and the second type is the stem-based plat (§5.3.3.2).

As a basis for the plats input data, I use all the possible verb measures and verb series that can be found as potential ICs, each named after an exemplar. Due to the lack of corpus and dictionary data, the representative items of the TLA system are chosen on the basis of the frequently exemplified verbs in different theses including (Owens, 1984; Harrama, 1993; Elgadi, 1986) and the language of daily use. Each chosen item is based on the verb measure and series phonological form as discussed in section 3.1 (Table 3.1 and Table 3.2) and section 4.3.

The list of the conjugations starts with the sound series followed by the defective series, then the double series and finally the hollow series as presented in Table 5.3. In each case, only the possible series-measure verbs are listed (see Table 3.2 in §3.1 for non-existing combinations).

The reason for considering measures as candidates for TLA ICs, is based on Aronoff's proposals for treating verb measures as inflectional conjugations rather than derivational ones. Likewise, the verb series in each possible corresponding measure is a potential representative of an IC, due to the varying stem allomorphy illustrated by different verbs (§4.3.1, §4.3.2, §4.3.3 and §4.3.4) which is crucial for the understanding of the stem-formative plats. However, different verb series and the stem allomorphy can also have an important role in the exponent-based plats in which the ablaut and imala based stem change is also included (a hyphen represents a consonant in Table 5.29 below). Listing the verb series as possible conjugations is also suggested by Owens in his classification of the ELA verb system (see §3.2.1). He considers the verb series in addition to the verb measures as conjugation classes identified by the stem syllable structures.

The input data for each plat type included each measure with their corresponding verb series. The reason for analysing all verb measures/series is mainly consistency and the aim for a neutral analysis of the system. In each framework used, we considered all measures in the system to provide a fairer comparison between the results of each approach. It is important to note that considering all the possible triliteral measures/series might seem to lead to potential redundancy among the conjugations, depending on the plat type. For example, the exponent-based plat lists *M*2 sound and double verbs although they have the same stem form (CVCCVC), thus, the same substems. Likewise, the same verbs are also listed in the stem-based plat although they share the same stem change (one indexed stem for the whole paradigm), but differ in terms of the verb series.

Nevertheless, this redundancy can be justified by the reasons Stump and Finkel (2013, p.186) provided for the plat representation of the French verb system. First, the PPA is designed to identify those lexemes that belong to the same IC. Therefore, the TLA plat (both in the exponent-based and the stem-based) includes verbs of different measures but which have the same stem-shape and stem alternation pattern in order to check if the PPA (regardless of the plat type) would classify these measures as one IC or different classes. For example, in §4.3.1 we showed that *M*2 sound series *kallim* 'talk' and *M*2 double series *haddid* 'threat' can be classified as lexemes of the same IC since they both have a single stem used across their whole paradigms. In fact, these two verbs are also listed in the exponent-based plat as belonging to the same IC (§5.3.3.1).

The TLA plats also include verb lexemes of the same series and the same measure, but different stem alternation patterns. For example, the verbs *maal* 'lean' and *baat* 'stay over' belong to the *M*1 hollow series, but they have different stem patterns: (AABCC) and (AABAA) respectively (Table 5.3 and Table 5.12). These verbs are included in the input plat to check if the PPA will reflect these differences and whether the programme will classify the verb *baat* as exceptional compared with other hollow verbs of the same series. However, the stem pattern of the verb *baat* can also be found in other measures of hollow series (*M*7 and *M*8).

It is also important to note that the TLA (exponent-based and stem-based) plats have the concrete representation. The different types of plat representation can affect the number of ICs and the possible verb lexemes of each IC (as was illustrated for French (§2.4.2.2)).[23] However, the reasons for choosing the hearer-oriented is based on the discussion presented in Chapter 4 which showed that it is not always possible to make decisions about morphophonologically motivated stems due to their semi-autonomous nature (§4.5.3).[24]

5.3.3.1 TLA Principal parts based on substems.

The principal-parts analysis of TLA can be subject to different types of data manipulation. One kind assumes a root-based approach and involves presenting the exponence of verb inflectional forms as discontinuous strings (components), encoding the vocalic pattern in an exponent-based plat. Under this approach, the stem consists

---

[23] The type of plat used for the analysis of principal parts of the Arabic of Bukhara has not been specified.
[24] I constructed the exponent-based plat which then was used as the input to generate the results. The stem-based plat was generated by the PPA tool. I am very grateful to Gregory Stump for confirming that both plats are constructed correctly and produce reasonable results.

of a set of substems, the consonantal root sequence, and a sequence of vowels. In plat representations the identity of the consonants is abstracted away from and they are represented as hyphens. For instance, the tri-consonantal verb *rukub-it* 'she rode' has three substems *rkb* and the vocalic exponence u-u; hence, the plat input of this form is -u-u-it.[25] Likewise, the principal parts analysis of the Arabic of Bukhara (Miller, 2014) and Icelandic (Stump & Finkel, 2013) is based on plats with multiple stem segments, which considers stems to be a subpart of the exponence. In this study, we follow the formalism suggested for the Arabic of Bukhara in encoding the inflectional properties of all the measures/series. However, the TLA plat template is based on the conjugation of the basic and derived verbs' inflectional systems whereas in Miller's (2014, p.224) work on the Arabic of Bukhara, the MP set is associated with the 'semantic, lexical and grammatical morphemes'. The analysis suggested for the Arabic of Bukhara seems to consider the derivative measures (such as M7) as part of the inflectional paradigm distributed across six ICs represented by the verb series. By contrast, in this study, we consider each verb measure and series as a separate lexeme and hence as potentially defining an IC; thus, each verb is listed with its conjugation pattern in the perfective and non-perfective templates. Consider the plat in Table 5.29 which represents *M*1 verb series ICs (the full plat is illustrated in Appendix B).[26]

[25] In this type of plat, the ablaut change and the monophthong *ee* are also regarded as subparts of the exponence. 'The exponence of a MPs in some word form is the full set of exponents of that property set in that word form.' (Stump & Finkel, 2013, p.21).

[26] The verbs *maal* 'lean', *kallim* 'talk', *saanid* 'support', *tesallif* 'borrow' and *tefaahim* 'understand', have the same stem patterns as that of the verbs *baaʕ* 'sell' (*M*1 hollow), *ʕallim* 'teach' (*M*2 sound), *ʕaamil* 'treat' (*M*3 sound), *teʕallim* 'learn' (*M*5 sound) and *teʕaamil* 'deal' (*M*6 sound) respectively.

CONJ	perf1Sg	perf2mSg	perf2fSg	perf3mSg	perf3fSg
getal	-e-a-t	-e-a-t	-e-a-ti	-e-a-	-u-u-it
nsee	-Ø-ee-t	-Ø-ee-t	-Ø-ee-ti	-Ø-ee-	-Ø-Ø-it
gall	-a-ee-t	-a-ee-t	-a-ee-ti	-a-Ø-	a-Ø-it
maal	-i-Ø-t	-i-Ø-t	-i-Ø-ti	-a:-	-a:-it
CONJ	perf1Pl	perf2mPl	perf2fPl	perf3mPl	perf3fPl
getal	-e-a-na	-e-a-tu	-e-a-tin	-u-u-u	-u-u-in
nsee	-Ø-ee-na	-Ø-ee-tu	-Ø-ee-tin	-Ø-Ø-u	-Ø-Ø-in
gall	-a-ee-na	-a-ee-tu	-a-ee-tin	-a-Ø-u	-a-Ø-in
maal	-i-Ø-na	-i-Ø-tu	-i-Ø-tin	-a:-u	-a:-in
CONJ	imperf1Sg	imperf2mSg	imperf2fSg	imperf3mSg	imperf3fSg
getal	nu-Ø-i-	tu-Ø-i-	tu-u-Ø-i	yu-Ø-i-	tu-Ø-i-
nsee	na-Ø-a-	ta-Ø-a-	ta-Ø-Ø-i	ya-Ø-a-	ta-Ø-a-
gall	n-i-Ø-	t-i-Ø-	t-i-Ø-i	y-i-Ø-	t-i-Ø-
maal	n-i:-	t-i:-	t-i:-i	y-i:-	t-i:-
CONJ	imperf1Pl	imperf2mPl	imperf2fPl	imperf3mPl	imperf3fPl
getal	nu-u-Ø-u	tu-u-Ø-u	tu-u-Ø-in	yu-u-Ø-u	yu-u-Ø-in
nsee	na-Ø-Ø-u	ta-Ø-Ø-u	ta-Ø-Ø-in	ya-Ø-Ø-u	ya-Ø-Ø-in
gall	n-i-Ø-u	t-i-Ø-u	t-i-Ø-in	y-i-Ø-u	y-i-Ø-in
maal	n-i:-u	t-i:-u	t-i:-in	y-i:-u	y-i:-in
CONJ	imper2mSg	imper2fSg			
getal	u-Ø-i-	u-u-Ø-i			
nsee	a-Ø-a-	a-Ø-Ø-i			
gall	-i-Ø-	-i-Ø-i			
maal	-i:-	-i:-i			
CONJ	imper2mPl	imper2fPl			
getal	u-u-Ø-u	u-u-Ø-in			
nsee	a-Ø-Ø-u	a-Ø-Ø-in			
gall	-i-Ø-u	-i-Ø-in			
maal	-i:-u	-i:-in			

Table 5.29 The TLA exponent-based plats.

Table 5.29 illustrates the input for the exponent-based plat for *M*1 verbs. The first column has to be labelled CONJ and lists the sound, defective, double and hollow verb series of *M*1. The rest of the columns represent the exponence of the perfective and non-perfective paradigm. In the intersection of the columns and the rows, the inflected forms are represented as discontinuous morphs. The 1SG and 3SGF perfective forms of sound verb *getal* are listed with three substems and their vocalic patterns -e-a-t and -u-u-it respectively. The suffixes are also represented and need to be separated from the stem with a hyphen as illustrated by the defective, double and hollow verb series in which each inflected form has two substems and varying vocalic pattern. The vocalism in these verbs is represented in line with that of the sound series as two components in every form. The absence of a vowel is represented by null Ø as illustrated by the 1SG and 3SGF of the defective verb *nsee* -Ø-ee-t and -Ø-Ø-it respectively. Likewise, the full plat lists every inflectional form of the perfective and non-perfective paradigm of different verb series and measure. The first column lists the (potential) ICs while the rest of the columns represent the set of MPs divided by multiple sections, as they do not all fit comfortably on one line. Each inflectional form is presented as a discontinuous string. In a plat with a single continuous string of exponents, the components are represented as 1A in the template's plats. The substems of biliteral root verbs has an empty component (in that it is a missing component that is present in other forms) represented by a null marker. It is worth noting that the exact input formatting of each form depends considerable on the ability of the programme to generate the correct forms from the plat representative forms. However, the essential representation includes discontinuous morphs separated by root radicals.

Computing this plat into the PPA tool identifies twenty-five ICs with two identical ones which include *kallim* 'talk' = *haddid* 'threat' and *tesallif* 'borrow' = *tesammim* 'be poisoned'. Associating these twenty-five ICs with MPs sets generates 24 MPs numbered 1 through 24 (Table 5.30). Of these property sets, there are 23 MPs which are unique ones, excluding the `perf2msg`[27] form which is identified as syncretic with the `perf1sg` form in all verbs. In addition, the 24 MPs fall into 10 distinct patterns that are crucial for deducing the whole paradigm of each verb. The ten distillations represent the remaining distinctive MP set (Table 5.31):

[27] For ease of cross referencing, in labelling MPs we use the formatting and typeface of the PPA when discussing its outputs, rather than the Leipzig glossing conventions.

1	perf1Sg	13	imperf2fSg
2	perf2mSg	14	imperf3mSg
3	perf2fSg	15	imperf3fSg
4	perf3mSg	16	imperf1Pl
5	perf3fSg	17	imperf2mPl
6	perf1Pl	18	imperf2fPl
7	perf2mPl	19	imperf3mPl
8	perf2fPl	20	imperf3fPl
9	perf3mPl	21	imper2mSg
10	perf3fPl	22	imper2fSg
11	imperf1Sg	23	imper2mPl
12	imperf2mSg	24	imper2fPl

Table 5.30 The MPs abbreviations.

distillation	
1	perf1Sg
2	perf3mSg
3	perf3fSg
4	imperf1Sg
5	imperf2fSg
6	imperf3mSg
7	imperf3mPl
8	imperf3fPl
9	imper2mSg
10	imper2fSg

Table 5.31 The TLA distillation of exponent-plat.

```
Distillation details ('*' marks isomorphic[28]; = marks identical)
      imper2fSg
            * imper2mPl
            * imper2fPl
      imper2mSg
      imperf1Sg
            * imperf2mSg
      imperf2fSg
            * imperf1Pl
            * imperf2mPl
            * imperf2fPl
      imperf3fPl
      imperf3mPl
      imperf3mSg
            * imperf3fSg
      perf1Sg
            = perf2mSg
            * perf2fSg
            * perf1Pl
            * perf2mPl
            * perf2fPl
      perf3fSg
            * perf3mPl
            * perf3fPl
      perf3mSg
```

[28] A MP can be isomorphic to another MP when any inferences or implicative relations from the former can equally be obtained from the latter. Therefore, to avoid redundancy, the two sets of MPs form a distillation.

The PPA computes the distillation of the inflectional system by analysing the conjugation patterns of each verb, determining the predictive and predictable patterns based on each MPs. The first MP (`perf1sg`) is by default regarded as the first distillation, since it is the first predictive MP set encountered by the program. Therefore, each new pattern of predictability is marked as a new distillation.

Understanding the distillation is essential for the three systems of principal parts analyses. In the exponent-based plat, selecting the request of computing all three possible schemes of principal parts provided in the menu of the online tool, the KATR output provides the results in (12). Our plat of TLA conjugations has nine of forty-five possible optimal static principal-parts analyses each of which has two principal parts. The distillations 2 {`perf3mSg`} and 4 {`imperf1Sg`} are selected as the best bit. Therefore, with these two distillations, PPA can recognize all the different 25 ICs.

(12)

Quick static principal parts
```
        2,4 (perf3mSg, imperf1Sg)
```

Best sets of static principal parts
```
        1,6 (perf1Sg, imperf3mSg)
        1,7 (perf1Sg, imperf3mPl)
        1,8 (perf1Sg, imperf3fPl)
        2,4 (perf3mSg, imperf1Sg)
        2,5 (perf3mSg, imperf2fSg)
        2,6 (perf3mSg, imperf3mSg)
        2,7 (perf3mSg, imperf3mPl)
        2,8 (perf3mSg, imperf3fPl)
        2,9 (perf3mSg, imper2mSg)
        2,10 (perf3mSg, imper2fSg)
       10 analyses with 2 principal parts
```

It is worth noting that the selection of the MPs in each static principal part analysis does not seem to be random and certain realized cells are selected. One principal part has to include a property set of the perfective that is either the 1SG or the 3SGM while the other principal part is most likely to be a property set that realizes the imperfective (any of the third person forms, but the other persons are also possible). However, there are two cases of principal parts realizing the {`imper.sg`}. In other words, the optimal static principal-parts analysis in TLA shows a general pattern in which certain cells are the best principal part ‘candidates’ in that they are clearly favoured as principal parts for determining the IC membership. Stump and Finkel distinguish two types of IC systems: ‘a morpho-syntactically focused’ IC such as Sanskrit (see Chapter 2, §4.2.4 for Sanskrit choice of distillations) that shows a general pattern in the static scheme. By contrast, other systems have morpho-syntactically unfocused ICs such as Tulu, in which the choice of the best candidate is quite random (2013, pp.72-74). According to Stump and Finkel (2013, p.385), ‘an IC system is morpho-syntactically focused to the extent that the distillations realized by its optimal static principal parts are constrained, an IC system that has only one optimal static principal-part set is maximally focused.’ By contrast, ‘an IC system is morpho-syntactically unfocused to the extent that the distillations realized by a lexeme’s optimal static principal parts are morpho-syntactically unconstrained, an IC system in which every distillation is realized by a principal part in one or another optimal static principal-part analysis is maximally unfocused.’

For TLA, the results in (12) show that the IC system in TLA tends to be constrained to coherence in the choice of the static principal parts; hence, TLA ICs exhibit morpho-syntactical focus on the third person forms in general and on the 3SGM perfective cell in particular both of which seem to be the locus of IC distinctions. According to Stump and Finkel (2013), a morpho-syntactically focused IC system does not necessarily reflect the system’s inflectional complexity. Two systems can have the same degree of morpho-syntactical focus, but different number of static principal-part sets and members. However, this characteristic of inflectional system is not necessarily central to IC transparency which will be further discussed below.

With regard to the adaptive principal parts, the results which are in the format of an if/then statement in (13), show that the TLA system requires two principal parts for determining IC membership under this scheme and these include distillation 6 (`imperf3mSg`) and distillation 1 (`perf1Sg`). If distillation 6 is assigned variant 1, then IC is *getal* 'kill'. If distillation 1 selects variant 10, then the IC is *naada* 'invite'.

(13)

```
We need 2 adaptive principal parts:

* if distillation 6 (imperf3mSg) has variant e14_17 (yi1St2Sa3SØ)
     the inflection class is irtadd
* if distillation 6 (imperf3mSg) has variant e14_4 (y1Ste2SaD3Si)
     the inflection class is tesallif
* if distillation 6 (imperf3mSg) has variant e14_23 (y1Sste2Si:3S)
     the inflection class is stefaad
* if distillation 6 (imperf3mSg) has variant e14_14 (y1Ssta2SØ3Sa)
     the inflection class is stanja
* if distillation 6 (imperf3mSg) has variant e14_11 (y1Ste2Sa:3Sa)
     the inflection class is telaaga
* if distillation 6 (imperf3mSg) has variant e14_15 (y1Si2SØ3S)
     the inflection class is gall
* if distillation 6 (imperf3mSg) has variant e14_22 (yi1St2Sa:3S)
     the inflection class is iltaam
* if distillation 6 (imperf3mSg) has variant e14_6 (yun1Su2Su3S)
     the inflection class is inkesar
* if distillation 6 (imperf3mSg) has variant e14_21 (yin1Sa:2S3S)
     the inflection class is injaab
* if distillation 6 (imperf3mSg) has variant e14_1 (yu1SØ2Si3S)
     the inflection class is getal
* if distillation 6 (imperf3mSg) has variant e14_13 (yi1St2Si3Sa)
     the inflection class is intisee
* if distillation 6 (imperf3mSg) has variant e14_16 (yin1Sa2SØ3S)
     the inflection class is ingall
* if distillation 6 (imperf3mSg) has variant e14_12 (yin1Si2Si3S)
     the inflection class is inkiwee
* if distillation 6 (imperf3mSg) has variant e14_18 (y1Sste2Si3SØ)
     the inflection class is stemarr
* if distillation 6 (imperf3mSg) has variant e14_10 (y1Ste2SaD3Sa)
     the inflection class is temanna
* if distillation 6 (imperf3mSg) has variant e14_20 (y1Sa:2S3S)
     the inflection class is baat
* if distillation 6 (imperf3mSg) has variant e14_2 (y1SaD2Si3S)
* * if distillation 1 (perf1Sg) has variant e1_2 (1SaD2Si3St)
       the inflection class is kallim
* * if distillation 1 (perf1Sg) has variant e1_9 (1SaD2See3St)
       the inflection class is bakka
* if distillation 6 (imperf3mSg) has variant e14_3 (y1Sa:2Si3S)
* * if distillation 1 (perf1Sg) has variant e1_3 (1Sa:2Si3St)
       the inflection class is saanid
* * if distillation 1 (perf1Sg) has variant e1_10 (1Sa:2See3St)
       the inflection class is naada
* if distillation 6 (imperf3mSg) has variant e14_19 (y1Si:2S3S)
     the inflection class is maal
* if distillation 6 (imperf3mSg) has variant e14_7 (yu1St2Su3Su)
```

```
     the inflection class is intefax
* if distillation 6 (imperf3mSg) has variant e14_8 (y1Ssta2SØ3Si)
     the inflection class is stanjid
* if distillation 6 (imperf3mSg) has variant e14_9 (ya1SØ2Sa3S)
     the inflection class is nsee
* if distillation 6 (imperf3mSg) has variant e14_5 (y1Ste2Sa:3Si)
     the inflection class is tefaahim
```

When we turn to the dynamic principal-part scheme, all verbs in TLA require a single member for the dynamic principal-part set to identify each conjugation, showing that the exponence of each IC is considerably different and IC predictive. All classes can have more than one possible option as a member of the dynamic principal-parts set. However, in each instance, the reference can only be made to a single distillation. For instance, in most ICs, the first possible option as a dynamic principal part is distillation 1 (`perf1Sg` cell). In the case of the sound, defective and double series of *M*7 and *M*8, the first possible option is distillation 2 (`perf3mSg`).[29] In *M*1 hollow verbs, on the other hand, the first choice of MP cell is distillation 4 (`imperf1Sg` cell) (Table 5.32 and Table 5.33).

[29] This principal-part choice may probably be motivated by the size of their template that in most cases can have only five inflectional forms.

IC	Gloss	dynamic Pp	Dist 1	Dist 2	Dist 3	Dist 4	Dist 5	Dist 6	Dist 7	Dist 8	Dist 9	Dist 10	cell predictor number
getal	kill	1	1	1	1	1	1	1	1	1	1	1	1.00
kallim	talk	1	1	1	1	1	1	1	1	1	1	1	1.00
saanid	support	1	1	1	1	1	1	1	1	1	1	1	1.00
tesallif	borrow	1	1	1	1	1	1	1	1	1	1	1	1.00
tefaahim	agree	1	1	1	1	1	1	1	1	1	1	1	1.00
inkesar	broken	2	2	2	2	2	2	2	2	2	2	2	1.00
intefax	bloated	2	2	2	2	2	2	2	2	2	2	2	1.00
stanjid	appeal	1	1	1	1	1	1	1	1	1	1	1	1.00
nsee	forget	1	1	1	1	1	1	1	1	1	1	1	1.00
bakka	upset	1	1	1	1	1	1	1	1	1	1	1	1.00
naada	call	1	1	1	1	1	1	1	1	1	1	1	1.00
temanna	wish	1	1	1	1	1	1	1	1	1	1	1	1.00
telaaga	meet	1	1	1	1	1	1	1	1	1	1	1	1.00
inkiwee	cauterized	2	2	2	2	2	2	2	2	2	2	2	1.00
intisee	forgotten	2	2	2	2	2	2	2	2	2	2	2	1.00
stanja	purify	1	1	1	1	1	1	1	1	1	1	1	1.00
gall	take	1	1	1	1	1	1	1	1	1	1	1	1.00
ingall	taken	2	2	2	2	2	2	2	2	2	2	2	1.00
irtadd	rejected	2	2	2	2	2	2	2	2	2	2	2	1.00
stemarr	continue	1	1	1	1	1	1	1	1	1	1	1	1.00
maal	lean	4	4	4	4	4	4	4	4	4	4	4	1.00
baat	stay	4	4	4	4	4	4	4	4	4	4	4	1.00
injaab	brought	1	1	1	1	1	1	1	1	1	1	1	1.00
iltaam	blamed	1	1	1	1	1	1	1	1	1	1	1	1.00
stefaad	benefit	1	1	1	1	1	1	1	1	1	1	1	1.00

Table 5.32 One of the possible optimal dynamic principal-part analyses for TLA verbs based on the exponent-plat.

Table 5.32 represents one of the possible analyses of the optimal dynamic principal-parts set for each IC that are listed on the vertical axis. The ten distillations (Dist) are represented in the horizontal column under which the rows lists a distillation number that represents the principal part required by each lexeme to deduce the cell in the lexeme's realized paradigm under this schematic analysis. The rightmost column is the cell predictor number (the average number of principal parts required to deduce a realized cell) (§2.4.2.4). The results show that each conjugation requires one optimal dynamic principal part to determine any cell realizing the ten distillations. Thus, each IC can be classified as a thin conjugation. In other words, the fact that a verb lexeme in each of these ICs, has a single dynamic principal part which is adequate for deducing the lexeme's paradigm, classifies TLA as a thin IC system (§2.4.2.4).

In fact, the conjugation system of TLA is as 'uniformly thin' as that of Tulu and Koasati, as in all of these systems the average cell predictor number and dynamic principal parts number equals 1.00 whereas the Icelandic system is considered to be thicker with a cell predictor number of 1.09 and dynamic principal part number of 1.56.

The results confirmed that the average cell predictor number and the dynamic principal number of TLA system is 1.00[30] (Stump & Finkel, 2013, pp.60-61). In addition, TLA ICs vary in terms of the number of the possible alternative dynamic principal-parts analyses, which affects the IC transparency based on the criterion C (§2.4.2.5) for a principal-parts based typological variation (Table 5.33).

Set	ICs	Dynamic Pp analysis	Ratio of p to the number of possible
A	getal 'kill', nsee 'forget', stanja 'purify', gall 'take', stefaad 'benefit', stemarr 'continue', stanjid 'appeal'	10	100.0% of 10 possible analyses
B	iltaam 'blamed', injaab 'brought',	8	80% of 10
C	maal 'lean', baat 'stay',	7	70% of 10
D	tefaahim 'understand', telaaga 'meet',	6	60% of 10
E	tesallif 'borrow', inkeṣar 'broken', temanna 'wish', intefax 'bloated', inkiwee cauterized', ingall 'taken', irtadd 'rejected', intisee, 'forgotten'	5	50% of 10
F	kallim 'talk', saanid 'support', bakka 'upset', naada 'invite'	2	20% of 10

Table 5.33 The number of optimal dynamic principal parts analyses for TLA conjugations.

The optimal dynamic principal-parts alternative analysis in Table 5.33 reveals that while the TLA ICs in set A are maximally transparent, the ICs in the rest of the sets deviate from maximal transparency to varying degrees. Consider, for example, the IC of the *M*1 double series *gall* 'take' which has the ten alternative principal-parts analyses represented in Table 5.34 in the format suggested by Stump and Finkel (2013). The

[30] Average number of principal parts: 1.00
Average lowest average of principal parts needed: 1.00

horizontal axis is headed by the ten distillations M (1-10) while each row shows an alternative possible dynamic principal part analysis listed as circled numbers. Therefore, for any verb that belongs to any IC of the set A, any cell of the verb paradigm allows the deduction of all the other cells. In other words, each component in each cell is unique across all ICs; hence, any cell can function as the verb's sole dynamic principal part. In other words, any IC in the set A might possibly be situated at the canonical extremes of IC transparency.

	distillations									
Alternative principal-part analyses	1	2	3	4	5	6	7	8	9	10
①	①	①	①	①	①	①	①	①	①	①
②	②	②	②	②	②	②	②	②	②	②
③	③	③	③	③	③	③	③	③	③	③
④	④	④	④	④	④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤	⑤
⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥	⑥
⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦	⑦
⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧	⑧
⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨	⑨
⑩	⑩	⑩	⑩	⑩	⑩	⑩	⑩	⑩	⑩	⑩

Table 5.34 The ten alternative optimal dynamic principal-part analyses for Conjugation *gall* 'take'.

The results also show that the ICs in sets B to F deviate from this canonical extreme of IC transparency to different extents. However, the nature and pattern of the deviation is the same, i.e. the deviation is based on varying the number of possible analyses for each IC, and none of them can have more than a single dynamic principal part for realizing a paradigm cell[31]. For instance, the conjugation of the *M*2 sound series *kallim* 'talk' has two alternative principal-parts analyses represented in Table 5.35, which illustrates that cell 1 `perf1Sg` or cell 2 `perf3mSg` has unique exponence, allowing the

[31] Unlike TLA, in Fur, some patterns of the deviation from IC transparency is based on not only the number of possible alternative analyses, but also on the absence of one informative optimal dynamic principal part. Therefore, in Conjugation IIIE, the sole adequate dynamic principal part requires simultaneous reference to two distillations (thick IC system).

deduction of the whole realized paradigm with no simultaneous reference to more than one principal part in each analysis.

Alternative principal-part analyses	distillations									
	1	2	3	4	5	6	7	8	9	10
①	①	①	①	①	①	①	①	①	①	①
②	②	②	②	②	②	②	②	②	②	②

Table 5.35 The two alternative optimal dynamic principal-part analyses for Conjugation of *kallim* ‘talk’.

The results in Table 5.35 show that none of the distillations 3 to 10 has any distinctive exponence that can distinguish this IC. Thus, the paradigms of the conjugations in set F have less predictive cells and each of these ICs is less transparent. Clearly, the complexity of TLA ICs can only be distinguished on the basis of criterion C of IC transparency according to which the ICs in set A are more transparent since they have more alternative possible analyses than that of the rest of the sets. Criteria (a) and (b) do not seem to provide a clear distinctive transparency pattern among the classes since all TLA ICs require a single dynamic principal part (§2.4.2.4).

Therefore, Stump and Finkel (2013) suggested a measure of **IC predictability** and **Cell predictability** both of which can provide clearer characterization of IC transparency. Applying these measures to all ICs in TLA gives the results in Table 5.36:

	1	2	3	4	5	6	7	8	9	10	Avg	IC predictability	
getal	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	1.000	(385 out of 385)
stanjid	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	1.000	(385 out of 385)
nsee	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	1.000	(385 out of 385)
stanja	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	1.000	(385 out of 385)
gall	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	1.000	(385 out of 385)
stemarr	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	1.000	(385 out of 385)
stefaad	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	1.000	(385 out of 385)
injaab	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.996	0.996	0.987	0.992	(382 out of 385)
iltaam	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.996	0.996	0.987	0.992	(382 out of 385)
maal	0.996	0.996	0.996	0.969	0.969	0.969	0.969	0.969	0.969	0.969	0.977	0.982	(378 out of 385)
baat	0.996	0.996	0.996	0.969	0.969	0.969	0.969	0.969	0.969	0.969	0.977	0.982	(378 out of 385)
tefaahim	0.938	0.938	0.996	0.938	0.996	0.938	0.996	0.938	0.938	0.996	0.961	0.961	(370 out of 385)
telaaga	0.938	0.938	0.996	0.938	0.996	0.938	0.996	0.938	0.938	0.996	0.961	0.961	(370 out of 385)
tesallif	0.879	0.879	0.996	0.879	0.996	0.879	0.996	0.996	0.879	0.996	0.938	0.922	(355 out of 385)
temanna	0.879	0.879	0.996	0.879	0.996	0.879	0.996	0.996	0.879	0.996	0.938	0.922	(355 out of 385)
inkesar	0.984	0.879	0.879	0.984	0.984	0.879	0.879	0.879	0.996	0.996	0.934	0.922	(355 out of 385)
intefax	0.984	0.879	0.879	0.984	0.984	0.879	0.879	0.879	0.996	0.996	0.934	0.922	(355 out of 385)
inkiwee	0.984	0.879	0.879	0.984	0.984	0.879	0.879	0.879	0.996	0.996	0.934	0.922	(355 out of 385)
intisee	0.984	0.879	0.879	0.984	0.984	0.879	0.879	0.879	0.996	0.996	0.934	0.922	(355 out of 385)
ingall	0.984	0.879	0.879	0.984	0.984	0.879	0.879	0.879	0.996	0.996	0.934	0.922	(355 out of 385)
kallim	0.363	0.363	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.870	0.579	(223 out of 385)
saanid	0.363	0.363	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.870	0.579	(223 out of 385)
bakka	0.363	0.363	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.870	0.579	(223 out of 385)
naada	0.363	0.363	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.870	0.579	(223 out of 385)
Avg	0.877	0.852	0.967	0.976	0.990	0.951	0.965	0.960	0.980	0.994	0.951	0.903	

Table 5.36 IC predictability and Cell predictability.

Table 5.36. presents the average cell predictably of each IC listed in decreasing order based on the results/figures scored. The two rightmost columns are the average cell predictability and IC predictability[32]. The results show that the implicative relations among cells in a realized paradigm of an IC lexeme member can vary (according to the type of plat employed). The cell predictability overall results reveal that Dist 10 cells have high predictability in all ICs. In other words, the MPs of `imper2fSg` is the most predictable feature-value set given the rest of the distillations (Table 5.29). In fact, the inferences about the exponence in the `imper2fSg` cell are motivated by Dist 1, 2 and 4 (Table 5.32). However, the cell predictability of this cell in *M*1 hollow verbs *maal* and *baat* is slightly lower than the rest of the conjugations with a score of 0.969. Dist 10 shares the same substems with that of Dist 4 (`imperf1Sg`), Dist 5 (`imperf2fSg`),

[32] The figures expressed as (A out of B) represent the ICs predictability which is the 'ratio of adequate (though not necessarily optimal) dynamic principal-part sets among all nonempty subsets of cells in L's realized paradigm' (Stump & Finkel, 2013, p.94). Hence 385 ÷ 385 = 1 (see Appendix B for the formula of the IC predictability measure).

Dist 6 (imperf3mSg), Dist 7 (imperf3mPl), Dist 8 (imperf3fPl) and Dist 9 (imper2mSg) as illustrated in Table 5.37 (see §4.3.4 for hollow series stem pattern). Thus, the predictability of the imperf1Sg cell on the basis of all of the possible sets of other cells which can completely determine the value of the realisation of the cell in question within that hollow verb inflection classes *maal* or *baat* is 0.969 predictable and Dist 4 is chosen as the dynamic optimal principal-part (Table 5.32).

	Dist 4	Dist 5	Dist 6	Dist 7	Dist 8	Dist 9	Dist 10
CONJ	**imperf1Sg**	**imperf2mSg**	**imperf3mSg**	**imperf3mPl**	**imperf3fPl**	**imperat2mSg**	**imperat2fSg**
maal	n-i:-	t-i:-	y-i:-	y-i:-u	y-i:-in	-i:-	-i:-i
baat	n-a:-	t-a:-	y-a:-	y-a:-u	y-a:-in	-a:-	-a:-i

Table 5.37 The TLA exponent-based plats for conjugations *maal* 'lean' and *baat* 'stay'

Similarity, in conjugations *tefaahim* and *telaaga* both of which are *M*6 verbs, the predictability of the realization of the cells in Dist 9, Dist 8 Dist 6, Dist 4 and Dist2 and Dist 1 is slightly low with a score of 0.938 since the vocalic exponence in the cells of are identical (the exponence of these cells in conjugation *tefaahim* is te-a:-i- and in *telaaga* te-a:-a-). In addition, the cell predictability results also show that none of the TLA conjugations has unpredictable cells. However, the degree of cell predictability can vary across ICs. Consider for example, the conjugation *naada* 'call' (or any conjugation in set F) which has a low cell predictability in distillation 1 (1 perf1Sg) and 2 (perf3mSg) (each cell scores 0.363). The perf1Sg and perf3mSg cells have the monophthong *ee* and the vowel final *a* as part of their vocalic exponence occurring stem finally. These two unique cells can provide inferences and predict the rest of the cells which, in turn, cannot absolutely determine the realization of the cells in Dist 1 or Dist 2. Hence, the optimal dynamic principal part for this conjugation can be either Dist 1 or Dist 2 (Table 5.32) (Appendix B). By contrast, the cells in question in conjugation *nsee* 'forget' have the highest cell predictability (each cell scores 0.996) since the realization of a given cell is highly predictability from any of the other cells that realize distinct distillations.

CONJ	**Dist 1**	**Dist 2**	**Dist 3**	**Dist 4**	**Dist 5**
	perf1Sg	**perf3mSg**	**perf3fSg**	**imperf1Sg**	**imperf2fSg**
naada	-a:-ee-t	-a:-a	-a:-Ø-it	n-a:-i-	t-a:-Ø-i
nsee	-Ø-ee-t	-Ø-ee-	-Ø-Ø-it	na-Ø-a-	ta-Ø-a-
	Dist 6	**Dist 7**	**Dist 8**	**Dist 9**	**Dist 10**
	imperf3mSg	**imperf3mPl**	**imperf3fPl**	**imper2mSg**	**imper2fSg**
naada	y-a:-i-	y-a:-Ø-u	y-a:-Ø-in	-a:-i-	-a:-Ø-i
nsee	ya-Ø-a-	ya-Ø-Ø-u	ya-Ø-Ø-in	a-Ø-a-	a-Ø-Ø-i

Table 5.38 The TLA exponent-based plats for conjugation *naada* 'call' and *nsee* 'forget'

The second observation is that the more alternative principal-parts analyses an IC has, the higher the predictability number is. Indeed, the TLA conjugation that associates with a high number of alternative dynamic principal-parts analyses is reflected in the higher IC predictability. For example, the conjugation *gall* 'take' (and any conjugation in set A (Table 5.33)) has ten alternative principal-parts analyses and it is maximally transparent; hence, the IC predictability is 1.00 whereas conjugation *kallim* 'talk' (and any conjugation in set F) with only two possible alternative analyses, has the predictability number of 0.579.

Overall, the ICs are divided into two main groups by cell and IC predictability. The first group includes all the conjugations in set A to E (Table 5.33) in which cells realizing any distillation have high cell predictability ranging from 0.879 to 0.996. By contrast, the second group is represented by the conjugation in set F in which distillation 1 and 2 score the lowest cell predictability in the system. Therefore, these distillations have to be selected as principal parts in the optimal dynamic scheme. In addition, these results further confirm the deviation of set F conjugations from the maximal transparency compared with the rest of the ICs. Nevertheless, none of the cells scored zero cell predictability, which indicates that the realization of each cell in each IC is necessarily predictable.

In summary, these two types of predictability measures can provide further information and a clearer picture of the IC transparency in TLA, confirming the overall results provided by computing the dynamic principal-parts analysis. The next section will discuss the IC system of TLA from a stem based prospective, following Stump and Finkel's (2013) approach to the conjugation system of French.

5.3.3.2 TLA Principal parts based on Themes and stem formative plats

The principal parts analysis of TLA can also be affected by another kind of data manipulation that assumes a stem-based approach and involves presenting only the stems of the inflectional forms of each IC while the affixes are irrelevant for the conjugation distinction. Under this approach, each verb paradigm is based on the stem forms which are divided into a theme accompanied with a stem formative (§2.4.2.3), either of which can be null (§5.3.1). For instance, the *M*1 double verb *bakka* 'upset' has the stems: *bakki* (3SGM.imperf), *bakk* (1PL.imperf), *bakkee* (1PL.perf) and *bakka* (3SGM.perf) for which the stem formatives would be /-i/, null, /-ee/ and /-a/ respectively whereas the lexeme's theme is *bakk*. Therefore, for the principal-parts analysis, the PPA input format of a chart of a paradigm has to represent each conjugation and be labelled as IC in the first column. The second column lists the gloss while the rest represents the indexed stem. Provided with the chart of all ICs, the PPA software can automatically generate the output plat listing the stem formatives. It can also provide the declarations of the lexeme by listing a name for that lexeme, theme and IC to which that lexeme belongs.

In §2.4.1 we showed that the stem space distinguishes TLA ICs, each of which has a set of stems with indices determining the stem distribution in the verb's realized paradigm. Likewise, the principal-parts analysis of French is based on plats with multiple stems. However, the number of the indexed stems depends on whether one assumes an abstract or a concrete analysis (§2.4.2.3). For TLA, we follow the formalism suggested for French in encoding the stem allomorphy in all the measures/series. However, the stems in the plat employed for TLA are morphologically stipulated based on a surface representation coded as concrete phonology. In other words, we will only consider the TLA concrete stem-formative plat, representing each verb measure and series as a separate lexeme and a potential IC for the reasons presented in section 5.3.3.

To construct the TLA plat for analyses of the principal parts, we provide the PPA with the chart in Table 5.39. We will also follow Stump and Finkel's analysis of French in employing a PPA generated plat in which the vertical axis lists indexed stems whose formatives are determined by the intersection of the relevant indexed stem with its IC membership (Table 5.40). The programme also provides the lexeme declarations and specifies the theme of each listed conjugation (Table 5.41). In Stump and Finkel analysis of French, the stem formatives join with the themes to produce the indexed stems inventory of each verb and they vary depending on the representation assumed, affecting the number of the indexed stems accordingly. However, for TLA verbs, most of the substrings of the stem form, which represent morphologically induced suppletion, seem to constitute the stem formatives. As a result, it seems that the majority of verb measures in TLA will belong to a unique IC.

CONJ	Gloss	Stem1	stem2	stem3	stem4	stem5
getal	kill	ugtil	ugutl	getal	getal	gutul
kallim	talk	kallim	kallm	kallim	kallim	kallm
saanid	support	saanid	saand	saanid	saanid	saand
tesallif	borrow	tesallif	tesallf	tesallif	tesallif	tesallf
tefaahim	understand	tefaahim	tefaahm	tefaahim	tefaahim	tefaahm
insemaʕ	heard	insimaʕ	insimiʕ	insemaʕ	insemaʕ	insimiʕ
irtefaʕ	carried	irtifaʕ	irtifiʕ	irtefaʕ	irtefaʕ	irtifiʕ
stanjid	appeal	stanjid	stanjd	stanjid	stanjid	stanjd
nsee	forget	ansa	ans	nsee	nsee	ns
bakka	upset	bakki	bakk	bakkee	bakka	bakk
naada	invite	naadi	naad	naadee	naada	naad
temanna	wish	temanna	temann	temannee	temanna	temann
telaaga	meet	telaaga	telaag	telaagee	telaaga	telaag
inkiwee	cauterized	inkiwi	inkiw	inkiwee	inkiwee	inkiw
intsee	forgotten	intisa	intis	intisee	intisee	intis
stanja	purify	stanja	stanj	stanjee	stanja	stanj
gall	take	gill	gill	gallee	gall	gall
haddid	threat	haddid	haddid	haddid	haddid	haddid
tesammim	poisoned	tesammim	tesamim	tesammim	tesammim	tesamim
ingall	taken	ingall	ingall	ingallee	ingall	ingall
irtadd	rejected	irtadd	irtadd	irtaddee	irtadd	irtadd
stemarr	continue	stemirr	stemirr	stemarree	stemarr	stemarr
maal	lean	miil	miil	mil	maal	maal
baat	stay	baat	baat	bit	baat	baat
injaab	brought	injaab	injaab	injib	injaab	injaab
iltaam	blamed	iltaam	iltaam	iltum	iltaam	iltaam
stefaad	benefit	stefiid	stefiid	stefad	stefaad	stefaad

Table 5.39 The PPA input chart for generating stem formative-plat.

IC	Gloss	stem1	stem2	stem3	stem4	stem5
getal	kill	u-til	u-utl	-etal	-etal	-utul
kallim	talk	-im	-m	-im	-im	-m
saanid	support	-id	-d	-id	-id	-d
tesallif	borrow	-if	-f	-if	-if	-f
insemaʕ	heard	-imaʕ	-imiʕ	-emaʕ	-emaʕ	-imiʕ
irtefaʕ	carried	-ifaʕ	-ifiʕ	-efaʕ	-efaʕ	-ifiʕ
nsee	forget	a-a	a-	-ee	-ee	-
bakka	upset	-i	-	-ee	-a	-
temmanna	wish	-a	-	-ee	-a	-
inkiwee	cauterized	-i	-	-ee	-ee	-
intisee	forgotten	-a	-	-ee	-ee	-
gall	take	-ill	-ill	-allee	-all	-all
haddid	threat	-	-	-	-	-
tesammim	poisoned	-mim	-im	-mim	-mim	-im
ingall	taken	-	-	-ee	-	-
stemarr	continue	-irr	-irr	-arree	-arr	-arr
maal	lean	-iil	-iil	-il	-aal	-aal
baat	stay	-aat	-aat	-it	-aat	-aat
injaab	brought	-aab	-aab	-ib	-aab	-aab
iltaam	blamed	-aam	-aam	-um	-aam	-aam
stefaad	benefit	-iid	-iid	-ad	-aad	-aad

Table 5.40 TLA output plat based on Stem formatives

LEXEME	kill	kill	1:g
LEXEME	talk	talk	1:kall
LEXEME	support	support	1:saan
LEXEME	borrow	borrow	1:tesall
LEXEME	understand	talk	1:tefaah
LEXEME	heard	heard	1:ins
LEXEME	carried	carried	1:irt
LEXEME	appeal	support	1:stanj
LEXEME	forget	forget	1:ns
LEXEME	upset	upset	1:bakk
LEXEME	invite	upset	1:naad
LEXEME	wish	wish	1:temann
LEXEME	meet	wish	1:telaag
LEXEME	cauterized	cauterized	1:inkiw
LEXEME	forgotten	forgotten	1:intis
LEXEME	purify	wish	1:stanj
LEXEME	take	take	1:g
LEXEME	threat	threat	1:haddid
LEXEME	poisoned	poisoned	1:tesam
LEXEME	taken	taken	1:ingall
LEXEME	rejected	taken	1:irtadd
LEXEME	continue	continue	1:stem
LEXEME	lean	lean	1:m
LEXEME	stay	stay	1:b
LEXEME	brought	brought	1:inj

LEXEME blamed	blamed	1:ilt
LEXEME benefit	benefit	1:stef

Table 5.41 The theme of TLA conjugations (the highlighted lexemes are ICs with identical conjugation pattern. The rightmost column represents the theme).

Table 5.39 represents the input plat based on the stem formative for twenty-one verbs out of twenty-seven different measures/series. I added the first column with the list of the verb lexemes whereas the plat as designed by the programme had the gloss of the ICs as the initial column. The rest of the columns represent the set of indexed stems while the stems in each IC is represented by the stem formative of the relevant indexed stem at the intersection of cells. Associating these twenty-seven verb lexemes (Table 5.39) with the list of stem formatives generates twenty-one inflection classes and five distillations that are needed to determine the IC membership. In other words, each indexed stem in TLA is required for the IC predictability.[33]

(14)

```
5 distillations
                stem1 stem2 stem3 stem4 stem5
```

The PPA computes the distillation of the inflectional system by analysing the conjugation patterns of each verb, determining the predictive and predictable patterns based on each indexed stem. The first predictive indexed stem encountered by the computer is stem1; hence, by default, it is regarded as a distillation. In fact, every other indexed stem is considered a new pattern of predictability and marked as a new distillation. When computing the static scheme of principal parts based on the plat of stem formatives, the output provided is shown in (15). The results reveal that the plat of TLA conjugation has only one possible optimal static principal-parts analysis that

[33] It is important to note that the number of the distillations can be affected by the type of plat representation assumed. However, the principal-parts analysis does not necessarily require the listing of all indexed stems as a distillation in other inflectional system. For example, in French, the principal-parts analysis selects 17 distillations of the 19 indexed stems since the indexed stems 6 and 9 are included in other distillations.

includes three essential indexed stems as principal parts including stem1 (that of the 3SGM.imperf), stem3 (that of the 1PL.perf) and stem4 (that of the 3SGM.perf).

(15)

Best sets of static principal parts

```
        1,3,4 (stem1, stem3, stem4)
     1 analyses with 3 principal parts
```

The reason for this limitation on other possible alternative best-set analyses of static principal parts might be the conditions of adequacy and minimality on any set of principal parts to be employed in the static scheme. Nevertheless, the resulting static scheme set of TLA tends to be based on the condition of distinguishing different ICs rather than the adequacy condition of determining all forms in any verbal paradigm. For instance, Dist stem5 for the *M*1 sound series is not listed in the static set of principal parts though it is necessary for deducing all the forms in that lexeme's verb paradigm. In fact, the choice of optimal set is particularly traceable to the seven conjugations listed in Table 5.42.[34] These ICs have similar sets of stem formatives, and it is necessary to employ three indexed stems of the five stem distillations in order to distinguish each of these conjugations. Consider the stem formatives in the following table:

IC	Gloss	stem1	stem2	stem3	stem4	stem5
nsee	forget	a-a	a-	-ee	-ee	-
bakka	upset	-i	-	-ee	-a	-
inkiwee	cauterized	-i	-	-ee	-ee	-
temmanna	wish	-a	-	-ee	-a	-
intisee	forgotten	-a	-	-ee	-ee	-
haddid	threat	-	-	-	-	-
ingall	taken	-	-	-ee	-	-

Table 5.42 The TLA conjugations with similar stem formatives

[34] This observation does not exclude any of the other listed ICs. However, unlike other ICs, the conjugations under consideration share the exact same stem formatives, the monophthong *ee* and the vowel final vowel (which undergoes ablaut change), in certain distillations. Nevertheless, they belong to different ICs. Therefore, the selected set as the optimal static principal part will have to be ADEQUATE and MINIMAL in distinguishing the similar conjugation patterns of these ICs.

The table above shows that Dist stem1 distinguishes neither conjugation *bakka* 'upset' and *inkiwee* 'cauterized' nor conjugations *temanna* 'wish' and *intsee* 'forgotten'. Likewise, Dist stem 2 can only determine IC *nsee* 'forget'[35] whereas Dist stem5 cannot differentiate any of the ICs in the list in question (i.e. *bakka*, *inkiwee*, *temanna* 'wish' and *intsee*). Dist stem3 can only determine the conjugation *haddid* 'threat' while Dist stem4 groups the listed conjugations (excluding *gall* and *stemarr*) into three patterns (with stem formatives `-ee`, `-a` and `–`); hence, fails to determine each IC individually. Therefore, the static principal-part set has to include {stem1, stem3, stem4} which is adequate as it can, for example, distinguish IC *nsee* by {stem1}, IC *bakka* by {stem1 and stem4}, and IC *ingall* by {stem3}. In other words, this set is minimal since the set {stem1 and stem4}, for instance, cannot determine IC *haddid* 'threat' and *ingall* 'taken'.

Turning now to the adaptive principal parts, the results show that the TLA system requires two principal parts for determining IC under this scheme, the root principal part Dist 1 (3SG.M.imperf /stem1), along with Dist 4 (3SGM.perf/ stem4) or Dist 3 (1PL.perf). If distillation 1 is assigned variant 1, then the IC is *getal* 'kill'. If distillation 4 selects variant 7, then the IC is *intsee* 'forgotten'. By contrast, the verb lexeme *ingall* 'taken' is identified as an IC by the adaptive principal part of stem 3 with the variant 7.

(16)

We need 2 adaptive principal parts:

```
We need 2 adaptive principal parts:
* if distillation 1 (stem1) has variant e1_14 (1Siil)
```

[35] Conjugation *nsee* shares the same stem formatives (monophthong *ee*) with conjugations *bakka, temanna, inkiwee, instee* and *ingall* in distillation stem3, and conjugation *inkiwee* and *intsee* in distillation stem 4. However, the stem1 and stem2 formatives are unique/distinctive in conjugation *nsee* and they can uniquely distinguish this IC, but this set cannot be adequate since stem2 does not distinguish the rest of the conjugations in this table.

```
    the inflection class is lean
* if distillation 1 (stem1) has variant e1_1 (u1Stil)
    the inflection class is kill
* if distillation 1 (stem1) has variant e1_2 (1Sim)
    the inflection class is talk
* if distillation 1 (stem1) has variant e1_13 (1Sirr)
    the inflection class is continue
* if distillation 1 (stem1) has variant e1_15 (1Saat)
    the inflection class is stay
* if distillation 1 (stem1) has variant e1_9 (1Sa)
* * if distillation 4 (stem4) has variant e4_7 (1See)
      the inflection class is forgotten
* * if distillation 4 (stem4) has variant e4_8 (1Sa)
      the inflection class is wish
* if distillation 1 (stem1) has variant e1_17 (1Saam)
    the inflection class is blamed
* if distillation 1 (stem1) has variant e1_11 (1S)
* * if distillation 3 (stem3) has variant e3_7 (1See)
      the inflection class is taken
* * if distillation 3 (stem3) has variant e3_9 (1S)
      the inflection class is threat
* if distillation 1 (stem1) has variant e1_7 (a1Sa)
    the inflection class is forget
* if distillation 1 (stem1) has variant e1_18 (1Siid)
    the inflection class is benefit
* if distillation 1 (stem1) has variant e1_16 (1Saab)
    the inflection class is brought
* if distillation 1 (stem1) has variant e1_4 (1Sif)
    the inflection class is borrow
* if distillation 1 (stem1) has variant e1_3 (1Sid)
    the inflection class is support
* if distillation 1 (stem1) has variant e1_10 (1Sill)
    the inflection class is take
* if distillation 1 (stem1) has variant e1_6 (1Sifaʕ)
    the inflection class is carried
* if distillation 1 (stem1) has variant e1_8 (1Si)
* * if distillation 4 (stem4) has variant e4_7 (1See)
      the inflection class is cauterized
* * if distillation 4 (stem4) has variant e4_8 (1Sa)
      the inflection class is upset
* if distillation 1 (stem1) has variant e1_12 (1Smim)
    the inflection class is poisoned
* if distillation 1 (stem1) has variant e1_5 (1Simaʕ)
    the inflection class is heard
```

In terms of the dynamic principal-parts analysis in the stem-based plats, the average number of dynamic principal parts to a given realized cell in this system is 1.2 and in most conjugations, the chosen dynamic principal part is any of the five distillations under this scheme, as represented in Table 5.43. The results reveal that all verbs in TLA are divided into two main groups: one which requires a single distillation in a dynamic principal-parts set to identify each conjugation and another which employs two distillations. For some classes, there can be more than one possible option as a member of the dynamic principal-parts set whereas others can have only one possible analysis (Appendix B). Consider the following table in which the first column lists the twenty-one conjugation classes while the third column represents an optimal dynamic set.

CONJ	Gloss	Optimal Dynamic Principal-part set	Stem1 Dist1	stem2 Dist2	stem3 Dist3	stem4 Dist4	stem5 Dist5	Average number
getal	kill	{1}	1	1	1	1	1	1.00
kallim	talk	{1}	1	1	1	1	1	1.00
saanid	support	{1}	1	1	1	1	1	1.00
tesallif	borrow	{1}	1	1	1	1	1	1.00
insemaʕ	heard	{1}	1	1	1	1	1	1.00
irtefaʕ	carried	{1}	1	1	1	1	1	1.00
stemarr	continue	{1}	1	1	1	1	1	1.00
nsee	forget	{1}	1	1	1	1	1	1.00
bakka	upset	{1,4}	1	1	1	4	1	1.00
temanna	wish	{1,4}	1	1	1	4	1	1.00
inkiwee	cauterized	{1,4}	1	1	1	4	1	1.00
intisee	forgotten	{1,4}	1	1	1	4	1	1.00
gall	take	{1}	1	1	1	1	1	1.00
haddid	threat	{3}	3	3	3	3	3	1.00
tesammim	poisoned	{1}	1	1	1	1	1	1.00
ingall	taken	{1,3}	1	1	3	1	1	1.00
maal	lean	{1}	1	1	1	1	1	1.00
baat	stay	{1}	1	1	1	1	1	1.00
injaab	brought	{1}	1	1	1	1	1	1.00
iltaam	blamed	{1}	1	1	1	1	1	1.00
stefaad	benefit	{1}	1	1	1	1	1	1.00

Table 5.43 One of the possible optimal dynamic principal-part analyses for TLA verbs in the stem-based plats.

Table 5.43 represents one of the possible analyses of an optimal dynamic principal-parts set for each IC that is listed on the vertical axis. The five distillations (Dist) are represented in the first row under which the rest of the rows list a distillation number. This distillation number represents the principal part required for each lexeme to deduce the cell in the lexeme's realized paradigm under this schematic analysis. The right most column is the cell predictor number. The results illustrate that most ICs require reference to a single principal part in each set. However, there are five conjugations that employ two distillations as principal parts. For example, the principal part of the conjugation *getal* ‘kill’ is the indexed stem1 whereas conjugation *bakka* ‘upset’ is listed with the set of schematic principal-part {1,4}. In terms of thin-thick system classification, despite the difference in the number of principal parts required in each of these conjugations, they are both considered as thin ICs, since determining any given cell of these two conjugations only need a single principal part.

In terms of IC transparency, the dynamic principal-parts results show that fourteen of twenty-two ICs are transparent ICs while the rest lie between the two extremes (Appendix B). For instance, in Conjugation *getal* ‘kill’ any indexed stem allows the deduction of the realized paradigm (Table 5.44). In other words, each of the five distillations is unique to this class and each cell is predictive of this IC. By contrast, for verbs that belong to conjugation *nsee* ‘forget’, either the stem1 formative or stem2 formative (/a-a/ and /a-/ respectively) can result in the deduction of the realized paradigm. On the other hand, the formatives of stem 3, stem 4 and stem 5 (/-ee/, /-ee/ and /-/ respectively) are not distinctive of this IC. Therefore, they cannot be employed as principal parts (Table 5.40 and Table 5.43). However, the paradigm of any lexeme that belongs to this IC shows some degree of transparency that would have been higher with more predicative stem formatives that can unambiguously identify the IC membership.

Alternative principal-part analyses	distillations				
	1	2	3	4	5
①	①	①	①	①	①
②	②	②	②	②	②
③	③	③	③	③	③
④	④	④	④	④	④
⑤	⑤	⑤	⑤	⑤	⑤

Table 5.44 The dynamic principal-part analysis for conjugation *getal* ‘kill’.

Alternative principal-part analyses	distillations				
	1	2	3	4	5
①	①	①	①	①	①
②	②	②	②	②	②

Table 5.45 The dynamic principal-part analysis for conjugation *nsee* ‘forget’.

Nevertheless, the rest of the TLA ICs exhibit different patterns of transparency deviation, mainly because the stem formatives of these ICs tend to be ‘uninformative’ and do not fully satisfy the conditions of optimal principal parts (Stump & Finkel, 2013, p.88). For instance, in the IC *ingall* ‘taken’ the stem formative of distillation 1 can identify distillation 2, 4, and 5, but it is uninformative for distillation 3. In other words, distillation 1 has limited capacity to fully function as an optimal principal part on its own. Therefore, this IC requires reference to two stem formatives, that of distillation 1 and 3, which have to be employed as the optimal principal parts for any verb that belongs to this conjugation (Table 5.44). A similar pattern of deviation is exhibited by IC *bakka* ‘upset’ in which the uninformativeness of the stem formative results in the stipulation of two principal parts (Table 5.45). In addition, the transparency of TLA ICs can also be weakened when the uninformativeness leads to the postulation of one and only one principal part for the realization of the lexeme’s paradigm. For example, the principal parts for conjugation *haddid* ‘threat’ has to be distillation 3 as illustrated in (Table 5.46).

Alternative principal-part analyses	distillations				
	1	2	3	4	5
①,③	①	①	③	①	①
③,④	④	④	③	④	③

Table 5.46 The dynamic principal part analysis for conjugation *ingall* ‘taken’.

Alternative principal-part analyses	distillations				
	1	2	3	4	5
①,④	①	①	①	④	①

Table 5.47 The dynamic principal-part analysis for conjugation *bakka* ‘upset’.

Alternative principal-part analyses	distillations				
	1	2	3	4	5
③	③	③	③	③	③

Table 5.48 The dynamic principal-part analysis for conjugation *haddid* ‘threat’.

Clearly, the transparency of TLA ICs can vary with different deviation patterns according to the four criteria of IC transparency presented by Stump and Finkel (see §2.4.2.4). Overall, the majority of the twenty-one TLA conjugations require only one dynamic principal part, but in a few cases, two dynamic principal parts are employed to determine the IC (Table 5.47).

TLA conjugation classes		Criteria			
		A	B	C	D
a	getal, kallim, saanid, tesallif, tefaahim, insemaʕ, irtefaʕ, maal, baat, injaab, iltaam, stefaad, gall, tesammim	1	1	5	100
b	nsee	1	1	2	40
c	bakka, temanna, inkiwee, intisee	2	1	1	10
d	ingall	2	1	2	20
e	haddid	1	1	1	20

Table 5.49 The results of TLA ICs according to the transparency criteria.

The table represents the results of computing the dynamic principal-part analysis that can be explained through the transparency criteria. Criterion A: the number of dynamic principal parts. Criterion B: the dynamic principal-parts number required to predict a cell in a paradigm. Criterion C: the number of the optimal dynamic analyses and criterion D: the ratio of the actual analyses to possible analyses (%). According to criterion A, the ICs in a, b and e are more transparent that those in c and d. Under criterion B, all ICs are equally transparent since determining any given cell of any given

IC requires reference to no more than a single principal part. Although identifying the IC of *bakka* ‘upset’ requires two dynamic principal parts, determining any cell of this class does not require simultaneous reference to both principal parts, only single distillation is required. Criterion C, on the other hand, classifies the conjugations in (a) as more transparent while conjugations in (c or e) have less transparency. The conjugation in (a) affords five alternative analyses whereas the conjugations in (e) can only have one possible analysis. Although conjugation sets in both (b and d) have two alternative principal-parts analyses, according to criterion D, the conjugation in (b) is more transparent with the ratio of actual analysis to possible analysis of 40% and the latter scores 20%. In addition, the various degrees of IC transparency exhibited by TLA conjugation can be further reflected in a more precise measure that is IC predictability that provides a measurable contrast of the differences that exist among these conjugations (Table 5.50).

	1	2	3	4	5	Avg	IC predictability [36]	
kill	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
talk	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
support	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
borrow	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
continue	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
heard	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
carried	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
take	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
lean	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
stay	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
brought	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
blamed	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
benefit	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
poisoned	0.938	0.938	0.938	0.938	0.938	0.938	1.000	(30 out of 30)
forget	0.500	0.500	0.875	0.750	0.938	0.713	0.767	(23 out of 30)
threat	0.750	0.875	0.000	0.750	0.938	0.662	0.500	(15 out of 30)
taken	0.500	0.750	0.000	0.500	0.938	0.537	0.367	(11 out of 30)
upset	0.000	0.750	0.750	0.000	0.938	0.487	0.233	(7 out of 30)
wish	0.000	0.750	0.750	0.000	0.938	0.487	0.233	(7 out of 30)
cauterised	0.000	0.500	0.750	0.000	0.938	0.438	0.233	(7 out of 30)
forgotten	0.000	0.500	0.750	0.000	0.938	0.438	0.233	(7 out of 30)
Avg	0.709	0.846	0.810	0.721	0.938	0.805	0.789	

Table 5.50 TLA IC predictability based on stem formative plat.

[36] See footnote 33.

The table above lists the average cell predictability of each IC in decreasing order based on the figures. The first column represents the twenty-one conjugations while the two rightmost columns are the average cell predictability and IC predictability. The rest of the columns show the five distillations along with the cell predictability in each conjugation. The results further confirm that the transparency criteria observation in Table 5.47. For instance, the high transparency of conjugation *getal* 'kill' (with one dynamic principal part) compared with conjugation *bakka* 'upset' (with two dynamic principal parts) as identified by criterion A is reflected in IC predictability of 1.00 for the former conjugation and of 0.233 for the latter. Likewise, Criterion B does not identify variation in IC transparency since any cell of any realized paradigm in any ICs can be deduced by reference to a single principal part. The results also reveal that TLA conjugations that are associated with a high number of alternative principal-parts analyses is reflected in a higher IC predictability. For example, the conjugation *kallim* 'talk' has five alternative dynamic principal-parts analyses and it is transparent; hence, the predictability is 1.00 whereas conjugation *haddid* 'threat' with only a single analysis has the predictability number of 0.500. In other words, the overall pattern shows that the more the alternative principal-parts analyses allowed, the higher the IC predictability number would be.

To provide a more precise content of the criterion D, Stump and Finkel (2013) employ the second measure of IC transparency, namely, cell predictability. This provides a measurable contrast between ICs based on the four criteria (§2.4.2.5), calculating the realization of each cell of any IC paradigm. The measures in Table 5.50 above illustrate that the cell predictability varies depending on the distillation and/or IC. Overall, the realization of distillation 5 cells shows high predictability across all ICs with an average of 0.938. The results of cell predictability also show that six of the

21 ICs have unpredictable cells scoring zero cell predictability. In fact, there are cells with zero cell predictability. Two of the six conjugations (*haddid* and *ingall*) have one unpredictable cell while the rest have two (*bakka*, *temanna*, *inkiwee* and *intisee*). In addition, the cell predictability in conjugation *nsee* 'forget' the cells associated with distillation 1 and 2 score nearly half less (0.500) than those above the dotted line whereas the cells associated with distillation 3, 4, 5 have a higher cell predictability.

This inflection system under the stem based plat has no more than two distillations whose realizations are totally unpredictable as revealed by the four conjugations *bakka*, *temanna*, *inkiwee* and *intisee* with two unpredictable cells and conjugations *haddid* and *ingall* with one. Each of these conjugations share the same stem formative(s) in one or more distillations and given the similarities in stem formatives, in some distillations it is not possible to predict the realisation of a feature-value set (Table 5.51).

IC	Gloss	stem1	stem2	stem3	stem4	stem5
haddid	threat	-	-	-	-	-
ingall	taken	-	-	-ee	-	-
bakka	upset	-i	-	-ee	-a	-
temmanna	wish	-a	-	-ee	-a	-
inkiwee	cauterized	-i	-	-ee	-ee	-
intisee	forgotten	-a	-	-ee	-ee	-

Table 5.51 The stems formatives in conjugations *haddid* 'threat', *ingall* 'taken', *bakka* 'upset', *temanna* 'wish', *inkiwee* 'cauterized' and *intisee* 'forgotten'.

Conjugation *haddid* is the conjugation in the system revealed with empty cells as it shares the same stem in all distillations including Dist 3 which in all the other conjugations is distinguished by a certain stem-formative. Both conjugations *haddid* and *ingall* exhibit the same plat input except for the third distillation. Likewise, the stems in Dist 1 scores zero since for each of the conjugation pair *bakka* and *inkiwee*, and the conjugation pair *temanna* and *intsee* exhibit the same stem formatives.

Likewise, the realization of Dist 4 is zero in these conjugations. In this case, conjugation *bakka* ‘upset’ is grouped with conjugation *temanna* ‘wish’ while conjugation *inkiwee* ‘cauterized’ is grouped with conjugation *intsee* ‘forgotten’ for sharing the same stem formative in the distillation in question (see the discussion on TLA conjugations with similar stem formatives in (Table5.42)).

5.4 Summary

This chapter has presented the verb IC system in TLA based on two approaches that combine insights gained from work on implicative relations with the stem space and principal parts. Section 5.2 presented a stem-based account, showing that inflectional complexity can be defined by stem sets modelled by the stem space and connected by ‘dependency relations’ listed by intermediate nodes of the inheritance hierarchy to minimize the amount of redundant information to be stored for a given lexeme. I proposed the TLA ‘stem space’ which consists of stem slots selected by five paradigm slots based on [+morphological] stems filling the stem space which is constrained by dependency relations of a varying complexity. Clearly, using the dependency relations modelled by the inheritance tree in filling the TLA stem space avoids stem redundancy in the lexicon and predicts six ICs for TLA. Section 5.3 was devoted to TLA principal-parts analysis based on different approaches to the inflectional system. One hypothetical option for the principal-parts analysis was based on assuming a segmentation that prioritizes maximizing the affix and differentiates ICs by the outer layer of the inflectional form. However, the principal-parts scheme can also be stem-based. An important aspect of this framework is that morphologically indexed stems play the role of principal parts and these are used as a measure for the inflectional system's degree of complexity and IC transparency. The theme-stem formative analysis for TLA seems considerably influenced by the plat sensitivity to data manipulations.

Therefore, following Stump and Finkel, we suggested that rather than dealing with themes in the analysis of the verb inflection, we employ plats based on indexed stems associated with a morphological index that determines the distribution of these stems (§5.3.2). Similar to the indexed stems of the stem space, the stem-based plat (without the data manipulations) showed that some verbs require one member in the dynamic principal-parts set while others can have up to four stems in same scheme. We also showed that the verb lexeme of each TLA IC has a single dynamic principal-parts set that is adequate for deducing the lexeme's paradigm. In addition, the average cell predictor number and dynamic principal-parts number in TLA equals 1.00. These results classify the TLA IC system as 'uniformly thin', just as that of Tulu and Koasati.

Chapter 6 Conclusion

In traditional generative descriptions, it has been assumed that variations in the shape of the lexical items in Semitic languages is governed by purely phonological rules and that Arabic verb morphology is autosegmental and morpheme-based. Therefore, the inflectional morphology of Arabic verbs was handled by dividing the inflectional forms into morphemes that have a specific meaning or morphosyntactic property. However, this model remains problematic when applied to arbitrary ablaut change within inflectional sub-paradigms (Chapter 2).

In this work, we have shown that TLA makes extensive use of verb stem variation that cannot always be reduced to pure phonological factors; hence, the morphomic patterning revealed in TLA conjugations poses serious empirical problems for any morpheme-based account. The morphomic approach to the TLA inflectional system combines insights gained from work on implicative relations with the stem space and principal parts both of which represent the morphological complexity of TLA verb paradigms through economy of description of paradigmatic inferences. However, each has different approaches to implicative structure (Chapter 4 and 5). The primary aim has been to analyse the paradigmatic behaviour in the verb inflectional morphology and provide an analytical framework for presenting the verb complexity. In contrast to a purely phonological account (autosegmental morphology), the verb inflectional paradigm in TLA was found to be subject to morphomic conditioning that can be sensitive to other components of the system.

This chapter will summarize the key findings of the main research questions of this study presented in each chapter. It will highlight the original contribution to the field of Arabic linguistics and morphology through using a stem-based approach. Finally, we will consider the limitations and some suggestions for future work.

6.1 Summary

Chapter 2 presented a general review of the related literature and discussed the issue of the status of Arabic morphology as morpheme-based compare with a stem-based account. These two distinct models result in different predictions. The morpheme-based root and pattern approach faces a number of conceptual hurdles and remains problematic. Recent work on Semitic languages has shown that using a stem approach is superior and essential for the study of lexeme formation. In addition, this work has also showed that the morphological alternations within the verb inflectional paradigm are best captured by an account which acknowledges the central role of stems in reflecting implicative relations and the organization of the system's ICs.

This study also showed that the choice between these two different frameworks is crucial for modelling other theoretical domains of paradigmatic implicative relations based on principal parts and for implementing the computational linguistics approach using the PPA tool. Another important issue, addressed in Chapter 2, is the stem morphomicity and its sensitivity to non-morphological components of the system ('extramorphological factors'), which have generated the debate over the existence of clear-cut boundaries for the stem. Therefore, in this study, we proposed handling TLA stem morphomicity using [+/-] referencing features.

Chapter 3 presented a general overview of the verb system. The relevant phonological, morphosyntactic features were introduced along with some verb paradigmatic patterns. In addition, previous analyses of verb morphology in two varieties of LA (ELA and JLA) were considered. Owens' (1984) root-based approach to the verb morphology was found to share some of the essentials of a word-based account, mainly categorizing the *M*1 verb series into different classes based on the syllable structure of certain selected forms in the perfective and imperfective paradigm. Nevertheless, the basis of his account seems to be the eclectic use of root and pattern for classifying simple verbs into two main types which are further categorized and sub-divided on different grounds. These include the number of consonants, the position of the vowel and the type of the paradigm (defectiveness). Likewise, Harrama (1993) provides a phonological approach using rules and processes to account for the morphological complexity in JLA inflectional paradigm. Although the proposed rules can handle stem alternation in a set of verbs, it leaves other sets of verbs as exception cases. Overall, none of these studies proposed the useful method of stem indexing for the economy of description and for capturing the essential paradigmatic implicative relations.

Chapter 4 provided a synchronic analysis based on the stem-account without the theoretical hurdles of the underlying form. Addressing the segmentation issue in TLA, a segmentation procedure was developed based on the Strictly Morphomic Stem hypothesis (Spencer, 2012) and taking into account the systematic syncretic forms across TLA verb paradigms. It was found that the complex stems of *M*1 sound verbs adhere to these two principles; hence, the FV is regarded as part of the stem.

This chapter also introduced the possible stem change patterns that can be encountered in the inflectional paradigm of the verb measures and verb series. The nature of the stem change pattern was shown to be morphological, but it is also dependent on extramorphological factors. It thus falls within semi-autonomous morphology. In some verbs, TLA shows stem alternations distinguishing distinctive cells of the perfect verb paradigms, (3PL/3SGF from 1/2 and the 3SGM forms), which has a morphomic distribution, requiring the listing of the individual stems (Stem 3 and Stem 5). By contrast, in verbs such as those of *M*2, the stem alternation is simply an instance of allomorphy (allostems), hence, describable using a single indexed stem (Stem 1). The morphomic status of the stem was modelled by using a representative method used for reflecting the morphological conditioning and the presence or absence of stem sensitivity to the morphosyntactic and/or (morpho)phonological components along with the morphological ones.

Chapter 5 provided a stem-based account, showing that inflectional complexity can be defined by stem sets provided by the stem space and connected by ‘dependency relations’ through intermediate nodes in inheritance hierarchies to minimize the amount of redundant information to be stored for a given lexeme. We proposed a stem space for TLA that consists of stem slots selected by five paradigm slots based on indexed stems. For instance, it was shown that the number of required stem slots can be established using *M*1, *M*2 and *M*7 sound, *M*1 double series and two verbs of *M*1 hollow series. The second part of this chapter was devoted to TLA principal-parts analysis based on different types of plats with varying input formatting. We considered a hypothetical option for modelling TLA principal parts based on the outer layer of the inflectional form through maximizing the affix by treating the FV and the monophthong as part of the exponence. Such a model can differentiate nine ICs. However, for economy of exponence, SPP is prioritized and the same affix allomorphy is shared by all verb measure/series. Therefore, the principal-parts analysis based on outer layer of the inflection word is not applicable.

Following Stump and Finkel’s (2013) typological classification of ICs and IC transparency, we provided two principal-parts analyses of the TLA inflection system, based on different representations of the plat input formatting. The first type is based on the exponence while the second type of plat is based on the stem formatives. Given the different input formatting, the different plats generated different results; however, they both seem to show sensitivity to data manipulations. Therefore, rather than only dealing with themes and/or substems in the analysis of the verb inflection, we presented plats based on indexed stems associated with a morphological index which determines the distribution of these stems. the stem-based plat (without the data manipulations) showed that some verbs require one member in the dynamic principal-part set while others can have up to four stems in same scheme. Despite the considerable differences in the plat representations/formatting, the overall results generally seem to classify TLA as a 'thin' IC system.

6.2 The contribution to knowledge

The work has established the system of stem allomorphy in triliteral TLA verbs in inflectional paradigms of each of the possible verb measures and/or series. This has highlighted a clearer distinctive feature of the system, distinguishing it from other LA systems and many Arabic dialects. The stem account provided some crucial information about the implicative relations and related networks as describe which have never been considered in the literature of LA systems. The stem inventory and stem morphomicity relies on the segmentation procedure and TLA segmentation algorithm

The status of semi-autonomous morphology was addressed by a novel descriptive device, that of referencing features for mapping the stem morphomicity boundaries, between autonomous morphology and other domains of the system. The data confirmed that it is not possible for stems in TLA to be only the result of synchronic phonological conditioning, but rather they are morphomic, albeit showing sensitivity in some cases to extramorphological factors. The significance of this work is abandoning the urge to seek a phonological account for the TLA verb morphology and showing that morphemehood can be rooted in the data only to a certain degree and sometimes not at all.

This present work also developed a stem-space account based on the morphologically conditioned stem change rather than the irregularity element induced/reflected by suppletion in some Romance languages. The algorithm for generating the TLA stem space first required following the segmentation principles as suggested by word-based accounts. Second, we identified stem alternations that induce stem indexing, using the inventory of stem change in six verbs as representatives of the required stem slots. Capturing the dependency relations among these stems showed a simple vs complex dependency relation in which the imperfective forms can be used as the default forms overridden by the perfective in some cases.

The typological classification of the TLA inflectional system based on principal parts provides an important contribution to Arabic linguistics in particular and to morphological typology in general. Even with the exponent-based plats, it was found that TLA is not radically different from other, non-Semitic, inflectional systems. Assessing the results against the inflectional class transparency criteria shows that TLA inflection can be characterized as ‘uniformly thin’, just as that of Tulu and Koasati.

In this work, I have demonstrated that the IC transparency and complexity within a single system can vary depending on the type of the plat employed, showing considerable sensitivity to the formatting of the input. I have evaluated four different plat representations for identifying the network of implicative relations in TLA. It was found that the plat based on the concrete and abstract phonological representation of verb outer-layer cannot always unequivocally identify each IC membership on its own, obscuring the essential implicative relations within the stem alternation patterns.

The IC transparency of the TLA system is considerably affected by the data manipulations presented in the exponent-based plat which assumes a root and pattern based system, resulting in a high number of ICs based on the vocalism of each verb provided in the input. Likewise, the results based on the stem-formative plat classifies TLA as a more complex system with a lower average IC predictability.

Given the two approaches to implicative relations, i.e. the stem-space and the principal parts tool, I proposed a novel approach to the TLA morphological complexity by employing a plat based on the indexed stems and the stem alternation pattern as represented in the stem-space chart. The advantage of this type of a plat is that the input does not have to be subjected to certain data manipulations which can influence the results. This approach provides a plausible account for the TLA verb inflectional system as a stem-based language.

6.3 Scope and suggestions for future work

This work addressed a number of important research questions which remain open for future research. The study mainly focused on the verb inflectional paradigm of triliteral verbs. Although we have partially considered the paradigmatic verb context when the object suffix is appended to 3SGM perfective form, further research should seek to establish the paradigmatic behaviour when the object pronouns are attached to all the possible inflectional forms in any sub-paradigm.

The complexity of nominal paradigms can also be a rich system of stem alternations, providing a good source for further work based on the stem space and principal parts model.

This synchronic account has highlighted the factors contributing to the morphomicity feature of TLA stems, the complexity of which is reflected by a feature referencing formalism. However, this account would benefit from the insights of diachronic investigation into how such factors may have developed and lead to the current change.

The stem space in this work is based on stem alternations with the [+morphological] feature, but a fully developed algorithm which specifies the essential principles for constructing the stem space in complex systems such as TLA would be an advantage. Such work should seek to provide specific boundaries for the notion of verb regularity. Likewise, future work should seek to develop clearer principles for constructing plats for Semitic based stems.

In semi-autonomous morphology the question of: ‘to index or not to index the stem?’ is essential and future study should clearly identify the indexing rules.

The present work has considered the implicative relation of TLA based on stem space and principal parts. Calculation of entropy measures would be an obvious extension of the current study.

Finally, while these questions have been left for future research, this work can hopefully advance a deeper knowledge of the TLA verb morphological complexity in inflectional paradigms.

Appendix A

A.1 Verb measures in Standard Arabic

M	form	example	gloss
*M*1	FaʕaLa	kataba	to write
*M*2	FaʕʕaLa	kassara	to shatter
*M*3	FaaʕaLa	kaataba	to write to sb.
*M*4	ʔaFʕaLa	ʔaktaba	to correspond
*M*5	taFaʕʕaLa	taḥaddata	to converse
*M*6	taFaaʕaLa	takaataba	to correspond
*M*7	inFaʕaLa	inkasara	to break
*M*8	iftaʕaLa	ijtamaʕa	to assemble
*M*9	ifaʕʕaLa	iḥmarra	to be red
*M*10	stFʕaLa	istaktaba	to dictate

A.2 Verb paradigms

(1) The paradigm of the sound verb *rekab* ‘ride’

MPs	Perfect	Imperfect	Imperative
1SG	rekab-(i)-t	n-arkub	
1PL	rekab-na	n-arukb-u	
2SG.M	rekab-(i)-t	t-arkub	arkub
2SG.F	rekab-ti	t-arukb-i	arukb-i
2PL.M	rekab-tu	t-arukb-u	arukb-u
2PL.F	rekab-tin	t-arukb-in	arukb-in
3SG.M	rekab	y-arkub	
3SG.F	rukub-it	t-arkub	
3PL.M	rukub-u	y-arukb-u	
3PL.F	rukub-in	y-arukb-in	

(2) The paradigm of the hollow verb *baaʕ* ‘sell’

	Perfective	Imperfective	Imperative
1SG	biʕ-(i)-t	n-biiʕ	
1PL	biʕ-na	n-biiʕ-u	
2SG.M	biʕ-(i)-t	t-biiʕ	biiʕ
2SG.F	biʕ-ti	t-biiʕ-i	biiʕ-i
2PL.M	biʕ-tu	t-biiʕ-u	biiʕ-u
2PL.F	biʕ-tin	t-biiʕ-in	biiʕ-in
3SG.M	baaʕ	y-biiʕ	
3SG.F	baaʕ-it	t-biiʕ	
3PL.M	baaʕ-u	y-biiʕ-u	
3PL.F	baaʕ-in	y-biiʕ-in	

(3) The paradigm of the hollow verb *gaal* ‘say’

MPs	Perfective	Imperfective	Imperative
1SG	gul-(i)-t	n-guul	
1PL	gul-na	n-guul-u	
2SG.M	gul-(i)-t	t-guul	guul
2SG.F	gul-ti	t-guul-i	guul-i
2PL.M	gul-tu	t-guul -u	guul-u
2PL.F	gul-tin	t-guul-in	guul-in
3SG.M	gaal	y-guul	
3SG.F	gaal-it	t-guul	
3PL.M	gaal-u	y-guul-u	
3PL.F	gaal-in	y-guul-in	

(4) The paradigm of the hollow verb *ġaab* ‘be absent’

MPs	Perfective	Imperfective	Imperative
1SG	ġub-(i)-t	n-ġiib	
1PL	ġub-na	n-ġiib-u	
2SG.M	ġub-(i)-t	t-ġiib	ġiib
2SG.F	ġub-ti	t-ġiib-i	ġiib-i
2PL.M	ġub-tu	t-ġiib -u	ġiib-u
2PL.F	ġub-tin	t-ġiib-in	ġiib-in
3SG.M	ġaab	y-ġiib	
3SG.F	ġaab-it	t-ġiib	
3PL.M	ġaab-u	y-ġiib-u	
3PL.F	ġaab-in	y-ġiib -in	

(5) The paradigm of the hollow verb *baan* ‘appear’

MPs	Perfective	Imperfective	Imperative
1SG	bin-(i)-t	n-baan	
1PL	bin-na	n-baan-u	
2SG.M	bin-(i)-t	t-baan	baan
2SG.F	bin-ti	t-baan-i	baan-i
2PL.M	bin-tu	t-baan-u	baan-u
2PL.F	bin-tin	t-baan-in	baan-in
3SG.M	baan	y-baan	
3SG.F	baan-it	t-baan	
3PL.M	baan-u	y-baan-u	
3PL.F	baan-in	y-baan-in	

(6) The paradigm of the defective verb *jree* 'run'

MPs	Perfective	Imperfective	Imperative
1SG	j(i)ree-t	ni-jri	
1PL	j(i)ree-na	ni-jr-u	
2SG.M	j(i)ree-t	ti-jri	i-jri
2SG.F	j(i)ree-ti	ti-jr-i	i-jr-i
2PL.M	j(i)ree-tu	ti-jr-u	i-jr-u
2PL.F	j(i)ree-tin	ti-jr-in	i-jr-in
3SG.M	j(i)ree	yi-jri	
3SG.F	j(i)r-it	ti-jri	
3PL.M	j(i)r-u	yi-jr-u	
3PL.F	j(i)r-in	yi-jr-in	

(7) The paradigm of the defective verb *kilee* 'eat'

MPs	Perfective	Imperfective	Imperative
1SG	k(i)lee-t	na-kil	
1PL	k(i)lee-na	na-kil-u	
2SG.M	k(i)lee-t	ta-kil	kuul
2SG.F	k(i)lee-ti	ta-kil-i	kuul-i
2PL.M	k(i)lee-tu	ta-kil-u	kuul-u
2PL.F	k(i)lee-tin	ta-kil-in	kuul-in
3SG.M	k(i)l-ee	ya-kil	
3SG.F	k(i)l-it	ta-kil	
3PL.M	k(i)l-u	ya-kil-u	
3PL.F	k(i)l-in	ya-kil-in	

(8) The paradigm of the defective verb *xdee* 'take'

MPs	Perfective	Imperfective	Imperative
1SG	x(i)dee-t	na-xid	
1PL	x(i)dee-na	na-xid-u	
2SG.M	x(i)dee-t	ta-xid	xuud
2SG.F	x(i)dee-ti	ta-xid-i	xuud-i
2PL.M	x(i)dee-tu	ta-xid-u	xuud-u
2PL.F	x(i)dee-tin	ta-xid-in	xuud-in
3SG.M	x(i)dee	ya-xid	
3SG.F	x(i)d-it	ta-xid	
3PL.M	x(i)d-u	ya-xid-u	
3PL.F	x(i)d-in	ya-xid-in	

(9) The paradigm of the defective verb *jee* ‘came/arrive’

MPs	Perfective	Imperfective	Imperative
1SG	jee-t	n-ji	
1PL	jee-na	n-j-u	
2SG.M	jee-t	t-ji	taʕaal
2SG.F	jee-ti	t-j-i	teʕaal-i
2PL.M	jee-tu	t-j-u	teʕaal-u
2PL.F	jee-tin	t-j-in	teʕaal-in
3SG.M	jee	y-ji	
3SG.F	j-it	t-ji	
3PL.M	j-u	y-j-u	
3PL.F	j-in	y-j-in	

(10) The paradigm of the defective verb *ree* ‘saw’

MPs	Perfective	Imperfective	Imperative
1SG	ree-t	n-ri	
1PL	ree-na	n-r-u	
2SG.M	ree-t	t-ri	šuuf
2SG.F	ree-ti	t-r-i	šuuf-i
2PL.M	ree-tu	t-r-u	šuuf-u
2PL.F	ree-tin	t-r-in	šuuf-in
3SG.M	ree	y-ri	
3SG.F	r-it	t-ri	
3PL.M	r-u	y-r-u	
3PL.F	r-in	y-r-in	

A.3 French System

A.3.1 Abstract (speaker-oriented) plat

```
% French verbs, abstracting MPSs by the stems that they employ.
% rules of sandhi to reduce IC differences; this is the speaker-oriented
plat.
% Greg Stump 2012

ABBR 1 1S2C

CONJ     Stem1 Stem2 Stem3 Stem4 Stem5 Stem6 Stem7 Stem8 Stem9 Stem10
TEMPLATE   1A    1A    1A    1A    1A    1A    1A    1A    1A    1A
ÊTRE      -sɥi  -ɛ    !sɔm  !ɛt   -sɔ̃   -et   -et   -fy   -fy   -sə
AVOIR     -ɛ    -a    -av   -av   -ɔ̃    -av   -av   -y    -y    -o
AIMER     -     -     -     -     -     -     -     -ɛ    -a    -ə
COLLER    -     -     -     -     -     -     -     -ɛ    -a    -ə
BEURRER   -     -     -     -     -     -     -     -ɛ    -a    -ə
DÉJEUNER  -     -     -     -     -     -     -     -ɛ    -a    -ə
ÉCROUER   -     -     -     -     -     -     -     -ɛ    -a    -
ÉCHOUER   -     -     -     -     -     -     -     -ɛ    -a    -
REFLUER   -     -     -     -     -     -     -     -ɛ    -a    -
REMUER    -     -     -     -     -     -     -     -ɛ    -a    -
LEVER     -     -     -     -     -     -     -     -ɛ    -a    -ə
CÉDER     -     -     -     -     -     -     -     -ɛ    -a    -ə
COPIER    -     -     -     -     -     -     -     -ɛ    -a    -
APPUYER   -     -     -j    -j    -     -j    -     -jɛ   -ja   -
BROYER    -     -     -j    -j    -     -j    -     -jɛ   -ja   -
ENVOYER   -wa   -wa   -waj  -waj  -wa   -waj  -wa   -wajɛ -waja -ɛ
FINIR     -     -     -s    -s    -s    -s    -s    -     -     -
HAÏR      !ɛ    !ɛ    -s    -s    -s    -s    -s    -     -     -
ALLER     !vɛ   !va   -     -     !vɔ̃   -     -     -ɛ    -a    @i
TENIR     -jɛ̃   -jɛ̃   -ən   -ən   -jɛn  -ən   -ən   -ɛ̃    -ɛ̃    -jɛ̃d
ACQUÉRIR  -jɛʁ  -jɛʁ  -eʁ   -eʁ   -jɛʁ  -eʁ   -eʁ   -i    -i    -eʁ
SENTIR    -     -     -t    -t    -t    -t    -t    -ti   -ti   -ti
VÊTIR     -     -     -t    -t    -t    -t    -t    -ti   -ti   -ti
COUVRIR   -ʁ    -ʁ    -ʁ    -ʁ    -ʁ    -ʁ    -ʁi   -ʁi   -ʁi   -ʁi
CUEILLIR  -     -     -     -     -     -     -     -i    -i    -ə
ASSAILLIR -     -     -     -     -     -     -     -i    -i    -i
BOUILLIR  -     -     -j    -j    -j    -j    -     -ji   -ji   -ji
DORMIR    -     -     -m    -m    -m    -m    -m    -mi   -mi   -mi
COURIR    -     -     -     -     -     -     -     -y    -y    -
MOURIR    -œʁ   -œʁ   -uʁ   -uʁ   -œʁ   -uʁ   -uʁ   -uʁy  -uʁy  -uʁ
SERVIR    -     -     -v    -v    -v    -v    -v    -vi   -vi   -vi
FUIR      -     -     -j    -j    -     -j    -     -     -     -
RECEVOIR  -wa   -wa   -əv   -əv   -wav  -əv   -əv   -y    -y    -əv
VOIR      -wa   -wa   -waj  -waj  -wa   -waj  -wa   -i    -i    -ɛ
POURVOIR  -wa   -wa   -waj  -waj  -wa   -waj  -wa   -y    -y    -wa
SAVOIR    -ɛ    -ɛ    -av   -av   -av   -av   -av   -y    -y    -o
DEVOIR    -wa   -wa   -əv   -əv   -wav  -əv   -əv   -y    -y    -əv
POUVOIR   -ø    -ø    -uv   -uv   -œv   -uv   -uv   -y    -y    -u
MOUVOIR   -ø    -ø    -uv   -uv   -œv   -uv   -uv   -y    -y    -uv
VALOIR    -o    -o    -al   -al   -al   -al   -al   -aly  -aly  -od
VOULOIR   -ø    -ø    -ul   -ul   -œl   -ul   -ul   -uly  -uly  -ud
ASSEOIR1  -wa   -wa   -waj  -waj  -wa   -waj  -wa   -i    -i    -wa
ASSEOIR2  -je   -je   -ɛj   -ɛj   -ɛj   -ɛj   -ɛ    -i    -i    -je
SURSEOIR  -wa   -wa   -waj  -waj  -wa   -waj  -wa   -i    -i    -wa
RENDRE    -     -     -d    -d    -d    -d    -d    -di   -di   -d
PRENDRE   -ɑ̃    -ɑ̃    -ən   -ən   -ɛn   -ən   -ən   -i    -i    -ɑ̃d
BATTRE    -     -     -t    -t    -t    -t    -t    -ti   -ti   -t
METTRE    -ɛ    -ɛ    -ɛt   -ɛt   -ɛt   -ɛt   -ɛt   -i    -i    -ɛt
```

PEINDRE	-ɛ̃	-ɛ̃	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲi	-ɛɲi	-ɛd̃
JOINDRE	-ɛ̃	-ɛ̃	-aɲ	-aɲ	-aɲ	-aɲ	-aɲ	-aɲi	-aɲi	-ɛd̃
CRAINDRE	-ɛ̃	-ɛ̃	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲi	-ɛɲi	-ɛd̃
VAINCRE	-	-	-k	-k	-k	-k	-k	-ki	-ki	-k
FAIRE	-ɛ	-ɛ	-əz	!fɛt	-ɔ̃	-əz	-əz	-i	-i	-ə
PLAIRE	-ɛ	-ɛ	-ɛz	-ɛz	-ɛz	-ɛz	-ɛz	-y	-y	-ɛ
CONNAÎTRE	-ɛ	-ɛ	-ɛs	-ɛs	-ɛs	-ɛs	-ɛs	-y	-y	-ɛt
NAÎTRE	-ɛ	-ɛ	-ɛs	-ɛs	-ɛs	-ɛs	-ɛs	-aki	-aki	-ɛt
REPAÎTRE	-ɛ	-ɛ	-ɛs	-ɛs	-ɛs	-ɛs	-ɛs	-y	-y	-ɛt
CROÎTRE	-wa	-wa	-was	-was	-was	-was	-was	-y	-y	-wat
CROIRE	-wa	-wa	-waj	-waj	-wa	-waj	-wa	-y	-y	-wa
BOIRE	-wa	-wa	-yv	-yv	-wav	-yv	-yv	-y	-y	-wa
CONCLURE	-	-	-	-	-	-	-	-	-	-
INCLURE	-	-	-	-	-	-	-	-	-	-
COUDRE	-	-	-z	-z	-z	-z	-z	-zi	-zi	-d
MOUDRE	-	-	-l	-l	-l	-l	-l	-ly	-ly	-d
SUIVRE	-	-	-v	-v	-v	-v	-v	-vi	-vi	-v
VIVRE	-i	-i	-iv	-iv	-iv	-iv	-iv	-ɛky	-ɛky	-iv
LIRE	-i	-i	-iz	-iz	-iz	-iz	-iz	-y	-y	-i
DIRE	-	-	-z	!dit	-z	-z	-z	-	-	-
RIRE	-	-	-j	-j	-	-j	-	-	-	-
ÉCRIRE	-	-	-v	-v	-v	-v	-v	-vi	-vi	-
CONFIRE	-	-	-z	-z	-z	-z	-z	-	-	-
CUIRE	-	-	-z	-z	-z	-z	-z	-zi	-zi	-

CONJ	**Stem11**	**Stem12**	**Stem13**	**Stem14**	**Stem15**
TEMPLATE	1A	1A	1A	1A	1A
ÊTRE	-swa	-swa	-swa	-swaj	-swaj
AVOIR	-ɛ	-ɛ	-ɛ	-ɛj	-ɛj
AIMER	-	-	-	-	-
COLLER	-	-	-	-	-
BEURRER	-	-	-	-	-
DÉJEUNER	-	-	-	-	-
ÉCROUER	-	-	-	-	-
ÉCHOUER	-	-	-	-	-
REFLUER	-	-	-	-	-
REMUER	-	-	-	-	-
LEVER	-	-	-	-	-
CÉDER	-	-	-	-	-
COPIER	-	-	-	-	-
APPUYER	-	-	-	-j	-j
BROYER	-	-	-	-j	-j
ENVOYER	-wa	-wa	-wa	-waj	-waj
FINIR	-s	-s	-	-s	-s
HAÏR	-s	-s	!ɛ	-s	-s
ALLER	!aj	-	!va	-	-
TENIR	-jɛn	-ən	-jɛ̃	-ən	-ən
ACQUÉRIR	-jɛʁ	-eʁ	-jɛʁ	-eʁ	-eʁ
SENTIR	-t	-t	-	-t	-t
VÊTIR	-t	-t	-	-t	-t
COUVRIR	-ʁ	-ʁi	-ʁ	-ʁ	-ʁ
CUEILLIR	-	-	-	-	-
ASSAILLIR	-	-	-	-	-
BOUILLIR	-j	-	-	-j	-j
DORMIR	-m	-m	-	-m	-m
COURIR	-	-	-	-	-
MOURIR	-œʁ	-uʁ	-œʁ	-uʁ	-uʁ
SERVIR	-v	-v	-	-v	-v
FUIR	-	-	-	-j	-j
RECEVOIR	-wav	-əv	-wa	-əv	-əv
VOIR	-wa	-wa	-wa	-waj	-waj

POURVOIR	-wa	-wa	-wa	-waj	-waj
SAVOIR	-aʃ	-aʃ	-aʃ	-aʃ	-aʃ
DEVOIR	-wav	-əv	-wa	-əv	-əv
POUVOIR	-ɥis	-ɥis	-ø	-uv	-uv
MOUVOIR	-œv	-uv	-ø	-uv	-uv
VALOIR	-aj	-al	-o	-al	-al
VOULOIR	-œj	-ul	-ø	-ul	-ul
ASSEOIR1	-wa	-wa	-wa	-waj	-waj
ASSEOIR2	-ɛj	-ɛ	-je	-ɛj	-ɛj
SURSEOIR	-wa	-wa	-wa	-waj	-waj
RENDRE	-d	-d	-	-d	-d
PRENDRE	-ɛn	-ən	-ɑ̃	-ən	-ən
BATTRE	-t	-t	-	-t	-t
METTRE	-ɛt	-ɛt	-ɛ	-ɛt	-ɛt
PEINDRE	-ɛɲ	-ɛɲ	-ɛ̃	-ɛɲ	-ɛɲ
JOINDRE	-aɲ	-aɲ	-ɛ̃	-aɲ	-aɲ
CRAINDRE	-ɛɲ	-ɛɲ	-ɛ̃	-ɛɲ	-ɛɲ
VAINCRE	-k	-k	-	-k	-k
FAIRE	-as	-as	-ɛ	-əz	!fɛt
PLAIRE	-ɛz	-ɛz	-ɛ	-ɛz	-ɛz
CONNAÎTRE	-ɛs	-ɛs	-ɛ	-ɛs	-ɛs
NAÎTRE	-ɛs	-ɛs	-ɛ	-ɛs	-ɛs
REPAÎTRE	-ɛs	-ɛs	-ɛ	-ɛs	-ɛs
CROÎTRE	-was	-was	-wa	-was	-was
CROIRE	-wa	-wa	-wa	-waj	-waj
BOIRE	-wav	-yv	-wa	-yv	-yv
CONCLURE	-	-	-	-	-
INCLURE	-	-	-	-	-
COUDRE	-z	-z	-	-z	-z
MOUDRE	-l	-l	-	-l	-l
SUIVRE	-v	-v	-	-v	-v
VIVRE	-iv	-iv	-i	-iv	-iv
LIRE	-iz	-iz	-i	-iz	-iz
DIRE	-z	-z	-	-z	!dit
RIRE	-	-	-	-j	-j
ÉCRIRE	-v	-v	-	-v	-v
CONFIRE	-z	-z	-	-z	-z
CUIRE	-z	-z	-	-z	-z

CONJ	**Stem16**	**Stem17**	**Stem18**	**Stem19**
TEMPLATE	1A	1A	1A	1A
ÊTRE	-etʁ	-et	-ete	-ete
AVOIR	-avwaʁ	-ɛj	-y	-y
AIMER	-e	-	-e	-e
COLLER	-e	-	-e	-e
BEURRER	-e	-	-e	-e
DÉJEUNER	-e	-	-e	-e
ÉCROUER	-e	-	-e	-e
ÉCHOUER	-e	-	-e	-e
REFLUER	-e	-	-e	-e
REMUER	-e	-	-e	-e
LEVER	-e	-	-e	-e
CÉDER	-e	-	-e	-e
COPIER	-e	-	-e	-e
APPUYER	-je	-j	-je	-je
BROYER	-je	-j	-je	-je
ENVOYER	-waje	-waj	-waje	-waje
FINIR	-ʁ	-s	-	-
HAÏR	-ʁ	-s	-	-
ALLER	-e	-	-e	-e
TENIR	-əniʁ	-ən	-əny	-əny

ACQUÉRIR	-eʁiʁ	-ɛʁ	-i	-iz
SENTIR	-tiʁ	-t	-ti	-ti
VÊTIR	-tiʁ	-t	-ty	-ty
COUVRIR	-ʁiʁ	-ʁ	-ɛʁ	-ɛʁt
CUEILLIR	-iʁ	–	-i	-i
ASSAILLIR	-iʁ	–	-i	-i
BOUILLIR	-jiʁ	-j	-ji	-ji
DORMIR	-miʁ	-m	-mi	-mi
COURIR	-iʁ	–	-y	-y
MOURIR	-uʁiʁ	-uʁ	-ɔʁ	-ɔʁt
SERVIR	-viʁ	-v	-vi	-vi
FUIR	-ʁ	-j	–	–
RECEVOIR	-əvwaʁ	-əv	-y	-y
VOIR	-waʁ	-waj	-y	-y
POURVOIR	-waʁ	-waj	-y	-y
SAVOIR	-avwaʁ	-aʃ	-y	-y
DEVOIR	-əvwaʁ	-əv	-y	-y
POUVOIR	-uvwaʁ	-uv	-y	-y
MOUVOIR	-uvwaʁ	-uv	-y	-y
VALOIR	-alwaʁ	-al	-aly	-aly
VOULOIR	-ulwaʁ	-ul	-uly	-uly
ASSEOIR1	-waʁ	-waj	-i	-iz
ASSEOIR2	-waʁ	-ɛj	-i	-iz
SURSEOIR	-waʁ	-waj	-i	-iz
RENDRE	-dʁ	-d	-dy	-dy
PRENDRE	-ɑ̃dʁ	-ən	-i	-iz
BATTRE	-tʁ	-t	-ty	-ty
METTRE	-ɛtʁ	-ɛt	-i	-iz
PEINDRE	-ɛ̃dʁ	-ɛɲ	-ɛ̃	-ɛ̃t
JOINDRE	-ɛ̃dʁ	-aɲ	-ɛ̃	-ɛ̃t
CRAINDRE	-ɛ̃dʁ	-ɛɲ	-ɛ̃	-ɛ̃t
VAINCRE	-kʁ	-k	-ky	-ky
FAIRE	-ɛʁ	-əz	-ɛ	-ɛt
PLAIRE	-ɛʁ	-ɛz	-y	-y
CONNAÎTRE	-ɛtʁ	-ɛs	-y	-y
NAÎTRE	-ɛtʁ	-ɛs	-e	-e
REPAÎTRE	-ɛtʁ	-ɛs	-y	-y
CROÎTRE	-watʁ	-was	-y	-y
CROIRE	-waʁ	-waj	-y	-y
BOIRE	-waʁ	-yv	-y	-y
CONCLURE	-ʁ	–	–	–
INCLURE	-ʁ	–	–	-z
COUDRE	-dʁ	-z	-zy	-zy
MOUDRE	-dʁ	-l	-ly	-ly
SUIVRE	-vʁ	-v	-vi	-vi
VIVRE	-ivʁ	-iv	-ɛky	-ɛky
LIRE	-iʁ	-iz	-y	-y
DIRE	-ʁ	-z	–	-t
RIRE	-ʁ	–	–	–
ÉCRIRE	-ʁ	-v	–	-t
CONFIRE	-ʁ	-z	–	-t
CUIRE	-ʁ	-z	–	-t

http://www.cs.uky.edu/~raphael/linguistics/morphologicalTypology/stems.french.abstract.data (Stump & Finkel, 2013)

A.3.2 Concrete (hearer-oriented) plat

% French verbs, abstracting MPSs by the stems that they employ.
% No rules of sandhi. This is the hearer-oriented plat.
% Greg Stump 2012

ABBR 1 1S2C

CONJ	**Stem1**	**Stem2**	**Stem3**	**Stem4**	**Stem5**	**Stem6**	**Stem7**
TEMPLATE	1A	1A	1A	1A	1A	1A	1A
ÊTRE	-sɥi	-ɛ	!sɔm	!ɛt	-sɔ̃	-et	-etj
AVOIR	-ɛ	-a	-av	-av	-ɔ̃	-av	-avj
AIMER	-	-	-	-	-	-	-j
COLLER	-ɔl	-ɔl	-ol	-ol	-ɔl	-ol	-olj
BEURRER	-œʁ	-œʁ	-øʁ	-øʁ	-œʁ	-øʁ	-øʁj
DÉJEUNER	-	-	-	-	-	-	-j
ÉCROUER	-	-	-	-	-	-	-j
ÉCHOUER	-u	-u	-w	-w	-u	-w	-uj
REFLUER	-	-	-	-	-	-	-j
REMUER	-y	-y	-ɥ	-ɥ	-y	-ɥ	-yj
LEVER	-ɛv	-ɛv	-əv	-əv	-ɛv	-əv	-əvj
CÉDER	-ɛd	-ɛd	-ed	-ed	-ɛd	-ed	-edj
COPIER	-i	-i	-j	-j	-i	-j	-j
APPUYER	-	-	-j	-j	-	-j	-j
BROYER	-	-	-j	-j	-	-j	-j
ENVOYER	-wa	-wa	-waj	-waj	-wa	-waj	-waj
FINIR	-	-	-s	-s	-s	-s	-sj
HAÏR	!ɛ	!ɛ	-s	-s	-s	-s	-sj
ALLER	!vɛ	!va	-	-	!vɔ̃	-	-j
TENIR	-jɛ̃	-jɛ̃	-ən	-ən	-jɛn	-ən	-ənj
ACQUÉRIR	-jɛʁ	-jɛʁ	-eʁ	-eʁ	-jɛʁ	-eʁ	-eʁj
SENTIR	-	-	-t	-t	-t	-t	-tj
VÊTIR	-	-	-t	-t	-t	-t	-tj
COUVRIR	-ʁ	-ʁ	-ʁ	-ʁ	-ʁ	-ʁ	-ʁij
CUEILLIR	-œj	-œj	-øj	-øj	-œj	-øj	-øj
ASSAILLIR	-	-	-	-	-	-	-
BOUILLIR	-	-	-j	-j	-j	-j	-j
DORMIR	-	-	-m	-m	-m	-m	-mj
COURIR	-	-	-	-	-	-	-j
MOURIR	-œʁ	-œʁ	-uʁ	-uʁ	-œʁ	-uʁ	-uʁj
SERVIR	-	-	-v	-v	-v	-v	-vj
FUIR	-	-	-j	-j	-	-j	-j
RECEVOIR	-wa	-wa	-əv	-əv	-wav	-əv	-əvj
VOIR	-wa	-wa	-waj	-waj	-wa	-waj	-waj
POURVOIR	-wa	-wa	-waj	-waj	-wa	-waj	-waj
SAVOIR	-ɛ	-ɛ	-av	-av	-av	-av	-avj
DEVOIR	-wa	-wa	-əv	-əv	-wav	-əv	-əvj
POUVOIR	-ø	-ø	-uv	-uv	-œv	-uv	-uvj
MOUVOIR	-ø	-ø	-uv	-uv	-œv	-uv	-uvj
VALOIR	-o	-o	-al	-al	-al	-al	-alj
VOULOIR	-ø	-ø	-ul	-ul	-œl	-ul	-ulj
ASSEOIR1	-wa	-wa	-waj	-waj	-wa	-waj	-waj
ASSEOIR2	-je	-je	-ɛj	-ɛj	-ɛj	-ɛj	-ɛj
SURSEOIR	-wa	-wa	-waj	-waj	-wa	-waj	-waj
RENDRE	-	-	-d	-d	-d	-d	-dj
PRENDRE	-ɑ̃	-ɑ̃	-ən	-ən	-ɛn	-ən	-ənj
BATTRE	-	-	-t	-t	-t	-t	-tj
METTRE	-ɛ	-ɛ	-ɛt	-ɛt	-ɛt	-ɛt	-ɛtj
PEINDRE	-ɛ̃	-ɛ̃	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ
JOINDRE	-ɛ̃	-ɛ̃	-aɲ	-aɲ	-aɲ	-aɲ	-aɲ
CRAINDRE	-ɛ̃	-ɛ̃	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ	-ɛɲ
VAINCRE	-	-	-k	-k	-k	-k	-kj

FAIRE	-ɛ	-ɛ	-əz	!fɛt	-ɔ̃	-əz	-əzj
PLAIRE	-ɛ	-ɛ	-ɛz	-ɛz	-ɛz	-ɛz	-ɛzj
CONNAÎTRE	-ɛ	-ɛ	-ɛs	-ɛs	-ɛs	-ɛs	-ɛsj
NAÎTRE	-ɛ	-ɛ	-ɛs	-ɛs	-ɛs	-ɛs	-ɛsj
REPAÎTRE	-ɛ	-ɛ	-ɛs	-ɛs	-ɛs	-ɛs	-ɛsj
CROÎTRE	-wa	-wa	-was	-was	-was	-was	-wasj
CROIRE	-wa	-wa	-waj	-waj	-wa	-waj	-waj
BOIRE	-wa	-wa	-yv	-yv	-wav	-yv	-yvj
CONCLURE	–	–	–	–	–	–	-j
INCLURE	–	–	–	–	–	–	-j
COUDRE	–	–	-z	-z	-z	-z	-zj
MOUDRE	–	–	-l	-l	-l	-l	-lj
SUIVRE	–	–	-v	-v	-v	-v	-vj
VIVRE	-i	-i	-iv	-iv	-iv	-iv	-ivj
LIRE	-i	-i	-iz	-iz	-iz	-iz	-izj
DIRE	–	–	-z	!dit	-z	-z	-zj
RIRE	-i	-i	-ij	-ij	-i	-ij	-ij
ÉCRIRE	–	–	-v	-v	-v	-v	-vj
CONFIRE	–	–	-z	-z	-z	-z	-zj
CUIRE	–	–	-z	-z	-z	-z	-zj

CONJ	**Stem8**	**Stem9**	**Stem10**
TEMPLATE	1A	1A	1A
ÊTRE	-fy	-fy	-sə
AVOIR	-y	-y	-o
AIMER	-ɛ	-a	-ə
COLLER	-olɛ	-ola	-olə
BEURRER	-øʁɛ	-øʁa	-œʁə
DÉJEUNER	-ɛ	-a	-ə
ÉCROUER	-ɛ	-a	–
ÉCHOUER	-wɛ	-wa	-u
REFLUER	-ɛ	-a	–
REMUER	-ɥɛ	-ɥa	-y
LEVER	-əvɛ	-əva	-ɛvə
CÉDER	-edɛ	-eda	-ɛdə
COPIER	-jɛ	-ja	-i
APPUYER	-jɛ	-ja	–
BROYER	-jɛ	-ja	–
ENVOYER	-wajɛ	-waja	-ɛ
FINIR	–	–	–
HAÏR	–	–	–
ALLER	-ɛ	-a	@i
TENIR	-ɛ̃	-ɛ̃	-jɛ̃d
ACQUÉRIR	-i	-i	-eʁ
SENTIR	-ti	-ti	-ti
VÊTIR	-ti	-ti	-ti
COUVRIR	-ʁi	-ʁi	-ʁi
CUEILLIR	-øji	-øji	-œjə
ASSAILLIR	-i	-i	-i
BOUILLIR	-ji	-ji	-ji
DORMIR	-mi	-mi	-mi
COURIR	-y	-y	–
MOURIR	-uʁy	-uʁy	-uʁ
SERVIR	-vi	-vi	-vi
FUIR	–	–	–
RECEVOIR	-y	-y	-əv
VOIR	-i	-i	-ɛ
POURVOIR	-y	-y	-wa
SAVOIR	-y	-y	-o
DEVOIR	-y	-y	-əv
POUVOIR	-y	-y	-u

MOUVOIR	-y	-y	-uv
VALOIR	-aly	-aly	-od
VOULOIR	-uly	-uly	-ud
ASSEOIR1	-i	-i	-wa
ASSEOIR2	-i	-i	-je
SURSEOIR	-i	-i	-wa
RENDRE	-di	-di	-d
PRENDRE	-i	-i	-ɑ̃d
BATTRE	-ti	-ti	-t
METTRE	-i	-i	-ɛt
PEINDRE	-ɛɲi	-ɛɲi	-ɛ̃d
JOINDRE	-aɲi	-aɲi	-ɛ̃d
CRAINDRE	-ɛɲi	-ɛɲi	-ɛ̃d
VAINCRE	-ki	-ki	-k
FAIRE	-i	-i	-ə
PLAIRE	-y	-y	-ɛ
CONNAÎTRE	-y	-y	-ɛt
NAÎTRE	-aki	-aki	-ɛt
REPAÎTRE	-y	-y	-ɛt
CROÎTRE	-y	-y	-wat
CROIRE	-y	-y	-wa
BOIRE	-y	-y	-wa
CONCLURE	-	-	-
INCLURE	-	-	-
COUDRE	-zi	-zi	-d
MOUDRE	-ly	-ly	-d
SUIVRE	-vi	-vi	-v
VIVRE	-ɛky	-ɛky	-iv
LIRE	-y	-y	-i
DIRE	-	-	-
RIRE	-i	-i	-i
ÉCRIRE	-vi	-vi	-
CONFIRE	-	-	-
CUIRE	-zi	-zi	-

CONJ	**Stem10a**	**Stem11**	**Stem12**	**Stem13**
TEMPLATE	1A	1A	1A	1A
ÊTRE	-səʁ	-swa	-swaj	-swa
AVOIR	-oʁ	-ɛ	-ɛj	-ɛ
AIMER	-əʁ	-	-j	-
COLLER	-oləʁ	-ɔl	-olj	-ɔl
BEURRER	-œʁəʁ	-œʁ	-øʁj	-œʁ
DÉJEUNER	-əʁ	-	-j	-
ÉCROUER	-ʁ	-	-j	-
ÉCHOUER	-uʁ	-u	-uj	-u
REFLUER	-ʁ	-	-j	-
REMUER	-yʁ	-y	-yj	-y
LEVER	-ɛvəʁ	-ɛv	-əvj	-ɛv
CÉDER	-ɛdəʁ	-ɛd	-edj	-ɛd
COPIER	-iʁ	-i	-j	-i
APPUYER	-ʁ	-	-j	-
BROYER	-ʁ	-	-j	-
ENVOYER	-ɛʁ	-wa	-waj	-wa
FINIR	-ʁ	-s	-sj	-
HAÏR	-ʁ	-s	-sj	!ɛ
ALLER	@iʁ	!aj	-j	!va
TENIR	-jɛd̃ʁi	-jɛn	-ənj	-jɛ˜
ACQUÉRIR	-eʁʁ	-jɛʁ	-eʁj	-jɛʁ
SENTIR	-tiʁ	-t	-tj	-
VÊTIR	-tiʁ	-t	-tj	-
COUVRIR	-ʁiʁ	-ʁ	-ʁij	-ʁ

CUEILLIR	-œjəʁ	-œj	-øj	-œj
ASSAILLIR	-iʁ	-	-	-
BOUILLIR	-jiʁ	-j	-j	-
DORMIR	-miʁ	-m	-mj	-
COURIR	-ʁ	-	-j	-
MOURIR	-uʁʁ	-œʁ	-uʁj	-œʁ
SERVIR	-viʁ	-v	-vj	-
FUIR	-ʁ	-	-j	-
RECEVOIR	-əvʁi	-wav	-əvj	-wa
VOIR	-ɛʁ	-wa	-waj	-wa
POURVOIR	-waʁ	-wa	-waj	-wa
SAVOIR	-oʁ	-aʃ	-aʃj	-aʃ
DEVOIR	-əvʁi	-wav	-əvj	-wa
POUVOIR	-uʁ	-ɥis	-ɥisj	-ø
MOUVOIR	-uvʁi	-œv	-uvj	-ø
VALOIR	-odʁi	-aj	-alj	-o
VOULOIR	-udʁi	-œj	-ulj	-ø
ASSEOIR1	-waʁ	-wa	-waj	-wa
ASSEOIR2	-jeʁ	-ɛj	-ɛj	-je
SURSEOIR	-waʁ	-wa	-waj	-wa
RENDRE	-dʁi	-d	-dj	-
PRENDRE	-ɑ̃dʁi	-ɛn	-ənj	-ɑ̃
BATTRE	-tʁi	-t	-tj	-
METTRE	-ɛtʁi	-ɛt	-ɛtj	-ɛ
PEINDRE	-ɛ̃dʁi	-ɛɲ	-ɛɲ	-ɛ̃
JOINDRE	-ɛ̃dʁi	-aɲ	-aɲ	-ɛ̃
CRAINDRE	-ɛ̃dʁi	-ɛɲ	-ɛɲ	-ɛ̃
VAINCRE	-kʁi	-k	-kj	-
FAIRE	-əʁ	-as	-asj	-ɛ
PLAIRE	-ɛʁ	-ɛz	-ɛzj	-ɛ
CONNAÎTRE	-ɛtʁi	-ɛs	-ɛsj	-ɛ
NAÎTRE	-ɛtʁi	-ɛs	-ɛsj	-ɛ
REPAÎTRE	-ɛtʁi	-ɛs	-ɛsj	-ɛ
CROÎTRE	-watʁi	-was	-wasj	-wa
CROIRE	-waʁ	-wa	-waj	-wa
BOIRE	-waʁ	-wav	-yvj	-wa
CONCLURE	-ʁ	-	-j	-
INCLURE	-ʁ	-	-j	-
COUDRE	-dʁi	-z	-zj	-
MOUDRE	-dʁi	-l	-lj	-
SUIVRE	-vʁi	-v	-vj	-
VIVRE	-ivʁi	-iv	-ivj	-i
LIRE	-iʁ	-iz	-izj	-i
DIRE	-ʁ	-z	-zj	-
RIRE	-iʁ	-i	-ij	-i
ÉCRIRE	-ʁ	-v	-vj	-
CONFIRE	-ʁ	-z	-zj	-
CUIRE	-ʁ	-z	-zj	-

CONJ	**Stem14**	**Stem15**	**Stem16**	**Stem17**	**Stem18**	**Stem19**
TEMPLATE	1A	1A	1A	1A	1A	1A
ÊTRE	-swaj	-swaj	-etʁ	-et	-ete	-ete
AVOIR	-ɛj	-ɛj	-avwaʁ	-ɛj	-y	-y
AIMER	-	-	-e	-	-e	-e
COLLER	-ol	-ol	-ole	-ol	-ole	-ole
BEURRER	-øʁ	-øʁ	-øʁe	-øʁ	-øʁe	-øʁe
DÉJEUNER	-	-	-e	-	-e	-e
ÉCROUER	-	-	-e	-	-e	-e
ÉCHOUER	-w	-w	-we	-w	-we	-we
REFLUER	-	-	-e	-	-e	-e
REMUER	-ɥ	-ɥ	-ɥe	-ɥ	-ɥe	-ɥe

LEVER	-əv	-əv	-əve	-əv	-əve	-əve
CÉDER	-ed	-ed	-ede	-ed	-ede	-ede
COPIER	-j	-j	-je	-j	-je	-je
APPUYER	-j	-j	-je	-j	-je	-je
BROYER	-j	-j	-je	-j	-je	-je
ENVOYER	-waj	-waj	-waje	-waj	-waje	-waje
FINIR	-s	-s	-ʁ	-s	-	-
HAÏR	-s	-s	-ʁ	-s	-	-
ALLER	-	-	-e	-	-e	-e
TENIR	-ən	-ən	-əniʁ	-ən	-əny	-əny
ACQUÉRIR	-eʁ	-eʁ	-eʁiʁ	-ɛʁ	-i	-iz
SENTIR	-t	-t	-tiʁ	-t	-ti	-ti
VÊTIR	-t	-t	-tiʁ	-t	-ty	-ty
COUVRIR	-ʁ	-ʁ	-ʁiʁ	-ʁ	-ɛʁ	-ɛʁt
CUEILLIR	-øj	-øj	-øjiʁ	-øj	-øji	-øji
ASSAILLIR	-	-	-iʁ	-	-i	-i
BOUILLIR	-j	-j	-jiʁ	-j	-ji	-ji
DORMIR	-m	-m	-miʁ	-m	-mi	-mi
COURIR	-	-	-iʁ	-	-y	-y
MOURIR	-uʁ	-uʁ	-uʁiʁ	-uʁ	-ɔʁ	-ɔʁt
SERVIR	-v	-v	-viʁ	-v	-vi	-vi
FUIR	-j	-j	-ʁ	-j	-	-
RECEVOIR	-əv	-əv	-əvwaʁ	-əv	-y	-y
VOIR	-waj	-waj	-waʁ	-waj	-y	-y
POURVOIR	-waj	-waj	-waʁ	-waj	-y	-y
SAVOIR	-aʃ	-aʃ	-avwaʁ	-aʃ	-y	-y
DEVOIR	-əv	-əv	-əvwaʁ	-əv	-y	-y
POUVOIR	-uv	-uv	-uvwaʁ	-uv	-y	-y
MOUVOIR	-uv	-uv	-uvwaʁ	-uv	-y	-y
VALOIR	-al	-al	-alwaʁ	-al	-aly	-aly
VOULOIR	-ul	-ul	-ulwaʁ	-ul	-uly	-uly
ASSEOIR1	-waj	-waj	-waʁ	-waj	-i	-iz
ASSEOIR2	-ɛj	-ɛj	-waʁ	-ɛj	-i	-iz
SURSEOIR	-waj	-waj	-waʁ	-waj	-i	-iz
RENDRE	-d	-d	-dʁ	-d	-dy	-dy
PRENDRE	-ən	-ən	-ɑ̃dʁ	-ən	-i	-iz
BATTRE	-t	-t	-tʁ	-t	-ty	-ty
METTRE	-ɛt	-ɛt	-ɛtʁ	-ɛt	-i	-iz
PEINDRE	-ɛɲ	-ɛɲ	-ɛ̃dʁ	-ɛɲ	-ɛ̃	-ɛ̃t
JOINDRE	-aɲ	-aɲ	-ɛ̃dʁ	-aɲ	-ɛ̃	-ɛ̃t
CRAINDRE	-ɛɲ	-ɛɲ	-ɛ̃dʁ	-ɛɲ	-ɛ̃	-ɛ̃t
VAINCRE	-k	-k	-kʁ	-k	-ky	-ky
FAIRE	-əz	!fɛt	-ɛʁ	-əz	-ɛ	-ɛt
PLAIRE	-ɛz	-ɛz	-ɛʁ	-ɛz	-y	-y
CONNAÎTRE	-ɛs	-ɛs	-ɛtʁ	-ɛs	-y	-y
NAÎTRE	-ɛs	-ɛs	-ɛtʁ	-ɛs	-e	-e
REPAÎTRE	-ɛs	-ɛs	-ɛtʁ	-ɛs	-y	-y
CROÎTRE	-was	-was	-watʁ	-was	-y	-y
CROIRE	-waj	-waj	-waʁ	-waj	-y	-y
BOIRE	-yv	-yv	-waʁ	-yv	-y	-y
CONCLURE	-	-	-ʁ	-	-	-
INCLURE	-	-	-ʁ	-	-	-z
COUDRE	-z	-z	-dʁ	-z	-zy	-zy
MOUDRE	-l	-l	-dʁ	-l	-ly	-ly
SUIVRE	-v	-v	-vʁ	-v	-vi	-vi
VIVRE	-iv	-iv	-ivʁ	-iv	-ɛky	-ɛky
LIRE	-iz	-iz	-iʁ	-iz	-y	-y
DIRE	-z	!dit	-ʁ	-z	-	-t
RIRE	-ij	-ij	-iʁ	-j	-i	-i
ÉCRIRE	-v	-v	-ʁ	-v	-	-t
CONFIRE	-z	-z	-ʁ	-z	-	-t

```
CUIRE     -z    -z    -ʁ    -z    -     -t
```

http://www.cs.uky.edu/~raphael/linguistics/morphologicalTypology/stems.french.concrete.data (Stump & Finkel, 2013)

A.4 Koasati System

```
% Koasati, a Muskogean language of Louisiana
% Data from: Kimball, Geoffrey D. 1991, Koasati Grammar. Lincoln & London:
% University of Nebraska Press, 56-89

 IC      1SgAff     2SgAff     3Aff     1PlAff       2PlAff
 1A      R-li       is-R       R        il-R         has-R
 1B      R-li       R⌊s⌋       R        R⌊l⌋         R⌊has⌋
 2Ai     R-li       R-ci       R        R-híli       R-háci
 2Aii    Xli-li     X-ci       Xli      X-híli       X-háci
 2B      Xlí:ci-li  X:ci⌊cí⌋   Xlí:ci   X:ci⌊hilí⌋   X:ci⌊hací⌋
 2C      R-li       R⌊ci⌋      R        R⌊li⌋        R⌊haci⌋
 3A.ka   Xka-li     X-híska    Xka      X-hílka      X-háska
 3A.ki   Xki-li     X-híska    Xki      X-hílka      X-háska
 3A.ko   Xko-li     X-híska    Xko      X-hílka      X-háska
 3B      R-li       R-íska     R        R-ílka       R-áska
 3Ci     R-li       R-tíska    R        R-tílka      R-táska
 3Cii    R-l-o      R-tísk-o   R-o      R-tílk-o     R-tásk-o

% vim:nospell nowrap textwidth=10000:
```

http://www.cs.uky.edu/~raphael/linguistics/morphologicalTypology/complexity.koasati.data
(Stump & Finkel, 2013)

A.5 Fur System

A.5.1 Fur plat

```
ABBR 1 1S/1C

% these first MPSs are the distillation
IC       Non3Subj Non3Perf Non3Pres 3sgSubj 3sgPerf 3sgPres 3plNonhSubj 3plNonhPerf 3plNonhPres

TEMPLATE 1A       1A       1A       1A      1A      1A      1A          1A          1A
I1A      LH+o     LH+ò     LH+èl    HH+o    HH+ò    HH+èl   HH+òl       HH+ùl       HH+èl+à/ì
I1B      LH+o     LH+ò     LF+Ø     HH+o    HH+ò    HF+Ø    HH+òl       HH+ùl       HH+è
I1C      LH+o     LH+ò     LH+ì     HH+o    HH+ò    HH+ì    HH+òl       HH+ùl       HH+è
I2A      HH+ò     HH+o     HH+èl    LL+o    LL+ò    LL+èl   LL+òl       LL+ùl       LL+èl+à/ì
I2B      HH+ò     HH+o     HF+Ø     LL+o    LL+ò    LL+Ø    LL+òl       LL+ùl       LL+è
I2C      HH+ò     HH+o     HH+ì     LL+o    LL+ò    LL+ì    LL+òl       LL+ùl       LL+è
II1A     LH+i     LH+i     LH+itì   HH+i    HH+i    HH+itì  HH+i+A(l)   HH+i+è      HH+iti+A(l)
II1B     LH+i     LH+i     LF+Ø     HH+i    HH+i    HF+Ø    HH+i+A(l)   HH+i+è      HH+è
II2A     HH+ì     HH+ì     HH+itì   LL+i    LL+i    LL+itì  LL+i+A(l)   LL+i+è      LL+iti+A(l)
II2B     HH+ì     HH+ì     HF+Ø     LL+i    LL+i    LF+Ø    LL+i+A(l)   LL+i+è      LL+è
IIIA     HH+ì     HH+à     HH+èl    LH+ì    LH+à    LH+èl   LH+è        LH+e        LH+èl+à
IIIB     HH+ò     HH+ò     HH+èl    LH+ò    LH+ò    LH+èl   LH+è        LH+e        LH+èl+à
IIIC     HH+ò     HH+ò     HH+èl    LF+Ø    LH+ò    LH+èl   LH+è        LH+e        LH+èl+à
IIID     HF+Ø     HH+à     HH+èl    LF+Ø    LH+à    LH+èl   LH+è        LH+e        LH+èl+à
IIIE     HF+Ø     HH+à     HH+èl    LF+Ø    LH+ò    LH+èl   LH+è        LH+e        LH+èl+à
IVA      HF+Ø     HH+ò     HH+èl    LF+Ø    LH+ò    LH+èl   LH+Al       LH+e        LH+èl+à
```

```
IVB      HH+ò      HH+ò      HH+èl     LH+ò     LH+ò     LH+èl    LH+Al       LH+e        LH+èl+à
IVC      HF+Ø      HH+à      HH+èl     LF+Ø     LH+à     LH+èl    LH+Al       LH+e        LH+èl+à
IVD      HH+à      HH+à      HH+èl     LH+à     LH+à     LH+èl    LH+Al       LH+e        LH+èl+à

% these 3 MPSs are isomorphic to previous ones
IC       3plHSubj  3plHPerf 3plHPres
TEMPLATE 1A        1A       1A
I1A      LH+òl     LH+ùl    LH+èl+à/ì
I1B      LH+òl     LH+ùl    LH+è
I1C      LH+òl     LH+ùl    LH+è
I2A      HH+òl     HH+ùl    HH+èl+à/ì
I2B      HH+òl     HH+ùl    HH+è
I2C      HH+òl     HH+ùl    HH+è
II1A     LH+i+A(l) LH+i+è   LH+iti+A(l)
II1B     LH+i+A(l) LH+i+è   LH+è
II2A     HH+i+A(l) HH+i+è   HH+iti+A(l)
II2B     HH+i+A(l) HH+i+è   HH+è
IIIA     HH+è      HH+e     HH+èl+à
IIIB     HH+è      HH+e     HH+èl+à
IIIC     HH+è      HH+e     HH+èl+à
IIID     HH+è      HH+e     HH+èl+à
IIIE     HH+è      HH+e     HH+èl+à
IVA      HH+Al     HH+e     HH+èl+à
IVB      HH+Al     HH+e     HH+èl+à
IVC      HH+Al     HH+e     HH+èl+à
IVD      HH+Al     HH+e     HH+èl+à
```

http://www.cs.uky.edu/~raphael/linguistics/morphologicalTypology/complexity.fur.data (Stump & Finkel, 2013)

A.5.2 Fur dynamic principal parts

```
          1       2       3       4       5       6       7       8       9
           Non3Su  Non3Pe  Non3Pr  3sgSub  3sgPer  3sgPre  3plNon  3plNon  3plNon
 -----------------------------------------------------------------------
      I1A
        3| 3       3       3       3       3       3       3       3       3      avg 1.00
        6| 6       6       6       6       6       6       6       6       6      avg 1.00
        9| 9       9       9       9       9       9       9       9       9      avg 1.00
      I1A: 1 principal parts; lowest average 1.00; 3 analyses; 33.3% of 9 possible analyses
   -----------------------------------------------------------------------
      I1B
      1,3| 1       1       3       1       1       3       1       1       3      avg 1.00
      1,6| 1       1       6       1       1       6       1       1       6      avg 1.00
      2,3| 2       2       3       2       2       3       2       2       3      avg 1.00
      2,6| 2       2       6       2       2       6       2       2       6      avg 1.00
      3,4| 4       4       3       4       4       3       4       4       3      avg 1.00
      3,5| 5       5       3       5       5       3       5       5       3      avg 1.00
      3,7| 7       7       3       7       7       3       7       7       3      avg 1.00
      3,8| 8       8       3       8       8       3       8       8       3      avg 1.00
      4,6| 4       4       6       4       4       6       4       4       6      avg 1.00
      5,6| 5       5       6       5       5       6       5       5       6      avg 1.00
      6,7| 7       7       6       7       7       6       7       7       6      avg 1.00
      6,8| 8       8       6       8       8       6       8       8       6      avg 1.00
       I1B: 2 principal parts; lowest average 1.00; 12 analyses; 33.3% of 36 possible analyses
 -----------------------------------------------------------------------
      I1C
        3| 3       3       3       3       3       3       3       3       3      avg 1.00
        6| 6       6       6       6       6       6       6       6       6      avg 1.00
      I1C: 1 principal parts; lowest average 1.00; 2 analyses; 22.2% of 9 possible analyses
 -----------------------------------------------------------------------
      I2A
        6| 6       6       6       6       6       6       6       6       6      avg 1.00
        9| 9       9       9       9       9       9       9       9       9      avg 1.00
       I2A: 1 principal parts; lowest average 1.00; 2 analyses; 22.2% of 9 possible analyses
 -----------------------------------------------------------------------
      I2B
        6| 6       6       6       6       6       6       6       6       6      avg 1.00
       I2B: 1 principal parts; lowest average 1.00; 1 analyses; 11.1% of 9 possible analyses
 -----------------------------------------------------------------------
      I2C
        3| 3       3       3       3       3       3       3       3       3      avg 1.00
        6| 6       6       6       6       6       6       6       6       6      avg 1.00
       I2C: 1 principal parts; lowest average 1.00; 2 analyses; 22.2% of 9 possible analyses
 -----------------------------------------------------------------------
     II1A
```

```
      3| 3      3      3      3      3      3      3      3      3      avg 1.00
      6| 6      6      6      6      6      6      6      6      6      avg 1.00
      9| 9      9      9      9      9      9      9      9      9      avg 1.00
     II1A: 1 principal parts; lowest average 1.00; 3 analyses; 33.3% of 9 possible analyses
 ------------------------------------------------------------------------
    II1B
    1,3| 1      1      3      1      1      3      1      1      3      avg 1.00
    1,6| 1      1      6      1      1      6      1      1      6      avg 1.00
    1,9| 1      1      1,9    1      1      1,9    1      1      9      avg 1.22
    2,3| 2      2      3      2      2      3      2      2      3      avg 1.00
    2,6| 2      2      6      2      2      6      2      2      6      avg 1.00
    2,9| 2      2      2,9    2      2      2,9    2      2      9      avg 1.22
    3,4| 4      4      3      4      4      3      4      4      3      avg 1.00
    3,5| 5      5      3      5      5      3      5      5      3      avg 1.00
    3,7| 7      7      3      7      7      3      7      7      3      avg 1.00
    3,8| 8      8      3      8      8      3      8      8      3      avg 1.00
    4,6| 4      4      6      4      4      6      4      4      6      avg 1.00
    4,9| 4      4      4,9    4      4      4,9    4      4      9      avg 1.22
    5,6| 5      5      6      5      5      6      5      5      6      avg 1.00
    5,9| 5      5      5,9    5      5      5,9    5      5      9      avg 1.22
    6,7| 7      7      6      7      7      6      7      7      6      avg 1.00
    6,8| 8      8      6      8      8      6      8      8      6      avg 1.00
    7,9| 7      7      7,9    7      7      7,9    7      7      9      avg 1.22
    8,9| 8      8      8,9    8      8      8,9    8      8      9      avg 1.22
     II1B: 2 principal parts; lowest average 1.00; 18 analyses; 50.0% of 36 possible analyses
 ------------------------------------------------------------------------
    II2A
      3| 3      3      3      3      3      3      3      3      3      avg 1.00
      6| 6      6      6      6      6      6      6      6      6      avg 1.00
      9| 9      9      9      9      9      9      9      9      9      avg 1.00
     II2A: 1 principal parts; lowest average 1.00; 3 analyses; 33.3% of 9 possible analyses
 ------------------------------------------------------------------------
    II2B
      6| 6      6      6      6      6      6      6      6      6      avg 1.00
     II2B: 1 principal parts; lowest average 1.00; 1 analyses; 11.1% of 9 possible analyses
 ------------------------------------------------------------------------
    IIIA
      4| 4      4      4      4      4      4      4      4      4      avg 1.00
    IIIA: 1 principal parts; lowest average 1.00; 1 analyses; 11.1% of 9 possible analyses
 ------------------------------------------------------------------------
    IIIB
    4,7| 4      4      4      4      4      4      7      4      4      avg 1.00
     IIIB: 2 principal parts; lowest average 1.00; 1 analyses; 2.8% of 36 possible analyses
 ------------------------------------------------------------------------
    IIIC
    1,4| 1      1,4    4      4      1,4    4      1,4    4      4      avg 1.33
    IIIC: 2 principal parts; lowest average 1.33; 1 analyses; 2.8% of 36 possible analyses
 ------------------------------------------------------------------------
    IIID
  1,5,7| 1      5      1      1      5      1      7      1      1      avg 1.00
  4,5,7| 4,5    5      4      4      5      4      7      4      4      avg 1.11
     IIID: 3 principal parts; lowest average 1.00; 2 analyses; 2.4% of 84 possible analyses
 ------------------------------------------------------------------------
    IIIE
    2,5| 2,5    2      2      2,5    5      2      2,5    2      2      avg 1.33
     IIIE: 2 principal parts; lowest average 1.33; 1 analyses; 2.8% of 36 possible analyses
 ------------------------------------------------------------------------
     IVA
    1,2| 1      2      1      1      2      1      1,2    1      1      avg 1.11
      IVA: 2 principal parts; lowest average 1.11; 1 analyses; 2.8% of 36 possible analyses
 ------------------------------------------------------------------------
     IVB
    1,7| 1      1,7    7      1,7    1,7    7      7      7      7      avg 1.33
    4,7| 4      4      4      4      4      4      7      4      4      avg 1.00
      IVB: 2 principal parts; lowest average 1.00; 2 analyses; 5.6% of 36 possible analyses
 ------------------------------------------------------------------------
     IVC
  1,2,7| 1      2      1      1      2,7    1      7      1      1      avg 1.11
  1,5,7| 1      5      1      1      5      1      7      1      1      avg 1.00
  2,4,7| 2,4    2      2      4      2,7    2      7      2      2      avg 1.22
  4,5,7| 4,5    5      4      4      5      4      7      4      4      avg 1.11
      IVC: 3 principal parts; lowest average 1.00; 4 analyses; 4.8% of 84 possible analyses
 ------------------------------------------------------------------------
     IVD
      1| 1      1      1      1      1      1      1      1      1      avg 1.00
      4| 4      4      4      4      4      4      4      4      4      avg 1.00
      IVD: 1 principal parts; lowest average 1.00; 2 analyses; 22.2% of 9 possible analyses
Average number of principal parts: 1.58
```

```
Average lowest average of principal parts needed: 1.04
Average number of alternative analyses: 3.26
Average ratio of actual to possible: 17.34%
```

http://www.cs.uky.edu/~raphael/linguistics/principalPartsFull.cgi (Stump & Finkel, 2013)

Appendix B

The input formatting and the results of the TLA exponent-based and stem-based plats. The data was converted into a plain text form

B.1 TLA exponent-based plat

B.1.1 The input formatting for the TLA exponent-based plat

% Noura Ramli

ABBR 1 1C1S2C2S3C3S4C % use for 1A for forms with all three stem consonants

CONJ	**perf1Sg**	**perf2mSg**	**perf2fSg**	**perf3mSg**	**perf3fSg**
getal	-e-a-t	-e-a-t	-e-a-ti	-e-a-	-u-u-it
kallim	-aD-i-t	-aD-i-t	-aD-i-ti	-aD-i-	-aD-Ø-it
saanid	-a:-i-t	-a:-i-t	-a:-i-ti	-a:-i-	-a:-Ø-it
tesallif	te-aD-i-t	te-aD-i-t	te-aD-i-ti	te-aD-i-	te-aD-Ø-it
tefaahim	te-a:-i-t	te-a:-i-t	te-a:-i-ti	te-a:-i-	te-a:-Ø-it
inkesar	!none	!none	!none	in-e-a-	in-u-u-it
intefax	!none	!none	!none	i-t-e-a-	i-t-u-u-it
stanjid	sta-Ø-i-t	sta-Ø-i-t	sta-Ø-i-ti	sta-Ø-i-	sta-i-Ø-it
nsee	-Ø-ee-t	-Ø-ee-t	-Ø-ee-ti	-Ø-ee-	-Ø-Ø-it
bakka	-aD-ee-t	-aD-ee-t	-aD-ee-ti	-aD-a-	-aD-Ø-it
naada	-a:-ee-t	-a:-ee-t	-a:-ee-ti	-a:-a-	-a:-Ø-it
temanna	te-aD-ee-t	te-aD-ee-t	te-aD-ee-ti	te-aD-a-	te-aD-Ø-it
telaaga	te-a:-ee-t	te-a:-ee-t	te-a:-ee-ti	te-a:-a-	te-a:-Ø-it
inkiwee	!none	!none	!none	in-i-ee-	in-i-Ø-it
intisee	!none	!none	!none	i-t-i-ee-	i-t-i-Ø-it
stanja	sta-Ø-ee-t	sta-Ø-ee-t	sta-Ø-ee-ti	sta-Ø-ee-	sta-Ø-Ø-it
gall	-a-ee-t	-a-ee-t	-a-ee-ti	-a-Ø-	-a-Ø-it
haddid	-aD-i-t	-aD-i-t	-aD-i-ti	-aD-i-	-aD-Ø-it
tesammim	te-aD-i-t	te-aD-i-t	te-aD-i-ti	te-aD-i-	te-aD-Ø-it
ingall	!none	!none	!none	in-a-Ø-	in-a-Ø-it
irtadd	!none	!none	!none	i-t-a-Ø-	i-t-a-Ø-it
stemarr	ste-a-ee-t	ste-a-ee-t	ste-a-ee-ti	ste-a-Ø-	ste-a-Ø-it
maal	-i-Ø-t	-i-Ø-t	-i-Ø-ti	-a:-	-a:-it
baat	-i-Ø-t	-i-Ø-t	-i-Ø-ti	-a:-	-a:-it
injaab	in-i-Ø-t	in-i-Ø-t	in-i-Ø-ti	in-a:-	in-a:-it
iltaam	i-t-u-Ø-t	i-t-u-Ø-t	i-t-u-Ø-ti	i-t-a:-	i-t-a:-it
stefaad	ste-a-Ø-t	ste-a-Ø-t	ste-a-Ø-ti	ste-a:-	ste-a:-it
CONJ	**perf1Pl**	**perf2mPl**	**perf2fPl**	**perf3mPl**	**perf3fPl**
getal	-e-a-na	-e-a-tu	-e-a-tin	-u-u-u	-u-u-in
kallim	-aD-i-na	-aD-i-tu	-aD-i-tin	-aD-Ø-u	-aD-Ø-in

saanid	-a:-i-na	-a:-i-tu	-a:-i-tin	-a:-Ø-u	-a:-Ø-in
tesallif	te-aD-i-na	te-aD-i-tu	te-aD-i-tin	te-aD-Ø-u	te-aD-Ø-in
tefaahim	te-a:-i-na	te-a:-i-tu	te-a:-i-tin	te-a:-Ø-u	te-a:-Ø-in
inkesar	in-e-a-na	!none	!none	in-u-u-u	in-u-u-in
intefax	i-t-e-a-na	!none	!none	i-t-u-u-u	i-t-u-u-in
stanjid	sta-Ø-i-na	sta-Ø-i-tu	sta-Ø-i-tin	sta-i-Ø-u	sta-i-Ø-in
nsee	-Ø-ee-na	-Ø-ee-tu	-Ø-ee-tin	-Ø-Ø-u	-Ø-Ø-in
bakka	-aD-ee-na	-aD-ee-tu	-aD-ee-tin	-aD-Ø-u	-aD-Ø-in
naada	-a:-ee-na	-a:-ee-tu	-a:-ee-tin	-a:-Ø-u	-a:-Ø-in
temanna	te-aD-ee-na	te-aD-ee-tu	te-aD-ee-tin	te-aD-Ø-u	te-aD-Ø-in
telaaga	te-a:-ee-na	te-a:-ee-na	te-a:-ee-tin	te-a:-Ø-u	te-a:-Ø-in
inkiwee	in-i-ee-na	!none	!none	in-i-Ø-u	in-i-Ø-in
intisee	i-t-i-ee-na	!none	!none	i-t-i-Ø-u	i-t-i-Ø-in
stanja	sta-Ø-ee-na	sta-Ø-ee-tu	sta-Ø-ee-tin	sta-Ø-Ø-u	sta-Ø-Ø-in
gall	-a-ee-na	-a-ee-tu	-a-ee-tin	-a-Ø-u	-a-Ø-in
haddid	-aD-i-na	-aD-i-tu	-aD-i-tin	-aD-Ø-u	-aD-Ø-in
tesammim	te-aD-i-na	te-aD-i-tu	te-aD-i-tin	te-aD-Ø-u	te-aD-Ø-in
ingall	in-a-ee-na	!none	!none	in-a-Ø-u	in-a-Ø-in
irtadd	i-t-a-ee-na	!none	!none	i-t-a-Ø-u	i-t-a-Ø-in
stemarr	ste-a-ee-na	ste-a-ee-tu	ste-a-ee-tin	ste-a-Ø-u	ste-a-Ø-in
maal	-i-Ø-na	-i-Ø-tu	-i-Ø-tin	-a:-u	-a:-in
baat	-i-Ø-na	-i-Ø-tu	-i-Ø-tin	-a:-u	-a:-in
injaab	in-i-Ø-na	in-i-Ø-tu	in-i-Ø-tin	in-a:-u	in-a:-in
iltaam	i-t-u-Ø-na	i-t-u-Ø-tu	i-t-u-Ø-tin	i-t-a:-u	i-t-a:-in
stefaad	ste-a-Ø-na	ste-a-Ø-tu	ste-a-Ø-tin	ste-a:-u	ste-a:-in
CONJ	**imperf1Sg**	**imperf2mSg**	**imperf2fSg**	**imperf3mSg**	**imperf3fSg**
getal	nu-Ø-i-	tu-Ø-i-	tu-u-Ø-i	yu-Ø-i-	tu-Ø-i-
kallim	n-aD-i-	t-aD-i-	t-aD-Ø-I	y-aD-i-	t-aD-i-
saanid	n-a:-i-	t-a:-i-	t-a:-Ø-i	y-a:-i-	t-a:-i-
tesallif	n-te-aD-i-	t-te-aD-i-	t-te-aD-Ø-i	y-te-aD-i-	t-te-aD-i-
tefaahim	n-te-a:-i-	t-te-a:-i-	t-te-a:-Ø-i	y-te-a:-i-	t-te-a:-i-
inkesar	!none	!none	!none	yun-u-u-	tun-u-u-
intefax	!none	!none	!none	yu-t-u-u-	tu-t-u-u-
stanjid	n-sta-Ø-i-	t-sta-Ø-i-	t-sta-i-Ø-i	y-sta-Ø-i-	t-sta-Ø-i-
nsee	na-Ø-a-	ta-Ø-a-	ta-Ø-Ø-i	ya-Ø-a-	ta-Ø-a-
bakka	n-aD-i-	t-aD-i-	t-aD-Ø-i	y-aD-i-	t-aD-i-
naada	n-a:-i-	t-a:-i-	t-a:-Ø-i	y-a:-i-	t-a:-i-
temanna	n-te-aD-a-	t-te-aD-a-	t-te-aD-Ø-i	y-te-aD-a-	t-te-aD-a-
telaaga	n-te-a:-a-	t-te-a:-a-	t-te-a:-Ø-i	y-te-a:-a-	t-te-a:-a-
inkiwee	!none	!none	!none	yin-i-i-	tin-i-i-
intisee	!none	!none	!none	yi-t-i-a-	ti-t-i-a-
stanja	n-sta-Ø-a-	t-sta-Ø-a-	t-sta-Ø-Ø-i	y-sta-Ø-a-	t-sta-Ø-a-
gall	n-i-Ø-	t-i-Ø-	t-i-Ø-i	y-i-Ø-	t-i-Ø-
haddid	n-aD-i-	t-aD-i-	t-aD-Ø-i	y-aD-i-	t-aD-i-
tesammim	n-te-aD-i-	t-te-aD-i-	t-te-aD-Ø-i	y-te-aD-i-	t-te-aD-i-
ingall	!none	!none	!none	yin-a-Ø-	tin-a-Ø-
irtadd	!none	!none	!none	yi-t-a-Ø-	ti-t-a-Ø-
stemarr	n-ste-i-Ø-	t-ste-i-Ø-	t-ste-i-Ø-i	y-ste-i-Ø-	t-ste-i-Ø-
maal	n-i:-	t-i:-	t-i:-i	y-i:-	t-i:-
baat	n-a:-	t-a:-	t-a:-i	y-a:-	t-a:-
injaab	nin-a:-	tin-a:-	tin-a:-i	yin-a:-	tin-a:-
iltaam	ni-t-a:-	ti-t-a:-	ti-t-a:-i	yi-t-a:-	ti-t-a:-
stefaad	n-ste-i:-	t-ste-a-i:-	t-ste-i:-i	y-ste-i:-	t-ste-i:-
CONJ	**imperf1Pl**	**imperf2mPl**	**imperf2fPl**	**imperf3mPl**	**imperf3fPl**
getal	nu-u-Ø-u	tu-u-Ø-u	tu-u-Ø-in	yu-u-Ø-u	yu-u-Ø-in
kallim	n-aD-Ø-u	t-aD-Ø-u	t-aD-Ø-in	y-aD-Ø-u	y-aD-Ø-in
saanid	n-a:-Ø-u	t-a:-Ø-u	t-a:-Ø-in	y-a:-Ø-u	y-a:-Ø-in
tesallif	n-te-aD-Ø-u	t-te-aD-Ø-u	t-te-aD-Ø-in	y-te-aD-Ø-u	y-te-aD-Ø-in

tefaahim	n-te-a:-Ø-u	t-te-a:-Ø-u	t-te-a:-Ø-in	y-te-a:-Ø-u	y-te-a:-Ø-in
inkesar	!none	!none	!none	yun-u-u-u	yun-u-u-in
intefax	!none	!none	!none	yu-t-u-u-u	yu-t-u-u-in
stanjid	n-sta-i-Ø-u	t-sta-i-Ø-u	t-sta-i-Ø-in	y-sta-i-Ø-u	y-sta-i-Ø-in
nsee	na-Ø-Ø-u	ta-Ø-Ø-u	ta-Ø-Ø-in	ya-Ø-Ø-u	ya-Ø-Ø-in
bakka	n-aD-Ø-u	t-aD-Ø-u	t-aD-Ø-in	y-aD-Ø-u	y-aD-Ø-in
naada	n-a:-Ø-u	t-a:-Ø-u	t-a:-Ø-in	y-a:-Ø-u	y-a:-Ø-in
temanna	n-te-aD-Ø-u	t-te-aD-Ø-u	t-te-aD-Ø-in	y-te-aD-Ø-u	y-te-aD-Ø-in
telaaga	n-te-a:-Ø-u	t-te-a:-Ø-u	t-te-a:-Ø-in	y-te-a:-Ø-u	t-te-a:-Ø-in
inkiwee	!none	!none	!none	yin-i-Ø-u	yin-i-Ø-in
intisee	!none	!none	!none	yi-t-i-Ø-u	yi-t-i-Ø-in
stanja	n-sta-Ø-Ø-u	t-sta-Ø-Ø-u	t-sta-Ø-Ø-in	y-sta-Ø-Ø-u	y-sta-Ø-Ø-in
gall	n-i-Ø-u	t-i-Ø-u	t-i-Ø-in	y-i-Ø-u	y-i-Ø-in
haddid	n-aD-Ø-u	t-aD-Ø-u	t-aD-Ø-in	y-aD-Ø-u	y-aD-Ø-in
tesammim	n-te-aD-Ø-u	t-te-aD-Ø-u	t-te-aD-Ø-in	y-te-aD-Ø-u	y-te-aD-Ø-in
ingall	!none	!none	!none	yin-a-Ø-u	yin-a-Ø-in
irtadd	!none	!none	!none	yi-t-a-Ø-u	yi-t-a-Ø-in
stemarr	n-ste-i-Ø-u	t-ste-i-Ø-u	t-ste-i-Ø-in	y-ste-i-Ø-u	y-ste-i-Ø-in
maal	n-i:-u	t-i:-u	t-i:-in	y-i:-u	y-i:-in
baat	n-a:-u	t-a:-u	t-a:-in	y-a:-u	y-a:-in
injaab	nin-a:-u	tin-a:-u	tin-a:-in	yin-a:-u	yin-a:-in
iltaam	ni-t-a:-u	ti-t-a:-u	ti-t-a:-in	yi-t-a:-u	yi-t-a:-in
stefaad	n-ste-i:-u	t-ste-a-i:-u	t-ste-i:-in	y-ste-i:-u	y-ste-i:-in

CONJ	**imper2mSg**	**imper2fSg**
getal	u-Ø-i-	u-u-Ø-i
kallim	-aD-i-	-aD-Ø-i
saanid	-a:-i-	-a:-Ø-i
tesallif	te-aD-i-	te-aD-Ø-i
tefaahim	te-a:-i-	te-a:-Ø-i
inkesar	!none	!none
intefax	!none	!none
stanjid	sta-Ø-i-	sta-i-Ø-i
nsee	a-Ø-a-	a-Ø-Ø-i
bakka	-aD-i-	-aD-Ø-i
naada	-a:-i-	-a:-Ø-i
temanna	te-aD-a-	te-aD-Ø-i
telaaga	te-a:-a-	te-a:-Ø-i
inkiwee	!none	!none
intisee	!none	!none
stanja	sta-Ø-a-	sta-Ø-Ø-i
gall	-i-Ø-	-i-Ø-i
haddid	-aD-i-	-aD-Ø-i
tesammim	te-aD-i-	te-aD-Ø-i
ingall	!none	!none
irtadd	!none	!none
stemarr	ste-i-Ø-	ste-i-Ø-i
maal	-i:-	-i:-i
baat	-a:-	-a:-i
injaab	!none	!none
iltaam	!none	!none
stefaad	ste-i:-	ste-i:-i

CONJ	**imper2mPl**	**imper2fPl**
getal	u-u-Ø-u	u-u-Ø-in
kallim	-aD-Ø-u	-aD-Ø-in
saanid	-a:-Ø-u	-a:-Ø-in
tesallif	te-aD-Ø-u	te-aD-Ø-in
tefaahim	te-a:-Ø-u	te-a:-Ø-in
inkesar	!none	!none

intefax	!none	!none
stanjid	sta-i-Ø-u	sta-i-Ø-in
nsee	a-Ø-Ø-u	a-Ø-Ø-in
bakka	-aD-Ø-u	-aD-Ø-in
naada	-a:-Ø-u	-a:-Ø-in
temanna	te-aD-Ø-u	te-aD-Ø-in
telaaga	te-a:-Ø-u	te-a:-Ø-in
inkiwee	!none	!none
intisee	!none	!none
stanja	sta-Ø-Ø-u	sta-Ø-Ø-in
gall	-i-Ø-u	-i-Ø-in
haddid	-aD-Ø-u	-aD-Ø-in
tesammim	te-aD-Ø-u	te-aD-Ø-in
ingall	!none	!none
irtadd	!none	!none
stemarr	ste-i-Ø-u	ste-i-Ø-in
maal	-i:-u	-i:-in
baat	-a:-u	t-a:-in
injaab	!none	!none
iltaam	!none	!none
stefaad	ste-i:-u	ste-i:-in

CLASS consonant b c d f g h j k l m n r s t w x

SANDHI D [:consonant:] ⇒ $1 $1 % doubling

The Sandhi doubling of the derived measures was suggested for the Arabic of Bukhara as a Semitic language.

B.1.2 The dynamic principal parts (exponent-based plat)

```
Distillation numbers
        1: perf1Sg
        2: perf3mSg
        3: perf3fSg
        4: imperf1Sg
        5: imperf2fSg
        6: imperf3mSg
        7: imperf3mPl
        8: imperf3fPl
        9: imper2mSg
        10: imper2fSg
```

```
          1      2      3      4      5      6      7      8      9      10
            perf1S perf3m perf3f imperf imperf imperf imperf imperf impera impera
 -------------------------------------------------------------------------------
       getal
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
          3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
          4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
          5| 5      5      5      5      5      5      5      5      5      5      avg 1.00
          6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
          7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
          8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
          9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
         10| 10     10     10     10     10     10     10     10     10     10     avg 1.00
     getal: 1 principal parts; lowest average 1.00; 10 analyses; 100.0% of 10 possible analyses
```

```
-----------------------------------------------------------------------------------
      kallim
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
   kallim: 1 principal parts; lowest average 1.00; 2 analyses; 20.0% of 10 possible analyses
-----------------------------------------------------------------------------------
      saanid
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
   saanid: 1 principal parts; lowest average 1.00; 2 analyses; 20.0% of 10 possible analyses
-----------------------------------------------------------------------------------



    tesallif
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
          4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
          6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
          9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
 tesallif: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
    tefaahim
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
          4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
          6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
          8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
          9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
 tefaahim: 1 principal parts; lowest average 1.00; 6 analyses; 60.0% of 10 possible analyses
-----------------------------------------------------------------------------------
     inkesar
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
          3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
          6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
          7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
          8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
  inkesar: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
     intefax
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
          3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
          6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
          7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
          8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
  intefax: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
     stanjid
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
          3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
          4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
          5| 5      5      5      5      5      5      5      5      5      5      avg 1.00
          6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
          7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
          8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
          9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
         10| 10     10     10     10     10     10     10     10     10     10     avg 1.00
  stanjid: 1 principal parts; lowest average 1.00; 10 analyses; 100.0% of 10 possible analyses
-----------------------------------------------------------------------------------
        nsee
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
          3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
          4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
          5| 5      5      5      5      5      5      5      5      5      5      avg 1.00
          6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
          7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
          8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
          9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
         10| 10     10     10     10     10     10     10     10     10     10     avg 1.00
     nsee: 1 principal parts; lowest average 1.00; 10 analyses; 100.0% of 10 possible analyses
-----------------------------------------------------------------------------------
       bakka
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
    bakka: 1 principal parts; lowest average 1.00; 2 analyses; 20.0% of 10 possible analyses
-----------------------------------------------------------------------------------
       naada
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
    naada: 1 principal parts; lowest average 1.00; 2 analyses; 20.0% of 10 possible analyses
-----------------------------------------------------------------------------------
     temanna
          1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
          2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
```

```
         4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
  temanna: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
    telaaga
         1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
  telaaga: 1 principal parts; lowest average 1.00; 6 analyses; 60.0% of 10 possible analyses
-----------------------------------------------------------------------------------


    inkiwee
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
  inkiwee: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
    intisee
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
  intisee: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
     stanja
         1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
         4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
        10| 10     10     10     10     10     10     10     10     10     10     avg 1.00
   stanja: 1 principal parts; lowest average 1.00; 10 analyses; 100.0% of 10 possible analyses
-----------------------------------------------------------------------------------
       gall
         1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
         4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
        10| 10     10     10     10     10     10     10     10     10     10     avg 1.00
     gall: 1 principal parts; lowest average 1.00; 10 analyses; 100.0% of 10 possible analyses
-----------------------------------------------------------------------------------
     ingall
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
   ingall: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
     irtadd
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
   irtadd: 1 principal parts; lowest average 1.00; 5 analyses; 50.0% of 10 possible analyses
-----------------------------------------------------------------------------------
    stemarr
         1| 1      1      1      1      1      1      1      1      1      1      avg 1.00
         2| 2      2      2      2      2      2      2      2      2      2      avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3      avg 1.00
         4| 4      4      4      4      4      4      4      4      4      4      avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5      avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6      avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7      avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8      avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9      avg 1.00
        10| 10     10     10     10     10     10     10     10     10     10     avg 1.00
  stemarr: 1 principal parts; lowest average 1.00; 10 analyses; 100.0% of 10 possible analyses
-----------------------------------------------------------------------------------
       maal
```

```
         4| 4      4      4      4      4      4      4      4      4      4       avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5       avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6       avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7       avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8       avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9       avg 1.00
        10| 10     10     10     10     10     10     10     10     10     10      avg 1.00
     maal: 1 principal parts; lowest average 1.00; 7 analyses; 70.0% of 10 possible analyses
  -----------------------------------------------------------------------------------
       baat
         4| 4      4      4      4      4      4      4      4      4      4       avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5       avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6       avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7       avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8       avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9       avg 1.00
        10| 10     10     10     10     10     10     10     10     10     10      avg 1.00
     baat: 1 principal parts; lowest average 1.00; 7 analyses; 70.0% of 10 possible analyses
  -----------------------------------------------------------------------------------
     injaab
         1| 1      1      1      1      1      1      1      1      1      1       avg 1.00
         2| 2      2      2      2      2      2      2      2      2      2       avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3       avg 1.00
         4| 4      4      4      4      4      4      4      4      4      4       avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5       avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6       avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7       avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8       avg 1.00
   injaab: 1 principal parts; lowest average 1.00; 8 analyses; 80.0% of 10 possible analyses
  -----------------------------------------------------------------------------------
     iltaam
         1| 1      1      1      1      1      1      1      1      1      1       avg 1.00
         2| 2      2      2      2      2      2      2      2      2      2       avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3       avg 1.00
         4| 4      4      4      4      4      4      4      4      4      4       avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5       avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6       avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7       avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8       avg 1.00
   iltaam: 1 principal parts; lowest average 1.00; 8 analyses; 80.0% of 10 possible analyses
  -----------------------------------------------------------------------------------
    stefaad
         1| 1      1      1      1      1      1      1      1      1      1       avg 1.00
         2| 2      2      2      2      2      2      2      2      2      2       avg 1.00
         3| 3      3      3      3      3      3      3      3      3      3       avg 1.00
         4| 4      4      4      4      4      4      4      4      4      4       avg 1.00
         5| 5      5      5      5      5      5      5      5      5      5       avg 1.00
         6| 6      6      6      6      6      6      6      6      6      6       avg 1.00
         7| 7      7      7      7      7      7      7      7      7      7       avg 1.00
         8| 8      8      8      8      8      8      8      8      8      8       avg 1.00
         9| 9      9      9      9      9      9      9      9      9      9       avg 1.00
        10| 10     10     10     10     10     10     10     10     10     10      avg 1.00
  stefaad: 1 principal parts; lowest average 1.00; 10 analyses; 100.0% of 10 possible analyses
Average number of principal parts: 1.00
Average lowest average of principal parts needed: 1.00
Average number of alternative analyses: 6.40
Average ratio of actual to possible: 64.00%
```

B.2 TLA stem-based plat

B.2.1 The input formatting for the TLA stem-based plat

(1) The PPA input chart for generating stem formative-plat.

CONJ	Gloss	Stem1	stem2	stem3	stem4	stem5
getal	kill	ugtil	ugutl	getal	getal	gutul
kallim	talk	kallim	kallm	kallim	kallim	kallm
saanid	support	saanid	saand	saanid	saanid	saand
tesallif	borrow	tesallif	tesallf	tesallif	tesallif	tesallf

tefaahim	understand	tefaahim	tefaahm	tefaahim	tefaahim	tefaahm
insemaʕ	heard	insimaʕ	insimiʕ	insemaʕ	insemaʕ	insimiʕ
irtefaʕ	carried	irtifaʕ	irtifiʕ	irtefaʕ	irtefaʕ	irtifiʕ
stanjid	appeal	stanjid	stanjd	stanjid	stanjid	stanjd
nsee	forget	ansa	ans	nsee	nsee	ns
bakka	upset	bakki	bakk	bakkee	bakka	bakk
naada	invite	naadi	naad	naadee	naada	naad
temanna	wish	temanna	temann	temannee	temanna	temann
telaaga	meet	telaaga	telaag	telaagee	telaaga	telaag
inkiwee	cauterized	inkiwi	inkiw	inkiwee	inkiwee	inkiw
intsee	forgotten	intisa	intis	intisee	intisee	intis
stanja	purify	stanja	stanj	stanjee	stanja	stanj
gall	take	gill	gill	gallee	gall	gall
haddid	threat	haddid	haddid	haddid	haddid	haddid
tesammim	poisoned	tesammim	tesamim	tesammim	tesammim	tesamim
ingall	taken	ingall	ingall	ingallee	ingall	ingall
irtadd	rejected	irtadd	irtadd	irtaddee	irtadd	irtadd
stemarr	continue	stemirr	stemirr	stemarree	stemarr	stemarr
maal	lean	miil	miil	mil	maal	maal
baat	stay	baat	baat	bit	baat	baat
injaab	brought	injaab	injaab	injib	injaab	injaab
iltaam	blamed	iltaam	iltaam	iltum	iltaam	iltaam
stefaad	benefit	stefiid	stefiid	stefad	stefaad	stefaad

(2) TLA output plat based on Stem formatives

IC	Gloss	stem1	stem2	stem3	stem4	stem5
getal	kill	u-til	u-utl	-etal	-etal	-utul
kallim	talk	-im	-m	-im	-im	-m
saanid	support	-id	-d	-id	-id	-d
tesallif	borrow	-if	-f	-if	-if	-f
insemaʕ	heard	-imaʕ	-imiʕ	-emaʕ	-emaʕ	-imiʕ
irtefaʕ	carried	-ifaʕ	-ifiʕ	-efaʕ	-efaʕ	-ifiʕ
nsee	forget	a-a	a-	-ee	-ee	-
bakka	upset	-i	-	-ee	-a	-
temmanna	wish	-a	-	-ee	-a	-
inkiwee	cauterized	-i	-	-ee	-ee	-
intisee	forgotten	-a	-	-ee	-ee	-
gall	take	-ill	-ill	-allee	-all	-all
haddid	threat	-	-	-	-	-
tesammim	poisoned	-mim	-im	-mim	-mim	-im
ingall	taken	-	-	-ee	-	-
stemarr	continue	-irr	-irr	-arree	-arr	-arr
maal	lean	-iil	-iil	-il	-aal	-aal
baat	stay	-aat	-aat	-it	-aat	-aat
injaab	brought	-aab	-aab	-ib	-aab	-aab
iltaam	blamed	-aam	-aam	-um	-aam	-aam

stefaad	benefit	-iid	-iid	-ad	-aad	-aad

(3) The theme of TLA conjugations (the highlighted lexemes are ICs with identical conjugation pattern. The rightmost column represents the theme).

LEXEME	kill	kill	1:g
LEXEME	talk	talk	1:kall
LEXEME	support	support	1:saan
LEXEME	borrow	borrow	1:tesall
LEXEME	understand	talk	1:tefaah
LEXEME	heard	heard	1:ins
LEXEME	carried	carried	1:irt
LEXEME	appeal	support	1:stanj
LEXEME	forget	forget	1:ns
LEXEME	upset	upset	1:bakk
LEXEME	invite	upset	1:naad
LEXEME	wish	wish	1:temann
LEXEME	meet	wish	1:telaag
LEXEME	cauterized	cauterized	1:inkiw
LEXEME	forgotten	forgotten	1:intis
LEXEME	purify	wish	1:stanj
LEXEME	take	take	1:g
LEXEME	threat	threat	1:haddid
LEXEME	poisoned	poisoned	1:tesam
LEXEME	taken	taken	1:ingall
LEXEME	rejected	taken	1:irtadd
LEXEME	continue	continue	1:stem
LEXEME	lean	lean	1:m
LEXEME	stay	stay	1:b
LEXEME	brought	brought	1:inj
LEXEME	blamed	blamed	1:ilt
LEXEME	benefit	benefit	1:stef

B.2.2 The dynamic principal parts (stem-based plats)

```
Distillation numbers
        1: stem1
        2: stem2
        3: stem3
        4: stem4
        5: stem5
                1      2      3      4      5
                stem1  stem2  stem3  stem4  stem5
 --------------------------------------------------
           kill
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
```

```
      kill: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------

           talk
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
      talk: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
        support
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
   support: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
         borrow
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
    borrow: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
          heard
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
     heard: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
        carried
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
   carried: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
         forget
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
    forget: 1 principal parts; lowest average 1.00; 2 analyses;
40.0% of 5 possible analyses
 --------------------------------------------------
          upset
           1,4| 1      1      1      4      1      avg 1.00
     upset: 2 principal parts; lowest average 1.00; 1 analyses;
10.0% of 10 possible analyses
 --------------------------------------------------
           wish
```

```
           1,4| 1      1      1      4      1      avg 1.00
      wish: 2 principal parts; lowest average 1.00; 1 analyses;
10.0% of 10 possible analyses
 --------------------------------------------------
     cauterized
           1,4| 1      1      1      4      1      avg 1.00
cauterized: 2 principal parts; lowest average 1.00; 1 analyses;
10.0% of 10 possible analyses
 --------------------------------------------------
      forgotten
           1,4| 1      1      1      4      1      avg 1.00
 forgotten: 2 principal parts; lowest average 1.00; 1 analyses;
10.0% of 10 possible analyses
 --------------------------------------------------
           take
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
      take: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
         threat
             3| 3      3      3      3      3      avg 1.00
    threat: 1 principal parts; lowest average 1.00; 1 analyses;
20.0% of 5 possible analyses
 --------------------------------------------------
       poisoned
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
  poisoned: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
          taken
           1,3| 1      1      3      1      1      avg 1.00
           3,4| 4      4      3      4      3      avg 1.00
     taken: 2 principal parts; lowest average 1.00; 2 analyses;
20.0% of 10 possible analyses
 --------------------------------------------------
       continue
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
  continue: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
           lean
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
      lean: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
```

```
            stay
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
      stay: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
        brought
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
   brought: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
         blamed
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
    blamed: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
 --------------------------------------------------
        benefit
             1| 1      1      1      1      1      avg 1.00
             2| 2      2      2      2      2      avg 1.00
             3| 3      3      3      3      3      avg 1.00
             4| 4      4      4      4      4      avg 1.00
             5| 5      5      5      5      5      avg 1.00
   benefit: 1 principal parts; lowest average 1.00; 5 analyses;
100.0% of 5 possible analyses
Average number of principal parts: 1.24
Average lowest average of principal parts needed: 1.00
Average number of alternative analyses: 3.76
Average ratio of actual to possible: 72.38
```

B.2.3 IC and cell predictability formulas

(1) IC predictability formula (Stump & Finkel, 2013, p.95 and p.99)

$$\mathrm{ICP}_{\mathrm{L}} = \frac{|_{m}[D_{\mathrm{L}}{}'] \,|}{|_{m}[P(D_{\mathrm{L}}) \setminus \varnothing] \,|}$$

(2) Cell predictability formula (Stump & Finkel, 2013, p.95 and p.99)

$$\mathrm{C{\scriptsize ELL}P}_{\langle w,\sigma\rangle} = \frac{|\, [_{m}[D_{\langle w,\sigma\rangle}]]_{-\langle w,\sigma\rangle} \,|}{|\, [_{m}[\mathrm{P}(D_{\mathrm{L}}) \setminus \varnothing]]_{-\langle w,\sigma\rangle} \,|}$$

P_L is the realized paradigm of lexeme L
M_L is the set of cells in P_L
D_L is any maximal subset of M_L none of whose members belong to the same distillation
D_L' is the set $\{N: N \subseteq D_L$ and N is a viable dynamic principal-part set for $P_L\}$
For any collection C is sets, $_m[C]$ represents $\{s \in C: |s| \leq m\}$
For any set S, P(S) is the power set of S (=the set of subsets of S)
For any sets S_1, S_2, $S_1 \backslash S_2$ is $\{x : x \in S_1$ and $x \notin S_2\}$

(Stump & Finkel, 2013, p.95 and p.99)

B.2.4 IC predictability

(1) The plat of TLA ICs based on verb exponents (affix-based segmentation)

IC	Example	ρ	υ
I	getal	a	q
II	ṣebaġ	b	q
III	ketab	c	q
IV	gaal	d	q
V	radd	d	r
VI	ns	e	s
VII	mš	f	s
VIII	rabb	g	t
IX	stanj	h	t

(2) IC predictability (TLA plat based on the outer layer of the inflected word)

```
Predictabilities
          ρ     υ     | Avg   IC predictability
      I 0.000 0.500 | 0.250 0.667
     II 0.000 0.500 | 0.250 0.667
    III 0.000 0.500 | 0.250 0.667
     IV 0.000 0.000 | 0.000 0.333
      V 0.500 0.000 | 0.250 0.667
     VI 0.000 0.500 | 0.250 0.667
    VII 0.000 0.500 | 0.250 0.667
   VIII 0.000 0.500 | 0.250 0.667
     IX 0.000 0.500 | 0.250 0.667
    Avg 0.056 0.389 | 0.222 0.630
```

(3) TLA plat based on the stem space

IC	ρ	σ	τ	υ	φ
I	A	A	A	A	A
II	A	A	B	A	A
III	A	A	B	B	B
IV	A	B	C	C	B
V	A	A	B	C	C
VI	A	B	C	C	D
stem space	S1	S2	S3	S4	S5

(4) IC predictability (TLA plat based on the stem space)

```
Predictability (based on m=4)

Predictabilities
          1     2     3     4     5   |  Avg  IC predictability
      I 1.000 0.875 0.000 0.750 0.750 | 0.675 0.500
     II 1.000 0.875 0.000 0.500 0.500 | 0.575 0.367
    III 1.000 0.750 0.625 0.375 0.500 | 0.650 0.700
     IV 1.000 0.625 0.625 0.750 0.000 | 0.600 0.433
      V 1.000 0.750 0.625 0.500 0.375 | 0.650 0.700
     VI 1.000 0.750 0.750 0.875 0.000 | 0.675 0.500
    Avg 1.000 0.771 0.438 0.625 0.354 | 0.637 0.533
```

Bibliography

Abumdas, A. (1985). *Libyan Arabic Phonology* (Doctoral dissertation). University of Michigan, Ann Arbor.

Abu-Haidar, F. (1985). Review of Jonathan Owens 'A short reference grammar of Eastern Libyan Arabic'. *Journal of the Royal Asiatic Society of Great Britain & Ireland, 117*, 192-193. doi:10.1017/S0035869X00138432.

Ackerman, F. Blevins, J., & Malouf, R. (2009). Parts and Wholes: Implicative Patterns in Inflectional Paradigm. In J. P. Blevins & J. Blevins (Eds), *Analogy in Grammar: Form and Acquisition*, (pp.54-82). Oxford, England: Oxford University Press.

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low entropy conjecture. *Language 89 (3)*, 429-64.

Ahmed, A. (2008). *Production and perception of Libyan Arabic vowels* (Doctoral dissertation). Newcastle University.

Al-Ageli, H. (1995). *Syllabic and metrical structure in Tripolitanian Arabic: A comparative study in standard and optimality theory* (Doctoral dissertation). University of Essex, Colchester.

Alderete, J., Tupper, P., Frisch, S., & Fraser, S. (2013), *Arabic Root Co-occurrence Restrictions Revisited: A Connectionist Approach to Phonological Constraint Induction*, University of South Florida. Retrieved from http://www.sfu.ca/~alderete/pubs/aldereteEtal2012_constraintInductionArabic.pdf.

Al-Najem, S. R. (2007). Inheritance-based Approach to Arabic Verbal Root-and-Pattern Morphology. In A., Soudi, A., Bosch, & N., Gunter (Eds.), *Arabic Computational Morphology: Knowledge-Based and Empirical Methods: Text, Speech and Language Technology* (Vol.38, pp.67-88). Netherlands: Springer.

Anderson, S. R. (1992). *A-morphous morphology*. Cambridge: CUP.

Anderson, S. R. (2008). Phonologically Conditioned Allomorphy in Surmiran (Rumantsch). *Word Structure, 1*, 109-34.

Anderson, S. R. (2011). Stress-conditioned Allomorphy in Surmiran (Rumantsch). M., Goldbach, M-O., Hinzeln, M., Maiden, & J. C., Smith (Eds), *Morphological Autonomy: Perspectives from Romance Inflectional Morphology* (pp.13-35). Oxford: Oxford University Press.

Anderson, S. R. (2013). Stem alternation in Swiss Rumantsch. In S., Cruschina, M., Maiden & J. C., Smith (Eds.), *The Boundaries of Pure Morphology: Diachronic and synchronic perspectives* (pp.8-23). Oxford: OUP.

Arkadiev P. M. (2012). Stems in Lithuanian verbal inflection (with remarks on derivation). *Word Structure, 5*(1), 7-27.

Aronoff, M. (1994). *Morphology By Itself*. Cambridge, MA: MIT Press.

Baerman, M. (2007). Morphological Typology of Deponency. In Baerman M., Corbett G., Brown D. & Hippisley A. (Eds.), *Deponency and Morphological Mismatches* (pp.1–19). Oxford, England: Oxford University Press

Baerman, M., & Corbett, G. (2012). Stem alternations and multiple exponence. *Word Structure, 5*(1), 52-68.

Bahloul, M. (2008). *Structure and function of the Arabic verb*. London: Routledge.

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory 12*, 571–596.

Benmamoun, E. (1992). *Functional and inflectional morphology: Problems of projection, representation and derivation* (Doctoral dissertation). University of Southern California, Los Angeles.

Benmamoun, E. (1999). Arabic Morphology: The central role of the imperfective. *Lingua*, *108*, 175-201.

Benmamoun, E. (2003). The role of the imperfective template in Arabic morphology. In J. Shimron (Ed.), *Language Processing and Acquisition in Languages of Semitic, Root-Based, Morphology* (pp. 99–114). Amsterdam: John Benjamins.

Blevins, J. P. (2001). Paradigmatic derivation. *Transactions of the Philological Society, 99*(2), 211-222.

Blevins, J. P. (2003). Stems and paradigms. *Language*, *79*, 737-767.

Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics, 42,* 531-573

Blevins, J. P. (2015). Inflectional paradigms. In M. Baerman (Ed.), *The Oxford Handbook of Inflection.* (pp. 87–111). Oxford, England: Oxford University Press.

Bloch, B. (1947). English verb inflection. *Language 23* (4), 399-418.

Bloomfield, L. (1926). A Set of Postulates for the Science of Language. *Language, 2*(3), 153-164.

Bloomfield, L. (1935). *Language*. London: Allen and Unwin.

Bohas, G., & Guillaume, J.-P. (1984). *Etudes des The´ories des grammairiens Arabes: I Morphologie et Phonologie*. Damascus, Syria.

Bonami, O., & Boyé G. (2002). Suppletion and dependency in inflectional morphology. In Van Eynde F., Hellan, L., & Beerman D. (Eds), *Proceedings of the HPSG '01 Conference* (pp. 51-70). Stanford: CSLI Publications.

Bonami, O., & Boyé G. (2006). Deriving inflectional irregularity. In S. Müller (Ed), *Proceedings of the HPSG '01 Conference* (pp.361-380). Stanford: CSLI Publications.

Bonami, O., & Boyé G. (2008). Paradigm shape is morphomic in Nepali. Paper presented at the *IMM 13*, Vienna.

Bonami, O. (2012). Discovering implicative morphology. *Paper presented at the Les Décembrettes 8*, Bordeaux. Retrieved from http://www.llf.cnrs.fr/sites/llf.cnrs.fr/files/biblio/Bordeaux-dec2012.pdf.

Borg, A. (1997). Cypriot Arabic Phonology. In A. S. Kaye (Ed.), *Phonologies of Asia and Africa (Including the Caucasus)*, (2, pp. 219-244). Winona Lake, Indiana: Eisenbrauns.

Boudelaa, S., & Marslen-Wilson, W. D. (2001). Morphological units in the Arabic mental lexicon. *Cognition*, *81*(1), 65-92.

Boyé, G., & Cabredo-Hofherr, P. (2006). The structure of allomorphy in Spanish verbal inflection. *Cuadernos de Linguistica del Istituto Universitario Ortega y Gasset, 13*, 9-24.

Boyé, G., & Hofherr, P.C. (2008). Defectivity as stem suppletion in French and Spanish verbs. Retrieved from http://www.smg.surrey.ac.uk/media/1301/boye-cabredo-hofherr-defective-paradigms.pdf.

Buckwalter, T. (1997). *The triliteral and quadriliteral roots of Arabic*. Retrieved from http://www.angelfire.com/tx4/lisan/roots1.htm.

Brame, K. (1974). The cycle in phonology: Stress in Palestinian, Maltese and Spanish. *Linguistic Inquiry* 5. 39–60.

Broselow, E. (1976). *The Phonology of Egyptian Arabic* (Doctoral dissertation). University of Massachusetts.

Brown, D. P., & Evans, R. (2010). Inflectional defaults and principal parts: an empirical investigation. In S. Müller (Ed.), *Proceedings of the HPSG10 Conference, Université Paris Diderot, Paris 7, France.* (pp.234-254). Stanford: CSLI Publications.

Cahill, L., & Gazdar, G. (1999a). German noun inflection. *Journal of Linguistics, 35*, 1–42.

Cahill, L., & Gazdar, G. (1999b). The PolyLex architecture: multilingual lexicons for related languages. *Traitement Automatique des Langues*, *40*(2), 5–23.

Camilleri, M. (2014). *The stem in inflectional verbal paradigms in Maltese* (Doctoral dissertation). University of Surrey, Surrey.

Cantineau, J. (1939). Remarques sur les parlers de sédentaires syro-libano-palesintinens. *Bulletin de la Société de Linguistique de Paris*, *40*. 80–88.

Cantineau, J. (1950). *Racines et Schèmes, Mélanges William Marçais* (pp. 119–124). Paris: Maisonneuve.

Carstairs, A. (1988). Some implications of phonologically conditioned suppletion. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology* (pp.67–94). Dordrecht: Kluwer Academic Publishers.

Carstairs, A. (1990). Phonologically conditioned suppletion. In W.U. Dressler, H.C. Luschützky, O.E. Pfeiffer & J. R. Rennison (Eds.), *Contemporary morphology* (pp.17–23). Berlin - New York: Mouton de Gruyter.

Cavalli-Sforza, V., Soudi, A., & Mitamura, T. (2000). Arabic morphology generation using a concatenative strategy. *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference* (pp. 86–93). Retrieved from http://dl.acm.org/citation.cfm?id=974317.

Chekayri, A., & Schee, T. (2003). The appearance of Glides in Classical Arabic defective verbs. *Folia Orientalia*, *3*, 5-34. Retrieved from http://www.phil.muni.cz/linguistica/art/scheer/Chek-ScheerAppGlides.pdf.

Chomsky, N. & Halle, M. (1968). *The sound patterns of English*. New York: Harper & Row

Corbett, G., & M. Fraser. (1993). Network Morphology: A DATR Account of Russian Nominal Inflection. *Journal of Linguistics*, *29*(1), 113–142. Retrieved from http://www.jstor.org/stable/4176210.

Corbett, G. & Mithun M. (1996). Associative formatives in a typology of a number of systems: Evidence from Yup'ik. *Journal of Linguistics*, *32*, 1-7. Retrieved from http://www.surrey.ac.uk/LIS/SMG/GREVS%20PUBLICATIONS/Associative%20forms%20in%20a%20typology%20of%20number%20systems.pdf.

Corbett, G., & Baerman M. (2006). Prolegomena to a typology of morphological features. *Morphology*, *16*, 231-246.

Corbett, G. (2007). Canonical typology, suppletion and possible words. *Language*, *83*, 8-42. Retrieved from http://www.surrey.ac.uk/LIS/SMG/suppletion-Lg-PREPRINT.pdf.

Corbett, G. (2009). Canonical Inflectional Classes. In: F. Montermini, G. Boyé, J. Tseng (Eds.), *Selected proceedings of the 6th Decembrettes* (pp. 1-11). Somerville: Cascadilla Proceedings Projects.

Corbett, G. (2015). Morphomic splits / Morphosyntactic complexity: a typology of lexical splits. *Language 91*(1):145-193. doi: 10.1353/lan.2015.0003.

Crysmann, B., & Bonami, O. (2012). Establishing Order in Type-based Realisational Morphology. In S. Muller (Ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar*, *Chungnam National University Daejeon* (pp. 123–143). Stanford, CA: CSLI Publications. Retrieved fromhttp://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2012/crysmann-bonami.pdf

Cuvalay-Haak, M. (1997). *The Verb in Literary Colloquial Arabic*. Mouton de Gruyter: Berlin/ New York.

Darden, B. (1992). The Cairene verb without form classes. In D. Bentari, G. N. Larson, & L. A Macleod (Eds), *The Joy of Grammar: A Festschrift in honor of James D. McCawley* (pp. 11-24). John Benjamins: Amsterdam/Philadelphia.

Elfitoury, A. (1976). *A Descriptive Grammar of Libyan Arabic* (Doctoral dissertation). University of Georgetown.

Elgadi, A. (1986) *Tripolitanian Arabic phonology and morphology: a generative approach* (Doctoral dissertation). Georgetown University.

Er-Rayyan, M.R. (1986) *Toward the construction of a temporal system for natural language in the light of the Arabic and English languages* (Doctoral dissertation). University of Pennsylvania.

Finkel, R., & Stump, G. (2007). Generating Hebrew verb morphology by default inheritance hierarchies. *In Proceedings of the Workshop on Computational Approaches to Semitic Languages* (pp.9–18). Retrieved from http://www.aclweb.org/anthology/W02-0502

Frisch, A., Pierrehumbert J., & Michael B. (2004). Similarity Avoidance and the OCP. *Natural Language and Linguistic Theory*, *22*, 179-228.

Gafos, A. (2003). Greenberg's asymmetry in Arabic: A consequence of stems in paradigms. *Language*, *79*, 317–355.

Goldsmith, J. (1976). *Autosegmental phonology* (Doctoral dissertation. MIT, York: Garland.

Greenberg, H. (1950). The Patterning of Root Morphemes in Semitic. *Word*, *5*, 162–181.

Griffini, E. (1913). *L'Arabo parlato della Libia*. Milan Librarie dell Real Casa.

Guerssel, M., & Lowenstamm, J. (1996). Ablaut in Classical Arabic measure I active verbal forms. In J. Lecarme, J. Lowenstamm, & Shlonsky, U. (Eds.), *Studies in Afroasiatic Grammar* (pp. 123-134). The Hague: Holland Academic Graphics.

Harrama, A. (1993). *Libyan Arabic morphology: Al-Jabal dialect* (Doctoral dissertation). The University of Arizona, Arizona.

Harris, Z. (1941). Linguistic structure of Hebrew. *Journal of the American Oriental Society, 61*, 143-167.

Harris, Z. (1942). Morpheme alternants in linguistics analysis. *Language 21*(3), 121-127.

Heath, J. (1987). *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. New York, Albany: State University of New York Press.

Heath, J. (1989). *From Code-Switching to Borrowing: foreign and diglossic mixing Moroccan Arabic.* London: Kegan Paul International.

Heath, J. (1997). Moroccan Arabic phonology. In A. S. Kaye & P. T. Deniels (Eds.), *Phonologies of Asia and Africa Inc Caucasus,* (vol. I, pp. 205–217). Indiana: Winona Lake.

Heath, J. (2003). Arabic derivational ablaut, processing strategies, and consonantal roots. In J. Shimron (Ed.), *Language processing and acquisition in languages of Semitic root-based, morphology,* (pp.100–115). Amsterdam: John Benjamins.

Hoberman, R. D., & Aronoff, M. (2003). The verbal morphology of Maltese: From Semitic to Romance. In J. Shimron (Ed.), *Language Processing and Language Acquisition in Languages of Semitic, Root-Based Morphology* (pp. 61–78). Amsterdam/Philadelphia: John Benjamins, *62*, 143–167.

Hockett, C. F. (1947). Problems of morphemic analysis. *Language 23* (4), 321-43.

Hudson, G. (1986). Arabic Root and Pattern Morphology without Tiers. *Journal of Linguistics*, *22*, 85-122.

Kenstowicz, M. (1980). Notes on Cairene Arabic Syncope. *Studies in the Linguistic Sciences*, *10*, 39–53.

Kiparsky, P. (2003). Syllables and moras in Arabic. In Caroline Fe´ry & Ruben van de Vijver (eds.), *The syllable in Optimality Theory*. Cambridge: Cambridge University Press, 147–182.

Kirchner, J., S. (2009). Ablaut and reduplication in Dakota: Revisiting the phonology-morphology relationship. In D. Rosenblum & C. Meeker (Eds.), *Santa Barbara Papers in Linguistics, Vol. 20: Proceedings from the 12th Annual Workshop on American Indigenous Language* (pp.1-15). Retrieved from http://www.linguistics.ucsb.edu/sites/secure.lsit.ucsb.edu.ling.d7/files/sitefiles/research/papers/Working%20Papers%20Vol%2020.pdf.

Maiden, M. (1992). Irregularity as a determinant of morphological change. *Journal of Linguistics*, *28*, 285-312.

Maiden, M. (2005). Morphological autonomy and diachrony. In by G Booji, J. van Marle (Eds.), *Yearbook of Morphology 2004* (pp.137-175). Dordrecht: Springer.

Maiden, M., (2009). From pure phonology to pure morphology the reshaping of the romance verb. *Recherches Linguistiques de Vincennes, 38*, 45-82.

Maiden, M., (2013). 'Semi-autonomous' morphology? A problem in the history of the Italian (and Romanian) verb. In S. Cruschina, M. Maiden, and J. C. Smith (Eds.), *The Boundaries of Pure Morphology: Diachronic and Synchronic Perspectives* (pp. 24-44). Oxford, England: Oxford University Press.

Maiden, M. (2011). Morphophonological persistence. In A. Ledgeway, M. Maiden, J. C. Smith (Eds), *The Cambridge History of the Romance Languages Cambridge*: CUP, 155-215.

Maiteg, T. (2013). *Prosodic constituent structure and anticipatory pharyngealisation in Libyan Arabic* (Doctoral dissertation). The University of Edinburgh, Edinburgh.

Matthews P. (1972). Inflectional morphology: a theoretical study based on aspects of Latin verb conjugation, number 6 in Cambridge Studies in Linguistics CUP, Cambridge.

Matthews, P. (1991). Morphology. Cambridge, England: Cambridge University Press.

Matthews, P. (1993). Grammatical Theory in the United States. CUP, Cambridge

McCarthy, J. (1979). *Formal problems in Semitic phonology and morphology* (Doctoral dissertation). University of Massachusetts, MIT.

McCarthy, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, *12*, 373-418.

McCarthy, J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry*, *17*, 207-263.

McCarthy, J. (1994). Non-Concatenative Morphology. *The Encyclopaedia of Language and Linguistics*, (Vol. 5, pp.2598-2600), Oxford: Pergamon Press.

McCarthy, J. (2000). The prosody of phase in Rotuman. *Natural Language and Linguistic Theory, 19*, 147-97

McCarthy, J. (2005). Optimal Paradigms. In L. Downing, T. A. Hall, & R. Raffelsiefen (Eds.), *Paradigms in Phonological Theory* (pp. 170-210). Oxford: OUP.

McCarthy, J. (2006). Prosodic Morphology. In K. Brown (Ed.), *Encyclopaedia of Language and Linguistics* (pp. 182-190). Oxford: Elsevier.

McCarthy, J., & Prince, A. S. (1986). *Prosodic Morphology*. MS., University of Massachusetts and Brandeis University.

McCarthy, J., & Prince, A. S. (1990a). Prosodic morphology and templatic morphology. In M. Eid & J.J. McCarthy (Eds.), *Perspectives on Arabic Linguistics, 2*, 1–54, Amsterdam: Benjamins.

McCarthy, J., & Prince, A. S. (1990b). Foot and word in prosodic morphology: the Arabic broken plural' *Natural Language and Linguistic Theory, 8*(2), 209–283.

McCarthy, J., & Prince, A. S. (1993). Prosodic Morphology I: Constraint interaction and satisfaction Rutgers Center for Cognitive Science, Technical Report No. 3

McOmber, M. L. (1995). Morpheme edges and Arabic infixation. In M. Eid (Ed.), *Perspectives on Arabic Linguistics*, (VII pp.173-189). Amsterdam/Philadelphia: John Benjamins.

Miller, K. (2014). The Arabic of Bukhara, A Principal parts analysis of the effects of contact influence on morphological typology. In R. Khamis-Dakwar & K. Froud (Eds.), *Perspectives on Arabic Linguistics*, (XXVI pp.213-240) Amsterdam/Philadelphia: John Benjamins.

Mitchell, T.F. (1952). The Active Participle in an Arabic Dialect of Cyrenaica. *Bulletin of the School of Oriental and African Studies*, *14*, 11-33.

Mitchell, T. F. (1960). Prominence and syllabification in Arabic. *Bulletin of the School of Oriental and African Studies*, *23*, 396-389.

Montermini, F., & Bonami, O. (2011). Stem spaces and predictability in verbal inflection. Ms. University of Toulouse and Paris Sorbonne. Retrieved from http://www.llf.cnrs.fr/sites/llf.cnrs.fr/files/biblio/MonterminiBonami2013.pdf

Montermini, F., & Boyé, G. (2012). Stem relations and inflectional class assignment in Italian. *Word Structure*, *5*, 69-87.

Morin, Y.-C. (1987). French data and phonological theory. *Linguistics 25*(5):815–843.

Owens, J. (1980). The syllable as prosody: a reanalysis of syllabification of Eastern Libyan Arabic. *Bulletin of The School of Oriental and African Studies*, *43*, 277-287.

Owens, J. (1983). Libyan Arabic Dialects. Orbis 23, 97-117.

Owens, J. (1984). *A Short Reference Grammar of Eastern Libyan Arabic*, Wiesbaden: Otto Harrassowitz.

Owens, J. (1993). A Grammar of Nigerian Arabic. Wiesbaden: Harrassowitz

Owens, J. (2006). *A Linguistic History of Arabic*. Oxford: Oxford University Press.

Panetta, E. (1943). *L'arabo Parlato in Bengazi*, Rome.

Pirelli, V., & Battista, M. (2000). The paradigmatic dimension of stem allomorphy in Italian verb inflection. *Rivista di Linguistica*, *12*(2), 307-379.

Ratcliffe, R. (1997). Prosodic templates in a word-based morphological analysis. In M. Eid, & Ratcliffe, R. (Eds.), *Perspectives in Arabic Linguistics X* (pp. 147-171). Amsterdam/Philadelphia: John Benjamins.

Ratcliffe, R. (1998). *The "broken" plural in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. John Benjamins: Amsterdam

Ratcliffe, R. (2003). Toward a universal theory of shape-invariant (templatic) morphology: Classical Arabic re-considered. In R. Singh & S. Starosta (Eds.), *Explorations in seamless morphology* (pp.212–269). London: Sage Publications.

Ravid, D. (2006). Word-level morphology: A psycholinguistic perspective on linear formation in Hebrew nominals. *Morphology*, *16*, 127-148.

Shimron, J (2003). *Language processing and acquisition in languages of Semitic, root-based morphology* (pp. 99-114). Amsterdam/Philadelphia: John Benjamins.

Soudi, A., Bosch, & N., Gunter (Eds.), *Arabic Computational Morphology: Knowledge-Based and Empirical Methods: Text, Speech and Language Technology* (Vol.38). Netherlands: Springer

Spencer, A. (1991). *Morphological Theory: An introduction to word structure in generative grammar*. Oxford: Wiley-Blackwell.

Spencer, A. (2012). Identifying Stems. *Word Structure*, *5*, 88–108.

Spencer, A. (2013a). Stems, the morphome, and meaning-bearing inflection. Retrieved fromhttps://www.academia.edu/3655885/Stems_the_morphome_and_meaning-_bearing_inflection.

Spencer, A. (2013b). Stems and Inheritance. Retrieved from https://www.academia.edu/4379008/Stems_and_inheritance.

Stump, G. (1993). Position classes and morphological theory. In G. E. Booij and J. van Marle (Eds.), *Yearbook of Morphology 1992* (129–180). Dordrecht: Kluwer.

Stump, G. (2001a). Inflection. In A. Spencer, & A. M. Zwicky (Eds.), *The Handbook of Morphology* (pp.13-43). Oxford: Blackwell.

Stump, G. (2001b). *Inflectional Morphology: A Theory of Paradigm Structure*, Cambridge, England: Cambridge University Press.

Stump, G., & Finkel, R. (2013). *Morphological Typology: From Word to Paradigm.* Cambridge, England: Cambridge University Press.

Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence. *Phonology*, *16*, 401–442.

Ussishkin, A. (2000). Root-and-pattern morphology without roots or patterns. In M. Hirotani, A. Coetzee, N. Hall, & J. Y. Kim (Eds.), *Proceedings of NELS 30* (pp. 655-610). Amherst: GLSA.

Versteegh, K. (2001). *The Arabic Language*. Edinburgh: Edinburgh University Press.

Watson, J. (2002). *The phonology and morphology of Arabic*. Oxford, England: Oxford University Press.

Watson, J. (2006). Arabic morphology: diminutive verbs and diminutive nouns in San'ani Arabic. *Morphology*, *16*, 189-204.

Watson, J. (2007). Syllabification patterns in Arabic dialects: long segments and mora sharing. *Phonology*, *24*, 335-356.

Watson, J. (2011). Word stress in Arabic. In M. Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology*, Vol. 5 (pp. 2990-3018). Oxford: Blackwell.

Wurzel, U. (1989). *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer.

Zwicky, M. (1985). How to describe inflection. *Proceedings of the Berkeley Linguistics Society meeting*, *11*, (pp.372-386).