

**The Arabic dialect of the Rwala tribe,
based on the ethnographic records of Alois Musil**

Volume I: Thesis



Joseph O'Hara
Wolfson College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
August 2019

Table of contents

Volume One Thesis

<i>Abstract</i>		vii
<i>Acknowledgements</i>		viii
<i>Transcription and transliteration</i>		ix
<i>List of abbreviations and symbols</i>		xi
<i>List of tables</i>		xiv
<i>List of figures</i>		xiv
1	Introduction	1
1.1	Musil and the Rwala tribe	1
1.1.1	Alois Musil	1
1.1.2	The Rwala tribe	4
1.2	The dialect of the Rwala	6
1.2.1	Sources of the Rwala dialect	8
1.3	The Rwala dialect in context	15
1.3.1	North-Arabian dialects	15
1.3.2	Classificatory characteristics of the Najdi groups	21
1.3.3	Sources of north-Arabian dialects	24
1.4	Methodology and aims	27
1.4.1	Musil's corpus	28
1.4.1.1	The origins of Musil's material	28
1.4.1.2	<i>yabgi</i> ~ <i>yabi</i> : a case study in variation	31
1.4.1.3	The prosody of bedouin oral poetry	33
1.4.2	Digitisation of the corpus	40
1.4.3	Objective 1: description of the dialect	42
1.4.4	Objective 2: evaluation of Musil's data	44
1.4.4.1	The status of Classical Arabic	47
1.4.5	Objective 3: creation of a glossary	49
2	Phonology	51
2.1	The sound system	51
2.1.1	Consonants	52
2.1.1.1	/ʔ/	54
2.1.1.2	/ğ/	63
2.1.1.3	/ḏ/	68

2.1.1.4	/g/ and /k/	69
2.1.1.5	Consonant variation	75
2.1.2	Vowels	77
2.1.2.1	Short vowel phonemes	81
2.1.2.2	Distribution of <i>e</i>	85
2.1.2.3	Distribution of <i>o</i>	86
2.1.2.4	Long vowels and diphthongs	87
2.2	The syllable and syllabification	89
2.2.1	Syllable types	95
2.2.2	Syllabification	99
2.2.2.1	Epenthesis	105
2.2.2.2	Prothesis	114
2.2.3	Syncope	120
2.2.3.1	Syncope of /i/ and /u/	122
2.2.3.2	Syncope of /a/	128
2.2.3.3	Syncope of the feminine marker /at/	130
2.2.3.4	Unsyncopated forms	132
2.2.3.5	Initial consonant clusters	136
2.3	Other phonological processes	141
2.3.1	The <i>ghawa</i> -syndrome	142
2.3.2	Assimilation	146
2.3.2.1	Place assimilation	147
2.3.2.2	Manner assimilation	148
2.3.2.3	Voice assimilation	149
2.3.2.4	Emphasis assimilation	150
2.3.3	Word-final vowel shortening	151
2.3.4	Pre-suffix /n/ gemination	155
2.3.5	V_V resolution	156
2.3.5.1	V deletion	157
2.3.5.2	Glide formation	158
2.3.5.3	Epenthesis of [ʔ]	159
2.3.6	Qualitative vowel processes	160
2.3.6.1	Low vowel raising	160
2.3.6.2	High vowel backing	170
2.3.6.3	High vowel lowering	171
2.4	Summary	172
3	Nominal morphology	177
3.1	Nouns	177
3.1.1	Gender	177

3.1.2	Noun templates	181
3.1.2.1	Minimal trilateral templates	182
3.1.2.2	Extended trilateral templates	183
3.1.2.3	Nouns with initial consonant clusters	184
3.1.2.4	Participles and verbal nouns	185
3.1.2.5	Quadrilateral and geminated templates	186
3.1.2.6	Nouns with more than four root letters	187
3.1.2.7	Suffixed noun templates	187
3.1.3	Semantic classes	190
3.1.4	Number	192
3.1.4.1	Internal plural templates	196
3.1.4.2	External plural forms	199
3.1.4.3	Dual	200
3.1.5	Diminutives	202
3.2	Adjectives	203
3.2.1	Adjectival templates	204
3.2.1.1	Plural adjectival templates	204
3.2.2	Colour adjectives	205
3.2.3	Elicative	205
3.2.4	Derivational morphemes	206
3.3	The construct state	210
3.4	The adnominal linker	212
3.5	Numerals	218
3.6	Summary	220
4	Verbal morphology	223
4.1	The Rwala verbal system in Cantineau and Prochazka	224
4.2	S-STEM in Musil	227
4.2.1	Suffixes of the S-STEM	227
4.2.2	Classes of the S-STEM	231
4.2.2.1	The A-CLASS	231
4.2.2.2	The I-CLASS	233
4.2.3	Stem vowel incidence of the S-STEM	235
4.2.4	Stem syllabification	237
4.3	P-STEM in Musil	243
4.3.1	Prefixes of the P-STEM	244
4.3.2	Suffixes of the P-STEM	247
4.3.3	Classes of the P-STEM	249
4.3.3.1	The paradigm	251
4.3.4	Vowel incidence of the P-STEM	252

4.4	Participles	256
4.4.1	Active Participle	257
4.4.2	Passive Participle	258
4.5	Imperative	259
4.6	Verbal noun	260
4.7	Passive forms	261
4.7.1	Internal passive of the S-STEM	261
4.7.2	Internal passive of the P-STEM	263
4.8	Derived forms	264
4.8.1	Morphology of the derived forms	266
4.8.1.1	FORM-II	266
4.8.1.2	FORM-III	269
4.8.1.3	FORM-IV	271
4.8.1.4	FORM-V	273
4.8.1.5	FORM-VI	276
4.8.1.6	FORM-VII	278
4.8.1.7	FORM-VIII	280
4.8.1.8	FORM-IX	281
4.8.1.9	FORM-X	283
4.8.2	Semantic correspondences of the derived stems	284
4.8.2.1	FORM-II	284
4.8.2.2	FORM-III	287
4.8.2.3	FORM-IV	288
4.8.2.4	FORM-V	290
4.8.2.5	FORM-VI	291
4.8.2.6	FORM-VII	291
4.8.2.7	FORM-VIII	293
4.8.2.8	FORM-IX	293
4.8.2.9	FORM-X	294
4.9	Phonological adjustments	295
4.9.1	C1 = guttural	295
4.9.2	C1 = /w/ or /y/	296
4.9.3	C2 = /ʔ/	297
4.10	Bilateral verbs	298
4.10.1	Deficient verbs	299
4.10.2	Hollow verbs	303
4.10.3	Final-weak verbs	307
4.10.4	Doubled verbs	311
4.11	Quadrilateral verbs	314
4.12	Summary	319

5	Closed classes	324
5.1	Particles	326
5.1.1	The vocative	326
5.1.2	Exclamations	327
5.1.3	Coordinating conjunctions	330
5.1.4	Subordinating conjunctions	332
5.1.5	Conditionals	333
5.1.6	Exceptives	334
5.1.7	Emphasisers	334
5.1.8	Restrictives	337
5.1.9	Negatives	338
5.1.10	Interrogatives	343
5.1.11	Question particle	347
5.1.12	Adverbs	348
5.1.13	Prepositions	352
5.1.14	Definite article	355
5.1.15	Indefinite article	359
5.1.16	Existential particle	359
5.2	Pronouns, demonstratives and relatives	361
5.2.1	Pronouns	361
5.2.1.1	Free independent pronouns	361
5.2.1.2	<i>īyyā-</i>	364
5.2.1.3	Bound objective and possessive pronouns	365
5.2.1.4	Bound pronouns on prepositions	371
5.2.1.5	Interrogative pronouns	375
5.2.2	Demonstratives	375
5.2.2.1	Presentatives	381
5.2.3	Relatives	383
5.2.3.1	With a definite antecedent	383
5.2.3.2	With an indefinite antecedent	385
5.2.3.3	With a zero antecedent	386
5.3	Nominal forms	387
5.3.1	Quantifiers	388
5.3.1.1	<i>kill</i>	388
5.3.1.2	<i>ḡamīʿ</i>	391
5.3.1.3	<i>baʿaḏ</i>	392
5.3.2	Reflexive pronouns	392
5.4	Summary	394

6	Conclusion	397
6.1	Attestation of classificatory characteristics	397
6.1.1	Attestation of bedouin-type characteristics	397
6.1.2	Attestation of north-Arabian characteristics	402
6.1.3	Attestation of central Najdi characteristics	406
6.2	Attestation of non-Najdi characteristics	411
6.3	Evidence of shared lexicon	415
6.3.1	Shared bedouin lexicon	416
6.3.2	Shared Najdi lexicon	419
6.3.3	Shared poetic lexicon	420
	Bibliography	426

Volume Two

Glossary of the Rwala dialect from Musil 1928

<i>Guidance notes for the glossary</i>	445
Glossary	462

Abstract

This thesis presents a linguistic description of the dialect of the Rwala tribe of northern Arabia as presented in Alois Musil's (1928) ethnographical work *The manners and customs of the Rwala bedouins*. A systematic presentation of the phonology and morphology is given, which compares the grammar of the language as Musil recorded it with other sources of the Rwala dialect, and with sources of north-Arabian dialects more broadly, in order to assess the fidelity of Musil's transcriptions. A complete glossary of Musil's work has been compiled according to the grammar as set out in the rest of the thesis.

Chapter 1 places the Rwala dialect in its north-Arabian context and considers the other sources useful for comparative work and the relevant isoglosses and classificatory characteristics of bedouin, north-Arabian, and Najdi dialects. In chapter 2 the phonology is presented, where it is discussed that Musil's system of transcription is unlikely to convey the true nature of the sound system of the dialect, and that several important phonological processes recorded for dialects of the type to which Rwala belongs are attested significantly less in Musil than in other accounts. Chapters 3 and 4 deal with the morphology of nominal and verbal forms, where idiosyncrasies in Musil are highlighted, although it is shown that in general his record of the language is in line with what would be expected. Chapter 5 deals with function words with similar conclusions to chapters 3 and 4, and in chapter 6 a preliminary lexical study is carried out which demonstrates the similarity of the dialect as presented by Musil to other Najdi dialects.

Acknowledgements

I would like to thank my supervisors for their support throughout the writing of this thesis. Michael Macdonald has consistently encouraged me in my research and I am sincerely grateful for how he has supported my pursuit of new avenues of study as the nature of the thesis has developed and has shared his knowledge and love of bedouin culture with me. I thank Jeremy Johns in particular for liaising with the department and the division in order to ensure that I could submit a dictionary as part of my thesis. Special thanks go to Daniel Burt of the Khalili Research Centre (Oxford) for creating the database according to my specifications (and changing it every time my demands evolved), without which this thesis could not have been written. My thanks go to Clive Holes, Ahmad Al-Jallad, and Bruce Ingham for being part of the process at the Transfer and Confirmation of Status assessments. I am especially grateful to Janet Watson and Clive Holes for agreeing to be examiners, for the care they took in reading my work, and for the many helpful comments they have given. Clive Holes is further thanked for his invaluable help with interpreting the material and for the considerable time he spent in doing so. I also wish to thank Nadia Jamil for her advice and support as the project neared completion. Finally, I am grateful to the Arts and Humanities Research Council and to the Khalili Research Centre for financial support during this project.

Transcription and transliteration

The transcription system used in *The manners and customs of the Rwala bedouins* (Musil 1928a) is idiosyncratic and different from the usual systems of transliteration used for dialects of this type. Although there are (mostly) small differences in many authors' systems, that of Musil has been changed in this work to one which should be mostly familiar to users of Arabic dialect dictionaries in general. The two systems can be seen in the following table, with changes in the shaded cells, followed by a short description.

Arabic	Musil	Thesis	Arabic	Musil	Thesis
ء	ʾ	ʾ	ط	ṭ	ṭ
ب	b	b	ظ	ẓ	ḏ
ت	t	t	ع	ʿ	ʿ
ث	ṭ	ṭ	غ	ḡ	ḡ
ج	ǧ	ǧ	ف	f	f
ح	ḥ	ḥ	ق	ḵ	g
خ	ḫ	ḫ	n/a	ẓ	ǧ
د	d	d	ك	k	k
ذ	ḏ	ḏ	n/a	č	ć
ر	r	r	ل	l	l
ز	z	z	م	m	m
س	s	s	ن	n	n
ش	š	š	ه	h	h
ص	ṣ	ṣ	و	w	w
ض	ḏ	n/a	ي	j	y

More detailed information on these consonants can be found in §2.1.1, but will here be described in short. *ḍ*, while listed by Musil in his transliteration chart, is never found in the work (as a result of the merger with *ḏ*, see §2.1.1.3). The symbol *ḏ* is used in transcriptions from other dialects, and when discussing roots which have /ḏ/ historically, but is not found in any transcriptions of the Rwala dialect. The transliteration *ḏ* for the Arabic letter *ḏ* is discussed in §2.1.13; the symbol is not universally used, perhaps owing to typographical constraints, and in other sources of the Rwala dialect it is only used in Cantineau (1936-37) (Ingham [1995] uses *ḏ*, although gives it the same phonetic description as *ḏ*, and Prochazka (1988) uses *ḏ*, which is theoretically the same as *ḏ*). The symbol *r* is almost never used in modern transcription of Arabic as a transliteration of Arabic *ḥ*, and *ḥ* is widely recognised in the field (the symbol *r* usually represents emphatic /r/ in modern studies). The representation of *ḡ* with *g* is discussed in §2.1.1.4, as are the symbols *ḡ* and *ḥ*; the symbol *č* is used in transcriptions of other dialects to represent IPA [tʃ] while *ḡ* and *ḥ* are used to represent IPA [ɟ] and [tʃ] respectively. Most descriptions of dialects related to the one in question use *y* to transcribe *ḡ*, and it should be noted that *ḡ* has not been replaced with *j* in order to avoid confusion with Musil's original text.

Vowels have been kept as Musil recorded them, with the exception that long vowels are represented with macrons (*ā, ē, ī, ō, ū*) rather than with circumflexes as they are found in Musil (*â, ê, î, ô, û*).

A table showing the probable phonetic values of each sound is found in §2.1.1.

List of abbreviations and symbols

A-CLASS	corresponds to CLA <i>faʿala</i> in the S-STEM and CLA <i>yafʿalu</i> in the P-STEM
C	common (not marked for masculine or feminine)*
C	consonant
C´	stray (unsyllabified) consonant
CLA	Classical Arabic
Co	coda
coll.	collective
const.	construct (usually the feminine form ending in <i>-t</i> in a construct phrase)
dim.	diminutive
f.	feminine*
F	feminine*
F	fricative
G	glide
FORM-I—FORM-X	indicates derived form of the verb
FORM-QI—FORM-QIV	indicates derived form of the quadriliteral verb
I-CLASS	corresponds to CLA <i>faʿila</i> in the S-STEM and CLA <i>yafʿilu</i> in the P-STEM
indecl.	indeclinable
IPA	International phonetic alphabet
L	liquid
lit.	literal meaning
m.	masculine*
M	masculine*
MSA	Modern Standard Arabic
N	nucleus; nasal
O	onset; obeys sonority hierarchy
OA	Old Arabic (the putative ancestor of the dialect)
P	plural*

P	sonority plateau
P-STEM	the prefix conjugation (or imperfect) of the verb
pl.	plural*
R	rhyme; sonority reversal
s	singular*
S	stop
sed.	sedentary dialect
S-STEM	the suffix conjugation (or perfect) of the verb
sb.	somebody
sg.	singular*
SPE	The Sound Pattern of English (Chomsky and Halle 1968)
SSP	Sonority Sequencing Principle
sth.	something
U-CLASS	corresponds to CLA <i>fa'ula</i> in the S-STEM and CLA <i>yaf'ulu</i> in the P-STEM
UR	underlying representation
V	vowel
μ	mora
σ	syllable node
<σ>	degenerate syllable
1-3	first, second or third person
.	syllable boundary
+	morpheme boundary
#	word boundary
/ /	indicates a phoneme or underlying representation
[]	indicates a pronunciation; [in transcriptions and translations] indicates an element which has been supplied
<>	indicates a spelling; [in syllable trees] extrametrical consonant
> or →	'becomes'
~	in free variation with
bold font	indicates a lexical root

superscript font used for the adnominal linker and other otiose word-final vowels

* unattested or reconstructed form

***** [in derivations] indicates the failure of a rule to apply

* Where forms in the table above are marked with an asterisk, note that the lower case forms are used in isolation (e.g. pl. = plural) whereas the small capitals are used for combinations, (e.g. 3MS = third person masculine singular).

Note that a more extensive set of abbreviations is used in the glossary, which can be found in Volume II.

List of tables

TABLE 1.1: Some major characteristics for classifying Najdi dialect groups	22
TABLE 2.1: Consonantal phoneme inventory for the Rwala dialect	53
TABLE 2.2: Syllable types	95
TABLE 4.1: Verbal stems in Cantineau and Prochazka	226
TABLE 4.2: Frequency of P-STEM prefix vowels	244
TABLE 5.1: Bound pronouns with prepositions (Musil)	372
TABLE 5.2: Bound pronouns with prepositions (Prochazka)	374
TABLE 5.3: Morpheme combinations of attested demonstratives	375

List of figures

FIGURE 2.1: Sonority peaks	90
FIGURE 2.2: Syllable structure	90
FIGURE 2.3: Syllable structure of open and closed syllables	91
FIGURE 2.4: Syllable structure of CV and CVV syllables	92
FIGURE 2.5: Syllable structure of geminate consonants	93
FIGURE 2.6: Weight-by-Position	93
FIGURE 2.7: Syllable structure of CVVC and CVCC syllables	95
FIGURE 2.8: Glottal stop epenthesis	96
FIGURE 2.9: Syllabification	101
FIGURE 2.10: Stray consonant syllabification	103
FIGURE 2.11: Stray consonant syllabification (2)	105
FIGURE 2.12: Epenthesis	106
FIGURE 2.13: Degenerate C syllables in onset and rhyme dialects	107
FIGURE 2.14: Adjunction-to-Mora	110
FIGURE 2.15: Complex coda formation	111
FIGURE 2.16: *Initial cluster epenthesis	115
FIGURE 2.17: Prothesis	115
FIGURE 2.18: FORM-VII verb template	116
FIGURE 2.19: Medial epenthesis between two stray consonants	117
FIGURE 2.20: Imperative prothesis	118

1 Introduction

This chapter addresses the source material for this study, namely the evidence of the Rwala dialect as recorded by Alois Musil in *The manners and customs of the Rwala bedouins* (Musil 1928a). The dialect of the Rwala tribe will be situated in terms of what is already known about it from previously published sources, and the Najdi dialect group is presented with reference to how it is defined in the existing literature and the sources that have been used in this thesis for comparative assessment of Musil's data. The objectives of the thesis are outlined and a methodology has been provided which, it is hoped, should explain the approach taken for the analysis in the following chapters.

1.1 Musil and the Rwala tribe

1.1.1 *Alois Musil*

Alois Musil (1868-1944) was a Czech explorer and orientalist, Catholic priest and prolific writer who spent a great amount of time in the north Arabian desert in the period from 1895 to 1918, and who travelled over thirteen thousand miles on camelback in order to carry out his investigations (see AAAS 1928: 290, AGS 1946: 686-7). Musil initially travelled to Palestine to study biblical languages at the newly established Dominican *École Biblique et Archéologique Française de Jérusalem*, where it is said that he started to learn Arabic (presumably the Palestinian dialect) from the locals (Košutová 2018: 11), but it is

unknown whether he had any training in Classical Arabic at this point. His first stay in the Middle East lasted from 1895-98, during which time he transferred his studies to Beirut and ultimately travelled with the Bani Şaḥar tribe who led him to Quşayr ʿAmrah, for the discovery of which he was to become best known in the Western world. After his reports on the site were met with scepticism from the academic community, Musil returned to Quşayr ʿAmrah twice, in 1900 and 1901, to photograph it and collect further evidence, after which he became a celebrated Orientalist. He returned to the Middle East at least six more times: in 1902 to survey ancient sites in Sinai, the Negev and around the Dead Sea; in 1908-09 when he joined the Rwala in order to survey north-Arabia and the Syrian desert; in 1910 when he mapped the area covered by the Damascus–Medina railway line at the request of the Ottoman government; and in three politically motivated trips with officials of the Austro-Hungarian empire, the most important of which was in 1914-15, at the beginning of the First World War, when the Rwala chief introduced him to the Saudis, Rashidis and Hashemites, whom he attempted to convince to support the interests of the Ottoman Empire (Kořutová 2018: 36).

As a result of the survey of 1908-09 Musil was able to present to the western world the detailed topography of an area which was relatively unknown at the time¹, a map of which is found in RGS (1910: 580). The success of this work was due in part to the close

¹ RGS (1910: 579) refers to the area Musil covered as a “blank space” on the map, although this is not entirely true, since the area had been visited by Euting and Huber, Lady Ann Blunt, and possibly Burckhardt.

relationship he cultivated during that time with the then chief of the Rwala tribe, an-Nūri ibn Ša‘lān, a tactic he had used previously among the Bani Şaḥar. Musil referred to himself as a “close friend of the Prince” (Musil 1927a: 300), and it is known in fact that he was accepted as one of the tribe, and even rode at its head on more than one occasion (Wright 1927: 177-206). His findings were published by the American Geographical Society in six volumes (Musil 1926, 1927a, 1927b, 1928a, 1928b and 1928c) which were noted for their accompanying maps, for their accuracy regarding topographical features, and for their contributions to the ethnography of the region.

While the first five of these volumes are designated “topographical itineraries” by the author, the sixth, *The manners and customs of the Rwala bedouins*, comprises Musil’s systematic presentation of the ethnographic data he collected during his travels. It was praised in 1946:

“The book on the Rwala Bedouins is a classic record of the material culture and daily habits of tribesmen who at the time of his sojourns among them were still wholly untouched by the influence of the modern world, living much as their ancestors had lived since the days of Abraham” (AGS 1946: 686).

The idealised conception that this writer had of Musil’s work says a lot about the time when he was writing, and it is ironic that Musil’s work was viewed in this way when he himself presents a far less romanticised portrayal of his subjects². What makes the work

² *Manners and Customs* (Musil 1928a) is arguably the work most susceptible to accusations of romanticism, since it puts forward the view which the Rwala wanted the outside world to have of them (cf. Lancaster and Lancaster 1988: 53-54). The other five books from the series are far more candid in their accounts of the realities of life among the bedouins.

unique – and the reason it forms the basis of this study – is the sheer volume of linguistic material recorded by Musil within it, comprising over six thousand individual poems, excerpts of conversation, proverbs and technical terms.

While Musil undoubtedly became proficient in probably several Arabic vernaculars, regrettably no sources that I am aware of shed light on his Arabic education (if he had a formal one), the timeline of which would be useful to know for a study of this sort. As mentioned above, it appears that he first learned Palestinian Arabic while living in Jerusalem, but it cannot be dismissed that he could have learned Classical Arabic before this. He claimed to have been fluent in fifteen languages (Harrigan 2009) and references to him being a very demanding Arabic lecturer (Harrigan, *ibid.*) imply that he was teaching Classical Arabic to students later in his career: he lectured at Charles University (Prague) 1920-1938 and at the newly inaugurated Czechoslovak Oriental Institute (which he helped to found) after 1927 (Košutová 2018: 38).

1.1.2 *The Rwala tribe*

The Rwala³ tribe is the largest and most powerful subsection of the ‘Aneze confederation: the latter is spread over the Arabian Peninsula, Iraq and the Levant while the former is located in the desert of southern Syria, northern Jordan, and north-western Saudi Arabia (cf. Lancaster and Lancaster 1995: 643). They have famously been described as “the

³ Various spelled Ruwala, Ruwayla, Roala, Ruwalla, Ruweilah, Rwalah in English sources. A member of the tribe is called a Rweyli (Rwaili, Ruwayli, etc.).

only true bedouin tribe of northern Arabia” (Musil 1928a: *xiii*), both by their neighbours and by scholars, although this attestation seems to refer only to their role (more apposite a century ago than it is today) of herding camels⁴. The two principal sources for the study of the Rwala tribe are the works of Musil and those of William and Fidelity Lancaster, all of whom spent considerable time undertaking ethnological and anthropological research while living among the Rwala. The Lancasters spent twenty years on and off living with the tribe and have published their findings in *The Rwala bedouin today* (Lancaster 1997), which is an important accumulation of ethnographic data⁵, and in numerous other articles and journal entries, including the *Encyclopaedia of Islam* entry on the Rwala. Their familiarity with the modern bedouin enables them to regularly bring balance to academic discussion by emphasising the bedouins’ own perspective.

The lives of the Rwala today are very different from those described by Musil. According to the Lancasters, they are now “an important group in the National Guard of Saudi Arabia, and are represented in the Army and Air Force; they play an active part in government, the professions and business in Saudi Arabia”. Furthermore, “they say they are

⁴ This definition of the term “bedouin” has been widespread: Lancaster and Lancaster (1995: 644) state that in 1908-1916 this is what “being Bedu” would have meant. It should be noted, however, that in modern usage this is no longer adequate. Lancaster notes: “nor do they accept that the camel is the only object that makes them bedouin; they explicate being bedouin as a moral and political attitude, almost an ideology, and they see the possession and use of the camel as enabling them to pursue that ideology rather than determining it” (Lancaster and Lancaster 1988: 53). Cf. Macdonald (2015: 53-55) and Lancaster (1997).

⁵ It is described as “one of the best modern ethnographies on Middle Eastern ethnic groups” (blurb), although it has also been criticised for its “unacademic” style (Tapper 1983: 153) and “fundamentally idealist view” (Sweet 1983: 212).

no longer Bedouin” on account of no longer depending on the desert as they once did, but it is noted that nonetheless “they maintain their strong tribal identity” (Lancaster and Lancaster 1995: 644).

1.2 The dialect of the Rwala

The Rwala speak an important but little studied dialect of Arabic, of which Cantineau remarked in 1936, “il serait utile de connaître à fond”⁶ (Cantineau 1936-37: 11). Since Cantineau, it has been classified as a north-Arabian (bedouin) dialect of the ‘Anazi branch (see below), and very little work has been done on it in the eighty years since then. Rosenhouse (1984: 5) more recently classified it as a Šammari dialect, although it seems likely that this was simply a mistake⁷; this discrepancy, however, is perhaps also a symptom of how little attention has been paid to this particular dialect, in spite of the fact that the tribe itself is both very large and (formerly) powerful. Cantineau noted that in his time the ‘Anazi and Šammari dialects were not well known in western scholarship:

“si les parlars de grands nomades étaient mieux connus, il est probable que beaucoup d’autres oppositions apparaîtraient. Une étude un peu sérieuse d’un dialecte ʿAnāze (Rwala, par exemple) est véritablement urgente.” (Cantineau 1936-37: 113)

Fifty years later, Al-Sweel was similarly bleak:

⁶ Since Cantineau managed to find a Rweyli informant on his second survey and included Rwala material in his second (1937) article, it may be that he considered this need to have been fulfilled only a year later.

⁷ Rosenhouse repeats the assertion that Rwala is a Šammari dialect in the *EALL*, however (Rosenhouse 2011).

“[Najdi Arabic] is the least studied among Arabic dialects in the framework of modern linguistics, and therefore references are less than one hand’s fingers in number.” (Al-Sweel 1987: 71)

This second appraisal is almost certainly an exaggeration, as several important studies predate Al-Sweel’s claim. There have been advances in the study of north-Arabian/Najdi dialects in recent decades, notably by Bruce Ingham, Theodore Prochazka and Marcel Kurpershoek, as well as numerous PhD studies. With the exception of this last category, available “sérieuse” sources of both the Rwala dialect and other related dialects are listed in §1.3.3, where I attempt to place the Rwala dialect in its linguistic context within the north-Arabian/Najdi group.

A matter of terminology may be clarified at this stage. Throughout the thesis the phrase “the Rwala dialect” refers abstractly to the language of speakers from the Rwala tribe, for which no single authoritative source exists. The ensuing description of the dialect based on Musil’s records is referred to more specifically as “the Rwala dialect according to Musil” or variations thereof. It goes without saying that evidence of the dialect provided by Musil and other writers may well be very different from the Rwala dialect as it exists today. Lancaster observes that “subsequent field research with the Rwala and other tribes has shown that ... the Rwala now often use a word different from that used in the 1970s for the same object” (Lancaster 1997: *vii-viii*). Furthermore, given the huge social changes that have affected the tribe in recent times, we may assume that the language one would record from them today would exhibit the well documented effects of sedentarisation on the dialects of

formerly nomadic groups, summarised by Rosenhouse (2011) as “many effects of borrowing, code mixing, koineization, and leveling”. It should also be noted that “the Rwala dialect according to Musil” is not monolithic, since the material comes from speakers from different social groups, from various social situations, and even from speakers of other dialects: this is discussed in §1.4.3.

1.2.1 Sources of the Rwala dialect

There are only a few published sources of the Rwala dialect, the largest of which is *The manners and customs of the Rwala bedouins* (Musil 1928a). As described above, this work is a remarkable and thorough ethnography of the tribe based on Musil’s observations from living among them in the early part of the twentieth century. It seems likely that most of the material was gathered during Musil’s stay with the Rwala in 1908-1909, although he does not specify that this was the case and it is possible that further material was collected on his political trip of 1914-1915 when it is known that he had further interactions with the chief of the tribe, but unknown whether he visited the tribe in general (cf. Košutová 2018). In addition to his descriptions of most aspects of their daily lives, the work contains a wealth of linguistic data including poems, proverbs, and excerpts of conversation, as well as a huge number of technical terms.

The work was very well received as an ethnography, but somewhat less so among Arabic dialectologists on account of his rather misleading system of transcription. Cantineau

said of Musil's records that, "bien que leur notation, et surtout leur vocalisme, soit souvent suspect, j'en ai tiré quelques recoupements intéressants" (Cantineau 1936-37: 123), showing that he had faith in the material itself but was suspicious about the way in which it was transcribed. He explicitly denounced the use of certain symbols where his language was somewhat stronger, e.g. "il ne faut certainement accorder aucune créance aux transcriptions [of *č* and *ž*] de Musil" (*ibid.*: 36); in addition to the symbols <č> and <ž> for the affricated allophones of *kāf* and *qāf*, which he described as "imaginaire" (*ibid.*: 36), he also rejected Musil's use of <ḵ> for *qāf* and <ḏ> for *ḏā*'. Unfortunately, however, he did not elaborate on the ways in which he considered Musil's *vocalisme* to be suspicious. Since Cantineau, most scholars who have used Musil's Rwala material (see examples below) have adapted his system of transliteration.

Some Rwala data was collected by Cantineau at a similar time and is published in his two-part article *Études sur quelques parlers de nomades arabes d'orient* (Cantineau 1936-37). Cantineau did not furnish us with a full description of the Rwala dialect, but rather his work is a very useful survey of the speech of members of almost twenty different tribal divisions synthesised into one work: this work remains the only source for the dialects of some of the smaller tribes he surveyed. Its presentation is such that many categories, e.g. the verbal system, are described in general terms for the whole region which is illustrated with examples from only one dialect, alongside a commentary which provides differing examples from other dialects where these occurred. As a result, the Rwala material is scattered

throughout the work and the coverage of this dialect is incomplete. Cantineau was also familiar with Musil's work and cites many examples from it: this is useful when Cantineau finds a form in Musil to differ from what he encountered on his survey but also means that in some cases Cantineau's data is no different from Musil's, e.g. about half of Cantineau's examples (for all dialects) of what is now termed the *ghawa*-syndrome are lifted straight out of Musil's Rwala material.

Many examples of isolated words in the Rwala dialect are found in *Saudi Arabian dialects* (Prochazka 1988), a work which, like Cantineau's, is a survey of a range of dialects. It deals almost exclusively with the verbal system, for which paradigms of most possible verbal forms are meticulously given in a comparative approach which lists each form from every dialect next to each other. The presentation is more complete than that found in Cantineau (1936-37), and while a particular form in a certain dialect may sometimes be missing, the Rwala forms are generally well attested. The work also has sections on pronouns and the suffixation of nouns which provide some evidence for forms outside the verbal system.

The final original source of Rwala material that I am aware of is *Texts in the dialect of the Rwalah of northern Arabia* (Ingham 1995), which is a transcription and annotated translation of six prose passages which were recorded by the Lancasters and given to Ingham to analyse. Ingham provides a very short description of what he perceived to be the

salient features of the dialect in the introduction, but the value of the article is the texts themselves which are prose narratives longer than any found in Musil.

It may be noted at this stage that there is very little useable linguistic content in Lancaster (1997). The Lancasters' work is purely anthropological and, in the words of one reviewer, "makes no attempt to transcribe words in an adequate way" (Stewart 1986: 245). In terms of transliteration the author claims to have "relied on commonsense" (Lancaster 1997: *xiii*) but the result is regrettably inscrutable, and Arabic terms, when they are given, are not always interpreted in any way that can be linguistically useful.

To the best of my knowledge there are only two secondary sources which have exclusively treated the Rwala material contained in these sources. One is *Literature and violence in north Arabia* (Meeker 1979), which is a literary study of a number of the poems recorded in Musil (1928a). The author does not deal with the language concerned except for commenting on certain stylistic features which are not a concern of this study. Similarly to my treatment in this thesis, Meeker decided to replace some of Musil's consonant symbols with more appropriate ones but left the vowels unchanged. It must be said that he did so imperfectly: for example, he replaced Musil's <ḡ> with <ḡ> and claims that this was based on Cantineau's criticism, when in fact Cantineau corrected it to <ḡ̣>.

The second is *The feature pharyngeal in Rwaili Arabic: a case for long distance multiple linking* (Parkinson 1992), a phonological study in which the author uses the Rwala dialect to argue the case for a cross-linguistic phenomenon termed Coronal Transparency (cf. Paradis

and Prunet 1989). It is clear from the examples cited that the work is entirely based on data from Prochazka (1988), although strangely the author does not explicitly acknowledge this in the article. The analysis is interesting and discussed where relevant in Chapter 2 (the author makes reference to the syncope and vowel-raising processes in the dialect which are treated in this thesis), although it should be noted that extreme care must be taken with this work. It is unfortunately riddled with errors which suggest that the author knows little about Arabic: as one example, while having /ǧ/ and /š/ missing from his consonant inventory could have been a palato-alveolar related oversight, he also lacks /t̪/ and incorporates /v/ as a voiced labial fricative⁸. More alarmingly, some of the forms he has taken from Prochazka have been changed, and in many cases key pieces of data on which his hypotheses rest are simply erroneous, to the point that it sometimes seems that they have been changed on purpose to lend support to the author's hypothesis. A particularly misleading example is his contrast of the first vowel in *samʿat* 'she heard' and *simʿ-ih* 'he heard it', on which basis he posits an entirely new rule termed Coronal Spreading (cf. §2.3.6.1), when in fact the forms given in Prochazka are *samʿat* and *samʿ-ih*, which have no difference in the first vowel and therefore provide no support for the rule he is trying to establish.

Apart from these two which focus on Rwala material, the dialect is sometimes referenced in works on other north-Arabian dialects. Many of these make reference to Musil

⁸ Ingham (2011) notes that /f/ may be sporadically voiced in some Najdi dialects, but it is almost certain that this was not Parkinson's motivation here.

for matters of vocabulary: a particularly notable case is *Voices from the desert : Glossary* (Kurpershoek 2005). This excellent glossary to the central-Arabian poetic material collected in the four other volumes of the same series (Kurpershoek 1994, 1995, 1999 and 2002) has been enriched with comparative evidence from many sources, with Musil featuring prominently. The author has adapted Musil's transcriptions to match his own system, which notably uses three short vowels (*a, i, u*) rather than Musil's five (*a, e, i, o, u*); it is admitted that "ambiguities" in Musil were replaced by "a good guess" (*ibid.*: xviii) which, for the most part, seems to refer to the author's practice of using editorial judgement in deciding whether a form with *e* in Musil should be rendered with *a* or *i* in that work, and the raising of some *a* vowels to *i* (see §.2.3.6.1 on low vowel raising).

Similarly, some of the poetry recorded by Musil has been reproduced and commented on by scholars working on bedouin oral poetry⁹. This includes *Praying mantis in the desert* (Kurpershoek 2015), which analyses several poems by Fayḥān ibn Zirībān and ʿAbdallāh ibn Subayyil and includes commentary on the extant textual sources. A short poem by the former which was recorded in Musil (1928a: 181-2) is reproduced with some emendations to make it metrically sound. Kurpershoek notes that Musil's versions of this poem and a poem by Ibn Subayyil which he also recorded (Musil 1928a: 292-4) differ from later written recensions of these same poems: these differences concern the overall length of the poems, the order of verses, and some vocabulary. He attributes this variability to the

⁹ Note that this list of sources using Musil's material in this way is probably not exhaustive.

probable existence in central Najd of written versions of the poems which could have supported a more stable transmission *vis à vis* the versions circulating in the north, and also to the fact that Musil may have misunderstood some of the language. Part of another poem recorded by Musil (*ibid.*: 306-7) is reproduced in *Metrical problems of the contemporary bedouin qaṣīda: a linguistic approach* (Palva 1993), and further excerpts (Musil 1928a: 81-2, 108, 300-1) are reproduced in *How interdialectal is peripheral oral bedouin poetry?* (Henkin 2009). In both of these studies, the material from Musil is considered representative of poetry from the core *nabaṭī* area where this poetry originated, in contradistinction to that of the more peripheral areas of the tradition such as Sinai, Negev, and northern Jordan. In Kurpershoek (2015) the author's treatment of Musil's transcriptions is the same as that just described for Kurpershoek (2005); Palva and Henkin, on the other hand, standardise the transcription of consonants but retain Musil's vowels.

It can be seen therefore that there are four primary sources for the Rwala dialect, which complement each other. While Musil is by far the most copious, Ingham's texts are helpfully of a different genre (narrative) and are particularly useful for matters of syntax. Prochazka comprehensively covers the verbal system, which is not so well described for the Rwala dialect in Cantineau, but Cantineau's work is well furnished with nominals and function words. For the purpose of assessing the reliability of Musil's material it is of tremendous use that we have examples of the Rwala dialect transcribed by somebody other than Musil himself; where Musil was first and foremost an ethnographer, the other three

were/are linguists, and as such their material may be critically appraised with slightly less scepticism. The records of Cantineau are particularly useful in this regard since he and Musil were more or less contemporaries. The secondary sources mentioned here essentially repeat the same material and do not feature prominently in the analysis, but there is a substantial amount of literature on related dialects which has been frequently consulted and is detailed in the next sections.

1.3 The Rwala dialect in context

1.3.1 *North-Arabian dialects*

North-Arabian is one of four major linguistic groups in the Arabian Peninsula. These dialect groups, excluding those of the Mediterranean littoral, were set out in Johnstone (1967a: 1) which remains, more or less, the scholarly standard. They are:

- (a) North Arabian
- (b) Hijazi
- (c) South-western Arabian
- (d) Omani

This classification has not been challenged substantially since Johnstone wrote in the sixties. We may compare a similar assessment over forty years later: “linguistically the Peninsula can be divided into four main areas: the north, including central and north-eastern areas, the Hijaz, the south-west, and the south” (Watson 2011a: 897).

Of these it is the north-Arabian group which concerns this study. The defining characteristics of this group are taken from Johnstone (1967a: 2-9 for i-iii) and Cantineau (1936-37: 222-4 for iv-viii):

- (i) the affrication of /k/ and /g/ (cf. §2.1.1.4)
- (ii) the effect of the gutturals on syllable structure
- (iii) the syllable structure of certain nominal and verbal forms
- (iv) “differentiation” between /i, u/-syncope and /a/-syncope (cf. §2.2.3)
- (v) the diphthongal plural S-STEM suffix *-ow* or *-aw* (cf. §4.2.1)
- (vi) the P-STEM suffixes *-īn* and *-ūn* (cf. §4.3.2)
- (vii) the lack of the *b*- imperfect
- (viii) usage of the adnominal linker (cf. §3.4)

Note that S-STEM refers to the suffix conjugation of the verb (the ‘perfect’ of Classical Arabic grammars) and P-STEM refers to the prefix conjugation (or ‘imperfect’). In (ii) Johnstone is referring to the phenomenon usually termed the *ghawa*-syndrome (cf. §2.3.1); the term first appeared in print in Blanc (1970), after Johnstone published this. His description of (iii) refers to the processes of syncope (the deletion of vowels in open syllables, cf. §2.2.3) and /a/-raising (the raising of /a/ to *i* in open syllables, cf. §2.3.6.1), which produce forms such as *bgara* ‘antelope’ (< **bagara*) and *kitab* ‘he wrote’ (< **katab*)¹⁰.

To this list may be added a number of distinguishing bedouin features: these are not shared by contiguous sedentary dialects, but are not unique to the north-Arabian group (list

¹⁰ In some analyses these rules are in a feeding relationship, whereby the application of /a/-raising creates the context for syncope, e.g. **bagara* > **bugara* > *bgara* ‘antelope’. This may also occur in a cycle, e.g. **katab* > *kitab* ‘he wrote’ > **kitab-eh* > **ktab-eh* > *ktib-eh* ‘he wrote it’. In this thesis I have proposed separate rules of /a/-syncope and /i, u/-syncope (cf. §2.2.3), the former of which derives these surface forms without reference to /a/-raising as an intermediary step. This analysis removes the need for /a/-raising to apply twice in the same cycle, e.g. **katabat* > **ktabat* > *ktibat* ‘she wrote’, rather than **katabat* > *kitabab* > *ktabat* > *ktibat*.

from Palva 2011); it will also be seen that the *ghawa*-syndrome occurs outside the north-Arabian group and is therefore not strictly a group characteristic in the way that Johnstone intended:

- (i) the voiced reflex of *q (cf. §2.1.1.4)
- (ii) the *ghawa*-syndrome (cf. §2.3.1)
- (iii) gender distinction in 2nd and 3rd plural forms (cf. §4.2, §4.3, §5.2)
- (iv) productivity of FORM-IV (cf. §4.8.1.3, §4.8.2.3)
- (v) stressable definite article *al-* (cf. §5.1.14) and relative pronoun *alli* (cf. §5.2.3)
- (vi) shared bedouin lexical items
- (vii) productive use of diminutives (cf. §3.1.5)

The north-Arabian group, while having the common characteristics outlined in these lists, has also been classified into subdivisions, and three models are outlined here. Firstly, the following divisions were established by Cantineau (cf. 1936-37: 226-235) who observed four groups:

- (i) ‘Anazi (Group A)
- (ii) Šammari (Group B)
- (iii) *les petits nomades* (Group C)
- (iv) Šammari dialects influenced by the *petits nomades* (Group Bc)

Contrasting with (iii), both (i) and (ii) were considered to be *grands nomades* since they predominantly (although there are exceptions) reared camels rather than sheep and goats. This classification was expanded by Johnstone (1967a: 1-2):

- | | | |
|-------|----------------------------|-----------------------|
| (i) | Syro-Mesopotamian dialects | (Cantineau’s Group C) |
| (ii) | Šammari dialects | (Cantineau’s Group B) |
| (iii) | ‘Anazi dialects | (Cantineau’s Group A) |
| (iv) | The east-Arabian dialects | |

Johnstone's classification importantly incorporates the east-Arabian (mainly Gulf) dialects: these dialects did not feature in Cantineau's studies of the dialects of north-Arabia and the Syrian Desert but are genetically related to them, having (for the most part) branched off from the 'Anazi group and subsequently become differentiated through separate developments more recently (in the last two hundred years). More recent study has demonstrated that the history of the east-Arabian dialects is more complicated than this brief outline, and that other dialects in eastern Arabia do not share this descent. However, the history of those dialects does not concern us here¹¹.

Subsequent research on other (non-eastern) varieties, and specifically greater knowledge of the central Najdi dialects encountered to the south of where Cantineau surveyed, has led to the rebranding of the north-Arabian group as the Najdi group, and the following general classification by Ingham (2011):

- (i) central Najdi dialects
the dialects of Central Najd and associated Bedouin tribes,
also the 'Anizah of the Syrian desert
- (ii) northern Najdi dialects
the dialect of Jabal Šammar and of the Šammar tribes of
northern Najd and the Jazīrah
- (iii) mixed northern-central dialects
the dialect of Qašim and of the Ḍafir tribe
- (iv) southern Najdi dialects
the dialect of Najrān and the Gḥaṭān tribe of the south and of
the Āl Murrah and 'Ājmān tribes of the east

¹¹ An introduction to this field of study can be consulted in Holes 2011b.

As Johnstone's system incorporated the east-Arabian dialects, so Ingham's system incorporates the southern Najdi dialects, which are similar to the central Najdi dialects but which share some features with non-Najdi dialects of south Arabia. In this classification, it is the northern and southern dialects which have defining characteristics (cf. Ingham 1982b: 69 where they are given as "idiosyncrasies") but the central group is left rather undefined. To my knowledge, his incorporation of the 'Anazi group (to which Rwala belongs) to the central group has not been explained, so the defining characteristics of this central Najdi group outlined in the next section largely draws from Cantineau's assessment of 'Anazi characteristics and may not be relevant to all central Najdi dialects.

A factor which influenced the classification of a Najdi rather than a north-Arabian group is the modern awareness of a northwest-Arabian dialect group, which spans the area formerly known as *Arabia Petrea*, i.e. the Sinai Peninsula, the Negev, the south of Jordan and the northwestern part of Saudi Arabia. This group has been classified on the basis of excellent and concentrated scholarship, notably by Blanc (for the bedouins of the Negev), de Jong (for the bedouins of Sinai) and Palva (for the bedouins of southern Jordan). Palva (summarised in Palva 2011) has demonstrated that these dialects, though they span a fairly wide area, show a remarkable affinity that clearly differentiates them from the Najdi (north-Arabian) group.

More recent study of the Najdi dialects themselves has also refined the Cantineau-Johnstone model and informed Ingham's classification. Cantineau's groups were based not

only on linguistic markers but also on the historical and contextual situation in which he encountered the dialects: while his treatment was excellent, it was also limited, in that he assessed almost twenty different dialects for his pioneering 1936-1937 study but did not manage to acquire complete descriptions of most of the dialects he surveyed. Cantineau, therefore, drew distinctions between the camel-rearing *grands nomades* of the powerful ʿAnazi and Šammari confederations, who were more recent arrivals to the Syrian Desert, and the sheep- and goat-rearing *petits nomades* who had preceded them. This was not done in error, and the distinction still stands: it has been noted that “for large bedouin tribes of the area ... dialect corresponds to tribe in an easily distinguishable way” (Ingham 1982a: 30). This model is insufficient, however, since the dialect geography also demonstrates that dialects from different sub-groups influence each other in ways that can be attributed to language contact. For example, it has been established (Johnstone 1967a: 2-6) that the occurrence of the [tʃ] and [dʒ] allophones of /k/ and /g/ in contradistinction to [ts] and [dʒ] does not vary on the basis of *petits nomades* and *grands nomades*, as Cantineau recorded, but rather represents a clear isogloss between central and peripheral dialects (cf. the isogloss map in Johnstone 1967a: 5).

This advancement in our awareness of the dialectal variation between the groups listed above has been consolidated by Ingham (1982a and elsewhere) who assessed the linguistic situation of the area from the perspective of sociolinguistics for the first time. Ingham, having new data from dialects of southern Najd and Mesopotamia at his disposal,

demonstrated that migration patterns (and in particular the need for the nomadic tribes of north Arabia to buy necessities such as coffee and tobacco from the sedentary populations) had created a dialect area in which the dialects of northern Arabia were part of a linguistic continuum with those of Mesopotamia and the Gulf. Ingham concludes that the easily distinguishable dialects of large bedouin tribes are most expediently described with the “family tree” (*Stammbaum*) approach to language variation, while the dialect geography of the sedentary populations of Arabia should be described with a “wave theory” (*Wellen Theorie*) approach which sees “linguistic features spread across geographical landscapes from centres of importance irrespective of the original geographical or ethnic origin of the population concerned” (Ingham 1982a: 30-31). The erosion of the importance of “bedouin” and “sedentary” labels in Arabia, since “bedouin” refers in many ways more to tribal affiliation than to the nomadic lifestyle, means that the interaction between these two different dialect geographies has produced a system of complex linguistic variation.

1.3.2 *Classificatory characteristics of the Najdi groups*

The Najdi dialect groups as proposed by Ingham, albeit affected by language contact in the way just outlined, still rely fairly heavily on the divisions observed by Cantineau, with some further modifications, the most important of which are synthesised in TABLE 1.1. I have included some of Cantineau’s characteristics of the Syro-Mesopotamian group, i.e. group (v) here, since they bear a marked difference to the other groups but are not included in

TABLE 1.1: Some major characteristics for classifying Najdi dialect groups

	(i) central Najdi	(ii) northern Najdi	(iii) northern-central(iv) southern Najdi	(v) Syro-Mesopotamian
realisation of /ğ/	[gʲ]	[gʲ]	[gʲ]	[ç]
affrication of /k/ and /g/	[ts], [dʒ]	[ts], [dʒ]	[ts], [dʒ]	[tʃ], [ç]
“imāla” of final /ā/	none	none	none	-ā > -e
“imāla” of f. ending /a/	none	-a > -eyh	none	-a > -e
diphthongs	maintained	maintained		reduced
<i>nāgat-i</i> / <i>nāgt-i</i>	<i>nāgat-i</i>	<i>nāgat-i</i>		<i>nāgt-i</i>
1CP independent pronoun	<i>henna</i>	<i>henna</i>	<i>henna</i>	<i>ahne</i>
1CS bound pronoun	<i>ni</i>	<i>-an</i>		
2MS bound pronoun	<i>-ak</i> / <i>-ik</i> / <i>-k</i>	<i>-ak</i>	<i>-ak</i> / <i>-iç</i> / <i>-k</i>	<i>-ik</i>
2FS bound pronoun	<i>-iç</i> / <i>-ç</i>	<i>-iç</i>		<i>-iç</i>
3MS bound pronoun	<i>-eh</i> , <i>-ih</i>	<i>-o</i> , <i>-ih</i>	<i>-o</i> / <i>-eh</i>	
reflex of <i>'akal</i>	<i>akal</i>	<i>kala</i>		<i>kala</i>
particle <i>gid</i>	<i>gid</i> , restricted	<i>gid</i> , restricted		<i>gid</i> , not restricted
nunation on proper names	none	none	none	frequent
reflex of fem. suffix <i>-ā'</i>	<i>-a</i>	<i>-a</i>		<i>-ā</i> / <i>-ā'</i>
<i>min</i> + <i>i</i>	<i>minni</i>	<i>mini</i>	<i>mini</i>	
<i>min</i> + <i>ak</i>	<i>mink</i>	<i>mnak</i>	<i>mink</i>	

Ingham's classification; the separate developments of the east-Arabian dialects are not included however. The table only includes phonological and morphological categories, but there is also an amount of shared vocabulary specific to each group which is sometimes used for the purposes of classification. These are discussed in Cantineau (1936-37: 226-235 and throughout) for groups (i), (ii) and (v), while information on vocabulary pertaining to group (iv) may be found in Ingham (1986: 90-100). Following Ingham (2011) we may consider groups (i) and (ii), i.e. the central Najdi (including 'Anazi) dialects and the northern Najdi (Šammari) dialects, to be the geographical core of the whole group, with groups (iii), (iv) and (v) representing mixed groups. The primary division between groups (i) and (ii) is that group (ii) attests a number of idiosyncratic features not found in group (i) (cf. Ingham 1982b: 69) and the others developed as follows: group (iii) attests features of both (i) and (ii) (cf. Ingham 1982b: 67); group (iv) predominantly matches group (i) except for the incorporation of some features of the dialects of south-Arabia (cf. Ingham 1986: 88); and the dialects of group (v) have acquired a number of features from the dialects of the sedentary populations around them (cf. Cantineau 1936-37: 232-233).

Within each group there remains considerable diversity, but the many nuanced ways in which a single dialect of any group shares features with those of another group (usually contact phenomena rather than genetic) are too many to be covered here, and the details of these are to be found in the dialect descriptions listed in the next section. The chart is

restricted to data which is available and therefore not all categories are attested for each dialect group.

1.3.3 *Sources of north-Arabian dialects*

This section attempts to align the dialect geography described above with the published sources which have been used as comparative evidence throughout this thesis. Given the effects of language contact already described, it goes without saying that the classification of evidence from secondary sources according to general types necessarily obscures the effects of levelling which may be observed between dialects of different groups. The divisions may not correspond to present geographical location, e.g. the dialect of the Muṭayr is listed as central Najdi because this is its dialect type, although its speakers are found in eastern Arabia. The following lists comprise a reference to the work in question, the tribe or dialect being described (the language of sedentary populations is marked “sed.”), although all the varieties listed here are typologically bedouin), and a note on the type of information it contains.

(i) central Najdi dialects

Abboud 1978	Sudayr	imperfect prefix
Cantineau 1936-37	various ¹²	scattered examples
Ingham 1979	Muṭayr	description
Ingham 1982a	Sba‘a	scattered examples, texts
Ingham 1994a	Sudayr, Riyadh	description, texts
Ingham 1995	Rwala	texts

¹² These are the ‘Anazi emigré of the Syrian desert: Ḥsane, Weld ‘Ali, Rwala, and Sba‘a.

Kurpershoek 2005	‘Utaybah, Dawāsir	vocabulary
Lehn 1967	Riyadh	vowels
Musil 1928a	Rwala	texts
Prochazka 1988	various ¹³	verbal morphology

(ii) northern Najdi dialects

Abboud 1964, 1979	Ḥāyil (sed.)	syntax; verbs
Cantineau 1936-37	various ¹⁴	scattered examples
Johnstone 1967b	‘Anayza (sed.)	description
Lehn 1967	‘Anayza (sed.)	vowels
Montagne 1935	Šammar of the Jazīrah	texts
Palva 1980	Bani Ṣaḥar	description
Prochazka 1988	Ḥāyil (sed.)	verbal morphology
Sowayan 1982	poetic material	texts
Sowayan 1992	poetic material	texts, glossary, description
Al-Sweel 1987, 1990	‘Anayza (sed.)	phonology

(iii.) mixed northern-central dialects

Cantineau 1936-37	ar-Rass (sed.)	scattered examples
Ingham 1982b	Ḍafir	description
Prochazka 1988	Qasim (sed.)	verbal morphology

(iv.) southern Najdi dialects

Ingham 1986	Āl Murrah	description
Johnstone 1961, 1964	Dōsiri (‘Ājmān)	description
Prochazka 1988	Najrān	verbal morphology

(v.) Syro-Mesopotamian dialects

Cantineau 1936-37	various ¹⁵	scattered examples
-------------------	-----------------------	--------------------

¹³ These are: Rwala, Sudayr, Riyadh, and Bīšah.

¹⁴ These are the Šammari sheep-rearing tribes: ‘Umūr, Slūt, Sirḥān, Bani Ḥālid, and Bani Ṣaḥar.

¹⁵ These are: Ḥadīdīn, Rogga, Mawāli, N‘ēm, and Faḍol.

Data from some east-Arabian dialects (other than Muṭayr and Dōsiri/Ājmān already listed) have also been consulted. These dialects are Najdi in origin but less reference has been made to them than to the other groups on account of their significant independent development (cf. Johnstone 1967a: 2). Nonetheless, regularly consulted works were:

Holes 2016	dialects of Bahrain ¹⁶	description
Johnstone 1967a	various ¹⁷	description

It should be added that reference has also been made to dialects outside the Najdi group. In the first instance are some of the northwest-Arabian dialects:

Bani Yasin and Owens 1984	Bduul	phonology
Blanc 1970	Negev bedouins	description
Henkin 2010	Negev bedouins	description
de Jong 2000,2011a	dialects of Sinai	description
Palva 1984-86	Ḥwēṭāt	description

Some dialects from the Hijazi group have also been consulted:

Al-Hazmi 1975	Ḥarb	description
Al-Mozainy 1981	Ḥarb	phonology
Prochazka 1988	various ¹⁸	verbal morphology

I have not made reference to dialects from the southwest-Arabian or Omani groups, with the exception of numerous examples from San’ani from Watson (2002), since this work has been consulted frequently on matters of phonology and morphology.

¹⁶ Regarding this work, it is the “A” dialects which have been used for comparison since they are Najdi in origin.

¹⁷ These are: Kuwait, Bahrain, Qatar, and the Trucial Coast.

¹⁸ These are Prochazka’s dialects of southern Hijaz and Tihāmah: Al-Qahabah, Rufaydah, Abha, Bal-Aḥmar, Tanūmah, Bal-Qarn, Ghāmid, Al-Qawz, and Ṣabyā.

This selection of secondary sources is not intended to be a thorough bibliography of sources of the north-Arabian or Najdi dialect group in general. A particularly notable absence is that of some older sources of ‘Anazi dialects, namely Wallin (1851-52), Wetzstein (1868), Socin (1900-01), de Landberg (1919, 1940), and Hess (1938). This is partly because the material they contain is largely poetic, and partly because in some cases the nature of these records means that they are as difficult to interpret linguistically as Musil’s, and to do justice to their evidence would require a careful treatment which must form part of another project.

1.4 Methodology and aims

In the light of the criticisms brought against Musil (1928a) from a linguistic point of view as outlined above, this thesis aims to assess the reliability of Musil’s Rwala material. In the first instance, this has necessitated a description of the dialect as Musil recorded it, the results of which have been compared with the dialect sources listed above in order to establish whether the dialect as recorded by Musil is in line with what could be expected of the Rwala dialect in terms of the other sources of the dialect (cf. §1.2.1) and its position within the north-Arabian/Najdi group (cf. §1.3.2-3).

1.4.1 *Musil's corpus*

The corpus used for the analysis in this thesis is entirely from Musil (1928a), which is one of the largest corpora of north-Arabian dialectal material but is not without limitations. In the first instance the disadvantages of working from a corpus are well known, and there are some cases where there are gaps in the corpus relating to certain matters the type of which could have been elicited from speakers by researchers undertaking fieldwork for the purpose of linguistic research. It is entirely due to the sheer size of the corpus that omissions of this sort have not frustrated the task of describing the dialect, although it will be seen that some morphological categories are regrettably unattested.

This section considers some issues in the corpus which pose methodological challenges, specifically: the origins of Musil's material; the linguistic nature of the poetic material (which is the most frequent source type by volume); and Musil's method of data collection *vis à vis* that of other researchers.

1.4.1.1 The origins of Musil's material

The linguistic material in Musil's corpus comes from a variety of sources, and not all speakers were members of the Rwala. Musil acknowledges this in the preface and describes the speakers as "Rwala, for the most part" (Musil 1928a: *xiii*), but his presentation of the material does not consistently supply the kind of metadata which would inform a truly sociolinguistic study: he provides no record of how the data were collected and does not,

with the exception of certain named poets and reciters, record who the sources of his material were. This means that it cannot be guaranteed that any single utterance in the corpus is definitively from a Rweyli, but nor is it possible to isolate the words and phrases of speakers of other dialects. Furthermore, Musil claims to have transcribed the dialect “exactly” as he heard it (Musil, *ibid.*), but some examples cast doubt upon this claim (see below).

There are exceptions where adequate referencing is provided, or at least hinted at.

One example is recorded in the preface, where Musil writes,

I have found it advisable ... to add some details which I learned from my companion Bleyhān eben [sic] Meṣreb, who was not a Rweyli (or a member of the Rwala), but hailed from the Gmuṣa division of the Sba‘a tribe. (Musil, *ibid.*)

As we will see below, there are instances where Musil recounts tales which were told to him by Bleyhān ibn Meṣreb in which the Arabic material provided as an illustration to the story differs from that of the rest of the corpus in certain linguistic aspects. The difficulty in assessing these cases is that Musil is rarely quoting directly from Bleyhān, but is rather quoting other people (whose identity and background may or may not be known) through the filter of Bleyhān’s idiolect. This introduces an element of dialectal variation to the corpus, as Musil did not hear all the material from Rweyli informants, and therefore the speech he recorded is not systematically representative of the speech of members of the Rwala at that time in terms of their linguistic norms.

Another clear case of this pertains to the oral poetry, which makes up the majority of the corpus. *Nabaṭī* poetry is a widespread art form found across much of Arabia and beyond¹⁹, and in many cases the poems which were recited to Musil by Rweyli reciters are in fact not Rweyli in origin. One example (Musil 1928a: 181) was composed by Fayḥān ibn Ziribān, a shaykh of the Dowsān clan of the ʿIlwā group of the large Najdi tribe Muṭayr (cf. Musil, *ibid.*; Ibn Subayyil 2018: 264), and a response to this poem also published by Musil (1928a: 292-297) was composed by ʿAbdallāh ibn Subayyil, a famed poet who lived a sedentary life in the small Najdi town of Nifi, and who was also of the Muṭayr²⁰. Another poem (*ibid.*: 108-109) contains several lines which appear in a poem claimed elsewhere to have been written by a bedouin in Sinai (Bailey 1991: 122); this would be an interesting example of *nabaṭī* poetry from a more peripheral area in Musil’s corpus, but Bailey contests the claim of authorship on account of the fact that the same lines are found in Musil and other sources. Musil attributes the lines to Muḥammad al-Gāḍī, said to be a judge from the town of ʿAnayza²¹.

¹⁹ “While *nabaṭī* poetry is traditionally that of the nomadic (now settled) Bedouin who live in the central regions of what became Saudi Arabia, poetry composed by the desert-dwellers of the contiguous regions of Sinai, the Negev, Jordan, Syria, and western Iraq is also usually included in the designation. The dialect poetry of the Gulf sometimes has the term loosely applied to it too, but some types of Gulf poetry are structurally, topically, and linguistically different and have a separate origin” (Holes 2011).

²⁰ These poems have been published in Kurpershoek 2015 and Ibn Subayyil 2018: 52-61; the former also contains Musil’s version.

²¹ It may be noted that the inhabitants of ʿAnayza are particularly proud of their *nabaṭī* poets (cf. Sowayan 1985: 11, Johnstone 1967b: 1). Henkin also casts doubt upon the Sinai origin of the lines, writing that she has identified the poem “in five Arabian and Syrian versions from as early as the 19th century in different sequences and in different narratives” (Henkin 2009: 264).

Much of the poetic content in Musil is referenced in this way and is known to have been composed by members of other tribes; even amongst that which is not, however, it seems likely that the poetry he recorded represents a mixture of poems composed by the Rwala and poetry from further afield. The specific difficulties of analysing this poetic material are discussed further in §1.4.1.3.

1.4.1.2 *yabġi* ~ *yabi*: a case study in variation

The implications of the ‘mixed’ nature of the corpus for the proposed study may be illustrated with an examination of one example of variation found in Musil’s data: the P-STEM of the verb *baġa* ‘he wanted’, which is attested both as *yabġi* and as *yabi* (the latter showing deletion of the second radical, *ġ*). These forms are attested in a ratio of 3:1, and while *yabġi* is the clear majority variant, the data nonetheless give one the impression that *yabi* is a significant variant (occurring as it does in a quarter of cases).

Closer examination reveals a different picture, however. Out of the thirteen instances of *yabi* forms, five (possibly six) occur in stories which were told to Musil by the aforementioned Bleyhān ibn Meṣreb (e.g. Musil 1928a: 404), who is of the Sba‘a tribe. The speaker in a sixth example, a certain Ḥulfa the daughter of Ḥleyf Āl Ḥdēreš, is also of the Sba‘a. The corpus demonstrates that Bleyhān consistently used *yabi* forms in the tales he recounted to Musil (assuming that Musil recorded these faithfully); this furthermore suggests that *yabi* forms may have been characteristic of the speech of speakers from the Sba‘a tribe.

Four further examples of *yabi* forms occur in poems which are known to have been composed by poets who were not members of the Rwala: three in a poem by Feḡṣal al-‘Angari, a sedentary villager from al-Baṣra (Musil, *ibid.*: 212-214), and one in the aforementioned poem by ‘Abdallāh ibn Subayyil (*ibid.*: 292-297)²². The sources of the remaining three examples of *yabi* forms are harder to pinpoint, but it is significant that at least ten of the thirteen attested *yabi* forms are traceable to speakers from outside the Rwala, or in poetry which was not Rweyli in origin.

The attestation of *yabġi* forms, which are the majority, appears to support the view that *yabġi* was the dominant form among the Rwala at the time of Musil’s sojourn among them. Of named poets, we may cite Zeydān ar-Rweyli as a poet who used *yabġi* (*ibid.*: 360-361), whose name unambiguously marks him as a member of the Rwala. Others are Meš‘ān al-Nseyri of the Āl Mur‘aḏ (*ibid.*: 607-613) and “a daughter of Ḥeġlān of the Duġmān clan” (*ibid.*: 632-633): both Āl Mur‘aḏ and Duġmān are factions of the Rwala²³. Furthermore, and in contrast to *yabi* forms, the vast majority of *yabġi* forms are not attributed to named speakers, probably precisely because the speaker was not an outsider to the tribe.

²² Interestingly, a *yabġi* form is also attested in this poem. Not enough is known about the transmission of this poem to draw firm conclusions about the significance of this fact, although it should be noted that the same hemistich reads *tabiy* in Kurpershoek’s edition of the same poem (Ibn Subayyil 2018: 58) as opposed to Musil’s *tabġi*. The reciter of the poem was one Ṭrād ibn Saṭṭām, a man who seems to have grown up among the Rwala but whose mother hailed from the Sirḥān tribe, and it is recorded that he learned poems from her (Musil 1928a: 46, 197). The limited evidence shows that Ṭrād recited both *yabġi* and *yabi* forms in the poems where they occur.

²³ Cf. Lancaster (1997: 193); the names are Murath and Doghman in Lancaster’s orthography.

1.4.1.3 The prosody of bedouin oral poetry

Bedouin oral poetry, or *nabaṭī* poetry, employs an elevated style of language which diverges from speech in several important ways. The first is its use of specific vocabulary, which is often obscure to the uninitiated. Holes writes,

Nor is it true that the *nabaṭī* poetic vocabulary has ever been identical with the everyday speech of those who compose it. Like that of many poetries, the vocabulary of *nabaṭī* poetry is part of an elevated diction acquired through a process of listening to and learning from established poets and imitating their style. Much of traditional *nabaṭī* poetry is unintelligible to the Arab layman, whether he is of Bedouin descent or not, such is the degree of its divergence from speech. (Holes 2011)

Another is that the poetry imposes formal constraints on the language. The commonest type of poem, the *gaṣīda* or ‘ode’²⁴, is composed of lines which are divided into two hemistichs, each of which has a different rhyme, which is regular throughout the poem. The metre is also regular throughout, and is chosen from many possible options; Sowayan (1985: 159), for example, recorded 51 unique metrical combinations. The metre is quantitative, and is constructed of feet which are made up of (usually three to four) syllables; the syllables may be heavy (C \bar{V} or CVC) or light (CV or C). The resulting syllabic matrix has certain important effects on the language, described by Holes as affecting:

- a) supra-segmental phonology: the words must fit (or be made to fit) into the fixed quantitative metrical matrix of long and short syllables that is maintained throughout the poem;
- b) syntax: given that the line is the basic sense unit, and each hemistich must rhyme, phrase structure is sometimes compressed in ways unusual in normal speech, and word order has to be flexible. (Holes 2011)

²⁴ Note that in Musil, as generally (cf. Holes 2011), the term *gaṣīda* is also used loosely of any *nabaṭī* poem and not just of those which have the specific structure described here.

To illustrate these considerations, the first three lines of a poem from Musil (1928a: 306-307) are given here along with Musil's translation. The poem is attributed to a black slave who was accompanying a trade caravan from Iraq to Ḥāyil, destined for the famed ruler Ibn Rašīd (who ruled 1869-1897, cf. al-Rasheed [2010: 24-26, 37]), in whose honour the poem was written.

- 1) *yā-heyh yā-hal ḥāliyāt al-mezāhib*
heḡn^{en} telāfa min ba'īd al-maḡibe
- 2) *šeyb²⁵ al-ḡawāreb sāyiḡāt al-maḥāḡīb*
elfēn mašdūdāt kill šī^{en} teḡīb-eh
- 3) *alli 'aley-hen muštāhīn al-ma'āzīb*
šāmow l-rā'ī musned^{en} yirta'ī b-eh

Glory to you, O riders! with your bags empty,
On female riding camels coming from far countries,
With shoulder blades grey and breast girths shifting;
Of saddled ones two thousand are bringing us goods of all kinds.
Those who ride them are yearning for hosts,
They look for him who leans against a cushion.

These lines scan almost perfectly in Musil's rendering, and serve as an illustration of the way in which bedouin oral poetry is structured²⁶. Specifically: the whole line is read as a series of syllables irrespective of word boundaries; word-final vowels are analysed as long (see *telāfa* in the second hemistich of the first line which scans as te.lā.fā); CVCC or CVVC syllables are broken into two, CVC.C or CVV.C respectively (see *heyh* in the first hemistich

²⁵ Note that *šeyb* here seems to be a variant of *šīb*, a plural of the adjective *ašyab* 'grey-haired', as the noun *šeyb* means 'grey hair'.

²⁶ Palva (1993: 82) also scans and comments on these verses from Musil, considering them to display a "relatively regular meter".

which scans as hey.h)²⁷. A full account of this framework is given in Sowayan (1985: 153).

This is shown in the following schematic rendering:

–	–	◡	–	–	–	◡	–	–	◡	–	–
yā	hey	h	yā	hal	ḥā	li	yā	tal	me	zā	hib
heḡ	nen	te	lā	fā	min	ba	‘ī	dal	ma	ḡī	beh
šey	bal	ḡa	wā	reb	sā	yi	ḡā	tal	ma	ḥā	ḡīb
‘el	fē	n	maš	dū	dā	t	kil	šin	te	ḡī	beh
‘al	lī	‘a	ley	hen	muš	ta	hī	nal	ma	‘ā	zīb
šā	mow	l	rā	‘ī	mus	ne	den	yir	ta	‘ī	beh

The metre is – – / ◡ – – – / ◡ – – / ◡ – –²⁸, which is sometimes referred to as the neo-*ṭawīl*²⁹ metre and is said to be the most common metre for modern *nabaṭī* poetry (Palva 1978: 13) (it is also by far the most common metre in Musil’s collection). In this case it is acephalic (i.e. missing the first syllable of the hemistich) and catalectic (i.e. missing the last

²⁷ Note that in Palva’s version (1993: 82) short vowels are (re)inserted in CVC.C and CVV.C sequences; e.g. the opening *yā heyh* is scanned as *yā.hey.hə*. There is some evidence that these short vowels are less likely to be heard in recited poetry than in that which is chanted or sung. Sowayan (1985: 152) cites Wallin (1852: 193) and de Landberg (1895: 17ff), who observed that short vowels which are elided in the spoken language were also elided when *nabaṭī* poetry was recited, but were heard when the poetry was chanted or sung. He writes, “the elision of these vowels gives rise to consonant clusters which are permitted in ordinary conversation and poetic declamation. To enunciate the elided vowels while reciting would make the poetry sound stilted, but it is almost impossible to sing or chant *nabaṭī* poetry without these vowels” (Sowayan, *ibid.*: 153).

²⁸ This corresponds to metre 48 in Sowayan’s inventory of 51 *nabaṭī* metres (Sowayan 1985: 159).

²⁹ This term refers to the *ṭawīl* metre of classical poetry, which had the metre ◡ – x / ◡ – x – / ◡ – x / ◡ – ◡ – (Wright [1967, vol.2: 364]; x indicates a syllable which may be long or short). The final foot of the second hemistich may also be ◡ – – or ◡ – – – in the classical *ṭawīl* (*ibid.*). Sowayan (1985: 139) notes that the neo-*ṭawīl* is also called *hlālī* on account of its popularity in the days of Banu Hilāl.

syllable of the hemistich)³⁰. The rhyme of the first hemistich is Cīb and the rhyme of the second hemistichs is beh³¹.

There are two instances of the transcription not fitting within this metrical structure, highlighted in bold in the above scheme. The first is *kill šien*, which must be read as *kil šin* if it is to adhere to the metre; the second is *muštāhīn*, in which the second syllable must be short (which is, in any case, in line with the usual pattern of FORM-VIII participles in this dialect, cf. §4.8.1.7). We may compare similar occurrences in a very structurally different poem (Musil 1928a: 12):

<i>erčebū-ni-l-ḥamra</i> <i>w-eṭ'emū-ni tamra</i> <i>allāh yeṭūl 'omr-ah</i> <i>yowm hom ḥallaṣū-ni</i>	Seat me upon a bay-mare, Feed me with one date. Oh, may Allāh prolong his life, For they have liberated me.
--	--

	-	◡	-	-	-	◡
	'er	će	bū	nil	ḥam	ra
	weṭ	'e	mū	nī	tam	ra
(al)	lā(h)	ye	ṭaw	wel	'om	ra(h)
	yow	m	hom	ḥal	ṣū	ni

³⁰ “As a rule, the new *ṭawīl* has a long syllable where the old *ṭawīl* permits either long or short ... the first syllable of each hemistich is optional in the new *ṭawīl*” (Palva 1978: 14). In other words, the neo-*ṭawīl* is (◡) -- / ◡ --- / ◡ -- / ◡ -- (-). Sawayan defines a foot as “a unique combination of one short syllable and two or three long syllables” (Sawayan 1985: 156) and therefore divides the acephalic and catalectic *ṭawīl* into feet differently: -- ◡ - / -- ◡ - / - ◡ -- (*ibid.*: 159), although he notes that this differs from *ṭawīl* only in the missing initial short syllable (*ibid.*: 162).

³¹ *maḡībe* scans as ma.ḡī.beh, as all of the second hemistichs in this poem rhyme with beh. The significance of this should not be overstated, since *-eh* for final *-a/-e* is fairly frequent in Musil.

Again, the transcription does not fit the metre in two instances. The first is in fact rather serious, as *yeṭūl* needs to be reconstructed as a different verb, *yeṭawwel*, in order to fit into the syllabic matrix. The occurrence of *yeṭūl* is also lexically problematic, as the FORM-I of the root **ṭwl** has only the stative sense of ‘to be long’, when what is required here (both metrically and for the sense) is the FORM-II verb *ṭawwal* ‘to make long’, i.e. ‘may God prolong his life’³².

The second is that *ḥallaṣū-ni*, which as transcribed conforms to the general pattern of FORM-II verbs in this dialect (cf. §4.8.1.1), scans as ḥal.ṣū.ni. This is not uncommon in the scansion of *nabaṭī* poetry, and we may compare Henkin’s treatment of Musil’s *delleṭ^{en}* ‘pot’ which scans as del.ten (Henkin 2009: 265). Henkin notes that “vowel reduction ... is not unusual in scansion of *nabaṭī* poetry, when necessary for the metre” (*ibid.*: note 42), and these two examples demonstrate that this applies not only to the deletion of vowels but also to the shortening of geminate consonants.

It is impossible to know whether it was the case that Musil heard *ḥalṣū-ni* but “corrected” it to *ḥallaṣū-ni*, or whether this was how the reciter pronounced it. As observed above, recited (as opposed to chanted or sung) poetry adheres more closely to the prosody of the spoken language, so it is possible that *ḥallaṣū-ni* is what he heard, despite its metrical

³² Cf. two expressions recorded by Sowayan: *ṭāl* ‘*umr-ak* and *ṭawwal allāh* ‘*umr-ak*, described as expressions “used by a narrator in addressing his audience as a sign of courtesy and decorum and as a means of establishing and maintaining rapport” (Sowayan 1992: 248).

irregularity. In the case of the unmetrical (and nonsensical) form *yeṭūl*, however, it seems likely that this was Musil’s mistake, as the version with *yeṭūl* makes no sense.

For a grammatical description of the dialect as recorded by Musil, it is therefore important to consider the ways in which issues such as these—which specifically pertain to the poetic material—could influence the data. We may consider a further example (Musil 1928a: 292)³³:

5) *beš-šadd waṭnāt w-be-l-maši ‘aḡlāt*
raff³⁴ al-mesāme‘ w-an-nwāḏer eḥdādī

At saddling they remain quiet; they hasten on the march,
 Carry their ears high, and their eyes keep sharp lookout.

–	–	∪	–	–	–	∪	–	–	∪	–	x
beš	šad	d	waṭ	nā	ten	w	bel	maš	y	‘aḡ	lāt
raf	fal	me	sā	me ^c	wan	n	wā	ḏer	ḥ	dā	di

An initial observation is that Musil’s text did not show the adnominal linker *-en* on the word *waṭnāt(en)*, although this *-en* is in fact necessary for the line to scan properly; the form with *-en* is attested in the edition of the same poem edited by Kurpershoek, however (Ibn Subayyil 2018: 54). Another is that the final word, *ḥdād*, has been recorded by Musil with prothesis (*eḥdād*) even though medial epenthesis (**ḥedād*) would fit the metre in this

³³ x has been given in the final syllable to denote a possible variable: in the first hemistich the rhyme is always a superheavy syllable CāC, while in the second hemistich it is always transcribed as a light syllable Ci. We have seen that word-final vowels are long in scansion, but this is not recorded in the transcription either line-internally or line-finally.

³⁴ Note that *raff* is misprinted in Musil as *ḡaff* (*ḡ* is represented by <ṛ> in Musil).

case. The final *-i* at the end of the second hemistich is otiose and occurs frequently in the poetry.

More important for the discussion here is that the form *waṭnāt* is, in fact, an active participle from the verb *waṭan* ‘be immobile, stay (in a place)’, the usual form of which (in Musil’s orthography) is **wāṭenāt* (see §4.4.1). The active participle in its regular form does not fit within the syllabic matrix of this poem, however, so the resulting form shows long-vowel shortening³⁵. This should of course not lead one to conclude that CaCCāt is a regular variant form of the FP active participle, but rather illustrates the type of variation in vowel length which is permitted in poetry when required by the metre.

The significance of all these illustrations is that the metrical constraints of the poetry (and poetic content makes up a majority of the data in Musil) has, in many cases, clearly introduced a measure of variation into the data in terms of the structure of attested forms.

We have seen the following types attested:

- (a) forms transcribed according to their regular dialectal form (in speech) which do not fit the metre (*ḥallaṣū-ni*, *kill šien*)
- (b) forms which fit the metre as transcribed but differ from their regular dialectal form (*waṭnāt*)
- (c) forms which fit neither the metre nor the regular dialectal patterns (*muštāhīn*)
- (d) forms which appear to be unambiguously erroneous (*yeṭūl*)

³⁵ It seems likely that the following form *‘aḡlāt* is also an active participle (< *‘āḡelāt*), mirroring *waṭnāt* (< *wāṭenāt*) in the first half of the phrase, although *‘aḡ(e)l* is also an attested adjective (see Glossary).

The effects of the poetic structure on much of the data must therefore be taken into account for the analysis proposed in this thesis.

1.4.2 *Digitisation of the corpus*

The amount of linguistic data in the corpus is huge, and has therefore been digitised in order to allow for the analysis presented in this thesis. This has been achieved through the use of a custom-made database created on FileMaker Pro 13 which has allowed for the entire corpus to be digitised with full grammatical tagging, which is described here.

Level one tagging: every prose passage, verse of poetry, proverb and isolated technical term was entered with its translation. The total number of entries was 6,531. At this level, each constituent was “grammar tagged” as one of the following categories: verb, substantive, pronoun, particle, adverb, adjective, numerical and onomastic.

Level two tagging: each grammar tag was then tagged for specific linguistic criteria. The total number of grammar tags was 23,069. Each tag was further specified with its own translation, lexical root and source type: conversation, proverb, poetry or technical term. The tags specific to each category were as follows:

Verb: form (FORM-I–FORM-X, quadrilaterals FORM-QI–FORM-QIV); conjugation (S-STEM, P-STEM); voice (active, passive); mood (indicative, subjunctive, jussive, imperative); person (first, second, third); gender (common, masculine, feminine); number (singular, dual, plural); active participles, passive

participle and verbal nouns were tagged as such; verbal suffixes were tagged either as “no suffix” or as the relevant person/gender/number, e.g. 3MS.

Substantive: definiteness (indefinite, defined by definite article, defined by suffixation, defined by construct); gender (as above); number (as above, but including collective); case (nominative, accusative, genitive, oblique, adverbial accusative); suffix (as above).

Pronoun: gender (as above); number (as verbs); type (free, bound).

Particle: class (conjunction, interjection, negative, preposition, vocative); suffix (as above).

Adverb: no further tags.

Adjective: as substantive, and also tagged as elative where applicable.

Numerical: no further tags.

Onomastic: no further tags.

Additional tags also allowed all entries to be tagged if: they had a case ending (this was mostly used for tagging forms with poetic suffixes); they were diminutive; or they had the adnominal linker (*tanwīn*).

This high level of tagging rendered the corpus fully searchable. This was of immeasurable value for producing the description of the language presented in the following chapters. It meant that, for example, all FORM-II verbal nouns or nouns with the 1CS suffix could be extracted immediately from the data, and the root search functionality meant that

all instances of specific classes, e.g. FORM-I imperatives of hollow verbs, could also be extracted.

1.4.3 *Objective 1: description of the dialect*

The digitisation of Musil's corpus has facilitated the presentation of a thorough description of the linguistic material set out according to the following categories: phonology; nominal morphology; verbal morphology; and closed classes (i.e. particles, pronouns, etc.). Syntax has not been included because it would have been an undertaking beyond what time has allowed for, and because the sections included were considered to be the most useful for comparative purposes and dialect classification.

In light of the discussion in §1.4.1.1, it is clear that the resulting description cannot be categorically defined as representing “the Rwala dialect” in any meaningful way. It is impossible to calculate how much of the material comes from speakers who, though Musil encountered them while living among the Rwala, hailed from other tribes and regions, and it is certain that the poetic material, whatever its origin, is composed in the elevated register of *nabaṭī* poetry and is therefore not representative of regular speech. Furthermore, it is important to stress that, even without these considerations, the Rwala represent a living speech community which must be expected to exhibit variation. In his preface, Musil wrote of those he observed, “I have not corrected their many inconsistencies in grammar and pronunciation” (Musil 1928a: *xiii*), indicating that he conceived of a “pure” version of the

dialect somehow corrupted in the mouths of its speakers: this would be an incorrect assumption about any speech community, as has been demonstrated by sociolinguistic research after Musil's time.

In addition to the dialectal variation already outlined, there are further social variables which could be expected to introduce variation between speakers into the corpus. Musil recorded speech from speakers of different ages, and from those whose social status within the tribe ranged from servant to shaykh. Men and women are also generally expected to use language differently, although in this corpus the data are almost exclusively from men. Furthermore, variation should be expected at the level of the individual speaker, since interactions between these members of these social groups will elicit different styles from them. The situational context of the recorded utterances ranges from humdrum conversations between family members to legal debates before a judge, and from popular riddles to lengthy retellings of important battles, culminating at the most sophisticated end of the spectrum with the performance of poetry. Any individual speaker will perform language differently in these varied contexts.

The ensuing discussion aims, therefore, to provide an account of a set of shared linguistic rules discernible from the corpus. On account of the known divergence of the poetic language from the patterns of the spoken language, the approach taken has been that the non-poetic material has been assessed first, and an appraisal of the poetic material has been carried out subsequently. Where the poetic material agrees with the rest of the

material on any given point, this has all been used as a source for the description of the dialect, but where the poetry attests forms which differ from what is found in the rest of the corpus this has been highlighted. It has been interesting to discover that in the majority of cases the poetic material agrees with the spoken material in matters of phonology and morphology. This approach has been preferred, since to completely separate the material into a description of the spoken language and a description of the poetic register would have necessitated frequent reference to the source type in a way which has proved to be unnecessary with regard to many, if not most, of the categories concerned.

1.4.4 *Objective 2: evaluation of Musil's data*

The dialect as recorded by Musil and described as set out in the previous section has been compared with the available source material previously listed. In this matter, the other sources of the Rwala dialect (cf. §1.2.1) have provided the first point of comparison, and the central Najdi sources (cf. §1.3.3) have been consulted with priority over the other sub-groups, since these should be the most closely related dialects to that of the Rwala. Comparison with dialects from other sub-groups has been made where Musil's material does not conform to the expected patterns of the central Najdi group, or where the issue being discussed is not sufficiently attested in the available material. Naturally, some matters are specifically addressed in secondary sources with regard to a less closely related dialect and in these cases comparison with these dialects has been a natural outcome. Likewise, in the

discussion of phenomena which are considered to be characteristic of the whole north-Arabian group or of bedouin dialects in general, comparison with central Najdi dialects has not been so prioritised.

It will be observed that the chapters of the thesis show different levels of engagement with the secondary sources, and this is on account of the type of information available. For example, Chapter 4, which deals with the verbal system, makes extensive use of the very thorough account of the verbal system of Rwala and other central Najdi dialects provided in Prochazka (1988) and the comparable material in Cantineau (1936-37), and these accounts are rich enough to mean that the verbal system can be appraised in comparison with these descriptions in almost all regards. In Chapter 5, on the other hand, which deals with the closed classes of particles and pronouns (among other things), it has been necessary to draw from a much wider range of secondary sources.

Two important considerations must be taken into account when making comparisons between Musil's material and that of the other listed sources. Firstly, and with the exception of Cantineau (1936-37), most of these studies were carried out considerably later than the date of Musil's explorations. This means that even those comparative examples which are specifically attributed to speakers of the Rwala tribe must be treated as representing a language variety after many decades of language change. This is compounded by the fact that the context in which the data were collected is often vastly different: Musil gathered his material from speakers who still lived a nomadic life in their tribal *dīra*, but Prochazka's

Rwala material (for example) was collected in Riyadh and London (Prochazka 1988: 10) from speakers who, it can be assumed, are distanced from Musil's not just in temporal terms but also with regard to background, lifestyle, and opportunities for language contact.

The second (related) consideration is the method in which the data were collected. Musil's ethnographic approach entailed years of continuous observation, with the researcher himself reaching a high degree of fluency in the language, which contrasts starkly with the sort of linguistic surveys on which most of the other studies are based. The probability of any particular grammatical category being attested in Musil is increased by the volume of data he collected, but other scholars were able to provide complete paradigms for the categories they examined since their method was direct elicitation. The usefulness of the latter approach is clear, but it must not be disregarded that this method can produce less natural data; with the exception of Ingham's (1995) texts (which were recorded by the Lancasters), it is rare outside Musil to find published examples of Rwala speech consisting of more than a few words.

This difficulty may be briefly illustrated by returning to the matter of *yabgi* ~ *yabi* variation. The variation in the P-STEM forms of this verb is well known from the dialectological literature: in the later sources, *yabi* is considered to be a Najdi feature (Ingham 2011, Ingham 1994: 26), and it is also the major variant in Šammari dialects according to Sawayan who notes that "the pharyngeal medial radical occasionally shows

up” (Sowayan 1992: 249)³⁶. Around the time of Musil, however, Cantineau recorded *yabi* only for the Šammari tribes³⁷ and recorded *yabġi* type forms for the ‘Anazi tribes, citing only Musil as evidence that *yabi* forms also occurred in ‘Anazi dialects. Prochazka, on the other hand, later recorded only *yabi* for the Rwala and related dialects (Prochazka 1988: 94-95).

This demonstrates the importance of temporal considerations, since Cantineau’s data appear to support the evidence from Musil that *yabġi* was the major variant among the Rwala at that time; this evidence should not be disregarded on the grounds that Prochazka recorded only *yabi* for the Rwala seventy years later. It further suggests that the variation attested in Musil’s corpus (which, we have noted, includes dialectal variation as well as internal variation) is more likely to be recorded through the type of long-term observation of multiple speakers which Musil undertook, as there is no reason for a single speaker in an interview situation to consciously provide evidence of this.

1.4.4.1 The status of Classical Arabic

On the matter of the comparison of Musil’s material with other sources from the dialectological literature, it should be noted that in some cases it has also proven useful to draw a comparison with Classical Arabic (CLA). It is important to stress that doing so is in

³⁶ *yabi* is furthermore widely attested along the coast of eastern Arabia (cf. Johnstone 1967a: 78, 125), and very early sources attest to its occurrence in southern Arabia (Landberg 1901: 433, Landberg 1905-13: 26, Reinhardt 1894: 246).

³⁷ Cantineau in fact claimed that he heard *yābi*, despite the records of previous accounts (Cantineau 1936-37: 198).

no way intended to imply that the Arabic recorded by Musil, or indeed any modern Arabic dialect, is descended from CLA. CLA as a variety of Arabic was codified in the eighth century by grammarians in southern Iraq, the sources they used being mainly the Qur'an, pre-Islamic Arabian poetry, and the spoken language of bedouins from Arabia, whom they considered to speak the purest form of Arabic. The fact that they codified it at all is evidence of dialectal variation within Arabic both at and before their time, and the once widely-held view that modern vernaculars are descended from—and even corruptions of—CLA has, in recent decades, been consigned to the status of a cultural myth³⁸.

In the description of Arabic dialects, the use of terminology which traditionally describes the structure of CLA can be misleading. As one example, some scholars writing about the Rwala verbal system have referred to verbs of the type *šireb/širib* 'he drank' as *fa'ila* (Prochazka 1988: 30) or *qatila* (Cantineau 1936-37: 79), in a way that could suggest that the authors felt that CLA was the ancestor dialect. I have referred to these verbs as the I-CLASS, which is arguably a more neutral term than *fa'ila* or *qatila*. This is an illustrative example, as in this thesis I have proposed the underlying representation /šarib/ for *šireb*, which outwardly resembles the CLA form *šariba*. Mine is a synchronic derivation, however, which is based on alternations in the surface forms *šireb* and *šarbat* 'she drank' (cf. §4.2.4) and which does not make any claim about historical changes; furthermore, it is in no way meant to imply that the CLA form is the one from which the modern forms have developed.

³⁸ Cf. Holes (2018: 2-9) for a fully referenced account of these arguments.

In other cases it has seemed less practical to change widely-used terminology, and some terms used to describe CLA are also used here. An example is the description of the derivational verbal system using labels such as FORM-I: in these cases, this has been felt to be the most useful way of describing these structures, and such labels should be read on their own terms without any assumption that the system described is identical to that of CLA. Similarly, the occasional use of Arabic grammatical terminology such as *maf'ūl mutlaq* (cf. §4.6) or *damir al-faṣl* (cf. §5.2.2) is purely meant to be illustrative, in cases where CLA has similar structures, and does not imply direct descent.

Occasional reference has been made in the thesis to Old Arabic (OA), which is an umbrella term used in reference to the earlier varieties of Arabic which are considered to be the putative ancestors of the modern dialects. Unlike CLA, for which we have direct evidence, OA must be reconstructed using the methods of historical linguistics.

1.4.5 *Objective 3: creation of a glossary*

Producing a glossary of the linguistic data from Musil (1928a) has been an aim of this project since its inception. It has been noted that several glossaries of related dialects make use of Musil's material (notably Kurpershoek [2005], Sawayan [1992] and Holes [2001]), but this is necessarily restricted, since cross-references are only made where a term occurs in another corpus. It has therefore been considered a desirable outcome of this work to produce a complete reference to Musil's material, ordered by root so that the

comparativist may easily make use of Musil's vast record of the dialect. In this way, the glossary (found in Volume II) performs something of a hybrid function for this project. The presentation of the material is based on the analysis from the following chapters in terms of the grammar of the language. At the same time, the glossary provides the source data on which that analysis is based; for this reason, the decision has been taken not to include unattested forms as headwords, even though this could have led to simpler presentation. More information on how this has been presented is provided in the Guidance Notes section of that Volume.

2 Phonology

This chapter begins by attempting to establish the likely sound system of the Rwala dialect, and by suggesting how Musil's transcriptions should be interpreted in the light of other evidence about the sounds of the language. The syllable structure of the language is then addressed, paying particular attention to the well-established processes of epenthesis and syncope attested for this dialect. Finally, other phonological processes attested in Musil's record are discussed, with special focus on the well-known process of the *ghawa*-syndrome and the matter of low vowel raising. Given that Cantineau once claimed that "il ne faut certainement accorder aucune créance aux transcriptions de Musil" (Cantineau 1936-37: 36), this chapter evaluates the fidelity of Musil's transcriptions by making reference to the other sources of the dialect of the Rwala available, as well as cross-dialectal evidence and recent discussion on phonological processes in Arabic dialects more generally.

2.1 The sound system

The following sections describe the consonant and vowel systems of the Rwala dialect as they were probably heard by Musil. It should be stressed at this point that Musil's records contribute very little by way of phonetic information, and the articulatory realisations discussed are mostly drawn from comparative evidence (predominantly from Cantineau) with a view to elucidating Musil's sometimes obscure system of transliteration.

His corpus does, however, inform us of the distributional contexts for certain phonological questions such as glottal stop deletion, affrication, and vowel-related processes.

2.1.1 Consonants

Bedouin dialects of the Arabian peninsula are known to be conservative in terms of their retention of most features of the Old Arabic (OA) phoneme inventory (Holes 2004a: 70, Watson 2002: 13-14), and the Rwala dialect is no exception. It exhibits twenty-six (possibly twenty-eight) consonantal phonemes at (probably) ten places of articulation (see the following table), compared to the OA twenty-eight, at nine places of articulation. The loss of phonemes is because of the merger of historic *ḏ and *ḏ̣, and because of the virtual collapse of the glottal stop as a distinct segment, but the list is augmented by the marginal phoneme /ć/. The additional place of articulation is palato-alveolar, which is represented by /š/ (and by the possible [čʃ] reflex of /ǧ/), which is a widespread development not particular to this dialect group.

The consonantal phonemes depicted in Table 2.1 are approximations of the sounds probably heard by Musil, based on descriptions of the dialect from Cantineau (1936-37: 128-145) and Prochazka (1988: 14-17). Where sounds are given in pairs, the left of the pair is voiceless and the right of the pair is voiced. Sounds in brackets () are allophones of the phonemes shown in superscript / / . *w* is listed in two columns in the table, as it is produced at two simultaneous places of articulation (labial-velar). The symbols are those used for

transliteration of words in the dialect throughout this work; in addition to the emphatic consonants (ʂ, ʈ and ɖ) which are marked with an under-dot, the following symbols adopted here depart from standard IPA notation (IPA shown in square brackets): ʾ [ʔ], ǰ [j], ʕ [ts], t [θ], ɖ [ð], ʃ [ʃ], ǧ [ɟ], ɸ [χ], ǧ [ɸ], ɦ [h], ʕ [ʕ], and y [j]³⁹.

TABLE 2.1: Consonantal phoneme inventory for the Rwala dialect

	bi-labial	labio-dental	inter-dental	dental/alveolar	palato-alveolar	palatal	velar	uvular	pharynx-geal	glottal
plosive	b			t, d		ǰ	k, g			ʾ
<i>emphatic</i>				ʈ						
affricate				(ʕ)/k/, (ǧ)/g/						
fricative		f	ʈ, ɖ	s, z	ʃ			ɸ, ǧ	ɦ, ʕ	h
<i>emphatic</i>			ɖ	ʂ						
nasal	m			n						
lateral				l						
tap				r						
glide	w					y	w			

Sections 2.1.1.1-4 focus solely on those consonants over which there is some dispute, and those which differ significantly from what was described by Musil. There are some departures since the author provides a transliteration chart in the preface (Musil 1928a: xiv) which does not lend itself to clear phonetic interpretation. The phonemes specifically addressed are: /ʾ/, which is routinely deleted; /ǰ/, the phonetic value of which is very uncertain; /ɖ/, which illustrates an important historical merger with OA *ɖ; and /k/ and

³⁹ A possible realisation of /ǰ/ is palato-alveolar [ɟ] but it is listed here as palatal because most realisations of this sound are palatal, and the sound does not form a group with other coronals in exhibiting assimilation of the definite article (see §2.1.1.2).

/g/, which have the allophones *ć* and *ǧ* in front environments (the realisations of these sounds are described in the relevant sections).

It should be mentioned at this point that while *ć* appears in brackets in the above chart, it does in fact have marginal phonemic status within the dialect. /*ć*/ is contrastive with /*k*/ in at least one lexeme:

ćeyf ‘how?’ *keyf* ‘satisfaction, pleasure’

Furthermore, the distinction between the 2_{MS} bound object/possessive pronoun *-(a)k* and the 2_{FS} pronoun *-(e)ć* has been grammaticalised so that /*k*/ and /*ć*/ are contrastive in identical phonological environments:

warā-ć ‘behind you (f.)’ *warā-k* ‘behind you (m.)’

2.1.1.1 /ʔ/

Some weakening of the glottal stop is a feature of the majority of modern Arabic dialects (Watson 2002: 18). Glottal stop weakening had already begun in OA (Kaye and Rosenhouse 1997: 227) and its loss is sometimes maximally considered to be a feature of all the modern dialects in distinction to CLA (Versteegh 1984: 21, Fischer and Jastrow 1980: 52). It has, however, been systematically retained in some Yemeni dialects (Behnstedt and Woidich 2005: 11) and its retention in bedouin-type dialects more readily than in urban dialects has also been noted (Cadora 1992: 13).

The glottal stop is deleted in all positions in the Rwala dialect according to Musil's data. The loss of non-radical glottal stops (i.e. in some conjunctions, pronouns and prepositions) has generally had no effect on adjacent segments: forms include *en* 'if, when', *ann(a)* 'that' and *ayy* 'which?' (compare the corresponding CLA forms *'in*, *'anna* and *'ayy^{un}*), although there has been an effect in some forms, e.g. *ćenn-eh* 'like' (not **ćayinn-eh*, compare CLA *ka-'inna-hu*). When the glottal stop is radical or morphemic, it has sometimes left phonological traces, as attested in the following positions:

Word-initially:

(a) with no effect on adjacent segments, resulting in a word beginning with a vowel, e.g. *ahl* < **'ahl* 'kin', or with a consonant in the case of words with onset clusters (usually as a result of prior *ghawa*-syndrome resyllabification, cf. §2.3.1), e.g. *hal* < **'hal* (< **'ahl*);

(b) with the loss of the whole initial syllable, and with a word-final vowel being inserted, leading to resyllabification, e.g. *ħa.ða* < **'a.ħaḏ* 'he took';

Word-medially:

(c) in pre-consonantal position, with compensatory lengthening of the preceding vowel, whether this is in a word-final CV³C syllable, e.g. *rās* < **ra's* 'head', *ḏīb* < **ḏi'b* 'wolf', or across a syllable boundary CV³.C, e.g. *yāḥoḏ* < **ya'ḥoḏ* 'he takes', *ḏība* < **ḏi'ba* 'she-wolf', *būra* < **bu'ra* 'ditch';

(d) in post-consonantal position, with no effect on adjacent segments, after a syllable break e.g. *ma.lak* < **mal.ʾak* ‘angel’ (leading to resyllabification), or in a word-final syllable, e.g. *ḏow* < **ḏow* ‘daylight’, although the glottal stop may also be replaced with a glide, e.g. *arwā-ni* < **arʾā-ni* ‘he showed me’⁴⁰;

(e) intervocalically, between two vowels of the same quality, which produces a long vowel, e.g. *tār* < **taʾar* ‘he avenged’, *rūs* < **ruʾūs* ‘heads’; if one vowel is already long then the triply-long vowel is reduced as in *rūs* above;

(f) intervocalically, between two vowels of different quality, where the loss of the glottal stop results in two adjacent varisyllabic vowels, or ‘hiatus’⁴¹, which is resolved by glide insertion, e.g. *swāl* < (**suwāl*) < **suʾāl* ‘request’, *ḥāyeṭ* < **ḥāʾeṭ* ‘long wall’; the resulting form may often be subject to syncope, e.g. *swāl* above, or compensatory lengthening, e.g. *miyye* (= **mīye*) < **miʾe* ‘a hundred’;

Word-finally:

(g) with final syllable weakening, which results in the loss of the coda and the shortening of the vowel, e.g. *zerga* < **zergā* ‘blue’⁴², *nebi* < **nebī* ‘prophet’;

⁴⁰ There is also one instance of glottal stop deletion in this position triggering gemination of the preceding consonant (*mettaḥedīn* < **metʾaḥedīn* ‘considering themselves [active participle]’), although this does not seem to be a frequent process in the dialect.

⁴¹ Hiatus elimination strategies are described in Trask and Millar (2015: 59), and further in §2.3.5.

⁴² Note that Musil recorded free variation of *zerga* ~ *zergaʾ*, but these forms with final *-aʾ* seem to be erroneous, as is discussed in the following paragraphs.

(h) possibly with the preceding vowel being replaced with *i*, e.g. *abdi* < **abda* ‘I begin’.

Rule (h) merits some explanation: rather than this being a strictly phonological process effected by the loss of the glottal stop, it seems that a fairly wide-reaching phenomenon in this and other north-Arabian dialects is the reinterpretation of “hamzated” roots (i.e. roots in which one consonant is the glottal stop) as weak roots containing a glide (cf. Sawayan 1992: 81 on the same development in the Šammari dialects). This is not restricted to C3 = /ʔ/ verbs but appears to have affected the whole system: C2 = /ʔ/ verbs pattern like C2 = /w, y/ verbs (cf. §4.9.3), including some rare examples where a glottal stop is hypothetically geminated (*layyemna* ‘we assembled’ < **laʔemna*, *tefawwelū* ‘take [it] as a good sign!’ < **tefaʔelū*), and C3 = /ʔ/ verbs pattern like C3 = /w, y/ verbs (see §4.10.3).

There are also cases where initial glottal stops are realised as a glide in the dialect. This has already been documented as a fairly general phenomenon in Arabic dialects (Kaye and Rosenhouse 1997: 277). For the Šammari dialects it has been recorded as occurring in restricted positions, where C1 = /ʔ/ > C1 = /w/ in verbs of forms I-III, although in non-passive FORM-I this is optional (Sawayan 1992: 80). This parallels the usage in Musil where the following forms occur: *wuḥid* < **uḥid* ‘it was taken’ (FORM-I passive), *wahḥar* < **aḥḥar* ‘push [it] back!’ (FORM-II), and *twahḥar* < **tʔahḥar* ‘it [the saddle] slid backwards’ (FORM-V). The change is not systematic, however, and forms with /ʔ/ occur e.g. *teʔaddūn* ‘you (pl.)

return sth.’ (not **tewaddūn*) and (with syncope of the P-STEM prefix) *t’awwel* ‘you (sg.) lead [us] back’ (not **twawwel*). Such variation in the form of C1 = /ʔ/ verbs is not unattested in related dialects, however, e.g. Bahraini *yi’akkil-hum* ‘he feeds them’ and *in’akal* ‘it is edible’ (Holes 2001: 14). It is also not only in verbal forms that this occurs, as shown by nouns such as *wanne* < **’anna* ‘lament’ and possibly by some particles such as *weyn* < **’eyn* ‘where?’.

In spite of processes (a)-(h) above, which are more or less regular, Musil does transcribe glottal stops and it appears that in many cases these are erroneous. Of his more than four hundred transcribed glottal stops (all of which are lexical as opposed to epenthetic, cf. §2.3.5.3), 90% occur in word-final *-a’*. All of these represent either roots in which the third consonant is /w, y, ʔ/, e.g. *ḡada’* ‘lunch’ from the root **ḡdḡ** (compare CLA *ḡadā’*) or words with the feminine suffix which is represented as *-ā’* in CLA, e.g. *ḡamra’* ‘red’ (CLA *ḡamrā’*). These glottal stops appear to be completely invented for two reasons. Firstly, they represent the only case of consonants inducing elision of following vowels in phonological phrases. Musil transcribes *ḡafa’ l-beyt* ‘the long wall of the tent’, where the elision of /a/ of the definite article creates a medial CVCC syllable (ga.**fa’**l.beyt); this occurs whenever *-a’* is followed by the definite article with only one exception. Such elision is not permitted for any other consonant: for example, *’atab al-beyt* ‘violation of the tent’, which has the same syllable structure, does not produce an illegal syllable (’a.ta.bal.beyt and not ’a.**tabl**.beyt). The elision is explained by the deletion of the glottal stop, since *ḡafa l-beyt*

syllabifies like every vowel-final word before the definite article (ga.fal.beyt). Secondly, it is arguable that in forms such as *gada'* and *hamra'* (which freely varies with the form *hamra*) the final vowel is short precisely because the glottal stop has been deleted. Final-vowel shortening is a regular process in the dialect (see §2.3.3), and it is conceivable that a final consonant (in this case /ʔ/) would have inhibited this: compare final VV *ʿaša* 'he rebelled' (the vowel length is attested before a suffix, *ʿaša-ni* 'he defied me') with final VVC forms which do not undergo vowel shortening, such as *nahār* 'day'. This second argument does suggest, however, that the (very rare) cases of final *-ā'* are genuine, e.g. *šedā'* 'a spirit which spooks unaccompanied animals and is fond of the taste of brains'; these words presumably carry final stress (stress being unrecorded by Musil, however) and do not undergo final-vowel shortening (> **šedā* > **šeda*).

The other 10% of Musil's transcribed glottal stops represents a variety of exceptions to the general rules of glottal stop deletion set out above, and account for only a small minority of potential glottal stops (approximately 3%, the rest having been deleted). Although this is hardly statistically significant enough to challenge the rule that glottal stops are deleted in all environments, the anomalies are considered in the following paragraphs.

In some instances, rule (c) above is violated by forms such as *ta'ti-h* 'you come to him', *ma'hūd* 'taken [passive participle]' and *yista'hel* 'he deserves', but the forms in this

category are rather mitigated by the attestation of nearly identical forms which do exhibit glottal stop deletion, i.e. *tātūn* ‘you (pl.) come’ and *māḥūḍ, yistāhel* (same meanings)⁴³.

Rule (d) is violated in very few instances, one example being *er’ow* ‘see! (pl.)’: the root *r’y* is certainly irregular by virtue of being “doubly weak” (i.e. both C2 and C3 of the root are one of /w, y, ’/) and *er’ow* contrasts with forms such as *tarow* ‘you (pl.) see’ (not **tar’ow*) and the aforementioned *arwā-ni* (not **ar’ā-ni*), where the glottal stop has been deleted in both instances, and *er^c-eh* ‘see him!’ where it has been replaced by /^c/ (see below). The glottal stop is also transcribed in the noun *ra’y* ‘opinion’ from the same root, a form which, according to rule (c), should be **rāy*, as it is recorded in other central Najdi dialects by Kurpershoek (2005: 106) and in the Šammari dialects by Sawayan (1992: 267). With this form being altogether anomalous, we may consider the doubly weak roots to exhibit anomalous forms in general. Elsewhere, we find roots in which the third consonant is historically /’/ realised in the CaCC template (see §3.1.2.1) only very rarely; one example may be *nab’* ‘piece of news, information’ but other forms with /’/ in post-consonantal position seem to be anomalous, e.g. *ḥal’* as a variant of *ḥelw* ‘sweet’ and *ćenn’* ‘like’ which varies with *ćenn*, neither of which have etymological glottal stops.

⁴³ Several of these forms occur in an interesting poem the metre of which has each line ending in Ca³-ha, e.g. *ma³-ha* ‘her water’, *lāma³-ha* ‘her lips’, *ṭanāya³-ha* ‘her teeth’. These forms are quite clearly contrived for a poetic style, and have no strong implications for the phonology since the second two examples would never have underlying glottal stops anyway. As with the examples already cited, forms without glottal stops are attested e.g. *mā-ha, ṭanāya-h* (masculine rather than feminine suffix).

Rule (f) is violated in certain isolated forms. This is slightly more difficult to account for since the attested forms do not always freely vary with other attested forms. In some instances, like *šī^{en} ~ šay^{en}* ‘thing’ the forms are in free variation with *šī^{en}*, *šīyye* and the sub-minimal form *ši*, which is undoubtedly a situation associated with words which would be sub-minimal on deletion of the glottal stop (compare *miyye* above), which generally come from doubly weak roots (the minimal word is discussed in §2.2.2). In other cases, however, Musil records such terms as *garrāⁱ* ‘snake charmer’⁴⁴ and *raⁱs* ‘chief’ which clearly violate the glottal stop deletion rules, when forms such as **garrāyi* and **rayīs* would be expected from the general pattern of the language. For this, a justification might be found in the statement that “in Einzellexemen ist in vielen Dialekten der Hamz bewahrt” (Behnstedt and Woidich 2005: 11). A subcategory of these exceptions to rule (f) is that of the verbs in FORM-II and FORM-V which have /ʔ/ as the first radical and have not substituted /w/, as was exemplified above by *te^ʔaddūn*.

In some cases of doubly weak roots Musil has recorded a final glottal stop, e.g. *gī^ʔ* ‘a captured colt’ and *ǧā^ʔ* ‘he came’. These forms are odd in that, assuming that the glottal stop was not pronounced, final-vowel shortening has not occurred (i.e. **gī*, **ǧa*), although it may be that the process cannot occur when its output would be a single light syllable (compare the non-shortening of the negative particles *lā* and *mā*, and of the vocative particle *yā*).

⁴⁴ For the use of the relational adjectival suffix in terms such as this see §3.1.2.7.

An interesting phenomenon which has been transcribed only in a couple of instances by Musil is that of the replacement of /ʔ/ by the pharyngeal fricative /ħ/ (compare the phenomenon termed *anʿana* by the Arab grammarians). Sawayan (1992: 80) claims that in peninsular dialects the intervocalic glottal stop, when it is retained, is replaced by /ħ/; this fact is also attested for Najdi Arabic in general (Ingham 1994a: 14, and referenced in Watson 2002: 18) where it is noted that the majority of modern dialects that retain the glottal stop intervocalically retain it only between two identical vowels (these studies compare *saʔal* ‘he asked’, which is attested in Cairene and other dialects, with *sāl* and *saʿal*, which are attested in Najdi). In Musil it is not restricted to intervocalic position, as we saw in *erʿ-eh* above, nor do the vowels have to be identical, as in the case of *gāʿile* ‘noon, the time of greatest heat’⁴⁵, which varies with *gāyile*, *gayyile*. Although it is barely attested in Musil’s corpus, the recorded spread of the phenomenon in related dialects begs the question of whether Musil’s *raʔis* above could, in fact, have been pronounced **raħis*, which would have been much more readily mistaken for *raʔis* than the expected **rayis* would have.

In summation, even discounting the ubiquitous forms ending in *-aʔ*, it appears that one should consider Musil’s transcription of the glottal stop to be erroneous. Apart from a few genuinely convincing instances (e.g. *raʔis*, although this could have been pronounced *raħis*) it appears that Musil was inferring the presence of the glottal stop from his knowledge

⁴⁵ Musil also records the phrase-initial form *ʿarbaʿin* ‘forty’ (for [ʔ]*arbaħin*), which seems less likely to exhibit *anʿana* (since it has only been attested elsewhere in modern dialects intervocalically) but is nonetheless interesting.

of CLA; this seems especially true of certain morphological forms such as the CLA feminine suffix *-ā'* and certain lexical forms where Musil would have been aware of the presence of /^ʔ/ in the consonantal root. Since glottal stop deletion is attested in all positions in the majority of cases, we may assume that the Rwala dialect did not regularly exhibit the *lexical* glottal stop except in isolated lexemes. For the epenthesis of a phonological glottal stop, see below (§2.3.5.3).

2.1.1.2 /ǧ/

Musil describes the sound associated with the letter *ǧīm* as having a variety of articulations, but since he transcribes it with the symbol <ǧ> throughout his work his records can shed no light on the phonological conditioning of these allophones (if that is what they are), or on their phonetic realisations in general. The matter of the Rwala pronunciation of *ǧīm*, however, is of general interest in the typology of Arabic dialects. Musil writes in his introduction that *ǧ* sounds:

“like *g* in *gem*, or *dsh*, *dj*, or *y* in *yoke*” (Musil 1928a: *xiv*).

It is not clear whether each of these is meant to represent a different sound, or whether some are alternative respellings of the same pronunciation. The first two, for example, “*g* in *gem*, or *dsh*”, may both signify the voiced palato-alveolar affricate (IPA [dʒ]); his “*dj*” and “*y* in *yoke*”, on the other hand, are both likely to be voiced palatal or palatalised consonants, even though the intended manner of articulation remains unclear (it

could range from approximant or glide, IPA [j], through fricative [ɟ], to plosive [gʲ], [j] or [dʲ]. In principle, however, Musil implies that there is both an affricate and a palatal/palatalised realisation of *ǧīm*.

Writing at a similar time to Musil, Cantineau (1936-37: 25) appears to record the Rwala *ǧīm* as IPA [j] or [dʲ]⁴⁶, which accounts for the palatal/palatalised realisation described by Musil but not for the affricate. Cantineau considered the articulation of *ǧīm* to be a classificatory characteristic for different social and dialect groups of the region: the ‘Anazi and Šammari dialects had the palatal/palatalised realisation (IPA [j] or [dʲ]); the *petits nomades* had an affricate (IPA [d͡ʒ]); and the dialects of the Sirḥān, Sardiye and of Ğōf had a glide (IPA [j]). His outline of these divisions has been taken up by later researchers (cf. Fischer and Jastrow 1980: 106), but it is important to note that the isogloss lines are not clear-cut: the Šammari dialects, for example, he admits may have all three types of

⁴⁶ It is necessary here to comment on the system of transcription used by Cantineau which, while more clearly elaborated than that of Musil, is not unproblematic. Regarding the plosives, Cantineau transcribes what he describes as a voiceless palatal plosive <k>, a voiced palatal plosive <g> and a voiceless velar plosive <q>. It will be immediately apparent that his palatals seem to correspond to what we would now term velar (IPA [k] and [g]) and his velar to a uvular (IPA [q]); his other velar sounds, the voiced and voiceless fricatives <ḳ> and <g̣> (Arabic *ḥāʾ* and *ǧayn*) could arguably be velar (IPA [x] and [ɣ]) or uvular (IPA [χ] and [ʁ]), but are described as uvular for the Arabian dialects recorded in Prochazka (1988). I propose that Cantineau’s “palatal” sounds should be considered velar for the following reasons: there is no evidence for the widespread palatal realisation of *kāf* (IPA [c]); it is hard to contextualise his use of the symbol *q* to represent what is undoubtedly a velar sound (IPA [g]); and his description of *gʷ* as a “palatalised *g*” (p.138) is inconceivable if *g* is already palatal. If this theory is correct then his description of “*g* prépalatal” for Rwala /ǧ/ would indicate a fronted velar sound which would correspond to a voiced palatal plosive (IPA [ɟ]), while his “*gy mouillé*” (palatalised *gʷ*) would indicate a palatalised velar. The latter is theoretically impossible since a palatalised velar is, phonetically, palatal, but the reason for his distinction between *gʷ* and fronted *g* may be found in his indication that *gʷ* is “difficile à distinguer de *dʷ*”; this explains it because a palatalised alveolar plosive is not the same as a palatal plosive, but the two are perceptually similar.

realisation. Furthermore he acknowledges that his data contradicts what others have recorded for the region: he placed Bani Şaḥar in the affricate group, for example, while Wetzstein had placed it in the glide group. Palva (1980: 114) records that the dialect of the Bani Şaḥar has [čʒ] and [gʲ] in free variation, and that the occurrence of both “vacillates noticeably in one and the same idiolect” (*ibid.*). This leads him to conclude that “the question is not one of the reflex of *ǧīm* in the dialect, but rather of its reflexes and how they are distributed” (*ibid.*: 115-6).

These comparative data show that it is probable that Musil’s description of multiple realisations of /ǧ/ reflects a situation in which /ǧ/ was pronounced differently even in the speech of individual speakers. Johnstone (1965: 241) presents a similar situation concerning the realisation of /ǧ/ as a pure glide [j], noting that it is not phonetically conditioned, that it does not occur in some common words, and that it is apparently reversible at will.

The glide pronunciation [j] has not, to my knowledge, been recorded for an ‘Anazi or central-Najdi dialect such as that of the Rwala. It has, however, been recorded for the Šammari and northern-Najdi dialects (Wetzstein 1868: 163), and it may be of importance that it was recorded for Sirḥān and Šararāt (*ibid.*, and confirmed by Cantineau 1936-37: 24-5). The wadi Sirḥān was in the *dīra* of the Rwala at the time of Musil’s visit (van Donzel 1997: 673) and the Šararāt paid *ḥuwwa* to the Rwala (Lancaster and Lancaster 1997: 316), and members of both groups are recorded by Musil as living among the Rwala while he was

staying with them. Considerations such as these may have increased the likelihood of Musil hearing [j] as an optional free variant realisation of /ǧ/.

Prochazka, at a later date, gives only the affricate pronunciation [ɟʒ]⁴⁷ as the majority variant for the dialects he studied, citing only a few exceptions where the pronunciation was palatal. The absence of the Rwala dialect from his list of exceptions implies that the pronunciation he heard among them was an affricate. It therefore appears that Musil's description of both an affricate and a palatal realisation of /ǧ/ is confirmed by two contradictory sources, Cantineau on the one hand and Prochazka on the other.

In a later study, Cantineau proposed that the historical development of the realisations of *ǧim* followed the pattern [g] > [gʲ] > [dʲ] > [ɟʒ] (Cantineau 1960: 57)⁴⁸. The second stage of this development ([gʲ] > [dʲ]) should probably be questioned, since [dʲ] represents an overshoot of palatalisation which is not required for the subsequent development to [ɟʒ]; furthermore, Cantineau (1936-37: 138, 1960: 58) admits that [gʲ] and [dʲ] are very difficult to distinguish and records them more or less as variants for the Rwala

⁴⁷ A note on Prochazka's terminology is also required. He describes the majority articulation of /ǧ/ as a voiced alveolar affricate (= IPA [ɟʒ]). It appears that what he meant was a voiced *palato*-alveolar affricate (= IPA [ɟʃ]), not only because [ɟʒ] seems an odd choice to cover almost the entirety of the vast area covered by his study, but also on the basis of the fact that he describes the affricated allophones of /k/ and /g/ (cf. §2.1.1.4) as coming in pairs of *č* and *j*, or *ć* and *ǧ* (<j> being his symbol for *ǧim*). It is well established that the pairs are [tʃ] and [ɟʃ], or [tʂ] and [ɟʂ], so it seems much more likely that he is describing the realisation of /ǧ/ *and* the voiced counterpart of *č* (both his <j>) as [ɟʃ]; furthermore, if <j> = [ɟʒ] then what does his <ǧ> correspond to?

⁴⁸ In this theory the glide pronunciation [j] also developed from [dʲ], and the pronunciation [ʒ] (typical of the Lebanese countryside and the urban centres of the Levant) developed from [ɟʒ].

rather than different stages of historical development. This does, however, indicate that the nuances of the articulatory realisations of the reflexes of *ǧim* are not always easy to perceive.

Naturally, the fact that Cantineau did not hear [ǧ] among the Rwala does not mean that it did not exist, and, while he undoubtedly had a broad grasp of the variation found within the whole region, the fact that Musil did claim to hear it (among other realisations), having spent so long with the Rwala, should not be discounted. At the same time, Musil's claim that /ǧ/ sounds "like g in gem" is not wholly convincing, considering that several of his other claims about the realisations of consonants are almost surely wrong (see below); there is also a possibility that, since the sound [ǧ] is not found in his native Czech⁴⁹, he could have conflated one foreign sound ([ǧ]) with another perceptually similar one ([dʒ]). Prochazka's record of hearing [ǧ] is also not quite convincing, on account of the fact that he generalises for almost the whole region. Furthermore, the fact that his Rwala informants were interviewed in Riyadh and London indicates that at the very least they would have been exposed to different varieties of Arabic that would contain [ǧ], and it is also not unreasonable to assume that they could have produced [ǧ] in an interview situation with a non-native interlocutor. Although none of these reasons should rule out the possibility of a Rwala [ǧ], Cantineau's description should probably be considered the most reliable on this matter.

⁴⁹ [ǧ] occurs in Czech only as a result of assimilation (Dankovičová 1999: 72) and in foreign borrowings (Šimáčková *et al.* 2012: 226).

2.1.1.3 /ḏ̣/

/ḏ̣/ is an emphatic voiced interdental fricative. Emphatic consonants in Arabic are traditionally described as velarised or pharyngealised (Ladefoged and Maddieson 1996: 365), although the term velarisation is rarely used in this context nowadays, and some researchers argue that the emphatics are uvularised (see Bellem and Watson 2014: 2 and note 3). Musil can shed no light on this question, but in literature on north-Arabian dialects in general (e.g. Ingham 1982a: *xxi*) and in Najdi dialects more specifically (e.g. Ingham 1994a: 14) the most commonly encountered description of the articulation of emphatics is that they are pharyngealised.

Bedouin dialects historically maintained the OA interdental fricatives, and outside the Arabian Peninsula the presence of interdentals tends to signify that a dialect is bedouin (Watson 2011a: 898), although in the Peninsula they are more widely attested than elsewhere in non-bedouin varieties as well. The merger of OA *ḏ and *ḏ̣ is attested for all recorded modern varieties of Arabic (Fischer and Jastrow 1980: 50). In bedouin-type dialects, only the interdental /ḏ̣/ remains, in contrast to urban-type dialects in which /ḏ/ is retained; this is on account of a historical restructuring of the OA system in which urban-type dialects merged the interdental fricatives with their corresponding dental stops, while the bedouin-type dialects maintained sets of both interdental fricatives and dental stops and merged only *ḏ with *ḏ̣ (Holes 2004a: 71).

Musil interestingly includes the symbol <ḏ> in his list of transliteration symbols but, as would be expected, it is never used in the work except in some proper names, and does not constitute a phoneme in the dialect. His use of the symbol <ḏ> to transcribe /ḏ/ is an odd choice and was criticised by Cantineau (1936-37: 18): taken at face value, the realisation of historical *ḏ as an emphatic voiced alveolar fricative (= IPA [z̤]) is very unlikely, and is only found in some words in Cairene Arabic (Watson 2002: 15) and in some northern Mesopotamian dialects (Fischer and Jastrow 1980: 50). It seems much more likely that his decision was etymological and that <ḏ> was employed as an abstract representation of /ḏ/ (CLA ḏā'), rather than an attempt to convey articulatory information, as is found in many standard systems of transliteration to this day. It is also likely that, in spite of Cantineau's criticism, <ḏ> may simply have been too difficult to produce typographically at the time (and we may further consider that Cantineau [1936-37] was published by l'Institut d'Études Orientales, which may have had more expertise in this area than the American Geographical Society which published Musil's work). An interesting curiosity recorded by Musil is that of an acrostic poem (Musil 1928a: 481-2) in which ḏ is given two verses, corresponding to ḏ and ḏ in the abjad order.

2.1.1.4 /g/ and /k/

These phonemes are articulated as voiced and voiceless velar plosives, which have affricated allophones in front vowel environments (see below). Bedouin dialects usually

have a voiced reflex of OA *q (Kaye and Rosenhouse 1997: 270) which has become the voiced counterpart of the voiceless /k/; this is sometimes described as a primary distinction between bedouin and sedentary dialects in the Arabian Peninsula, since the sedentary reflex of *q is (though not without exception) voiceless (Watson 2011a: 898). The Rwala realisation of *q as a voiced velar plosive is supported by Cantineau (1936-37: 28, 139) and Prochazka (1988: 14), although it cannot be ruled out that the realisation may be a voiced uvular plosive (IPA [ɢ]) in some cases (cf. Watson 2011a: 898).

Musil's transcription of the plosive allophone of /g/ as <ḳ> is somewhat difficult to account for, and was publicly disregarded by Cantineau shortly after its original publication (Cantineau 1936-37: 36). The symbol is not defined in articulatory terms at any point in his work, and is not problematic from a phonological point of view (its distribution with the affricated allophone is regular), but the choice of symbol does suggest that Musil was implying that the consonant was voiceless, which it certainly cannot have been. Since Musil did not give a definition to the symbol <ḳ>, it is probably the case that he was using it for its abstract representational value (associating it with CLA *qāf*) rather than using it to describe the phonetic facts of the language as he heard it, as in the case of <ḏ> above: we may compare the use of <ḳ> throughout Wright's grammar (Wright 1967 vol.1: 2ff, but first edition 1859) as part of a tradition in which *qāf* is considered to be the emphatic equivalent of *kāf* and represented as such. Since voiced and voiceless are not very clear-cut categories, and their definitions and perception vary from language to language, it remains possible

that Musil did conceive of the sound as voiceless, but speculation of this sort does not seem necessary. The affricated allophone attested for /g/ can certainly only have developed from a voiced consonant, so the voiced realisation of the plosive allophone of /g/ is not in doubt.

It is a well known fact that many of the Arabic dialects of the Arabian Peninsula have affricated⁵⁰ allophones of *kāf* and *gāf* which generally occur in the environment of front vowels (Kaye and Rosenhouse 1997: 270, 273). It appears that the phenomenon first came to the attention of western scholars in an article by Wallin (1855) which commented on certain dialects of the Najdi group; the subject was later extensively addressed by Cantineau (1936-37: 27-39), whose work was then expanded on by Johnstone (1963). The affricated allophones may be either [tʃ] and [dʒ] as a pair or [ts] and [dʒ] as a pair: Cantineau recorded that the *petits nomades* had the former and that the ‘Anazi, Šammari, and sedentary groups had the latter. Johnstone, noting that some of the *petits nomades* (such as the ‘Awāzim and the pariah Šlēb⁵¹) had the latter pair as well, convincingly argues that

⁵⁰ The affricated allophones are frequently referred to as such in the relevant literature (e.g. Johnstone 1963). This terminology requires comment, simply because it could misleadingly imply that the affrication of the velar plosives [k] and [g] produces velar affricates, e.g. [kx] and [gɣ]. The process which yielded the affricated allophones of [k] and [g] in the north-Arabian dialects is one of palatalisation with subsequent affrication. In the case of /ǧ/ above we saw a weak form of palatalisation producing [dʒ]: in these cases a palatalised consonant retains its primary place of articulation (e.g. bilabial, alveolar, velar) but also develops a secondary place of articulation as the tongue rises slightly towards the hard palate in anticipation of a following front vowel: [b] > [bʲ], [d] > [dʲ], [g] > [gʲ] (cf. Ashby 2011: 131-133). In the case of /k/ and /g/ we see a stronger type of palatalisation which involves a non-palatal consonant, in this case velar, becoming palato-alveolar (or first palatal and then palato-alveolar). When a plosive becomes palato-alveolar, the change in place of articulation is typically also accompanied by a change in manner – affrication: [t] or [k] > [ts] or [tʃ], [d] or [g] > [dʒ] or [dʒ] (cf. Trask and Millar 2015: 56).

⁵¹ These are the Šolubba, cf. Betts (1989) and Simpson (1994).

the variation is areal and does not coincide with other dialect boundaries or with social status: [tʃ] and [ɟʒ] are found in peripheral north-Arabian dialects while [ts] and [dʒ] are found in the interior dialects (Johnstone 1967a: 6, and see map, p.5).

Cantineau's observation that in the Rwala dialect these allophones were realised as [ts] and [dʒ] has been confirmed by examples from Prochazka (1988: 16) and Ingham (1995: 124)⁵². It is difficult, therefore, to account for Musil's transcriptions which use the symbols <č> and <ž>, and which imply that the realisations are [tʃ] and [ʒ], the latter of which, to my knowledge, has never been attested in a bedouin dialect. The symbols <č> and <ž> are sparingly described by Musil as "ch" and "like z in azure" (Musil 1928a: xiv). Daniel Jones's 1917 pronouncing dictionary only recorded the word "azure" with [ʒ] at a time close to when Musil was writing (Jones 1917: 28), and the same work records [tʃ] for words starting with "ch", so it appears that Musil did indeed mean to associate these sounds with the symbols he used (incidentally, these sounds have the same orthographic representation in the Czech alphabet, which was Musil's native language).

This apparent mistake on the part of Musil was already noted by Cantineau who asked concerning this matter, "comment expliquer l'erreur de Musil qui a vécu si longtemps chez Rwala?" (Cantineau 1936-37: 36). Musil was certainly aware of the different articulations of these allophones since he insisted elsewhere that the Šammari tribes had [ts]

⁵² To this list it may be added that Lancaster (1997: 195) transcribes Musil's *kwācbe* (one of the Rwala tribal divisions, cf. Musil 1928a: 700) as *kwatzbe*. Although Lancaster's transcriptions are not generally very useful, this one seems to point to the [ts] realisation and not [tʃ] as suggested by Musil.

and [ɟ] (<č> and <ǧ> in his orthography) while the Rwala had [tʃ] and [ʒ] (Musil 1928b: 18-19), so it is difficult to argue that his transcriptions can be explained as some sort of abstraction or generalisation as in the cases of <ḵ> and <ẓ> above. Likewise, it seems unlikely that the explanation could be that Musil did not find [ɟʃ] and [ɟʒ] perceptually distinctive⁵³, since he has argued for their distinctiveness in the comparison of realisations between the Rwala and the Šammar. Finally, it is difficult to argue that he could have been correct in recording [ʒ], not only because this articulation is completely unattested in other accounts of dialects from the region but also because such a situation is not supported typologically. A realisation of the affricated allophone of /g/ as [ʒ] is not impossible, since deaffrication is a viable sound change, but even if this were the case, [ʒ] would only be possible as the deaffricated reflex of [ɟʃ], and not of [ɟʒ], but by all accounts [ɟʒ] is the expected Rwala pronunciation. Further support for this conclusion may come from the instances in which the affricated allophone of /g/ was transcribed as <ǧ> by Musil: examples are *ǧāsi* ‘hard’ (root **qsw**), *mistaǧellāt* ‘freely-moving FP (haunches)’ (root **qll**) and *mweyǧāt* ‘peeping at FP’ (root **wyq**). If we are to believe that Musil transcribed exactly what he heard (cf. Musil 1928a: *xiii*) then it seems possible that when speakers did have the [ɟʃ]

⁵³ [ɟʒ], like [ɟʃ] mentioned above, only occurs in Czech as a result of assimilation (Dankovičová 1999: 72), but unlike [ɟʃ] it does not occur in foreign borrowings (Šimáčková *et al.* 2012: 226).

pronunciation of this allophone then Musil transcribed it as <ǧ>, suggesting that his <ž> was not intended to represent the same sound⁵⁴.

In terms of distributional context, it is generally true that the affricated allophone occurs in the environment of front vowels; as in many dialects, however, this rule is not clearly effected in practice⁵⁵. This is demonstrated in the following assessment of /k/ and /g/ in the contiguity of high vowels, which shows a strong trend. Back vowels inhibit the occurrence of the affricated allophones: *u*, *ū*, *o* and *ō* occur adjacent to the plosive allophones in over 96% of instances. The reason for the affricated allophones occurring in the remaining <4% appears to be levelling in many cases. It is common, for example, that P-STEM or imperative verbs with the affricated allophone as the final consonant retain it upon the suffixation of *-ū(n)*, e.g. *yešarreǧ*, *yešarreǧūn* ‘he goes/they go into the inner desert’: here the affrication of the unsuffixed stem has presumably been analogically levelled within the paradigm. There are some examples which suggest that the affricated allophone alone may have come to be associated with some roots to the complete exclusion of the plosive allophone, e.g. **kdb** from whence *céḏūb*, *caḏḏāb* and *cuḏbān*, all meaning ‘liar’, although this

⁵⁴ It may be noted that the opposite situation is also attested, although less frequently, where /ǧ/ is transcribed as <ž>. This occurs, for example, in the verb *hāǧ*, *yehīǧ* ‘to rut (camel)’, which is recorded as *hāǧ*, *yihīǧ* in Kurpershoek (2005: 341). It is difficult to speculate from the scant evidence as to possible types of variation which could have produced the pronunciations Musil claims to have heard, although this is further evidence that /ǧ/ and the affricated allophone of /g/ could have had the same pronunciation for some speakers. A further example of this is the variation between roots **qḏ** and **ǧḏ** (see Glossary), both of which seem to be related to the root **ḏǧ** in CLA.

⁵⁵ The distributional context is not black and white in any recorded dialect. We may consider two of the section headings of Johnstone’s comparative article on the subject as evidence of this: “the non-occurrence of these affricates in contiguity with the front vowels” (Johnstone 1963: 219) and “the occurrence of affricates in contiguity with back vowels” (*ibid.*: 220).

is speculative. Both of these examples may suggest that the rule conditioning the affrication of /k/ and /g/ is no longer operative (Holes takes this view for an east-Arabian dialect, cf. Holes 2016: 52).

In the environment of the front vowels *i* and *ī* the affricated allophones are surprisingly only attested in 74% of cases. One reason for the high occurrence of the plosive allophones in contiguity with the front vowel *i* is the fairly frequent variation of /u/ ~ /i/ word-medially (see below) which does not always seem to affect the conditioning of /k/ and /g/: compare *kull* ~ *kill* (but not **ćill*) ‘all, every’ and *kuṭr* ~ *kiṭr* ~ *ćiṭr* ‘great number’. This is particularly noticeable in the paradigm of hollow verbs where /i/ has been partially levelled across the first and second person S-STEM forms (cf. §4.10.2): from the root **qwl** we find no affrication in the forms *gult* ~ *gilt* ‘I said’ whereas in the form *ǧīl* ‘it was said’ /g/ is affricated. The absence of affrication in *gilt* (not **ǧilt*) perhaps indicates that affrication is lexicalised to an extent, since the *u* ~ *i* variation does not trigger it while *ǧīl*, which does not exhibit such variation, is always attested with affrication.

2.1.1.5 Consonant variation

Musil’s material shows the following minor variations, none of which are attested more than a handful of times in the corpus:

/t̥/ is sometimes realised as *f*, e.g. *telǧ* ~ *felǧ* ‘snow’. An interesting form is *ḥatāfīl* ‘the last drop of sth.’, which is presumably a plural form from a geminated singular (e.g.

**hattāl*) and which exhibits both *t* and *f* in the same word. The reverse is also attested, as the term *šafra* ‘edge’ is variously recorded as *šatra* and even as *šadra*.

/g/ is sometimes realised as *ġ*, e.g. *ġāder*, *dīret al-ġašmeriyye*, and *ġār*. The first two are places, namely Qāder and the land of Qašmeriyye, which parallels Montagne’s record of */g/ > ġ* in place names amongst the Šammar: *feġāra* for Feqāra and *ġāmešliye* for Qāmešliye (Montagne 1935: 75). That this phenomenon is mostly attested in place names suggests that this is not a regular process in the dialect but possibly an attempt to pronounce the sound [q], which probably exists in some sedentary dialects of the region (the same view is expressed by Cantineau 1936-37: 140). Musil’s record of *ġār* is more interesting, since he does not translate it (citing it only as an example of this sound change); it does not appear to be a place name, and Cantineau translates it as ‘tar’, in which case it is hard to account for why such a word should be pronounced with a “borrowed” pronunciation⁵⁶. Alternatively, it may be possible that the place name al-Qāra was meant, which is also attested in Musil.

/ħ/ is sometimes realised as *h*, e.g. *bi-mdāhel* (< **medāhel*) ‘on the borders of’, but for this realisation Musil provides sociolinguistic information, describing it as a pronunciation typical of speakers of the Dahāmše.

⁵⁶ *ġār* meaning ‘tar’ is not attested in Musil, but as an alternative interpretation, *ġār* is attested meaning ‘he hunted by lying in wait’ or ‘he hunted by catching the game in snares’; these options, again, seem apposite to bedouin life and not ideal candidates for borrowed pronunciation.

In one case *ć* is voiced and occurs as *ǰ*, namely in *turǰiyye* ‘earring’ which has the plural form *tarāći*; it is uncertain whether voice assimilation could have occurred in this case.

An interesting example of variation is recorded in the speech of one speaker who (consistently, according to Musil) pronounced *ǰeǰāmīr* ‘thick logs of wood cut from palm stems’ as *ǰeṭāmīr*. In this case *ǰ* ([ǰ]) is devoiced to *ṭ* ([θ]) and *ǰ* is either fronted (to [ǰ]) or possibly deaffricated (to [ʒ]). While I argued above that Musil’s symbol <ǰ> does not represent [ʒ] (as he intended it) but rather [ǰ], in this particular case it is not the affricated allophone of /g/ and both realisations are equally difficult to explain.

In only one case /ǰ/ is realised as *d*, *serāhīd* ‘well fed (pl.)’ (not **serāhiǰ*), but this was clearly only in order to force a rhyme in a poem in which every line ended with the sequence al-CVCāCīd.

In one poem, /ǰ/ is realised as *z*, *be-z-zeraf* ‘on a tray’ (not **be-ǰ-ǰeraf*).

There is variation between *ǰ* and *ḍ* for the verb *ǰāg yidūg/ḍāg yidūg* ‘to taste’ (root *ǰwq*), although it seems likely that this variation is restricted to this verb; the same variation is recorded for Bahrain (Holes 2001: 193).

2.1.2 Vowels

Musil’s system of transliteration accounts for: five short vowels, *a*, *e*, *i*, *o* and *u*, of which probably only /a/, /i/ and /u/ are phonemic (see §2.1.2.1); five long vowels, *ā*, *ē*, *ī*,

ō and *ū*, all of which are phonemic; and two phonemic diphthongs, /aw/ and /ay/, which are realised as *aw*, *ow*, *ō* and *ay*, *ey*, *ē* respectively. The vowel *e* is a variant of /a/ or /i/ and *o* is a variant of /u/. Compared to his treatment of consonants which, with the exception of the allophones *č* and *ǰ*, is essentially a phonemic transcription, his treatment of vowels is somewhat broader, having *e* and *o* at his disposal to record a greater range of vocalic variation. This contrasts starkly, however, with the other published texts of the Rwala dialect which, with the exception of Cantineau (who used a phonetic transcription system), have only used *a*, *i* and *u* to record short vowels in the dialect (cf. Prochazka 1988 and Ingham 1995)⁵⁷. In this regard, the records of the dialect can be ordered on a spectrum on which we have Prochazka and Ingham at one end, using a categorical (and loosely phonemic) system, and Cantineau at the other end, using a phonetic system. Musil lies somewhere in the middle, as his five-vowel system is more gradient than that of Prochazka and Ingham, but not nearly as gradient as Cantineau's eighteen vowel system⁵⁸.

To explain Musil's system, it must be assumed that he desired to record the pronunciation of the dialect more or less as he heard it; from this perspective, it is perhaps useful to consider that the system he has used is identical to the inventory of Czech vowel phonemes, /i:, ɪ, ε:, ε, a:, a, o:, o, u:, u/ (Dankovičová 1999: 72). The implication of this is

⁵⁷ In a few isolated examples Prochazka has also used *ə* to represent an epenthetic vowel, e.g. *bāytəkum* 'your tent' (Prochazka 1988: 199).

⁵⁸ The system in Cantineau (1936-37: 13) is also based on a five vowel system (also *a*, *e*, *i*, *o* and *u*) but is elaborated with diacritics marking openness, closeness (including a distinction between close and very close), frontness and backness.

that it seems probable that he recorded each vowel in relation to his own perception of it; in the dialect it is clear that /a/ and /i/ occupy large (and possibly even overlapping) portions of the vowel space, but the difference between the more close or fronted realisation of /a/ and the more open realisation of /i/ is obscured by Musil's association of both sounds with what, for him, would have been a contrastive sound, *e*. It is therefore impossible to speculate as to what the phonetic quality of any sound transcribed with <e> might represent, since it may range from a front low variant of /a/ to a non-low (mid-) vowel, and possibly even to a high vowel. His treatment of *o* is similar, although it is less problematic since *o* does not vary with any other phoneme and therefore it is safe to assume that *o* is always a variant of /u/.

In Generative Phonology, speech sounds, or segments, are not the smallest units of phonological structure. Segments are believed to be composed of smaller units, known as features. Features have an autonomous ontological status, combine simultaneously in the structure of segments, and specify segments for various attributes, such as voicing, place and manner of articulation, lip rounding, and so on. In the classical SPE model (Chomsky and Halle 1968), which is still widely used today—unaltered or in numerous adapted versions—features are binary: a segment is assigned a positive (+) or negative (-) value for each feature in its structure (referred to as its 'feature matrix'). Segments that share the same specification for one or more features form a 'natural class'.

In the SPE framework of features, major contrasts in vowel quality are encoded by a combination of three tongue body features, [\pm HIGH, \pm LOW, \pm BACK], and one lip position feature, [\pm ROUND]. [+HIGH] indicates that the tongue body is raised from the neutral position, [+LOW] that it is lowered from the neutral position, and [+BACK] that it is retracted from the neutral position. In addition, [+ROUND] characterises a sound for which the lips are rounded. The so-called ‘corner vowels’, which occupy the corners of a quadrilateral model of the vowel space, [i, æ, ɑ, u], have the following SPE featural specifications:

[i]	[+HIGH, -LOW, -BACK, -ROUND]
[u]	[+HIGH, -LOW, +BACK, +ROUND]
[æ]	[-HIGH, +LOW, -BACK, -ROUND]
[ɑ]	[-HIGH, +LOW, +BACK, -ROUND]

Mid vowels, such as [e] and [o], which are of an intermediate height between ‘high’ and ‘low’, are specified as [-HIGH, -LOW]:

[e]	[-HIGH, -LOW, -BACK, -ROUND]
[o]	[-HIGH, -LOW, +BACK, +ROUND]

In the majority of SPE-type frameworks, the symbols <a> and <ɑ> are used interchangeably for a [+LOW, +BACK, -ROUND] vowel, even though [a] is defined as a front (i.e. [-BACK]) vowel in the International Phonetic Alphabet (IPA Chart 2018). With regards to /a/ in this dialect, I assume that it has [+BACK] and [-BACK] allophones: Musil’s <e> probably represents a [-HIGH, -LOW, -BACK] allophone, but it is considered that his <a> probably covers both [-HIGH, +LOW, -BACK] and [-HIGH, +LOW, +BACK] and is therefore

treated as [\pm BACK]. In addition to these vowel quality features, I have occasionally used [\pm LONG] in formulations to distinguish vowel length.

The following section deals only with the vowel phonemes and their transcription; phonological alternations are addressed in §2.3.6.

2.1.2.1 Short vowel phonemes

The phonemic contrast between the low vowel /a/ and the high vowels /i, u/ is not in doubt, and minimal pairs are plentiful. Examples include:

<i>ġā-na</i>	‘he came to us’	<i>ġā-ni</i>	‘he came to me’
<i>ayyā-k</i>	‘beware!’	<i>iyyā-k</i>	‘you, with you’
<i>dabbe</i>	‘ridge of hills’	<i>dibbe</i>	‘spear with a supple shaft’
<i>ḍahr</i>	‘back’	<i>ḍuhr</i>	‘noon’
<i>fakk</i>	‘he saved’	<i>fukk</i>	‘save!’
<i>bakra</i>	‘young she-camel’	<i>bukra</i>	‘tomorrow’

The contrastive status of /i/ and /u/, however, is tenuous. In word final position minimal pairs are plentiful: *hu* ‘he’ and *hi* ‘she’, *abu* ‘father of’ and *abi* ‘I want’ and, although they are not attested for the same verb in this corpus to provide a minimal pair, the S-STEM suffixes *-ti* (2FS) and *-tu* (2MP) are contrastive. Word medially, however, it is very difficult to find minimal pairs: one possible candidate is ‘*edd* ‘spring, well’ (probably from underlying /‘idd/, compare ‘*idd* in Kurpershoek 2005: 209) and ‘*odd* ‘tell!’. This situation is not uncommon in the dialects of Arabic: Prochazka claims that /i/ and /u/ have merged in the dialects of Saudi Arabia (Prochazka 1988: 17), although this cannot be completely true, and

Ingham refers to a “distributional overlapping between *i* and *u*” (Ingham 1994a: 14) in Najdi Arabic in general. Furthermore, the specific difficulty of finding minimal pairs for these phonemes is discussed for other dialects, e.g. by Al-Wer (2011) for Jordanian Arabic from ‘Ammān and by de Jong (2011a: 29) for a bedouin dialect from Sinai. From Prochazka’s own data it is clear that the phonemes have not fully merged: it would be more accurate to say that the functional load of /i/ and /u/ is very low, with the result that in non-final position there is often free variation in the same word⁵⁹. Examples from Musil include:

<i>kill</i> ~ <i>kull</i>	‘all, every’
<i>šif(e)t</i> ~ <i>šuf(e)t</i>	‘I saw’
<i>gil</i> ~ <i>gul</i>	‘say!’
<i>ğibb</i> ~ <i>ğubb</i>	‘deep artificial well’

This variation is also found in the participle prefix /mV/, e.g. *mitmašleḥ^{en}* ~ *mutmašleḥ^{en}* ‘having the sleeves thrown back over the head and tied’. /i/ and /u/ appear to be in free (rather than allophonic) variation in non-final position since no distributional context exists. /u/ certainly occurs slightly more frequently in the environment of labials, liquids, velar stops, or emphatic consonants (i.e. /b, m, l, r, k, g, t̤, s̤, d̤/) but this trend is not statistically strong enough to posit allophonic variation⁶⁰.

⁵⁹ By way of analogy we might compare the variation that occurs between phonemes of low functional load for some speakers of English, e.g. between /θ/ and /ð/ in *booth* or *paths*, and between /ʃ/ and /ʒ/ in *Persian* or *Asia*.

⁶⁰ In the verbal system (see §4.3.4), this context does appear to be distributional for the verbal stems of FORM-I verbs in the P-STEM, but this apparent allophony is limited to that one *locus*.

The lack of minimal pairs for *e* and *o* with /a, i, u/ indicate that they do not have phonemic status. Minimal pairs for /a/ and *e* occur generally when *e* is demonstrably a variant of /i/, e.g.

<i>bal</i>	‘in fact’	<i>bel ~ bil</i>	‘camels’
<i>ann(a)</i>	‘that’	<i>enn(a) ~ inn(a)</i>	[introduces main clause]

There is also a number of minimal pairs where *i* never occurs, but these are in very restricted contexts. They are: in the object/possessive suffixes (*beyt-eh* ‘his tent’, *beyt-ah* ‘her tent’); in FORM-I between the 1CS P-STEM (or FORM-IV 3MS) and the imperative (*ab‘ad* ‘I am far away’, *eb‘ad* ‘be far away!’; *agbal* ‘he approached’, *egbel* ‘accept!’); in the 3MS S-STEM and imperative of FORM-I geminate or FORM-II verbs (*dann* ‘it tinkled’, *denn* ‘tinkle!’; ‘*ammar* ‘he filled’, ‘*ammer* ‘fill!’); and between the 3MS S-STEM of FORM-III and the MS FORM-I active participle (*māna*‘ ‘he asked for pardon’, *māne*‘ ‘a pardoner’). It is tempting to view all of these forms with *e* as having /i/ in their underlying representation, and indeed some do vary with *i* (although only very rarely), e.g. the MS active participle *nāyim ~ nāyem* ‘sleeping’. On account of the fact that *i* almost never surfaces in the verbal paradigms and in the 3MS object/possessive suffix, it may be necessary to consider *e* as having marginal phonemic status. This may be comparable to the status of /é/ which, as has been established, is a contrastive phoneme only in very restricted contexts⁶¹.

Minimal pairs for /i/ and *e* occur, likewise, where *e* is a realisation of /a/. Examples include:

⁶¹ Cf. Al-Wer (2011) on the contested phonemic status of *o* in the dialect of ‘Ammān.

<i>ǧebile</i> ~ <i>ǧebīla</i>	‘tribe’	<i>ǧebil-i</i>	‘my opponent’
<i>en-ni</i> ~ <i>an (+ni)</i>	‘that I am’	<i>in-ni</i>	‘(behold) I am’

As before, there are a restricted number of minimal pairs where *e* occurs to the exclusion of /a/:

<i>‘aleyye</i>	‘to me’	<i>‘aliyye</i>	‘high (f.)’
<i>ayye</i>	‘which? (f.)’	<i>ayy-i</i>	‘woe is me!’
<i>ente</i>	‘you (MS)’	<i>enti</i>	‘you (FS)’

These examples, however, do not suggest that *e* is a contrastive phoneme. In *‘aleyye*, the *ey* is a variant of the diphthong /ay/ (see below), and in *ayye* the *e* is the feminine marker which is plentifully attested as *a*, despite the fact that the form **ayya* does not occur; *ente*, although it most probably has the underlying representation /inta/, is only ever recorded with final *e* and so may support the phonemic status of /e/.

The vowels *u* and *o* mostly occur as variants in the same word, e.g.:

<i>kufɾ</i> ~ <i>kofɾ</i>	‘infidels’
<i>ħull-i</i> ~ <i>ħoll-i</i>	‘my darling’
<i>kurkub</i> ~ <i>korkob</i>	‘a plant from which yellow dye is made’
<i>ħurr</i> ~ <i>ħorr</i>	‘thoroughbred’
<i>fursān</i> ~ <i>forsān</i>	‘horses’

As such, the minimal pairs of *o* with /a/ and /i/ are the same as those of /u/, and there are no minimal pairs of /u/ and *o*. There are, however, certain morphemes where either *u* or *o* occurs to the exclusion of the other: for *u*, this is the 2MP S-STEM suffix *-tu*, while for *o* these are the verbal stem of non-geminate FORM-I verbs in the P-STEM, the independent pronouns *hom* ‘they’ and *entom* ‘you (pl.)’ and the corresponding bound object/possessive

pronouns *-hom* ‘them, their’ and *-kom* ‘you (pl.), your (pl.)’. These being non-contrastive, however, the case for the phonemic status of *o* is weak.

Finally, there are a number of morphemes which are attested with both /a/ and /i/. These are: the adnominal linker *-an* ~ *-en*, where *-en* seems to have the underlying form /in/ on the basis of cross-dialectal evidence (see §3.4); the allomorph of the possessive suffix *-i*, which is *-ya* ~ *-ye* ~ *-yi* following a vowel (see §5.2.1.3); and the FORM-I P-STEM prefix *ya-* ~ *yi-* (see §4.3.1). There are also a few instances of the epenthetic vowel *a* ~ *e* attested as *i* (see §2.2.2.1), although these cases seem anomalous. The reason why /a/ and /i/ are possibly non-contrastive in these cases is not always clear and is discussed in the relevant sections.

2.1.2.2 Distribution of *e*

The fact that *e* represents both a raised (or fronted) /a/ and a lowered /i/ means that it is difficult to describe the distributional context of the variation. There is some evidence that /a/ raising (or fronting) is blocked by a class of “back” consonants, namely the glottals /ʔ, h/, the pharyngeals /ʕ, ħ/, the uvulars /ħ, ġ/, and the emphatics /ṣ, ṭ, ḍ/, with the velars /g, k, w/ possibly exerting the same effect. This is demonstrable from the realisations of the feminine marker *-a* (the CLA *tāʾ marbūʿa*): its variant *-e* is only attested once after a glottal consonant (/h/), once after a pharyngeal consonant (/ʕ/), never after a uvular consonant, sparsely after a velar consonant (once after /g/ and five times after /k/

and /w/) and only rarely after an emphatic consonant (three times after /ʃ/, and never after /ṭ/ or /ḍ/). In all of these cases, the variant *-a* is regularly attested. For all other consonants (apart from /n/ and /r/) the distribution of *-a* and *-e* is either even⁶² or there is a preference for *-e*. It is unclear whether the lowering of /i/ to *e* is conditioned by the same environment, although *i* is statistically less likely to occur after members of the “back” group: the average occurrence of *i* in a CV context (where V is any short vowel) is 13%, while after the back consonants (excluding velars) it is only 6%.

2.1.2.3 Distribution of *o*

In much the same way, the context for the variation of *u* and *o* is unclear. The only potential distributional context supported by the data is that *o* occurs after a preceding *a* more regularly than *u* does: a study of unique headwords in the corpus shows 50 instances of *a-o* and only 14 instances of *a-u*. The latter group consists of verbs having the 2MP suffix *-tu* (e.g. *šatteytu* ‘you (pl.) grazed in the winter’), nouns with consonantal /w/ (e.g. *sa‘aluwwa* ‘a female desert spirit taking the form of a greyhound’), geminate verbs having *u* as their stem vowel in the P-STEM (*arudd* ‘I return sth.’, but compare *yerodd* ‘he returns sth.’) and the gentilic terms *abu* ‘father of’ and *aḥu* ‘brother of’. Of these, the *-tu* suffix is invariable, as are the forms *abu* and *aḥu*, but *o* is by far the more common variant in P-STEM

⁶² Since the overall ratio of *a:e* in the corpus is approximately 3:2, I have assumed that a distribution of this ratio after each consonant represents an “even” distribution.

verbal stems (almost without exception, *u* only occurs in geminate verbs), and *o* does not occur before /w/ except in the diphthong *ow*, which represents /aw/.

2.1.2.4 Long vowels and diphthongs

There are five long vowels, *ā*, *ē*, *ī*, *ō* and *ū*, the phonemic status of which is not in doubt. A few examples will demonstrate that the four types of variation attested for the short vowels are always contrastive for the long vowels:

<i>zān</i>	‘spear shaft’; ‘it was good’	<i>zēn</i>	‘good, beautiful’
<i>bīḏ</i>	‘white (pl.)’	<i>bēḏ</i>	‘eggs’
<i>gōm</i>	‘followers of a smaller chief’	<i>gūm</i>	‘get up!’
<i>īd</i>	‘festival’	<i>ūd</i>	‘stalk’

In the first three examples, the long pure vowels *ē* and *ō* represent the underlying diphthongs /ay/ and /aw/ respectively, and the monophthongised forms given vary correspondingly with *zeyn*, *beyḏ* and *gowm*, as detailed in the following paragraphs.

The long pure vowel *ē* is a reflex of OA /ay/ which varies in this dialect between *ay*, *ey* and *ē*. Of these, *ey* is the major variant (57% of occurrences) followed by *ay* (24%) and *ē* (19%). Furthermore, *ay* is rather restricted in occurrence: it occurs almost to the exclusion of *ey* and *ē* when /y/ is geminate (80% *ayy*, and <1% *ēy*), and in fact the sequence *ayy* makes up 70% of the occurrences of *ay* overall. *ay* almost always appears in closed syllables and rarely in open syllables, at a ratio of 9:1, while *ey* and *ē* appear in both open and closed syllables, with much closer ratios of 7:6 and 3:4 respectively. This may suggest that a rule initially raised /ay/ to *ey* in open syllables only, before being generalised across the lexicon.

Similarly, the long pure vowel \bar{o} is a reflex of OA /aw/ which varies in this dialect between *aw*, *ow* and \bar{o} . We find a similar pattern of distribution to the /ay/ diphthongs, namely that the monophthongised variant \bar{o} is the least attested (only 4%, while *aw* is 42%, and *ow* is 54%). Conversely to *ay*, however, *aw* has a strong preference for occurring in open syllables (at a ratio of 11:1) while *ow* and \bar{o} are almost evenly split between open and closed syllables. This may suggest a parallel development that /aw/ was raised to *ow* in closed syllables only, before being generalised across the lexicon.

In the S-STEM the 3MS suffix $-\bar{u}$ varies with *-aw* and *-ow* (but not $-\bar{o}$) in the unsuffixed forms, although before a suffix $-\bar{u}$ is mostly attested ($-\bar{u}$ also occurs in both contexts in the imperatives). It seems therefore that a rule has diphthongised the underlying / \bar{u} / suffix in the S-STEM only, and only in word-final position. Without suffixes the ratio of $-\bar{u}$, *-ow*, and *-aw* is fairly even (3:3:2) and when suffixes are added the ratio is 12:2:1, but the few examples of *-aw* and *-ow* before a suffix probably represent later levelling. Neither Cantineau nor Prochazka have the $-\bar{u}$ variant for *Rwala* in the S-STEM, which indicates that the sound change / \bar{u} / > /aw/ must have been complete, rather than partial as suggested by Musil's data. There are no other instances of *aw* or *ow* varying with \bar{u} in Musil's corpus, which shows that this does not represent a general sound change in the language but rather a development which is restricted to just one morpheme: indeed, as noted in §1.3.1, these diphthongal plural S-STEM suffixes are a well known characteristic of north-Arabian dialects.

Ingham's Rwala data does show variation of *ū* with the monophthongised reflex *ō* (*yōm* ~ *yūm* 'when', *tōb* ~ *tūb* 'shirt', Ingham 1995: 125-126), but this is unattested in Musil.

2.2 The syllable and syllabification

The theoretical model followed in the analysis of syllable structure is described here, followed by a description of syllables and syllabification in the Rwala dialect from §2.2.1 onwards.

Linguists have not reached a consensus on the precise definition of a syllable, but one productive approach starts from the observation that peaks in prominence or sonority correspond to the number of syllables. Sonority is a measure of the degree of openness of the vocal tract, which correlates with perceived relative loudness. Sounds can therefore be ordered on a scale from less sonorous to more sonorous, based on openness; this is referred to as the sonority hierarchy, one model of which is given here:

SONORITY HIERARCHY:

oral stops < fricatives < nasals < liquids < glides < vowels (Foley 1972, cited in Clements 1990: 286)

It follows that the flow of speech consists of sonority peaks and sonority troughs, in which a syllable may be defined as the organisation of sounds around sonority peaks. This is demonstrated in Figure 2.1, where the syllable [brænd] 'brand' has one sonority peak, while a rearrangement of its constituent segments as [rbædn] has three peaks, and would potentially be perceived as three syllables if one were to try to pronounce it.

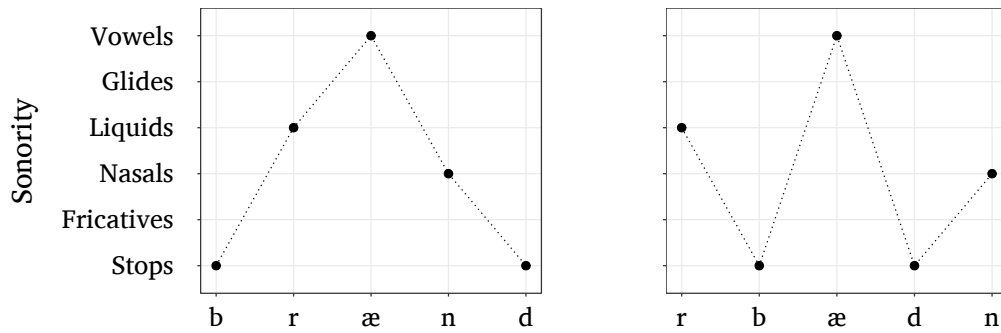


FIGURE 2.1: sonority peaks

The organisation of a syllable is governed by what has been termed the Sonority Sequencing Principle.

SONORITY SEQUENCING PRINCIPLE (SSP):

Between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted. (Clements 1990: 285)

The peak forms the nucleus (N) of the syllable (σ) and is usually a vowel, but may be a syllabic consonant if that is the most sonorous segment. Less sonorous segments (usually consonants) on either side of the nucleus make up the onset (O) and the coda (Co); the nucleus and the coda together make up the rhyme (R), as demonstrated in Figure 2.2.

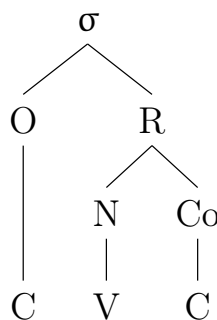


FIGURE 2.2: syllable structure

The nucleus is the only obligatory constituent of a syllable, although different languages may impose further constraints. Some languages (including Arabic) have obligatory onsets, and others disallow codas or certain types of coda, but the opposite constraints are not attested and no language bans onsets or imposes obligatory codas. Syllables with a coda are closed, while those without a coda are open; possible syllable types are shown in Figure 2.3.

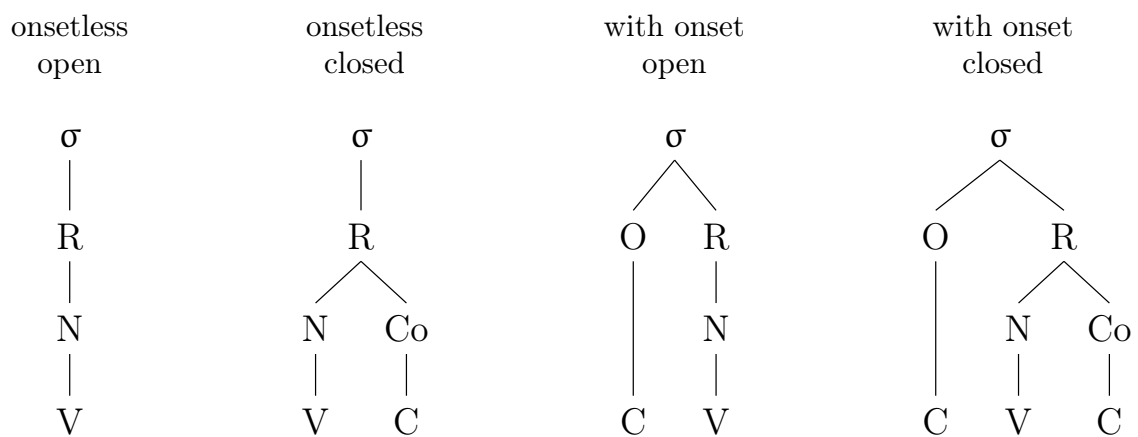


FIGURE 2.3: syllable structure of open and closed syllables

It has been observed in many languages that the composition of the rhyme, but not the onset, is the most important factor in determining phenomena such as stress allocation and the size of the minimal word (Hayes 1995, cited in Broselow 2018: 35). The composition of the rhyme thus determines the weight of the syllable, which is expressed by an abstract unit: the mora (μ). The mora is a sub-syllabic constituent which exists within a hierarchy: segments are grouped into moras, which in turn are grouped into syllables. Larger domains than the syllable have been proposed, resulting in the following hierarchy in

which each constituent must contain at least one unit from the constituent below it (cf. Zsiga 2013: 341):

PROSODIC HIERARCHY
 Utterance (U)
 Intonational phrase (I)
 Phonological phrase (Φ)
 Phonological word (ω)
 Foot (F)
 Syllable (σ)
 Mora (μ)

It follows, therefore, that a syllable must contain at least one mora. In most versions of moraic theory, short vowels are associated with one mora only, while long vowels are associated with two moras (Figure 2.4). When these moras are grouped into syllables, the monomoraic syllable is ‘light’ while the bimoraic syllable is ‘heavy’.

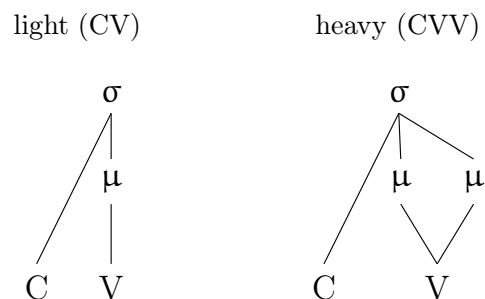


FIGURE 2.4: syllable structure of CV and CVV syllables

In many versions of moraic theory, geminate consonants are also associated with a mora and are thus distinguished from non-geminate consonants (cf. Hayes 1989: 257). A geminate consonant therefore has an association with the last mora of the syllable it closes and with the onset of the next (Figure 2.5).

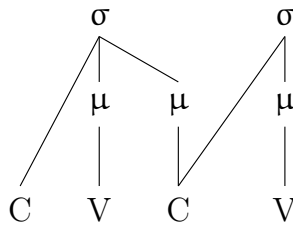


FIGURE 2.5: syllable structure of geminate consonants

Research has shown that in many languages CVV and CVC syllables pattern together, which is to say that CVC syllables are also heavy. Non-geminate coda consonants are therefore assigned a mora by the Weight-by-Position condition (Hayes 1989: 258) as shown in Figure 2.6.

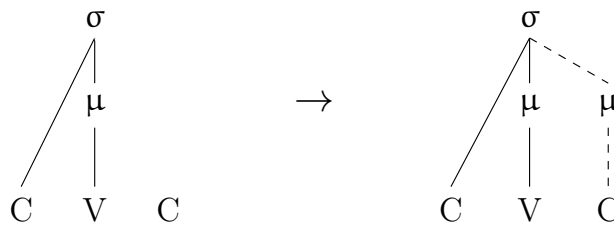


FIGURE 2.6: Weight-by-Position

These principles of Hayes's version of moraic theory are summed up by Watson:

SOURCES OF SYLLABLE WEIGHT:

- (a) Short vowels contribute one mora and long vowels two moras (universal).
- (b) Geminate consonants contribute one mora (universal).
- (c) Weight-by-Position: a 'coda' consonant is assigned a mora in the course of syllabification (parametric). (Watson 2002: 54)

It will be seen that the heavy syllables CVV (Figure 2.4) and CVC (Figure 2.6) are both bimoraic. It has been proposed (Broselow 1992a: 10) that a bimoraic syllable is the maximum weight of a syllable in many varieties of Arabic, as expressed in the bimoraicity constraint:

BIMORAICITY CONSTRAINT

Syllables are maximally and optimally bimoraic. (Broselow 1992a: 10)

This analysis explicitly proscribes trimoraic syllables. Furthermore, it stipulates that two moras is not only the maximum weight of a syllable but also the preferred weight. This predicts that the language will employ syllabification processes which are motivated by grouping moras into bimoraic syllables, referred to as ‘maximising bimoraicity’. It will be explored in the following discussion of syllables in the Rwala dialect whether this preference for bimoraic syllables is supported.

The syllables CVVC and CVCC may be ‘superheavy’ or trimoraic in some languages (McCarthy 1979b: 446). Assuming the Arabic bimoraicity constraint, however, CVVC and CVCC syllables must be constructed in such a way that they do not add a third mora to the syllable. In many dialects of Arabic, these syllables are restricted to domain-final position, where the domain is language specific; for example, CVCC is restricted to phonological word-final position in San’ani Arabic but to utterance-final position in Cairene Arabic (Watson 2002: 58). Research on Arabic has proposed that the final C is extrasyllabic and is not incorporated into the adjacent syllable at any stage in the derivation (Watson, *ibid.*, McCarthy and Prince 1990: 14). The superheavy syllables therefore consist of one heavy syllable and one extrasyllabic C (Figure 2.7).



FIGURE 2.7 syllable structure of CVVC and CVCC syllables

The non-moraic C linked directly to a syllable node in Figure 2.7 does not constitute a well-formed syllable, since the prosodic hierarchy stipulates that a syllable must contain at least one mora: this syllable is therefore described as degenerate (cf. Kager 1995: 376). The degenerate syllable is tolerated in peripheral position according to the language constraint (word-final, phrase-final, and so on), but when the degenerate syllable is not domain-final, syllable repair processes may be triggered (as dealt with in §2.2.2).

2.2.1 Syllable types

Dialects of the Arabian Peninsula are divided between phonologically conservative ones, in which superheavy syllables (CVVC and CVCC) are restricted to word-final position, and innovative dialects in which superheavy syllables may also occur in non-final position (Watson 2011a: 900). Table 2.2 lists the syllable types attested in Rwala.

TABLE 2.2: Syllable types

	open	closed	doubly closed
light	CV	—	—
heavy	CVV	CVC	—
superheavy	—	CVVC	CVCC

It will be seen that this inventory of syllable types excludes onsetless syllables, which are not found in the dialect. There is, however, a limited number of vowel-initial morphemes in the dialect, including: a large number of particles such as the relative pronoun *alli*, the definite article *al-*, and conjunctions such as *elya* ‘when’ and *en* ‘if’; a number of pronouns such as *ana* ‘I’ and *ent(e)* ‘you (MS)’; words with prefix morphemes such as *a-* which is the 1CS marker in the P-STEM and also part of the template of the FORM-IV S-STEM; words with prothesis of a vowel such as the imperatives, e.g. *erkab* ‘mount!’, and the S-STEM of some derived verbs, e.g. *enḏabaḥ* ‘he was killed’ (FORM-VII); a small number of historically vowel initial content words such as *is(e)m* ‘name’ and *ib(e)n* ‘son’; some loan words such as *enklēsi* ‘a variety of Mauser revolver’; and words in which the deletion of /ʔ/ has left the first syllable onsetless, such as *ahl* ‘kin’, *arḏ* ‘earth’ and *iḏen* ‘ear’.

When these morphemes occur in utterance-initial position, the restriction on onsetless syllables triggers epenthesis of [ʔ] (although this is not transcribed by Musil), as shown in Figure 2.8.

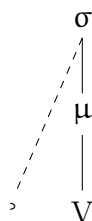


FIGURE 2.8: glottal stop epenthesis

Within the utterance, however, if the onsetless vowel is preceded by a consonant then this coda consonant may be resyllabified as an onset to the following onsetless syllable, and if it is preceded by a vowel then a hiatus avoidance strategy will occur (cf. §2.3.5).

The northern and central bedouin dialects of Arabia are classified as part of the innovative group which allows non-final superheavy syllables (Watson 2011a: 900). Musil's records also place the Rwala dialect in this group, but barely so: the superheavy syllables are indeed attested non-finally, but they only account for <2% of all non-final syllables. Non-final superheavy syllables occur in the innovative group for two reasons: the dialects have pervasive syncope rules which create such syllables (cf. §2.2.3), or they do not have the pervasive syllable repair processes which block such syllables in the conservative group (cf. §2.2.2).

The non-final superheavy syllables CVVC and CVCC occur as a result of syncope in the following Rwala examples from Musil:

/šāḥeb + i/	>	šā.ḥe.b-i	>	šāḥ.b-i	'my friend'
/gamma + eh/	>	ḡam.ma.t-eh ⁶³	>	ḡamm.t-eh	'its footboard'

As can be inferred from the very low attestation of non-final superheavy syllables, syncope is shown to be much less active in Musil's data than would be expected of an 'innovative' dialect: note that the forms *šāḥeb-i* and *ḡamm.t-eh* are attested in both cases. Non-final CVVC and CVCC also occur as a result of the concatenation of suffixes in the

⁶³ The ordering of the affrication rule /g/ > ḡ is not relevant to this example.

following examples, in which a repair strategy (epenthesis is given as an example) has not taken place:

/rās + ha/	>	<i>rās.-ha</i>	‘her head/hair’	(not * <i>rā.s-[e].ha</i>)
/kill + hom/	>	<i>kill.-hom</i>	‘all of them’	(not * <i>kil.l-[e].hom</i>)

In both of these cases, the epenthesised forms *rās-eha* and *kill-ehom* are also attested.

This is discussed in more detail below, but, in general terms, the relative paucity of non-final superheavy syllables in Musil’s data stems from the fact that syncope occurs less often than would be expected and epenthesis occurs more often than would be expected.

While the dialect does not have a restriction therefore on the occurrence of superheavy syllables, there is a restriction on the occurrence of CVV which may not occur in word-final position. This is on account of a word-final vowel shortening rule, and the situation is described in more detail in §2.3.3.

Some super-superheavy syllables (i.e. CVCCC or CVVCC) are recorded. I consider the few examples of CVCCC to be completely anomalous: the only two clear examples are *galb-k* ‘your heart’ (but note that *galb-ak* is attested four times) and *wašf-k* ‘your description’. Of the other examples of word-final CVCCC, two arguably exhibit Musil’s completely arbitrary glottal stop insertion (*ćann’ mā b-i-š* ‘as if nothing ailed me’ and *ćenn’ hal^{en} l-i* ‘like kin to me’) and the others occur within the domain of the phonological phrase, which means that they are not in fact super-superheavy at all (see next section). CVVCC is much more frequent, but usually occurs as a result of syncope within the phonological phrase, which again means that they are not truly super-superheavy syllables. The exception is the active

participle of the FORM-I doubled verb which has the template $C\bar{V}C_iC_i$, which is fairly frequently attested. As is discussed below, there are phonological reasons why final CVVCC should be more common than final CVCCC, just as word-medial CVVC is more common than word-medial CVCC.

The dialect tolerates word-initial consonant clusters which occur as a result of syncope, which makes CCV another possible syllable type. As shown in §2.2, syllables are formed around sonority peaks, so in words like *brakat* ‘she knelt’ the first syllable, .bra., has ascending sonority and constitutes a CCV syllable while other clusters, such as .wle. in *wled-i* ‘my boy’, have two sonority peaks and probably represent two syllables, the first of which is degenerate: .w.le. Onset clusters such as these are considered further in §2.2.3.5.

2.2.2 *Syllabification*

Syllabification takes place according to the following algorithm, adapted from Hayes (1989: 258-9) and Watson (2002: 63):

- (a) moras are assigned to vowels and geminate consonants (i.e. the underlying representation, cf. Figures 2.4 and 2.5)
- (b) word-final consonants are extrametrical
- (c) moras which dominate vowels are associated with a syllable node
- (d) single prevocalic consonants are adjoined to the syllable node
- (e) moras are assigned by Weight-by-Position (cf. Figure 2.6)
- (f) unsyllabified moras are adjoined to the syllable node
- (g) an extrametrical consonant is adjoined to preceding syllable node if the syllable governs only one mora

Syllabification is demonstrated in Figure 2.9, illustrated by /gaṭa/ ‘sand grouse’, /tamra/ ‘date’, /ḍabaḥ/ ‘he killed’, /ḍabbāḥ/ ‘murderer’, and /gilt + ha/ ‘I said it (f.)’.

Concerning (b), word-final consonant extrametricality has been proposed for a number of languages, including Arabic, to account for the fact that CVC syllables are heavy (bimoraic) word-internally but light (monomoraic) word-finally (cf. Hayes 1995: 58-60). Evidence for this principally comes from the assignment of word stress, which Musil’s records shed no light on. His data do, however, show that a single CVC syllable is not a permitted word. The minimal word is either monosyllabic CVVC, e.g. *bāb* ‘gate’, or monosyllabic CVCC, e.g. *bint* ‘daughter’, or disyllabic CVCV, e.g. *gaṭa* ‘sand grouse’. Monosyllables of the form CVC are disallowed as content words and are restricted to clitics, such as *min* ‘from’. I therefore assume that CVC is monomoraic unless it is word-medial (and therefore assigned a mora by Weight-by-Position), and include (b) and (g) in the above algorithm.

The concept of a minimal word may be expressed in terms of the prosodic hierarchy (cf. §2.2), which stipulates that a prosodic constituent contains at least one unit from the tier directly below it. Since the phonological word governs the foot, it is also assumed that a foot is minimally bimoraic. In the following analysis, the minimal word is bimoraic, and monomoraic words are seen as sub-minimal. The Rwala dialect has a small number of sub-minimal content words, but these are usually lengthened in various ways. In §2.1.1.1 it was

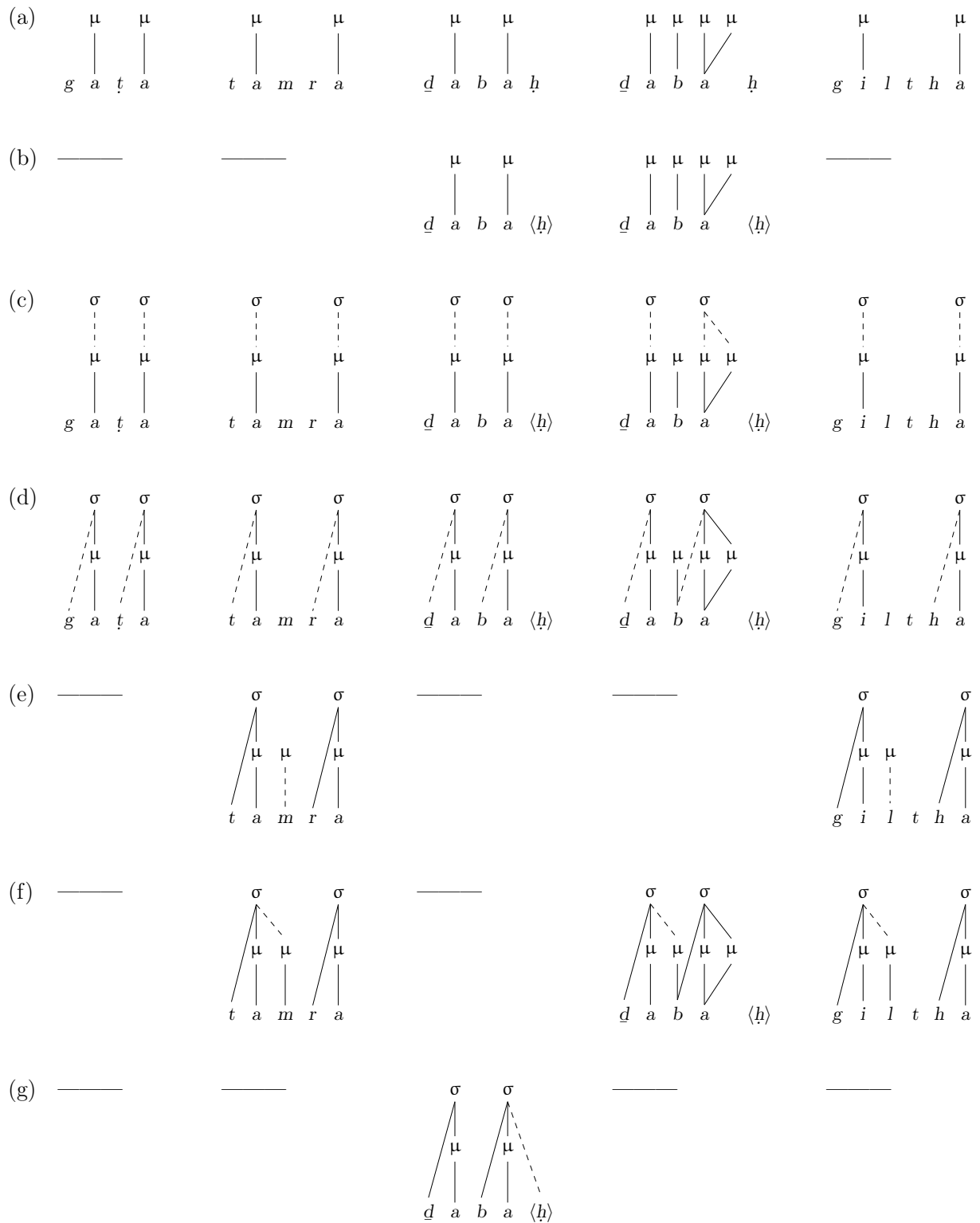


FIGURE 2.9: syllabification

seen that the sub-minimal word *šay*, *ši* ‘thing’ is produced by at least some speakers at *šiyye*, which meets the minimal word requirements. Similarly, *yad* ‘hand’ is more regularly

attested as *īd*. Some sub-minimal words are lengthened in their suffixed forms, such as *ab* ‘father’ but *abū-k* ‘your (MS) father’, and *ġan* ‘they (f.) came’ but *ġann-ah* ‘they (f.) came to her’ (cf. §2.3.4). It may be noted that a CVV syllable does meet the bimoraic requirement for a minimal word, but none of these words (e.g. *lā* ‘no’, *yā* ‘[vocative particle]’) are content words; this is not a question of minimality but rather breaches a separate requirement, namely that Arabic stems must end in a consonant (McCarthy and Prince 1990: 18).

With regard to (e), it is recalled that Weight-by-Position (Figure 2.6) occurs only when the preceding vowel is dominated by one mora only, thus eliminating the possibility of Weight-by-Position creating trimoraic syllables (CVVC or CVCC). The stipulation in (g) ensures that only monomoraic CVC is syllabified word-finally, to the exclusion of CVVC. This is necessary since extrametrical consonants are adjoined to the preceding syllable node and thus distinguished from the final C in CVVC and CVCC, which is extrasyllabic and is not incorporated into the preceding syllable at any stage (cf. §2.2).

It is observed that the syllabification algorithm leaves consonants unsyllabified in the forms *ḍab.baa.ḥ* and *gil.t.ha*: such consonants are commonly referred to as ‘stray’ consonants (C[˘]) (Hayes 2009: 265). Stray consonants are subject to deletion in many languages, a process which is termed Stray Erasure (consider Spanish /eskulp + tura/ > es.kul.p.tu.ra > [eskultura]). Arabic, however, does not generally permit the deletion of consonants, possibly because consonants carry most of the lexical information in Arabic (cf. Broselow 1992a: 30). The data show that stray consonants occur frequently word-finally, in words such as *ḍabbāḥ*

and many others, e.g. *bāb* (baa.b) ‘gate’ and *bint* (bin.t) ‘daughter’. I therefore add a further step to the syllabification algorithm:

- (h) a stray consonant is adjoined to a new syllable node, $\langle\sigma\rangle$

In Figure 2.10 this is illustrated for /*ḍabbāḥ*/ and /*gilt+ha*/ only, since /*gaṭa*/, /*tamra*/ and /*ḍabaḥ*/ do not have stray consonants.

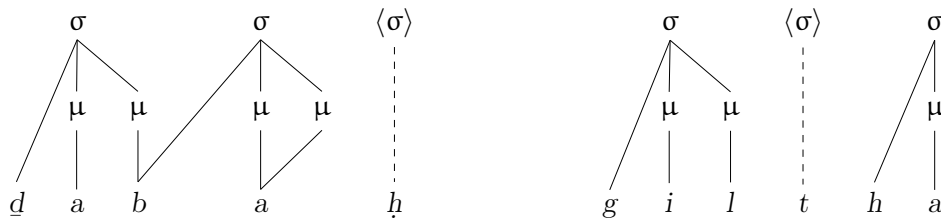


FIGURE 2.10: stray consonant syllabification

The sequences CVVC and CVCC can thus be seen to conform to the syllable structure of these syllables given in §2.2 (Figure 2.7), where it was proposed that ‘superheavy’ syllables are in fact composed of a single bimoraic syllable (CVV or CVC) and a degenerate syllable. These degenerate syllables are permitted word-finally, but where they occur word-medially (*gil.t.ha*) they are not tolerated and resolved by a syllable repair strategy (cf. §2.2.2.1).

With no data on the stress patterns in Musil’s material, it is difficult to establish the domain of syllabification with any certainty. Some examples make it possible to see the extent to which syllabification extends beyond the syntactic word, however:

(1a)	<i>walad + i</i>	>	wle.di	‘my boy’
(1b)	<i>tara + ni</i>	>	ta.rā.ni	‘look at me!’
(2a)	<i>be + miṣlāb</i>	>	bemṣ.lāb	‘with a fire poker’
(2b)	<i>ğara + l-i</i>	>	ğe.rā.li	‘it happened to me’
(2c)	<i>balċi + tiswi</i>	>	bal.ċits.wi	‘perhaps you will prepare’
(3a)	<i>taṭla^c + aš-šams</i>	>	taṭ.la.‘aš(.šams)	‘the sun rises’
(3b)	<i>enna + al-leyl</i>	>	en.nal(.leyl)	‘the night is’
(3c)	<i>rā^c al-hawa</i>	>	rā.‘al.ha.wa	‘the beloved’
(3d)	<i>rā^ċ l-beyt</i>	>	rā.‘il.beyt	‘the master of the tent’
(4a)	<i>habba + ar-rīḥ</i>	>	habb.tar.rīḥ	‘blowing of the wind’ ⁶⁴
(4b)	<i>ḥamsa + ar-nāg</i>	>	ḥams.tar.nāg	‘five spices’

The examples in (1) show that suffixes form part of the phonological word, as exemplified by the fact that syllabification processes such as syncope occur (1a) and that final-vowel shortening does not (1b). The examples in (2) show that this is not limited to suffixes, since prepositions (2a), prepositional phrases (2b) and even some conjunctions (2c) appear to form part of the phonological word. Examples (3a) and (3b) demonstrate that the definite article has two allomorphs, /al/ following a consonant and /l/ following a vowel⁶⁵; the contrast in (3c) and (3d) however demonstrate that before a stressed definite article (in 3c, and see discussion in §5.1.14) a preceding vowel is deleted, while before an unstressed definite article (3d) the same vowel is maintained. While the examples in (3), which are all in the construct state, do not prove that construct phrases are a single syllabifiable domain, the examples in (4) display the kind of syncope which would be associated with one. Part of the motivation for this theory is the need to avoid super-superheavy syllables: arguably the

⁶⁴ This idiom implies a favourable wind that brings success if it blows upon you.

⁶⁵ In the traditional view, the /l/ here assimilates to the following consonant in the case of coronals and is /l/ before all other consonants (cf. §5.1.14). Though not transcribed by Musil, [ʔal] is assumed to occur utterance-initially only, and the [ʔ] is epenthetic.

syncope that derives *ḥamst* < *ḥamsat* would never occur were the word not phonologically part of a larger unit which permits this (compare also *ṭamānt ayyām* ‘eight days’ and *ḥamsta‘aš* ‘fifteen’).

2.2.2.1 Epenthesis

As has been seen in the previous section, epenthesis of *a* or *e*⁶⁶ may occur when concatenation causes non-final superheavy syllables (CVVC or CVCC) to occur, as in the following examples:

<i>‘eyaal + na</i>	>	<i>‘e. yaa. la. na</i>	<i>‘our children’</i>
<i>gilt + ha</i>	>	<i>gil. ta. ha</i>	<i>‘I said it (f.)’</i>
<i>zaml + na</i>	>	<i>zam. la. na</i>	<i>‘our pack camels’</i>

In §2.2.2 it was proposed that stray consonants which are not syllabified by the syllabification algorithm are adjoined to a syllable node, in order to incorporate them into prosodic structure and thus protect them from Stray Erasure. Stray consonants are thus represented as in Figure 2.11.



FIGURE 2.11: stray consonant syllabification (2)

⁶⁶ There are ten examples of epenthesis of *i* in the corpus but this tiny group does not seem statistically significant. There may be some evidence that epenthetic *i* is somewhat motivated by vowel harmony, since it occurs after /i/ or /ī/ (e.g. *ḥarīb-ina* ‘our enemy in war’, *ḡīt-ini* ‘you came to me’) but the evidence is too scant to verify this with any certainty.

It is recalled that the prosodic hierarchy (cf. §2.2) stipulates that each higher constituent contains at least one unit of the element below it, which is to say that a syllable must contain at least one mora. The syllable to which a stray consonant is adjoined does not, and is therefore degenerate. The Rwala dialect repairs this through epenthesis, which supplies the degenerate syllable with a vocalic nucleus and therefore a mora (Figure 2.12).

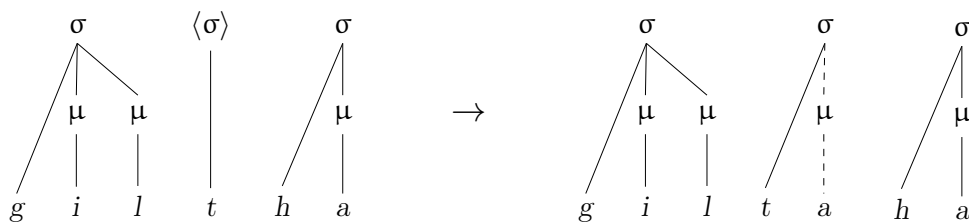


FIGURE 2.12: epenthesis

The minimal syllable for a stray consonant in this dialect is shown in Figure 2.13(a) (cf. Broselow 1992a: 30). This ensures that epenthesis will occur to the right of an unsyllabified segment.

The position of epenthesis has attracted considerable attention in Arabic dialect typology. Selkirk (1981) observed that in some dialects the stray consonant becomes the onset of the epenthetic nucleus (as in this dialect, illustrated by *gil.ta.ha*), while in other dialects it becomes the rhyme. The latter can be illustrated by Iraqi *kitab-it-l-a* ‘I wrote to him’:

ki. tab. t. la > ki. tab. it. la > ki. ta. bit. la⁶⁷

⁶⁷ Example from Broselow (2018: 42).

Broselow (1992a) terms the former type ‘onset’ dialects and the latter type ‘rhyme’ dialects. In Broselow’s analysis, onset dialects link a stray consonant directly to a syllable node (as in Figure 2.11) while rhyme dialects link it to a mora. This contrast is shown in Figure 2.13.

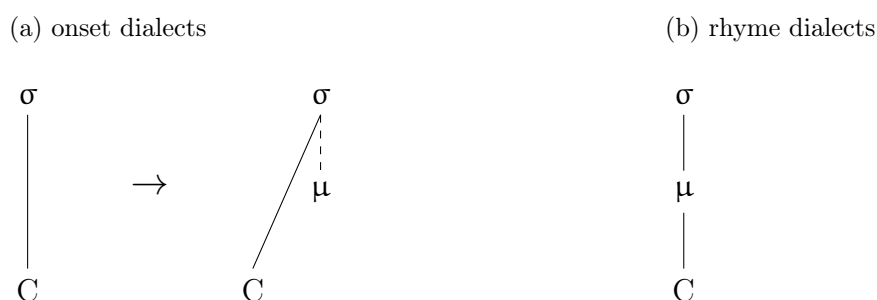


FIGURE 2.13: degenerate C syllables in onset and rhyme dialects

It follows that in rhyme dialects, i.e. those in which the stray consonant is moraic, two options are viable. The C syllable may not require epenthesis, since a vowel-less syllable is not proscribed by the prosodic hierarchy (cf. 2.2), in which case a CCC cluster will maintain. Alternatively, if epenthesis occurs, it will be to the left of the stray consonant with the C forming a moraic coda, and CCC is broken into CvCC. The non-moraic C of onset dialects must always trigger epenthesis, however, with CCC broken into CCvC. These three paths are illustrated with examples of the phrase ‘I/you (m.) said to him’ from three dialects (examples from Kiparsky 2003: 150, my emphasis):

Rhyme dialects			
Moroccan	<i>qəltlu</i>		(qəl.t.lu)
Iraqi	<i>gʕlitla</i>		(gi.lit.la)
Onset dialect			
Cairene	<i>ʕultʕlu(h)</i>		(ʕul.ti.lu[h])

This analysis is supported by further evidence from the Rwala dialect, as unsyllabified moraic consonants do occur in restricted morphological contexts (cf. §2.2.2.2). In these cases, clusters are resolved by prothesis to the left of the consonant, rather than rightwards epenthesis as occurs with non-moraic consonants.

Thus far it has been argued that a stray consonant will always trigger epenthesis. The data show that this is not always the case, however, since the dialect does allow non-final superheavy syllables, suggesting that epenthesis is optional:

<i>‘eyaal + na</i>	>	<i>‘e. yaal. na</i>	‘our children’
<i>darb + na</i>	>	<i>darb.na</i>	‘our way’

Of these, epenthesis after CVVC is completely optional while epenthesis after CVCC occurs in the vast majority of cases, for reasons addressed below. While it may be considered that the innovative group (i.e. those dialects which permit non-final superheavy syllables) would not exhibit a high degree of epenthesis, the fairly extensive occurrence of epenthesis as shown in Musil’s data is in line with other evidence for the same dialect. Prochazka (1988: 199, 201) records *bāyt-ana* ‘our tent’ for Rwala, which corresponds to Musil’s *beyt-ena*, but contrasts with the form *bēt-na* from Riyadh or Bīshah (Prochazka *ibid.*), which are also from the innovative group. Similarly, his *’iḥt-ana* ‘our sister’ corresponds to Musil’s *bint-ena* ‘our daughter’, and contrasts with *’iḥt-na* from Riyadh and *bint-ne* from Hofuf. Prochazka’s data thus confirms that the Rwala dialect is not so much a part of the innovative group in terms of epenthesis, which implies that Musil transcribed these forms faithfully.

The optional nature of epenthesis after CVVC is not unattested in the dialects of Arabic. Broselow (2018: 38) lists Sudanese, the dialects of Upper Egypt, and the dialects of the Levant and of the Gulf as varieties that allow non-final CVVC after suffixation (i.e. dialects in which *Rwala* ‘*eyāl-na* would be grammatical), and Watson (2007: 348) lists San’ani, al-Hudayda, Yaafi^q, Yarīmi, Ibbi, Middle Egyptian, and Meccan. Furthermore, other dialects show the same variation between forms with vowel insertion and forms without, such as San’ani *kitāb-anā* ~ *kitāb-nā* ‘our book’ (Watson 2002: 69).

The question of whether the tolerance of non-final CVVC appears to permit trimoraic syllables has been addressed by Broselow (1992a), where a mora-sharing option (called Adjunction-to-Mora) is suggested to explain this. Mora-sharing proposes that the coda consonant of CVVC is allowed to share the second mora of the preceding vowel, preserving the bimoraic status of the syllable (Figure 2.14). Non-final CVVC is therefore contained within a single syllable (as opposed to final CVVC which is one syllable and one degenerate syllable), and the final consonant does not add weight to the syllable on account of being allowed it to share the second mora of the preceding vowel, rather than adding a (third) mora of its own (Broselow 2018: 39). Mora-sharing is considered to be able to produce single CVVC syllables on account of sonority: since the sonority distance between VC is greater than that between CC (cf. the sonority hierarchy, §2.2), VC is able to share the final mora of a syllable whereas CC is not (Broselow 1992a: 15). This accounts for why word-medial CVVC occurs much more frequently than word-medial CVCC.

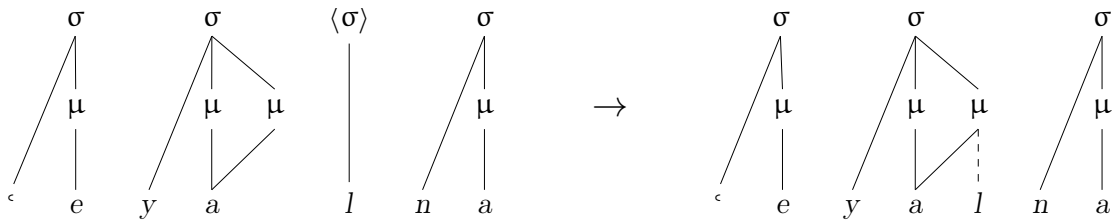


FIGURE 2.14: Adjunction-to-Mora

This theory is supported by acoustic studies: in Broselow *et al.* (1995, 1997) it has been demonstrated that in Jordanian, Syrian and Lebanese varieties of Arabic (which allow, e.g., *kitāb-na*) the long vowel of CVVC is shorter in duration than in CVV, though still longer than CV, and that the coda consonant of CVVC is shorter than that of CVC. Other methods by which a trimoraic syllable could be avoided are not attested for Rwala: the dialect does not employ removing a mora from the long vowel (i.e. closed syllable shortening, e.g. **eyal-na* < *eyāl-na*) as would occur in, for example, Cairene Arabic, nor does it employ deletion of the coda consonant (e.g. **eyā-na* < *eyāl-na*), a theoretical strategy not attested in any Arabic dialect⁶⁸ (Broselow 2018: 40). It seems, therefore, that the Rwala dialect provides evidence of mora-sharing in operation.

As opposed to the variation seen with regard to CVVC syllables after suffixation, the dialect tends not to allow the same variation with CVCC syllables: forms such as *eyāl-na* above occur frequently but forms such as *darb-na* are rare. This pattern is not exclusive to this dialect, but in fact many dialects which allow word-medial CVVC do not allow word-

⁶⁸ An exception to Broselow's assertion that consonant deletion is not attested in any Arabic dialect as a syllable repair strategy may be found in Holes (2016: 73) where the forms *id-na* (< **ind-na*) 'we have' and *bit-na* (< **bint-na*) 'our daughter' in some Baḥārna village dialects (especially in women's speech), although it is noted that /l/ and /n/ are particularly prone to deletion in this position (*ibid.* note 37), e.g. *git* (< **gilt*) 'I said' and *kit* (< **kint*) 'I was'. This affects CVCC syllables only, however.

medial CVCC (Broselow *et al.* 1995: 122). It has been established that this may be because, while CVVC syllables may permit the sharing of a mora between a vowel and a consonant, a similar sharing of the mora between two consonants is not permitted⁶⁹. As such, any resulting CCC sequence must be resolved by epenthesis, as in the case of *zaml-ana*. Where epenthesis does not occur, I propose that the stray consonant is adjoined directly to the syllable node dominating the preceding vowel, thus preserving the bimoraic status of the syllable (Figure 2.15).

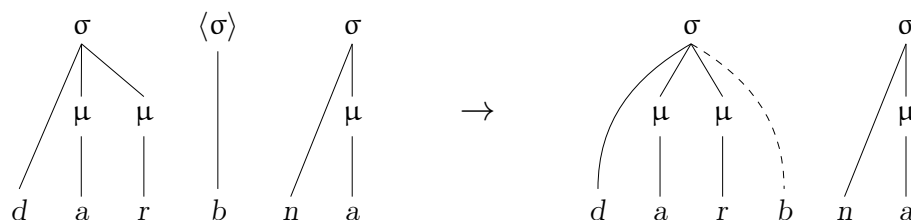


FIGURE 2.15: complex coda formation

This appears to be in violation of the syllabification algorithm given in §2.2.2, where coda consonants are only adjoined by Weight-by-Position, which effectively restricts syllables to one coda consonant (i.e. CVC). As illustrated by the examples *zaml-ana* and *darb-na*, however, such complex codas are restricted to clusters of falling sonority (cf. §2.2): /ml/ as a nasal-liquid cluster has rising sonority and is therefore not a permitted coda, while /rb/ as a liquid-stop cluster is sometimes permitted. Complex coda formation has not been

⁶⁹ “Under the analysis in which languages that tolerate non-final CVVC syllables do so by assigning them a bimoraic analysis in which the coda consonant shares a mora with the preceding vowel, CVCC syllables can be ruled out by a ban on the sharing of a single mora among two segments that are close together on the sonority scale.” (Broselow 2018: 40)

included in the syllabification algorithm on account of its infrequency (it is attested only 23 times in the entire corpus).

The exception to this sonority-based explanation is that epenthesis may also not be induced following a geminate consonant, e.g. *fakk-na* 'liberate us!' (not **fakk-ana*), in which case CCC is more likely to occur. This may be because geminate consonants are also able to share a mora.

Assuming that the epenthetic vowel *e* is a variant of /a/, the rule for optional epenthesis insertion can therefore be expressed as:

$$\emptyset \rightarrow a / C' _$$

This formulates that epenthesis occurs to the right of a stray consonant (C'); the rule is optional since Adjunction-to-Mora occurs as an alternative repair strategy in VVC' sequences (and complex coda formation occurs infrequently in CC' sequences). It must be noted that the other sources of the Rwala dialect indicate that a slightly more specific epenthesis rule is needed, namely:

$$\emptyset \rightarrow a / C' _ +$$

This formulates that epenthesis occurs between a stray consonant and a morpheme boundary (represented by +), and this needs to be included in the formulation because there is only very limited evidence of epenthesis occurring without a morpheme boundary in that position. As discussed in more detail in §2.2.3, syncope rules regularly produce word-medial CCC and VVCC which may not be repaired by epenthesis; e.g. Prochazka (1988: 34,

47) records the forms *yakbūn* (< **yaktibūn*) ‘they write’ and *tsāfrīn* (< **tsāfirīn*) ‘you (f.) travel’ as regular verbal forms in which epenthesis has not operated. Parallel forms in Musil would be **yaktebūn* and **tsāferīn*, although this may be interpreted in different ways: it may represent a failure of syncope to delete the vowel in that position, but may equally represent an interaction of syncope and epenthesis in which the vowel is deleted and subsequently replaced in the same position. Since there is some evidence supporting the latter interpretation, the inclusion of morpheme boundaries in the above rules does not seem necessary to produce the correct output for Musil’s material.

One final possible occurrence of epenthesis is not explained by the above rules, which is that monosyllabic words having the template CVCC may exhibit epenthesis and surface as CVCeC. It is considered that the template of these words is monosyllabic because these are attested as variants, e.g. *şaḥen* ~ *şaḥn*, and also because suffixation produces forms such as *matn-aha* ‘her waist’, where arguably epenthesis would not have been triggered if the underlying stem were disyllabic (**meten-ha*). The following examples also have the form CVCC in their CLA cognates:

CaCC	<i>şaḥen</i> ‘shallow copper pan’, <i>meten</i> ‘waist’, <i>semen</i> ‘melted butter’
CiCC	<i>tiben</i> ‘straw’, <i>ḥeber</i> ‘ink’, <i>meleḥ</i> ‘salt’
CuCC	<i>ḥubez</i> ‘bread’, <i>goşen</i> ‘twig’, <i>ḥesen</i> ‘beauty’

In most of these examples, the final consonant has a higher sonority than the consonant which precedes it, with nasals and liquids being the most common (*meleḥ* is an exception to this, and *semen* has a sonority plateau). Ingham (2011) records the same

situation for Najdi dialects in general, although restricts its application to word-final consonant clusters where the second element of the cluster is /l, r, w, y, n/; arguably, however, examples such as *hubez* indicates that sonority in general could be the reason for this epenthesis. An alternative explanation would be that in word-final consonant clusters with ascending sonority the final consonant forms its own syllabic peak, and therefore that the process may not be one of epenthesis *per se* but rather the transcription CVCeC indicates Musil's perception of the disyllabic structure of the consonant cluster. This would be impossible to judge from Musil's records alone, but perhaps it is safe to follow Ingham who also describes this phenomenon as epenthesis⁷⁰. The syllabic status of consonant clusters is discussed in more detail in §2.2.3.5 below.

2.2.2.2 Prothesis

It follows from the analysis of onset dialects and rhyme dialects in §2.2.2.1 that word-initial vowel insertion (prothesis), if it occurs, should be much more restricted in occurrence in the former than in the latter. This may be illustrated by examples from a study of the pronunciations of the English words 'three' and 'floor' by Arabic speakers: Egyptian (onset) speakers uniformly gave [θirii] and [filoor] while Iraqi (rhyme) speakers gave [iθrii] and [ifloor] (Broselow 1992b: 74). Accordingly, since the Rwala dialect is an onset dialect,

⁷⁰ To be precise, Ingham refers to it as anaptyxis, but we are referring to the same process.

the sites of prothesis are restricted to a few morphological categories, namely the FORM-I imperative and the S-STEM of several of the derived forms of the verb.

Prothesis therefore acts as a syllable repair process only in predefined contexts. In §2.2.2.1 it was observed that an unsyllabified consonant which is non-moraic triggers epenthesis to the right (cf. Figure 2.13a) while epenthesis to the left is attested in rhyme dialects where the unsyllabified consonant is moraic. In the Rwala dialect, FORMS VII-X of the verb exhibit prothesis, illustrated by FORM-VII *enḏabaḥ* < /nḏabaḥ/ ‘he was killed’. According to the syllabification algorithm so far established, /n/ would form a degenerate syllable, and epenthesis (if it occurred) would occur to the right (Figure 2.16).

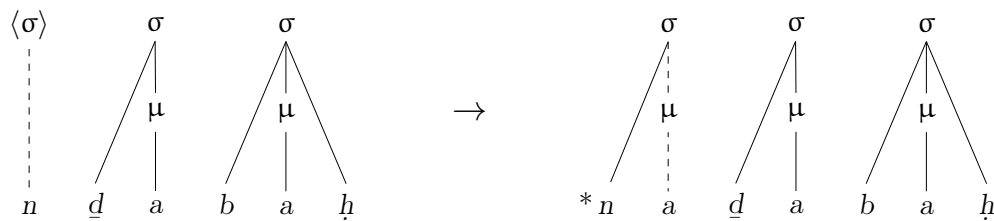


FIGURE 2.16: *initial cluster epenthesis

This does not occur, however, and the correct output is produced only when /n/ is assigned a mora (Figure 2.17).

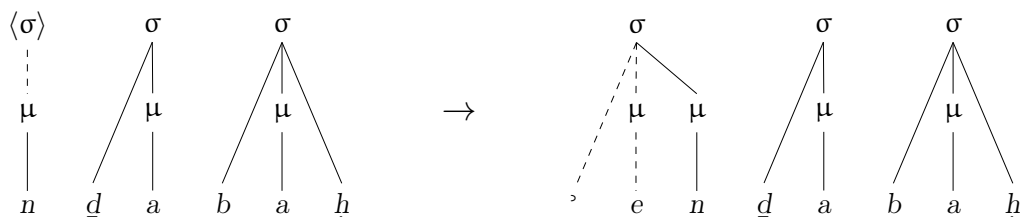


FIGURE 2.17: prothesis

We assume, then, that /n/ of /ndabaḥ/ is linked to a mora in its underlying representation. It has been argued elsewhere (McCarthy and Prince 1990: 13; Watson 2002: 134-135) that this is accounted for by the Prosodic Morphology Hypothesis:

PROSODIC MORPHOLOGY HYPOTHESIS

Templates are defined in terms of the authentic units of prosody: the mora, the syllable, the foot, and the phonological word. (McCarthy and Prince 1990: 3)

In this analysis, the FORM-VII verb template is represented not as a string of C and V segments (nCaCaC) but rather as a prosodic template in which the minimal unit is not the segment but rather the minimal prosodic unit, which is the mora (Figure 2.18, and cf. the Prosodic Hierarchy, §2.2). FORM-VII is therefore explained as derived from the bimoraic verbal stem through the prefixation of a mora, which is then associated with /n/⁷¹.

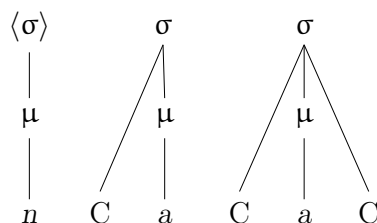


FIGURE 2.18: FORM-VII verb template

The mora associated with /n/ in the underlying representation thus accounts for prothesis. Since the /n/ is moraic, it cannot associate with the non-moraic consonant to its right and thus forms the coda of a degenerate syllable, which is syllabified through prothesis

⁷¹ The analysis of this /n/ as a prefixed morpheme and therefore moraic follows McCarthy and Prince (1990). Since the derived forms VII-X have in common ‘a reduction or minimization of the valence of the underlying verb’ (*ibid.*: 38), the initial incomplete syllable is analysable as a separate morpheme, which they call a detransitivising prefix.

unless preceded by a vowel (hence the forms *enḏabaḥow* ‘they were killed’ but *alli-nḏabaḥow* ‘those who were killed’).

The association of /n/ with a mora also explains why these CCV clusters are resolved at all, rather than remaining as a permitted word-initial cluster. Word-initial CCV clusters are formed in the dialect through syncope (cf. §2.2.3) and the initial C is not usually moraic, hence the cluster is not usually resolved by prothesis. A superficially similar process of vowel insertion does occur before syncope-derived clusters, exemplified by *reḥāb* > *rḥāb* > *erḥāb* ‘riding camels’, but this reflects a different situation and occurs only where syllabification produces two adjacent stray consonants. These are resolved by medial epenthesis, which is therefore not strictly prothesis in the sense discussed here. This is illustrated in Figure 2.19 for the phrase *tamān rḥāb* ‘eight riding camels’; in this example it can be seen that the mora linked to /r/ is assigned by Weight-by-Position following epenthesis, and is not underlying.

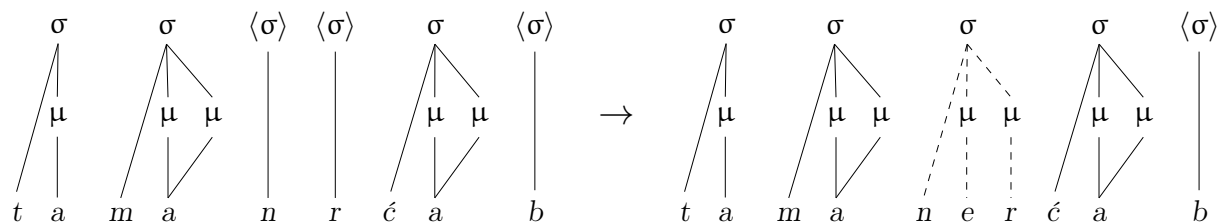


FIGURE 2.19: medial epenthesis between two stray consonants

The prothesis on imperatives is conceptually the same as for FORM-VII, although the initial mora is not explained by prefixation according to the Prosodic Morphology Hypothesis. It seems likely that the initial mora can be explained through the derivation of

the imperative as a truncated part of the P-STEM, since moras tend to be conserved (cf. Hayes 1989). This is illustrated in Figure 2.20, where the form *erkab* ‘ride! (MS)’ is derived from the form *tarkab* ‘you ride (2MS)’ (Figure 2.20a) and would, after the deletion of *ta-*, retain the mora of the coda of the first syllable (Figure 2.20b). As we have seen elsewhere, this would mean that the mora-bearing consonant would be unable to syllabify with the non-moraic consonant to its right and therefore prothesis would be induced.

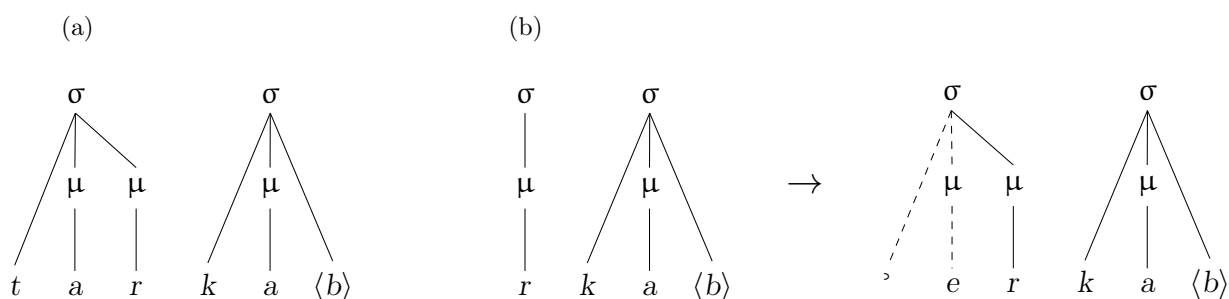


FIGURE 2.20: imperative prothesis

Prothesis can therefore be considered restricted to words with initial consonant clusters in their underlying representation, the first of which must be moraic. As discussed above, this effectively rules out the possibility of prothesis on any initial CC sequence that has been created via syncope (§2.2.3), since the cluster is not present in the underlying representation. The quality of the prothetic vowel differs from that of the epenthetic vowel (which is always a low vowel /a/), as illustrated by the following forms from the FORM-I imperative paradigm:

	/a/-stem	/i/-stem	/u/-stem
MS	<i>erkab</i> ‘mount!’	<i>ebšer</i> ‘rejoice!’	<i>ogʻod</i> ‘sit!’

These forms demonstrate a type of vowel harmony not seen elsewhere in the dialect. This is partly why the initial vowel must be analysed as prothetic and not simply as part of the template (ʔVCCVC), since in templates the vocalic melody forms one level of the non-concatenative morphology: e.g. an active participle must have the melody /a-i/ and a passive participle must have the melody /a-u/. For the imperatives, V2 is lexical, and V1 is governed partially by V2, since the available melodies are /i-a/, /i-i/, and /u-u/; this effectively means that the imperative has no morphologically marked vowel melody. Furthermore, the contrast between *ogḏob* ‘grasp!’ and *yegḏob* ‘he grasps’ shows that vowel harmony occurs only in the imperative.

Such a vowel must therefore be inserted, rather than templatic, and it must be subject to a phonological rule. Since prothesis is quite rare, the rule can be quite specific:

$$\emptyset \rightarrow V \left[\begin{array}{l} \alpha \text{ ROUND} \\ - \text{LOW} \end{array} \right] / \# \text{ } \overset{\mu}{\text{CCV}} [\alpha \text{ ROUND}]$$

This rule establishes that when prothesis occurs the prothetic vowel will firstly be round (i.e. /u/) if the following vowel is round, and not if it is not, and secondly that it will not be low, which precludes it from being /a/. Effectively, the prothetic vowel will be /u/ if followed by /u/ and otherwise /i/. This situation can be expressed even more simply in terms of autosegmental phonology: if the prothetic vowel is specified as [-LOW] but unspecified for roundness, the roundness feature of the following vowel spreads to this

underspecified segment. Right to left vowel harmony (or regressive vowel harmony), while fairly rare cross-linguistically, is attested for several dialects of Arabic (Monahan 2011).

This rule applies in every case of prothetic vowel insertion attested in the dialect, but since the verbal stems of the derived forms which induce prothesis all have /a/ as the stem vowel the output in these cases is invariable.

It must be mentioned that a rule of prothesis incorporating vowel harmony is not attested for other records of Rwala. Prochazka's account has a fixed prothetic vowel for Rwala, which is *i*; his southern dialect group exhibits vowel harmony, however, e.g. *ʔuktub*, *ʔiksir*, *ʔašrab* 'write! break! drink!'. Interestingly, Musil's system exactly matches that described by Palva for the Bani Šaḥar (Palva 1980: 128), who gives *ʔidrib*, *ʔuktub*, *ʔišrab* 'hit! write! drink!' in the imperative paradigm. While the dialect of the Bani Šaḥar is of the northern-Najdi (Šammari) type⁷², and its speakers are therefore not as close to the Rwala linguistically as they are geographically, it is nonetheless evidence that a system such as that presented by Musil's data is feasible.

2.2.3 *Syncope*

Broselow (2018: 43) defines syncope as a metrically conditioned process not motivated by restrictions on possible syllables but constrained by them, in that the output must be syllabifiable. Syncope is therefore not a syllable repair process in the way that

⁷² Cantineau placed it in Group Bc, which he used to describe dialects which are of the Šammari type (B) but which are heavily influenced by the Syro-mesopotamian sheep-rearing semi-nomads (C).

epenthesis and prothesis are, but rather an independently motivated process: for example, Watson (2002: 70) shows that one motivation for syncope is to reduce the number of monomoraic syllables in order to maximise the number of optimal bimoraic syllables. Most of the north-Arabian bedouin dialects, including the Rwala dialect, exhibit at least two separately motivated syncope processes, which apply separately to the high vowels /i, u/ and to the low vowel /a/. This distinction was first described by Cantineau for the dialects in question, who elaborated on it at length (Cantineau 1936-37: 49-65). A primary source of evidence for this division of syncope is found in the following S-STEM forms (Rwala forms from Prochazka 1988: 29, 31):

<i>ga'ad</i>	'he sat'	<i>simi'</i>	'he heard'
<i>g'adat</i>	'she sat'	<i>sam'at</i>	'she heard'
<i>ga'adt</i>	'I sat'	<i>simi't</i>	'I heard'

In the 3_{FS} forms two different syncope rules can be observed. In *sam'at* (< /sami'at/) a high vowel has been targeted and a sequence of two monomoraic syllables has been reduced to one bimoraic syllable (CV.CV. > CVC.). In the example *g'adat* (< /ga'adat/) a low vowel in the same position has not been targeted, and the reduction of monomoraic syllables has not produced a bimoraic syllable (CV.CV. > CCV.). Cantineau's observation of the requirement for separate syncope rules to explain this (and other) phenomena has been confirmed in more recent research (cf. Al-Mozainy 1981: 45-51, 57-64, Al-Mozainy *et al.* 1985, Kenstowicz 1983 [cited in Al-Mozainy *et al.* 1985], de Jong 2011b).

2.2.3.1 Syncope of /i/ and /u/

The syncope of high vowels in an unstressed open syllable is well attested in dialects across the Arabic-speaking world. Cantineau described the Rwala dialect as being ‘différentiel’ (Cantineau 1936-37: 154), indicating the group of dialects which delete /i/ and /u/ in certain contexts but retain /a/ in the same position (i.e. they differentiate between the high and low vowels), as opposed to dialects in which this syncope rule applies regardless of the quality of the vowel in question (*ibid.*: 49).

The following are some examples of the rule in operation from Musil’s corpus. Forms which show the deleted vowel (which are either unsyncoated variants or references to the paradigm) are given in brackets:

(1)	<i>blād</i>	‘country’	(~ <i>bilād</i>)
	<i>wšu</i>	‘what?’	(~ <i>wušu</i>)
	<i>lhūd</i>	‘blisters’	(~ <i>luhūd</i>)
	<i>mnaššer</i>	‘wearing’	(~ <i>munaššer</i>)
	<i>rweyli</i>	‘of the Rwala’	(~ <i>ruweyli</i>)
	<i>rgubt-ak</i> ⁷³	‘your neck’	(~ <i>rugubt-eh</i> ‘his neck’)
	<i>gḍa</i>	‘judges’	(contrasts with <i>gaḍa</i> ‘he settled’)
(2)	<i>fāhmīn</i>	‘knowing (pl.)’	(~ <i>fāhemīn</i>)
	<i>ṭālb-eh</i>	‘asking for it’	(cf. unsuffixed <i>ṭāleb</i>)
	<i>labsaw</i>	‘they wore’	(cf. 3MS <i>labes</i>)
	<i>menāsf^{en}</i>	‘hutches’	(cf. usual pl. * <i>menāsef</i>)
	<i>nesm-eh</i> ⁷⁴	‘his soul’	(~ <i>nesem-eh</i>)
(3a)	<i>āḥḍīn-kom</i>	‘capturing (pl.) you’	(cf. unsuffixed <i>āḥeḍīn</i>)
(3b)	<i>muṭṭālbīn</i>	‘quarrelling (pl.)’	(cf. usual form * <i>mutaṭālebīn</i>)

⁷³ The elision of the /a/ of the feminine ending is dealt with separately (§2.2.3.3).

⁷⁴ This example is more tenuous since it is unclear what the underlying vowel is. The unsuffixed form is always disyllabic (*nesem, niseṣ*), although this could also show epenthesis from an underlying CVCC form.

The examples in (1) demonstrate where syncope has occurred in the word-initial syllable, and the examples in (2) show medial syncope.. It is apparent that the examples in (2) have the motivation of maximising bimoraicity: in *labsaw* and *nesm-eh* two monomoraic syllables (CV.CV.) have become one bimoraic one (CVC.) and in the other examples the deletion of the vowel has produced a medial superheavy syllable (CVVC.) which, assuming Adjunction-to-Mora (cf. §2.2.2.1), may also be considered bimoraic. The items in (1), however, have not increased the number of bimoraic syllables at all, although the number of monomoraic syllables has been reduced; here we see a targeting of monomoraic syllables at the beginning of the phonological word (cf. Watson 2002: 72-73) which is, in fact, the most frequent site for syncope attested in Musil. The examples in (3) have been added because they show multiple processes which have targeted monomoraic syllables: in (3a) the syncope of /i/ is accompanied by the lack of epenthesis before the suffix (compare the hypothetical **ā.ḥe.ḍī.na.kom*) and in (3b) syncope of two vowels has truly maximised the bimoraicity of syllables (compare **mu.ta.ṭā.le.bīn*). The apparent targeting of /a/ by this high-vowel syncope rule is discussed below.

The rule that accounts for this deletion may be expressed as:

$$V \left[\begin{array}{l} + \text{HIGH} \\ - \text{LONG} \end{array} \right] \rightarrow \emptyset / _]_{\sigma}$$

This rule states that high, short vowels before a syllable boundary (i.e. *i* or *u* in an open syllable) will be deleted. It should first be noted that there are some apparent exceptions to this rule. The first case is that of word-final vowels, although this is explained

by the analysis that final vowels are underlyingly long; as I have demonstrated elsewhere, this can be seen after suffixation. Accordingly, *taʿti* ‘you (f.) will give’ (suffixal form *taʿti-ni* ‘you (f.) will give me’) does not syncopate to **taʿt* despite seeming to fulfil the requirements for syncope in its surface form. A second category is that of words in which the syllable is only open on account of epenthesis (see §3.1.2.1), where, e.g. *tiben* ‘straw’ (underlyingly /tɪbn/) does not syncopate to **tben*. It also regularly fails to apply to nouns having the form CuCuC, e.g. *hoʃon* ‘stallions’ (not **hʃon*). This exception was also noted by Cantineau (1936-37: 50) who considered it a confusion between nouns of the templates CuCuC and CuCC; this analysis seems to be supported by Musil’s data since *hoʃn* occurs in free variation with *hoʃon*, and since the CVCC forms (such as *tiben*) do not trigger syncope even when epenthesis has rendered the first syllable open.

Discounting these regular exceptions, however, the rule is still only partially applied in Musil, and forms such as those listed above occur alongside forms where syncope would be expected to occur such as *gʻurūb* ‘basin’ (not **gʻrūb*), *lisān* ‘tongue’ (not **lsān*), *šeʻibān* ‘valleys’ (~ **šeʻbān*) and *fāhemīn* ‘understanding [active participle]’ (~ *fāhmīn*). Furthermore, word-medial deletion is not attested nearly as frequently as deletion in word-initial syllables. Since syncope rules are shown to be exceptionless in other descriptions of the dialect, and there are mitigating factors for why Musil did not transcribe word-medial syncope (see below, §2.2.3.4), it seems possible that the syncope in word-initial syllables is

better attested in his transcriptions simply because it was more conspicuous than the word-medial syncope where he could have been convinced that he heard a vowel.

It is necessary to consider the source type in assessing this observed failure to syncope, since the metrical requirements of poetry have been identified as a possible factor in phonological variation in this corpus (cf. §1.4.1.3). Taking an example from the previous paragraph, for instance, one finds that *fāhemīn* is attested in poetry while the syncopated form *fāhmīn* is attested in conversation. This appears to support the hypothesis that the structure of poetry causes surface forms to differ from the phonological norms of the spoken language in some cases. Such an assumption is too sweeping, however. *fāhemīn* scans, in accordance with the metre of the poem in which it occurs, as – ʊ – (fā.he.mī.[n-]), but a similar example of a syncopated form from the list above, *menās^{en}*, scans (again, according to the metre of its poem) as x – ʊ – (me.nā.s.fen), showing that in both forms the short/syncopated vowel scans as a light syllable, regardless of whether or not Musil recorded the syncope in his transcription. This situation is restricted to CVVC syllables since, in spite of the argument in §2.2.2.1 that Adjunction-to-Mora allows bimoraic CVVC syllables to occur, I have not found an example where a CVVC syllable is accepted as a single heavy syllable in poetry for the purposes of scansion. Similarly, the syncopated form *ṣḥūn* ‘pans’ scans as ʊ – (ṣ.ḥū.[n-]) in exactly the same way as the unsyncopated form *ṣuyūḥ* ‘chiefs’ (ṣu.yū.[ḥ-]), meaning that in many cases /i, u/-syncope makes little difference to the scansion of poetry, whether transcribed or not. Where syncope creates a CVC syllable,

however, this is frequently taken advantage of in poetry, as exemplified by *labsaw* from the list above, which scans as – – (lab.saw), where the unsyncopated form **labesaw* would not fit the metre. It seems possible that the performance of poetry could have been one factor influencing the failure to syncopate, but this is necessarily speculative, as these examples show. Apart from *fāhemīn*, the other unsyncopated examples in the previous paragraph (*ġurūb*, *lisān* and *še^hibān*) all occur in conversation, which suggests that the metrical structure of poetry should not be considered a primary *explicans* of these phonologically surprising forms in Musil.

It is interesting to note that, in a minority of instances, this syncope rule appears to target the low vowel /a/; this occurs when the regular conditions for /a/ syncope (as described in the next section) are not met. Forms include *ṭwīl* ‘long’ (~ *ṭawīl*) and *ṣwīb-ak* ‘your wounded person’ (unsuffixed form *ṣawīb*); the form *muṭṭālbīn* from the above list may also come into this category, although it is unclear to what extent this particular paradigm has been lexicalised (data for the FORM-V-VI participles is scant, and the underlying form of the prefix could be considered either *muta-/mita-* or just *mut-/mit*). Although some Arabic dialects do target all vowels for syncope in this context (Teepie 2011, and Cantineau 1936-37: 49 where it is claimed that the Maghrebi dialects and the Lebanese dialect of Kfar ‘Abīda “ne sont pas *différenciels*”), these cases nonetheless stand out as odd in a dialect that has been recorded as applying this syncope rule regularly to high vowels only. The most likely explanation seems to be that these /a/ vowels which are deleted in open syllables

have been the target of a vowel raising process (cf. §2.3.6.1) and that the process by which syncope occurs is /ṭawīl/ > ṭiwīl > ṭwīl, where syncope does in fact target a high vowel. This could also explain why /i, u/-syncope targeting /a/ is not widely attested, since low vowel raising is contextually conditioned (it is blocked in the contiguity of guttural consonants or when followed by certain sonorants), as is discussed in §2.3.6.1. It will be seen, however, that low vowel raising is not recorded systematically by Musil, meaning that an argument such as this one has the potential to become rather circular.

There is a very small number of cases where syncope has occurred in a closed syllable, e.g.:

/be + miṣlāb/ >	<i>be-mṣlāb</i>	‘with a fire poker’
/bālci + tiswi/ >	<i>bālci tswi</i>	‘perhaps you will prepare’
/waḡn + āt/ >	<i>wḡnāt</i>	‘cheeks’

Cases like these are unpredictable and are not well attested, and it is perhaps best to view them as exceptions to the rule that only vowels in open syllables may be deleted. The first two examples also occur in poetry, but these surprising forms do not conform to the metrical requirements in the poems where they occur, and the final example is not poetic. It is also uncertain how these forms are syllabified; specifically, whether the coda of the targeted syllable syllabifies with the preceding consonant (bems.lāb, cf. Figure 2.15) or forms a degenerate syllable which may cluster with the following consonant (bem.s.lāb/bem.slāb).

2.2.3.2 Syncope of /a/

There is a second wide-reaching syncope rule well attested for this dialect group, namely that the sequence CaCaCV is frequently realised as CCaCV. Many dialects treat the sequence CaCaCV differently: Prochazka (1988: 22) observes that it is variously realised as CaCaCV, CaCCV or CCVCV in the dialects of Saudi Arabia, of which Rwala falls into the last category, and Cantineau (1936-37: 164) observed that all the bedouin dialects of the Syrian desert realise the sequence as CCVCV. The general occurrence of this process in (predominantly eastern) bedouin dialects has been referred to as the *bšala/zlima*-pattern (Rosenhouse 1982: 19, 1984: 14)⁷⁵.

The application of this rule as it applies to sequences of /a/ in open syllables can be seen in the following examples. As opposed to the examples in the previous section, this syncope rule is most frequently triggered by some form of suffixation and, as such, parallels with the unsyncopated form are found in the corpus more frequently:

<i>barak</i>			‘kneel’ [verbal stem]
<i>barak</i> + <i>at</i>	>	<i>brakat</i>	‘she knelt’
<i>kasab</i>			‘capture’ [verbal stem]
<i>kasab</i> + <i>ow</i>	>	<i>ksabow</i>	‘they captured’
<i>ṭalab</i>			‘he asked for’
<i>ṭalab</i> + <i>eh</i>	>	<i>ṭlab-eh</i>	‘he asked for it’
<i>walad</i>			‘boy’
<i>walad</i> + <i>i</i>	>	<i>wled-i</i>	‘my boy’

⁷⁵ In contrast to the corresponding term “*ghawa*-syndrome” coined by Blanc, however, this term coined by Rosenhouse does not seem to have caught on.

šāğar		‘trees’
šāğar + a	>	šğara
		‘tree’

The distinguishing features of this syncope rule are that, firstly, only /a/ is targeted in this sequence: it occurs in forms such as *bgara* (< **bagara*) ‘female antelope’, but not *‘elem-eh* (which does not become **‘lem-eh*) ‘his reputation’. Secondly, the syncope is dependent on a following short /a/ in an open syllable and is blocked if the following syllable is closed. It can be seen that this rule eschews many of the standard assumptions about syncope: it is very unlikely to maximise bimoraic syllables, it certainly does not target syllables in which the onset could join the coda of a preceding syllable, and it is not restricted to producing syllabifiable outputs. It has been suggested that the motivation for the process is “a fairly abstract kind of dissimilation, eliminating an underlying configuration of two successive identical nuclei by deleting the first of them” (Al-Mozainy *et al.* 1985: 142). The rule may be formulated as follows:

$$V \begin{bmatrix} +\text{LOW} \\ -\text{LONG} \end{bmatrix} \rightarrow \emptyset / _CV \begin{bmatrix} +\text{LOW} \\ -\text{LONG} \end{bmatrix}]_{\sigma}$$

Alternatively, since /a/ is the only low short vowel, we may formulate it more simply as:

$$a \rightarrow \emptyset / _Ca]_{\sigma}$$

This rule expresses that a short /a/ vowel is deleted when it occurs in an open syllable followed by another open syllable also with a short /a/ vowel. The rule cannot apply when the following syllable is word-final, again because these vowels are underlyingly

long, e.g. *bana* ‘he built’ and *banā-ha* ‘he built it’; this contrasts with forms like *gḏa* ‘judges’ which are syncopated under the high vowel syncope rule, which is blind to following syllables. The rules similarly contrast when the following vowel is long, with syncope attested for the high vowels (*kbār* ‘big [pl.]’) but not for the low vowel (*saḥāb* ‘clouds’, *saḥāba* ‘a cloud’).

As with the high vowel syncope rule, the application of this rule is very frequent but not totally consistent. One explanation for the inconsistency, discussed in more detail in §2.2.3.4-5, is that there are very few examples of its application producing a sequence that would violate the Sonority Sequencing Principle (cf. §2.2). The SSP tells us that the consonants in an onset cluster must have increasing sonority values; the above example *wled-i* is one of the few exceptions. The SSP is not the determining factor, however, since there are examples from the verbal system of forms in which syncope has not occurred, but for which the resulting cluster would not violate SSP, such as *daraḡow* ‘they rolled out [the tent]’ (not **draḡow*). In general terms Musil’s record of this syncope process may be described as sporadic; e.g. *wled-i* varies with the unsyncopated form *walad-i*, and the rule is applied to *ḏbaḥat* ‘she killed’ (< *ḏabaḥat*) but not to *ḏabaḥow* ‘they killed’.

2.2.3.3 Syncope of the feminine marker /at/

A third type of syncope attested frequently in the dialect is that of the FS marker /a/, which has the allomorph /at/ in construct and upon suffixation. This is attested in the

following possible syntactic constructions (the forms in brackets are the form of the word when unsuffixed or not in a construct phrase):

pronominal suffix	<i>šoft-eh</i>	‘his view’	(<i>šowfa</i>)
dual suffix	<i>beyramtēn</i>	‘two pieces of fabric’	(<i>beyrama</i>)
construct phrase	‘ <i>aširt ar-rwala</i>	‘the Rwala tribe’	(‘ <i>ašire</i>)

This phenomenon is also recorded by Cantineau (1936-37: 54) and Prochazka (1988: 196); the latter also records it for the FS S-STEM verbal suffix /at/ (*ħzimt-ih* ~ *ħzimit-ih* ‘she tied him’, *ibid.*: 135) although Cantineau (1936-37: 158) claims that there is no syncope of this verbal suffix in the Rwala dialect.

On account of its targeting of a low vowel not followed by another low vowel in an open syllable, Cantineau considered this to require a separate syncope rule. The data from Musil do not explicitly require a separate rule, since this process shares the motivation of the high vowel syncope rule of eliminating monomoraic syllables; while it is unusual for low vowels to be targeted in this position, the phenomenon of /a/ being subject to syncope under this rule has already been noted (e.g. *ṭwīl* ~ *ṭawīl*). One reason for positing a separate rule is that it seems to be restricted in occurrence to this one morpheme. A second reason, although impossible to corroborate from Musil, is the evidence from Al-Mozainy (1981: 192-3) who demonstrates that an identical case of syncope in the Ḥarb dialect appears to be differently ordered *vis-à-vis* the high vowel syncope rule in terms of stress assignment: while he concludes that the peculiarities he cites may be “mere exceptions” (*ibid.*), his data appear to show that the cyclic assignment of stress requires a different rule for the syncope of the

/at/ suffix. Since we have no data on stress placement from Musil, it is impossible to conduct further investigation into this phenomenon, but nonetheless the possibility of the requirement for a separate rule should not be disregarded.

2.2.3.4 Unsyncopated forms

One of the more striking aspects of Musil's Rwala transcriptions is the number of cases in which syncope is not attested in contexts in which the previously described syncope rules would be expected to operate. Precise figures on the frequency of syncopated *versus* unsyncopated forms are impossible to compute from such a large data set: for example, *šarbet* 'she drank' (syncopated, < **šaribet*) has the same consonantal structure and homomorphic suffix as *wağnet* 'cheek [construct form]' (unsyncopated), so figures for these syncopated forms cannot be extracted easily from the entire corpus. Some general observations can be made, however. 12% of all unique forms in the corpus are cluster initial (1,647 tokens), all of which must attest to one of the syncope rules: arguably this shows that the operation of syncope is fairly frequent, especially considering that the number of forms exhibiting medial syncope will increase this statistic. The absence of syncope can be assessed in individual paradigms, however; for example, there are 92 tokens of word-initial Ca.Ca. which is a notably high number of cases for syncope to have not occurred (with no clear phonological motivation for syncope to be blocked).

An important example is found in the P-STEM of the FORM-I strong verb (see discussion in §4.3.3), where, according to the evidence of Cantineau and Prochazka, the syncope of /i, u/ but not of /a/ produces the differently syllabified plural forms *yīšrabūn* ‘they drink’ but *yiktbūn* ‘they write’. In Musil, the vowel *e* has been levelled across the paradigm, as in the following examples which all have different vowels in the verbal stem:

<i>yarćab</i>	‘he rides’	<i>yarćebūn</i>	‘they ride’
<i>yanšed</i>	‘he asks’	<i>yenšedūn</i>	‘they ask’
<i>yagtol</i>	‘he kills’	<i>yeǵtelūn</i>	‘they kill’

Following Cantineau and Prochazka, the expected 3MP forms here would be **yarćabūn* but **yanšdūn* and **yeǵtlūn*. I suggest that this is a point where the phonological (and phonetic) facts of the dialect have been obscured by Musil’s transcription style, and propose several interpretations of this data. The first is that Musil’s <e> might not always represent a full vowel, as it seems highly likely that in a syncopated form such as **yeǵtlūn* epenthesis of a super-short vowel will be induced. Cantineau also implies that this is the case since he describes the *yaktbūn* form, which he represents as <yakt•būn>, as having “un simple point vocalique” (Cantineau 1936-37: 186): this terminology may describe something as small as the audible obstruent release. Therefore Musil’s transcription of *yeǵtelūn* could in fact be better rendered as *yeǵte^llūn*, which could be supported by the limited evidence from Prochazka of the Rwala dialect having a short, schwa-like epenthetic vowel, e.g. *bāytakum* ‘your tent’ (Prochazka 1988: 199).

A second interpretation is that *e* does represent a full vowel, and that the forms demonstrate the interaction of syncope and epenthesis. Broselow (1992a: 36) considers an example from a rhyme dialect (Iraqi):

UR	/yilbas + ūn/	
syncope	yilbsūn	
epenthesis	yilibsūn	
output	<i>yilibsūn</i>	‘they wear’

Here a syncope rule has deleted the vowel *a* and syllabification has triggered epenthesis of *i* in a different position, but Broselow notes that if syncope created an unsyllabifiable sequence in onset dialects the epenthetic vowel would be inserted in the same position as the deleted vowel⁷⁶. She predicts that onset dialects may exhibit changes of vowel quality in open syllables if the quality of epenthetic vowels is not identical to that of the vowels targeted by syncope, and this prediction is arguably supported by Rwala examples such as *yagtol*, *yeǵtelūn*, where the only evidence of syncope or epenthesis is the change in vowel quality:

UR	/yagtol + ūn/	
syncope	yagtlūn	
epenthesis	yagtelūn	
fronting	yeǵtelūn	
output	<i>yeǵtelūn</i>	‘they kill’

It might be argued that in the A-CLASS, since forms such as **yarćabūn* are not attested, the variation in vowel quality between *yarćab* and *yarćebūn* is also evidence that syncope

⁷⁶ “On this account, then, syncope need not look ahead to see if its output will be syllabifiable; but if it does apply in an environment not allowing full syllabification, its effects will be undone by epenthesis.” (Broselow 1992a: 36)

and epenthesis have occurred, but given that the relationship between *a* and *e* is so uncertain this can be no more than speculation. While the descriptions in Cantineau and Prochazka do not attest syncope of /a/ in this paradigm, we have already seen that in Musil /a/ in an open syllable is also likely to be subject to the general (i.e. /i, u/) syncope rule, possibly on account of low vowel raising.

A third interpretation is simply that this paradigm has been systematised by Musil in a way that completely obscures the nature of what he actually heard. Such a situation is paralleled by that of the active participle, where we consistently find forms of the type *šāreb*, *šārebīn* ‘drinking’ in Musil, as opposed to the expected *šāreb*, **šārbīn* with syncope of the *e* (see discussion in §4.4.1). Once again, we could consider the *e* in the plural form to be an undefined super-short sound (*šār^ebīn*), which is a solution that Cantineau considers in order to explain the same phenomenon in the texts of Montagne (Cantineau 1936-37: 155-156). There is no evidence for the interaction of syncope and epenthesis in this example since the targeted vowel is always *e* which is the same as the epenthetic; I would suggest however that epenthesis is less likely to have occurred here since, as was noted in §2.2.2.1, epenthesis is much more likely to occur in cases of medial CVCC (*zaml-ana*, compare *yeǧtelūn*) than in cases of medial CVVC (*eyāl-ana* ~ *eyāl-na*, compare *šārebīn*). The case of the participle forms, however, is just as likely to have been misrecorded by Musil in general, since he could have assumed uniformity in the paradigm *šāreb*, *šārebīn* as being akin to the CLA forms *qātil^{un}*, *qātilūn^a*.

2.2.3.5 Initial consonant clusters

Where syncope produces initial consonant clusters in Arabic dialects which tolerate such syllables and do not repair them, it is not uncommon for the resulting clusters to violate the Sonority Sequencing Principle. As expressed in §2.2, the SSP asserts that consonants in the syllable onset should only increase in sonority. The sonority hierarchy is repeated here:

SONORITY HIERARCHY:

oral stops < fricatives < nasals < liquids < glides < vowels (Foley 1972, cited in Clements 1990: 286)

In a list of unique forms, the following clusters which either obey or violate the sonority hierarchy are attested in Musil's *Rwala* corpus:

obey	722
violate	589
(of which reversal)	411
(of which plateau)	178

Every viable combination of stop (S), fricative (F), nasal (N), liquid (L) and glide (G) is attested, and may be divided into those which obey the sonority hierarchy (O), those which reverse it (R) and those which form sonority plateaus (P)⁷⁷. The most and least frequently attested are:

⁷⁷ In this analysis I follow Watson's treatment of initial clusters in San'ani Arabic (Watson 2002: 74-76).

1.	N-F	148	(R)
2.	S-F	138	(O)
3.	F-G	114	(O)
4.	F-F	112	(P)
5.	F-L	111	(O)
6.	S-L	98	(O)
7.	F-S	94	(R)
...			
20.	G-S	6	(R)
=	L-N	6	(R)
=	N-N	6	(P)
23.	G-G	3	(P)
24.	G-N	2	(R)
25.	L-L	0	(P)

The absence of L-L in the corpus underscores an important feature of sonority plateaus, which is that they are only permissible when there is distance between the place of articulation of both consonants (Watson 2002: 76 note 18). Since the two liquids /l/ and /r/ share a place of articulation (dental-alveolar), initial clusters of /lr/ or /rl/ are not predicted to occur, which the data confirm to be the case⁷⁸. It may also be predicted, then, that when clusters are formed from two consonants of equal sonority (i.e. the same manner of articulation) and the same place of articulation, the first consonant would be syllabified with what precedes, which is true in every case, e.g.:

<i>elya tdiǧǧ</i>	‘when you pound [coffee beans]’	(el.yat.diǧǧ)
<i>mā tṭawwel</i>	‘she does not lengthen [sth.]’	(māt.ṭaw.wel)
<i>alli tdir</i>	‘[a bird] which circles’	(al.lit.dir)

⁷⁸ It is also the case that Arabic roots do not permit consonants which share a place of articulation (cf. Zemánek 2011b), so these clusters do not occur also for lexical reasons.

The SSP may account for some of the idiosyncrasies in Musil’s transcriptions. Clusters which obey the sonority hierarchy are more common overall than those which don’t, and although this is not surprising from a cross-linguistic perspective, it differs with the data from other Arabic dialects which permit initial CC, such as San’ani in which sonority-reversing clusters are more frequent (Watson 2002: 76). Although this is only indirect evidence, it may suggest that the apparent irregularity of the syncope processes according to Musil’s transcriptions could be partially explained by his being less likely to perceive syncope when the resulting cluster reversed the sonority hierarchy.

In Musil’s data, glide-initial clusters (all of which are reversals or plateaus by default) make up only 3% of the total number of clusters: this is surprising since /y/ is a frequently attested P-STEM prefix, which is recorded by Prochazka (1988: 36, 46-48) as forming clusters in FORM-I before guttural stems and in forms II, III, V and VI. The forms in Musil generally do not have clusters in the same paradigms:

	Prochazka	Musil
FORM-I	<i>yḥamil</i> ‘he carries’	<i>yehadar</i> ‘he goes down’
FORM-II	<i>ykassir</i> ‘he smashes’	<i>yebaddel</i> ‘he swaps sth.’
FORM-III	<i>ysāfir</i> ‘he travels’	<i>yirādef</i> ‘he rides behind’
FORM-V	<i>ytiċallam</i> ‘he speaks’	<i>yetaġassel</i> ‘he scrubs himself’
FORM-VI	<i>ytaḥārab</i> ‘he fights’ ⁷⁹	<i>yitaḥāšer</i> ‘he shares out’

Here, the FORM-I example represents the low vowel syncope rule (which is only applied in the case of C1 = guttural verbs on account of the *ghawa*-syndrome, see §2.3.1) and

⁷⁹ Prochazka does not provide Rwala examples for FORM-VI, but this form is given for the dialects which pattern like Rwala in all the other examples given.

the others represent the high vowel syncope rule. Syncope does not seem to have applied in Musil, but it turns out that this only applies to the forms with the /y/ prefix; compare:

3MS	<i>yəḥadar</i>	1CP	<i>nḥadar</i>
3MS	<i>yəbaddel</i>	2MS	<i>təbaddel</i>
3MS	<i>yirādef</i>	2MS	<i>trādef</i>

The discrepancy between Musil's recording of syncope after a nasal /n/ or a stop /t/ but not after a glide /y/ may be explained by the relative sonority of these sounds. The clusters with /y/ would all be reversals of the sonority hierarchy, while the other forms are more varied:

3MS	<i>*yḥadar</i>	(R)	1CP	<i>nḥadar</i>	(R)
3MS	<i>*yabaddel</i>	(R)	2MS	<i>təbaddel</i>	(P)
3MS	<i>*yirādef</i>	(R)	2MS	<i>trādef</i>	(O)

There are three possible interpretations of these data. The first is that Musil's transcriptions reliably recorded that syncope is less likely to occur when the sonority reversal in the resulting cluster is greater (or, alternatively, that the sonority gap between the consonant and the target vowel is smaller). This would not fully resolve the question, however, since his data show only a slight preference for clusters which obey the sonority hierarchy and this interpretation would at best articulate a general trend. In accordance with this interpretation it would be possible to posit that syncope interacts with epenthesis which effectively cancel each other out in sonority-reversing clusters (*yibaddel* > *yabaddel* > *yebaddel*) but again the data does not support this, as the relatively high frequency of sonority reversing clusters renders the motivation for epenthesis too undefined.

The second interpretation is that Prochazka correctly discerned that syncope had taken place and Musil was less likely to recognise this when the resulting clusters had greater sonority reversals. There would be very good reasons for this, because in consonant clusters which reverse the SSP the initial consonant forms a sonorous peak which is essentially syllabic. In Watson's analysis the first consonant of such a cluster is both moraic and attached to an independent syllable node, e.g. *ša.fuṭ* > *š.fuṭ* 'šfuṭ [a Yemeni dish]' and *ḥa.tab* > *ḥ.tab* 'wood' (Watson 2002: 76). The sonorous peak of the first consonant would therefore explain why Musil did not discern the syncope as readily in these cases.

The third interpretation is that syncope is not a categorical (phonologised) process but a gradient (phonetic) one, and that Prochazka was too categorical in his interpretation of the data. The question of how far the syncope process was phonologised is a very relevant one: syncope is presented as an exceptionless process in the relevant literature, but an acoustic study of a north-Arabian dialect which could shed light on how much this is a gradient process is lacking. It is evident, furthermore, that Prochazka's interpretation of his data was categorical, as is demonstrated more generally by his representation of the dialects he recorded with only three short vowels, /a, i, u/, all of which are contrastive, compared to Musil's gradient system which blurs phonemic contrasts by associating vowels with less abstract representation (*a, e, i, o, u*).

These possible interpretations overlap with those from the previous section. If syncope is not phonological but rather a gradient process of vowel reduction, then *yebaddel*

could represent a reduced sound (*yebaddel*) as in the case of *šārebīn* above. In this case, there is no real discrepancy between Musil's data and those of Prochazka and others, but it is merely a situation in which Musil's system of transcription is overly broad while that of the others is overly phonemic. If, however, Musil's records are to be taken at face value, then it is likely that an interaction of syncope and epenthesis must be posited to account for forms such as *yeǧtelūn* and *yebaddel*. This process has been attested in other dialects, but seems unlikely to be the case here because nobody else has recorded such a situation. Finally, if Musil had systematised the grammar in his head to the point that he was not discerning syncope when it occurred, then the natural conclusion is that his forms *yeǧtelūn* and *yebaddel* really represent **yeǧtlūn* and **ybaddel*, which match the transcriptions of the dialect from Prochazka, Ingham and Cantineau.

2.3 Other phonological processes

This section deals with other phonological process attested for the dialect in Musil's records. The first, the *ghawa*-syndrome, is a well-documented process which is common to all the bedouin dialects of northern and central Arabia and beyond, and the last section, qualitative vowel processes, similarly describes processes common to many north-Arabian dialects; the others are more general processes, some of which are encountered in some form or other in possibly every dialect of Arabic.

2.3.1 *The ghawa-syndrome*

The term “*ghawa-syndrome*”⁸⁰ was coined by Blanc (1970: 125-127)⁸¹ to describe a phonological process which has been shown to be prolific in bedouin-type dialects, particularly in the eastern dialects and becoming “less and less frequent among the Bedouin dialects the further west we turn” (Rosenhouse 1984: 14). It is essentially a process of vowel insertion triggered by the sequence of /a/ and a guttural consonant (i.e. a uvular, pharyngeal or glottal fricative) /ğ, ħ, ḥ, ʕ, h/⁸². The rule may be formulated as:

$$\emptyset \rightarrow a / aX_]_o C \quad (X = \dot{g}, \dot{h}, \dot{h}, \dot{c}, h)$$

The output of this rule creates the sequence (C)a.Ca. which, in dialects which have the low vowel syncope rule described in §2.2.3.2, is targeted for syncope, ultimately producing the output CCa. Examples for each consonant from Musil include:

⁸⁰ A discrepancy between the spellings “*gahawa-syndrome*” and “*ghawa-syndrome*” exists, which must have arisen on account of the low vowel syncope rule: since it is named after the word for coffee (which exhibits the process), a dialect without a low vowel syncope rule similar to the one elaborated for the Rwala dialect would have the form *gahawa* (such as the dialect of the Negev bedouins recorded by Blanc) while in a dialect which does delete /a/ in the sequence Ca.Ca. the same output of the *ghawa-syndrome* (i.e. *gahawa*) would become the target of syncope (> *ghawa*). The *ghawa-syndrome* can therefore be seen to be in a ‘feeding’ relationship with /a/-syncope.

⁸¹ It should be noted that the process had already been described by other scholars before Blanc gave it this name, e.g. by Cantineau (1936-37: 65-66, 167-170), Mitchell (1960: 388) and Johnstone (1964: 80, 1967a: 6-7).

⁸² Note that this list is usually limited to fricatives, and the only non-fricative guttural consonant /ʔ/ is not usually included; Johnstone (1967b: 6, 13[i]) does tentatively include it with the reservation that it triggers the *ghawa-syndrome* “with less regularity” (p.6). As we have seen in §2.1.1.1, however, the likelihood of /ʔ/ being deleted in the required context for the *ghawa-syndrome* (CaX.) renders this unlikely to occur. de Jong’s (2011b) entry in the *EALL* includes /r/ on the list but, since /ʕ/ is not included and /r/ does not occur in any of his examples, I can only conclude that this is an error and that he meant to include /ʕ/. There is one case in Musil where the process appears to be triggered by /r/ (**marwa* > *mrawa* ‘watering place’ but this appears to be anomalous.

<i>hal</i> ~ <i>ahal</i>	(< * <i>ahl</i>)	‘kin’
<i>shala</i>	(< * <i>saḥla</i>)	‘tiny date’
<i>ghawa</i> (~ <i>gahwa</i>)	(< <i>gahwa</i>)	‘coffee’
<i>ḥamar</i>	(< * <i>aḥmar</i>)	‘red’
<i>ḥaḍar</i>	(< * <i>aḥḍar</i>)	‘green’
<i>mḡalūt</i>	(< * <i>maḡlūt</i>)	‘mad (dog)’
<i>y‘arfūn</i>	(< * <i>ya‘refūn</i>)	‘they know’

The derivation can be seen as follows:

UR	/saḥla/	
<i>ghawa</i> -syndrome vowel insertion	saḥala	
low vowel syncope	shala	
output	<i>shala</i>	‘tiny date’

In synchronic terms, the rule might also be formulated as one of metathesis (cf. Al-Mozainy 1981: 193):

$$CaX]_oC \rightarrow CXa]_oC$$

It seems, however, that from a diachronic perspective this would be a mistake, since the rule is a feature of a large number of bedouin dialects, some of which do not have the low vowel syncope rule, and as such the metathesis rule does not describe the facts of those dialects. The difference between syncopated and unsyncopated *ghawa*-forms can be seen in a comparison of terms from Musil and Blanc (1970: 126, for the Negev):

Rwala	Negev	
<i>ghawa</i> (~ <i>gahwa</i>)	<i>gahawah</i>	‘coffee’
<i>hal-eh</i>	<i>ahal-ih</i>	‘his kin’
<i>ḥamar</i>	<i>aḥamar</i>	‘red (m.)’

There are certain contexts in which the *ghawa*-syndrome does not occur. According to de Jong (2011b), derived forms of the verb and quadriliteral verbs are usually unaffected,

and in Musil we find FORM-IV *aḥdar* ‘he cast down’ (not *[a]ḥadar) and FORM-X *ista‘ğelat* ‘she induced labour’ (not *ist‘ağelat < *ista‘ağelat). This generalisation may be too broad, however, as the *ghawa*-syndrome occurs in derived forms in some dialects, e.g. *āḥatirm-ah* ‘I respect him’ (< *aḥtarim-ah) in the speech of the Bahraini ‘Arab (Holes 2016: 200). Musil’s corpus attests ‘aṭa (< *a‘aṭa < *a‘ṭa), effectively collapsing the distinction between FORM-I and FORM-IV in this verb, a usage which is widespread throughout Arabia. It does seem, however, that quadrilateral verbs are unaffected, e.g. *zağzağat* ‘she whispered’ (not *zğazağat < *zağazağat). It has been observed that the *ghawa*-syndrome does not occur “in cases where the -aX sequence terminates a base and precedes a pron[ominal] suff[ix]” (Blanc 1970: 127), which explains why in Musil we find /ḍabaḥ + na/ > *ḍabaḥna* ‘we killed’ (not *ḍbaḥana < *ḍabaḥana) and /ma^c + hom/ > *ma^c-hom* ‘with them’ (not *m^ca-hom < *ma^ca-hom). de Jong (2011b) notes that borrowings from CLA may often be unaffected, and in Musil we find that Islamic terms, which may be in this category, are exempt, e.g. *ar-raḥmān* ‘Allāh the Most Gracious’ (not *ar-rḥamān < *ar-raḥamān). The *ghawa*-syndrome may affect names generally (e.g. *bğadād* < *bağdād*), although this is not attested in Musil.

There are, however, contexts where *ghawa*-forms are absent from Musil even though they would be predicted to occur. The following examples show where Musil did and did not record *ghawa*-forms, alongside Rwala examples from Cantineau⁸³ (1936-37: 168-169)

⁸³ These are of the same lemma if possible, and otherwise of the same template.

which demonstrate that Musil recorded *ghawa*-forms in many but not all of the same contexts (X = ġ, ħ, ḥ, ʕ, h):

	Cantineau		Musil	
CaXCa	<i>ṣhala</i>		<i>ṣhala</i>	‘tiny date’
	<i>ghawa</i>		<i>ghawa/gahwa</i>	‘coffee’
aXCaC	<i>ḥamar</i>		<i>ḥamar</i>	‘red’
	<i>ḥaḍar</i>		<i>ḥaḍar</i>	‘green’
maXCaC(a)	n/a		<i>maḍar</i>	‘cheek’
	n/a		<i>maḥnaga</i>	‘coral necklace’
maXCāC(a)	n/a		<i>maḡrāfa</i>	‘trough’
yaXCVC	<i>iḥaḍa</i> ⁸⁴		<i>yaḥḍi</i>	‘he brings’
	<i>iḥalb</i> ⁸⁵	‘he milks’	<i>yiḥarg</i> ⁸⁶	‘he pulls down’
maXCūC(a)	<i>mʕarūf</i>	‘known’	<i>mḡalūt</i>	‘mad (dog)’
	<i>mḥatūṭa</i>	‘placed’	<i>maḥsūba</i> ⁸⁷	‘counted’

It can be seen that, in at least two categories (yaXCVC and maXCūCa), Cantineau recorded *ghawa*-forms where Musil did not. The inclusion of *gahwa* in the above list is anomalous, as forms from the CaCCa template are almost always *ghawa*-forms in Musil. Conversely, the forms *yiḥarg* and *mḡalūt* are also somewhat anomalous: *ghawa*-forms of verbals are attested in Musil’s corpus, as the examples in this table indicate, but with much

⁸⁴ Cantineau gives *iḥaḍi* for Ḥsane (a group very closely related to Rwala) which also has *i* in the final syllable, like Musil’s example.

⁸⁵ This example is not from Rwala, for which Cantineau gives only one example (*iḥaḍa*) but from the closely related Ḥsane dialect.

⁸⁶ The form *yiḥarg* is in itself interesting: the underlying form must be /yaḥrVg/, the /a/ being needed to trigger the *ghawa*-syndrome vowel insertion, so the derivation is *yaḥrVg* > *yaḥarVg* > *yḥarVg* > *yḥarg*. I consider Musil’s prefix *yi-* as representing the articulation of syllabic /y/ which is, essentially, a vowel (as in Cantineau’s transcription of it as *i*), but the syncope of the final vowel is not a regular process in the dialect (although the form is widely attested in others, as shown by Cantineau).

⁸⁷ I have supplied the final /a/ of the feminine suffix, since the attested form is in construct: *maḥsūbt*.

less regularity than nominals. The data show, for example, that the FORM-I P-STEM of verbs with a guttural first radical are attested with the non-*ghawa*-form (e.g. *yaʿrefūn* ‘they know’) five times more frequently than with the *ghawa*-form (e.g. *yʿarfūn* [same meaning]). The attestation of *ghawa*-forms defies easy generalisation; it seems to be the case that Musil recorded them without exception when the initial consonant was a root consonant, whereas in the case of derived forms (including all verbals) he did not record them in the majority of cases. If this observation is causative rather than correlative, it indicates that Musil’s preconceptions about the structure of the language could have led him to ignore the *ghawa*-syndrome with fairly high frequency.

2.3.2 Assimilation

The consonants in Musil’s system of transcription map very closely onto the phoneme inventory of the dialect, meaning that details of phonetic variation are largely obscured. Nonetheless, a number of assimilation patterns are recorded, mostly in cases of total assimilation which are easily rendered in Musil’s system. These are subdivided into assimilation of place, manner, voice and emphasis in the following sections. A final type of assimilation frequently cited for Arabic dialects is the assimilation of the definite article /*(a)l*/ to a following coronal (cf. §5.1.14).

2.3.2.1 Place assimilation

Nasal place assimilation is attested in the corpus, as shown in the following examples of /n/ being transcribed as <m> by Musil:

<i>gumfoḍ</i>	<	* <i>gunfoḍ</i>	‘hedgehog’
<i>ḡamb</i>	<	* <i>ḡamb</i>	‘side’
<i>mumbaṭaḥ</i>	<	* <i>munbaṭaḥ</i>	‘lying on one’s stomach’
<i>al-mal‘ūm mesta‘ezz</i>	<	* <i>al-mal‘ūn mesta‘ezz</i>	‘the accursed one thinks too much of himself’

All of these examples attest to the assimilation of /n/ before the bilabials /b, m/ and the labio-dental/f/. The <m> in Musil’s transcriptions is not phonetically precise, since it is highly likely that /n/ > [m̥] before the labio-dental and /n/ > [m] before the bilabials: i.e. *gumfoḍ* represents *gu[m̥]foḍ* while *ḡamb* represents *ḡa[m]b*. Cross-linguistic evidence tells us that nasal consonants are the most likely to assimilate to the place of articulation of a following consonant (Gordon 2016: 128), and that /n/ as a coronal nasal is more likely to be assimilated given that coronals are typologically more likely to be assimilated than velars or labials (*ibid.*: 131). Given that /n/ is shown to assimilate to labials in this dialect, it is very likely that it also assimilates to consonants with places of articulation other than labial, but that the system of transcription lacked the nuance to record it. With a following velar, for example, one would expect nasal place assimilation to render *min-kom* ‘from you (pl.)’ as *mi[ŋ]-kom*.

2.3.2.2 Manner assimilation

Musil specifically describes one process of manner assimilation, writing:

“the words ‘*azwet^{en} li* [‘person of the same descent as I’], if spoken quickly, sound like ‘*azwet elli*, because the *en* tends to double the following syllable, *li*” (Musil 1928a: 263).

It is unclear what exactly he was suggesting here⁸⁸, but the assimilation of /n/ > /l/ is not unattested in Arabic, even being recorded in the Qur’an (**hud^{an} li-l-muttaqīna* > [*hudallilmuttaqiina*] ‘a guide for the righteous’, Qur’an 2: 2, cf. Zemánek 2011a). Two other cases of manner assimilation of /n/ are attested, both from the particle *min*: from the relative pronoun *min* ‘he who’ we have assimilation before /l/ in the phrase *mel l-eh ḥalīle* ‘whoever has a wife’ and from the preposition *min* ‘from’ we have assimilation before /r/ in the phrase *mir rās* ‘of heads [partitive usage]’. This assimilation, however, is not marked by Musil in two out of the three phrases in which he specifically mentioned it, showing that his preference is for a phonemic rather than phonetic approach to transcription (at least for consonants). It is unknown, therefore, whether forms in similar contexts, such as *tenleḡem* ‘you put on the bridle’, were also pronounced with assimilation (i.e. *te[l]leḡem*) in fast speech. Interestingly, a reverse type of assimilation is also once recorded, where Musil claims that *fāl^{en}* ‘omen’ was pronounced *fān^{en}*.

⁸⁸ For instance, whether by “*en*” he meant to specify the adnominal linker morpheme or simply used this instance as a marker of /n/ in general. Presumably by “double the following syllable” he meant that the following consonant was doubled.

Other cases of manner assimilation are recorded, such as *mettāri* ‘avenger’ < **mettāri*, but these forms are not regularly attested (contrast *tettāba* ‘you make sure of your lineage’⁸⁹).

2.3.2.3 Voice assimilation

The verbal system shows some evidence of anticipatory devoicing before suffixes, as in the following examples:

<i>gasatt</i>	<	<i>gasadt</i>	‘I composed a poem’
<i>aḥatt</i>	<	<i>aḥadt</i>	‘I took’

This occurs in all but one instance (*ḥešedt* ‘you [MS] gathered’), but is restricted to final clusters, as shown by *aḥatt* but *aḥadt-ah* ‘I took (married) her’. This type of assimilation is well known from grammars of Classical and Standard Arabic.

There is some sporadic evidence of voice assimilation, e.g. *azġar* ‘smaller’ < *ašġar*, although this is either not a regular process or not regularly recorded. It is interesting to note that the same root shows assimilation in other Arabian dialects: *zġayyir* ‘small’ is given by Rosenhouse (2011) for bedouin dialects in general, and it is attested in Yemeni dialects, with *z(i)ġir* pl. *zġār* ‘small’ (Piamenta 1990-91: 283) and San’ani *zuġayr*, *zuġayrī*, *zuġayyir* ‘small’ (Watson 2006: 196), which are not examples of bedouin usage.

⁸⁹ Although this has no bearing on Musil’s transcription of /t̤/ without assimilation, the form in itself is difficult to account for: the sense as Musil translated it seems to be possibly related to the root **tbt**, although no /t/ is present.

2.3.2.4 Emphasis assimilation

The subject of emphasis spreading has received much scholarly attention, but since Musil's transcriptions provide almost no evidence about it the process cannot be dealt with here in depth. Emphasis spreading differs from the other types of assimilation in this section in that emphasis, which refers not only to the secondary articulation of pharyngealisation but may encompass other articulatory correlates such as lip rounding (Bellem and Watson 2014: 2 and notes 7 and 8), is a non-primary feature of a broad group of segments which may spread both locally and long-distance (Watson 2002: 268). In general terms, most Arabic dialects exhibit emphasis spreading to some degree, although this effect varies in terms of domain, which may be restricted to the syllable, the whole-word, or beyond, and there may be differences in the segments or morpheme boundaries which block emphasis spread (Zemánek 2011a). Many dialects have a set of secondary emphatics, often the labials /b, m/ and the liquids /l, r/, which also partake in this process; this has been confirmed for the Rwala dialect by Cantineau (1936-37: 129, 135) who also includes /f/ (labio-dental) and /w/ (labial-velar).

The only examples of emphasis spreading which Musil recorded in his transcriptions involve /t/ assimilating the emphasis of an adjacent emphatic coronal:

<i>taḡāmaṭṭ</i>	<	* <i>taḡāmaṭṭ</i>	'I covered myself'
<i>muṭṭālbīn</i>	<	* <i>mutṭālbīn</i>	'quarreling [active participle MP]'
<i>eṣṭebeḥ</i>	<	* <i>eṣṭebeḥ</i>	'he had breakfast'

The last example is from FORM-VIII, which is a frequent site for assimilation (Zemánek 2011a), not only of emphasis but also of voice (cf. CLA *izdaḥama* ‘it was crowded’ < **iztaḥama*); in Musil’s corpus, however, there are no FORM-VIII verbs attested which would demonstrate voice assimilation, so it is impossible to know whether Musil would have recorded it.

2.3.3 *Word-final vowel shortening*

It was mentioned in §2.2.1 that CVV is not permitted in word-final position. The general pattern in the dialect is that final vowels are always short, but that most of these short vowels are long when the word is suffixed, e.g.:

<i>tara</i>	‘behold!’	<i>tarā-ni</i>	‘look at me!’
<i>yeḡi</i>	‘he comes’	<i>yeḡi-ć</i>	‘he comes to you (f.)’

This rule may be extended to include prepositional phrases which may be considered as forming part of the phonological word, e.g.

<i>ḡara</i>	‘it happened’	<i>ḡerā l-i</i>	‘it happened to me’
-------------	---------------	-----------------	---------------------

This situation is found in many Arabic dialects, and therefore the issue of word-final vowel shortening has been much discussed. A distinction may be made between scholars who analyse these dialects as having underlying final CVV syllables alongside a process which shortens them word-finally, and those who analyse them as having underlying final CV syllables and a process which lengthens them before a suffix (McCarthy 2005: 1-2). Essentially both views are adequate from a purely descriptive point of view, although

McCarthy (*ibid.*) argues that the former option is to be preferred in the light of general language typology.

There is some evidence from Musil's Rwala material, albeit not very widely attested throughout the corpus, which also suggests that word-final vowels are long in their underlying representation and subject to a rule which shortens them. This evidence is the existence of word-final vowels which do not lengthen before suffixes. Although rare, these are represented by the bound object pronouns which, in the few cases that more than one suffix is concatenated, do not lengthen upon suffixation:

<i>tefaṭṭen-ni-h</i>	'she reminds me of it'	(not * <i>tefaṭṭen-nī-h</i>)
<i>tanṭey-na-h</i> (for * <i>taṭṭi-na-h</i>)	'you give it to us'	(not * <i>tanṭey-nā-h</i>)

If all final vowels are short in their underlying representation in this dialect then a rule lengthening these vowels before a suffix would be required to produce the form *tarā-ni* from an underlying /*tara + ni*/ (cf. the analysis of Cairene Arabic in Watson 2002: 201-203). This rule predicts, however, that there will be no short vowels before suffixes, which is not supported by these data because this rule would also produce the form **taṭṭi-nā-h* from an underlying /*taṭṭi + na + h*/, where pre-suffix vowel lengthening would apply twice. The variation of *tara* and *tarā-ni*, therefore, must be explained by a rule which shortens word-final vowels only:

$$CVV \rightarrow CV / _ \#$$

To produce the correct output, both long and short final vowels must be permitted in the underlying representation. If it is a requirement that final vowels may only be either

short or long in their underlying representation then the incorrect output is attained, as shown in the following derivations (forms in asterisks are incorrect outputs):

(1) underlyingly short

/tara/	/tara + ni/	UR
_____	_____	word-final vowel shortening
<i>tara</i>	<i>**tara-ni**</i>	output

/ta ^ʕ ti/	/ta ^ʕ ti + na + h/	UR
_____	_____	word-final vowel shortening
<i>ta^ʕti</i>	<i>**ta^ʕti-na-h**</i>	output

(2) underlyingly long

/taraa/	/taraa + nii/	UR
tara	taraani	word-final vowel shortening
<i>tara</i>	<i>tarā-ni</i>	output

/ta ^ʕ tii/	/ta ^ʕ tii + naa + h/	UR
ta ^ʕ ti	_____	word-final vowel shortening
<i>ta^ʕti</i>	<i>**ta^ʕtī-nā-h**</i>	output

(3) no restriction on underlying length

/taraa/	/taraa + ni/	UR
tara	_____	word-final vowel shortening
<i>tara</i>	<i>tarā-ni</i>	output

/ta ^ʕ tii/	/ta ^ʕ tii + na + h/	UR
ta ^ʕ ti	_____	word-final vowel shortening
<i>ta^ʕti</i>	<i>ta^ʕtī-na-h</i>	output

Of these derivations only (3), in which the verbs have long final vowels and the suffixes have short final vowels in their underlying representation, produces the correct output.

The rule of final vowel shortening is almost exceptionless in Musil's record of the Rwala dialect. Word-final CVV only occurs in two contexts: the first is some particles such as the negative particles *lā* and *mā*, the pronoun *mā*, the vocative *yā*, although these may be interpreted by means of a restriction on the application of final-vowel shortening producing an output that is not a minimal word (cf. §2.2.2). The second is in two morphemes, namely the suffixes of the imperative which are regularly transcribed as having the long vowels *-ī* (FS) and *-ū* (CP). It is worth noting that these final long vowels are absent from the imperative forms given by Cantineau (1936-37: 187-188) and those in Ingham's Rwala material (e.g. *warridu* 'water your herds!', Ingham 1995: 125, 132), and that Prochazka (1988: 38-40) does not record final long vowels in the imperative paradigm for any of the Arabian dialects he describes, although the Rwala dialect stands out in his survey as one of only two dialects that have diphthongal suffixes in the imperative paradigm⁹⁰. It is impossible to reconstruct whether there is a connection between Prochazka's atypical Rwala diphthongs and Musil's atypical Rwala final long vowels, but with no other evidence it cannot be disregarded that Musil may have systematised his imperatives to a paradigm that does not conform to the phonology of the dialect. A possible explanation could be that, since the S-STEM plural suffix is usually a diphthong (*-aw/-ow*, cf. §4.2.1) and the P-STEM plural suffix is usually consonant-final (*-ūn*, and the FS suffix may either be short *-i* or long *-īn*, cf. §4.3.2), Musil may have been anticipating a heavier suffix in the imperative than the

⁹⁰ The other is Hofuf, and there is one anomalous instance from Hāyil.

expected *-i* and *-u*, and recorded them as long vowels, *-ī* and *-ū*. This solution is not ideal, however, since he did record short *-i* as the FS suffix in some P-STEM forms. Some Arabic dialects which generally disallow final CVV syllables, such as Cairene, allow final CVV in particular circumstances (McCarthy 2005: 10, Watson 2002: 56 note 1), but these tend to be borrowings (e.g. Cairene *gatō* ‘cake’) or concatenated forms and should therefore be more readily considered as exceptions than the Rwala imperative paradigm. It may be considered that the Rwala imperatives regularly had final stress (perhaps by analogy with the *-ūn* and *-īn* endings of the P-STEM), to which we may compare the forms where a final glottal stop was not deleted after *ā* as described in §2.1.1.1. Alternatively, it is possible that Musil transcribed these forms incorrectly but, in either case, word-final vowel shortening seems to be a firmly established phonological rule in the dialect.

2.3.4 *Pre-suffix /n/ gemination*

A process, termed *n*-strengthening by Watson (2002: 205-206), geminates the word-final /n/ before suffixation in the following forms:

/min + eh/	>	<i>minn-eh</i>	‘from him’
/‘an + eh/	>	‘ <i>ann-eh</i>	‘to him’
/ǧan + ha/	>	ǧ <i>ann-ah</i>	‘they (f.) came to her’

Since *min*, ‘*an* and ǧ*an* are all sub-minimal words (cf. §2.2.2), Watson’s analysis that gemination is triggered by the requirement for suffixes to be affixed to a form at least the size of the minimal phonological word accounts for these data too. The forms vary with non-

geminated variants: *minn-eh* but also *min-eh* and *min-ha* ‘from her’, ‘*ann-eh* but also ‘*an-ha* ‘to her’. The variation of geminated and non-geminated forms is unpredictable, as demonstrated by the only attested second person forms of ‘*an* which are ‘*an-k* ‘to you (m.)’ and ‘*ann-eć* ‘to you (f.)’ (cf. §5.2.1.4). The gemination is only attested for vowel-initial suffixes.

A superficially similar but probably unrelated phenomenon is that of the gemination of /n/ before a suffix in FP verbs with the suffix *-en*, e.g.:

<i>gaššenn-eh</i>	‘they (f.) strike against him’	(< /gaššen + eh /)
<i>yerā’enn-eh</i>	‘they (f.) watch over him’	(< /yerā’en + eh /)
<i>đabaḥenn-eh</i>	‘they (f.) killed him’	(< /đabaḥen + eh /)

These forms in Musil are only attested in poetry, although this is probably just a reflection of the content of the recorded material. Such forms are attested in other dialects in normal speech, e.g. Omani *đarbann-uh* ‘they (f.) hit him’ (Holes 2011c: §2.1.3.6).

2.3.5 *V_V resolution*

On account of a restriction on onsetless syllables, any sequence of V.V is unacceptable in the language and resolved by one of three repair strategies: V deletion, glide insertion, and epenthesis of [ʔ]. At the phrase level all three strategies occur; Musil’s data do not shed light on what causes this variation, but it is likely that the process adopted depends on factors such as the speed of delivery and stress (see Watson 2002: 228). Where concatenation creates a V.V sequence at the word level, only V deletion or glide formation may occur.

2.3.5.1 V deletion

V deletion is by far the most common hiatus elimination strategy. Examples at the phrase level include:

two short vowels

<i>lena-bn</i>	<	/le-na ibn/	‘we have a son’
<i>tāli l-leyl</i>	<	/tāli al-leyl/	‘the period between midnight and sunrise’
<i>wuddi-rʿab</i>	<	/wudd-i arʿab/	‘I want to mount’
<i>an-uḥalliʿ</i>	<	/ana uḥalliʿ/	‘I give you (f.) away’

one long vowel and one short vowel

<i>yā-lli</i>	<	/yā alli/	‘O you who...’
<i>mā-rīd</i>	<	/mā arīd/	‘I do not want’
<i>lā-nt</i>	<	/lā int/	‘you (MS) are not’
<i>yā-mm</i>	<	/yā umm/	‘O mother of...’

As these lists demonstrate, a long vowel tends to be maintained regardless of the quality of the following vowel, but the situation with two short vowels is more varied. In terms of hierarchy, the high vowels /i, u/ tend to be maintained causing the loss of the low vowel /a/, although phrases such as *lena-bn* are an exception to this trend.

V deletion also takes place at the word level via concatenation:

<i>laḥalli</i>	<	/la + aḥalli/	‘I shall cause [it] to happen’
<i>lagat</i>	<	/laga + at/	‘she met’

2.3.5.2 Glide formation

In some cases a phrase level hiatus is eliminated via glide formation, as illustrated by the following examples:

<i>ġuwi</i>	‘strong’
<i>ġuwiyy al-‘ezūm</i>	‘steadfast in decisions’
<i>ġarhadi</i>	‘destructive; a destroyer’
<i>ġarhadiy an-nowm</i>	‘sleep is a destroyer’

Glide formation is restricted to contexts in which the left-most vowel is high /i, u/, since the glide is formed via the association of the features of the high vowel with the onset of the next syllable. All the clear examples in the corpus are of /i/, and although glide formation of /w/ is almost certainly attested, Musil’s convention of transcription records many cases of word-final <w> which are probably vocalic:

<i>hw</i>	‘he, it’
<i>hw ašwa</i>	‘it would be better’ (represents [h(u)wašwa]?)
<i>ġazw</i>	‘raiding troop’
<i>ġazw eben rašīd</i>	‘Ibn Rašīd’s raiders’ (represents [ġaz(u)weben]?)

At the word level, glide insertion may also eliminate hiatus, as seen in the feminine relational adjective (cf. §3.2.4):

<i>/fedawi + a/</i>	>	<i>fedawiyye</i>	‘free servants’
---------------------	---	------------------	-----------------

Within individual stems, glide insertion is found regularly in some templates but this is a different process. Common sites include:

where /ʔ/ has been deleted (see §2.1.1.1)	
<i>swāl</i> < <i>suwāl</i> < <i>suāl</i> < /suʔāl/	‘question’

in some plural forms (see §3.1.4.1)

ʿawāref (sg. *ʿārefa*)

‘hereditary native judges’

ḥarāyeġ (sg. *ḥarīġe*)

‘shallow wells’

in the diminutive pattern (see §3.1.5)

twēreš (non-dim. *ṭāreš*)

‘little pilgrim’

2.3.5.3 Epenthesis of [ʔ]

This hiatus elimination strategy is only attested at the phrase level. It is not specifically transcribed by Musil, but there are 215 instances of V_V in the corpus where I assume that, since neither glide formation nor V deletion took place, glottal stop epenthesis has occurred. Examples are:

haḍōla agrab

‘these people are closer’

hāḍa illa

‘this, but’

yaḥma umm-eh

‘he protects its mother’

yeġeddi al-mezāher

‘he prepares the way for the camels carrying loads’

The assumption of glottal stop epenthesis stems from the restriction in most varieties of Arabic on onsetless syllables (cf. §2.2.1 and Figure 2.8):

haḍōla agrab

ha.ḍō.la.ʔag.rab

This assumption applies also to vowel-initial words in phrase-initial position, where epenthesis of the glottal stop is predicted to occur:

alli yākol

ʔal.li.yā.kol

‘He who eats’

it

ʔit.

‘Go!’

2.3.6 *Qualitative vowel processes*

2.3.6.1 Low vowel raising

Many of the Arabic dialects of north Arabia have a rule which raises /a/ > i in an open syllable. This was first elaborated in a detailed study by Johnstone (1967b) for the dialect of ‘Anayza. Johnstone observed (*ibid.*: 4) that in that dialect *i* occurred to the exclusion of *a* in open syllables except under certain conditions. The conditioning environments he recorded are:

- (a) when the vowel of the following syllable is *a*, *ā*, *ē* or *ō*: *a* occurs if it is preceded by or followed by one of the guttural consonants *ġ*, *ħ*, *ʕ*, *h*, *h* or followed by one of *l*, *r*, *n*, *w*
- (b) when the vowel of the following syllable is *ī*, *u* or *ū*: *a* occurs if it is preceded by one of the guttural consonants *ġ*, *ħ*, *ʕ*, *h*, *h* (Johnstone, *ibid.*: 6, 11)

Research in subsequent decades has shown that this rule applies widely to bedouin-descended dialects of northern, central, and eastern Arabia. The extent to which the dialectological literature implies that the Rwala dialect should have this rule is uncertain, however. In many cases the rule is much more vaguely expressed, which is perhaps best shown by Ingham’s assertion that in the Najdi dialects in general, /a/ “does show fronting to a sort of *e* sound in front environments” (Ingham 1994a: 15). Cantineau does not define the nature of the process either, simply noting that /a/ has a wide range of realisations (extending to both *i* and *u*) in the environment of unspecified neighbouring consonants and unspecified vowels in neighbouring syllables (Cantineau 1936-37: 48).

These imprecise descriptions are inadequate, however, as it appears that in the data of both Ingham (specifically the Rwala material in Ingham 1995) and Cantineau fairly regular processes can be observed; furthermore, the data show a situation almost identical to that described by Johnstone for ‘Anayza. Prochazka adapted Johnstone’s model in his description of the Rwala dialect and others closely related to it, which may be demonstrated by his fourfold presentation of S-STEM verbs which all have CaCaC as their underlying representation (forms from Prochazka 1988: 28-29, my emphasis):

<i>kitab, ktibat</i>	‘he/she wrote’
<i>ḍibaḥ, ḍbaḥat</i>	‘he/she killed’
<i>ḥalab, ḥlibat</i>	‘he/she milked’
<i>ga‘ad, g‘adat</i>	‘he/she sat’

In this system, the raising of /a/ to *i* is dependent both on its position in an open syllable and on the consonantal environment: namely, /a/ in an open syllable is realised as *i* unless it is in the environment of a guttural fricative /ğ, ḥ, ʕ, ḫ, h/ or if the following consonant is /l, r, n/ (*ibid.*: 28). This description finds close support in Ingham’s Rwala texts, where e.g. *nizal, nzalow* ‘he/they camped’ (Ingham 1995: 126) corresponds to *ḍibaḥ, ḍbaḥat* above.

The following rule may therefore be posited for /a/-raising⁹¹:

$a \rightarrow [-\text{LOW}] / X_]_{\sigma} Y$	X = not /ğ, ḥ, ʕ, ḫ, h/
	Y = not /ğ, ḥ, ʕ, ḫ, h, l, r, n/

⁹¹ It will be observed that this formulation does not account for all of Johnstone’s data, notably that this allows consonantal environment to block the raising of *a* > *i* when the vowel of the following syllable is *i*, which is ruled out by Johnstone (also cf. Johnstone 1967b: 11). A solution to this is addressed below.

Since Prochazka's evidence is restricted almost exclusively to inflected verbal forms, I have attempted to address this limitation by comparing his results with Ingham's Rwala data, which cover all lexical categories. Like Prochazka, Ingham uses only three vowels so the data are comparable. The following are apparently cases of /a/-raising in Ingham (my emphasis):

<i>sibil-ah</i>	'her pipe'	<i>ribi^c</i>	' <i>rabi^c</i> [a season]'
<i>yimīn</i>	'left'	<i>dibaḥ</i>	'he killed'
<i>gīṣidah</i>	'poem'	<i>tuwīl⁹²</i>	'tall'
<i>nizal</i>	'he pitched a tent'		

/a/-raising is attested in Ingham's texts in almost every context where it could have occurred, which lends support to the theory. The clear parallel between Prochazka's data and Ingham's therefore strongly suggests that low vowel raising is a regular process in the dialect; furthermore, the data from Cantineau seems to conform to these patterns as well, although he did not comment on it except in his remark already cited that /a/ had a very wide range of realisations extending even to *i* (Cantineau 1936-37: 48). To evaluate the extent to which Musil recorded the raising of low vowels in his corpus, the above examples from Ingham may be compared to the same forms in Musil:

⁹² The inclusion of *tuwīl* rests on an assumption that /a/ would be raised before being backed: /ṭawīl/ > tiwīl > *tuwīl*. Backing is addressed in §2.3.6.2.

	Ingham	Musil	
(1)	<i>sibīl-ah</i>	<i>sebīl</i>	‘[her] pipe’
	<i>yimīn</i>	<i>yemīn</i>	‘left’
(2)	<i>gīšīdah</i>	<i>gašīde</i>	‘poem’
	<i>nizal</i>	<i>nazal</i> (~ <i>nazel</i>)	‘he pitched a tent’
	<i>ribī^c</i>	<i>rabi^c</i>	‘ <i>rabi^c</i> [a season]’
(3)	<i>ḍibaḥ</i>	<i>ḍabaḥ</i> ~ <i>ḍebaḥ</i>	‘he killed’
	<i>ṭuwīl</i>	<i>ṭawīl</i> ~ <i>ṭwīl</i>	‘tall’

Here there are clearly no examples in Musil of /a/ raising all the way to *i*, but the examples in (1) show where /a/ is transcribed as *e*, which could possibly be an indication of raising (although it could merely indicate fronting); this is conceivable since Ingham elsewhere described the raised vowel as “a sort of *e* sound”, as cited above. The examples in (2), on the other hand, show no raising at all, and in (3) both options occur. It can therefore be observed that /a/-raising is only attested in Musil if we assume that *e* represents a raised /a/, and in cases such as *ṭwīl* where the raised vowel has been subject to syncope, although the latter is more tenuous evidence.

There are some cases where the rule given above does not seem to account for the data. Examples from Ingham are:

	Ingham	Musil	
(1)	<i>‘iṣrīn</i>	<i>‘aṣrīn</i>	‘twenty’
(2a)	<i>wiridtu</i>	<i>wariden</i>	‘you (pl.)/they (f.) came to water’
(2b)	<i>yġīrūn</i> ⁹³	<i>yaġīrūn</i>	‘they (m.) attack’

⁹³ Similarly to *ṭwīl* above, the inclusion of *yġīrūn* rests on an assumption that /a/ would be raised before being subject to the high vowel syncope rule.

In the case of (1), this is not explained by the posited rule on account of the fact that raising seems to occur in a closed syllable. In (2) the following consonant is of the group /ğ, ɸ, ʕ, ɣ, h, l, r, n/ and raising also should not have occurred. The case of (2a) is special because it is an I-CLASS verb (cf. §4.2.2.2), and whether these are subject to a synchronic raising rule (i.e. the underlying representation is /CaCiC/) or not (i.e. the underlying representation is /CiCiC/) is addressed below.

One possible solution to this situation has been proposed by Parkinson (1992: 107-108, 110-111), although it is important to stress (as mentioned in the introduction) that Parkinson's article contains many errors which mean that his theories must be approached with caution. Parkinson sees two separate raising rules as necessary to account for the Rwala data recorded by Prochazka, alongside the two syncope rules generally accepted for the dialect (cf. §2.2.3). These are a rule similar to that of /a/-raising already described, and a second rule which I here term /i/-spreading⁹⁴. In his view, the former only occurs if the following vowel is also /a/, and he sees it as a dissimilatory process in which a sequence *a-a* becomes *i-a* if the first syllable is open. Naturally, some of the cases of raising already given would not be subject to /a/-raising in this framework since the following vowel is not /a/. He describes /i/-spreading as an assimilatory process in which a sequence *a-i* becomes a sequence *i-i* and which is blind to syllable structure, which may be expressed as follows:

⁹⁴ Parkinson refers to these rules respectively as Raising and Coronal Spreading: the latter term is apposite in Parkinson's treatment, which is mostly concerned with feature geometry, but I have used terminology which does not make reference to features, as I have not elaborated features in this work.

a → [-LOW] / _ C₁i

In this formalisation, C₁ refers to any number of consonants. He demonstrates the application of these rules in the verbal system (Parkinson 1992: 116):

UR	/katab/	/katabat/	/sami ^ʕ /	/sami ^ʕ at/
/i,u/-syncope	_____	_____	_____	sam ^ʕ at
/a/-syncope	_____	katabat	_____	_____
/i/-spreading	_____	_____	simi ^ʕ	_____
/a/-raising	kitab	ktibat	_____	_____
output	<i>kitab</i>	<i>ktibat</i>	<i>simi^ʕ</i>	<i>sam^ʕat</i>
	‘he wrote’	‘she wrote’	‘he heard’	‘she heard’

As is apparent from these derivations, the existence of only /a/-raising and not including /i/-spreading would also produce the correct output. The only evidence for two rules is that /i/-spreading is not blocked by /ğ, ħ, ^ʕ, ħ, h, l, r, n/ while /a/-raising is (examples from Prochazka):

/i/-spreading	/šarib/	>	šīrib	‘he drank’
	/ta ^ʕ ib/	>	tī ^ʕ ib	‘he became tired’
/a/-raising	/ḍarab/	>	ḍarab	‘he hit’
	/naḥal/	>	naḥal	‘he sifted’

In the /a/-raising examples, raising has been blocked by the following /r/ and /ħ/, while in the /i/-spreading example the consonants /r/ and /^ʕ/ have not had the same effect. Parkinson also argues that /i/-spreading is blind to syllable structure, as opposed to /a/-raising which occurs only in open syllables, but on this point the data he extracted from Prochazka is incorrect and must be discounted. Interestingly, however, this aspect of his rule does receive some support from Ingham’s form *ʕišrīn*.

Apart from the examples *išřin* and *yġirūn*, which may be anomalous and do not provide strong support for the theory in themselves, the major locus of /i/-spreading is in the S-STEM I-CLASS, given here from both Prochazka and Musil:

Prochazka		Musil		
3MS	3FS	3MS	3FS/3MP	
<i>širib</i>	<i>šarbat</i>	<i>šireb</i>	<i>šarbet</i>	‘drink’
<i>libis</i>	<i>labsat</i>	<i>labes</i>	<i>labsaw</i>	‘wear’

Here it may be observed that while Prochazka has *i* in the first syllable of both of the 3MS forms, Musil does not. Arguably this discrepancy can only be explained if the underlying representation is /CaCiC/, and the variation in Musil between surface forms CaCeC ~ CiCeC is a result of the partial application of a synchronic rule which raises /a/ > *i*⁹⁵. This rule would have to be different from /a/-raising since it is not blocked by /ħ, ġ, ɣ, ʕ, h, l, r, n/, which supports Parkinson’s hypothesis. The defining feature which separates this class of verbs from the A-CLASS is that the /i/ in the underlying representation /CaCiC/ is the subject of syncope when vowel-initial suffixes are added: this occurs in Musil’s forms as well as in Prochazka’s. The raising of /a/ to *i* in forms with a zero or consonant-initial suffix only partially occurs in Musil, however, but is not necessary in order to differentiate the two classes because the syncope has already sufficed to do so.

Terms from Ingham which appear to show /i/-spreading are:

⁹⁵ Other descriptions (cf. Palva 1980: 124) view the I-CLASS forms as resulting from a historical change rather than a synchronic one; this is addressed in more detail in §4.2.4.

	Ingham	Musil	
(1)	<i>širib</i>	<i>šireb</i>	‘he drank’
(2)	<i>rićib</i>	<i>raćebna</i>	‘he/we rode’
	<i>wiridtu</i>	<i>wariden</i>	‘you (pl.)/they (f.) came to water’
	<i>‘išrīn</i>	<i>‘ašrīn</i>	‘twenty’
	<i>yġirūn</i>	<i>yāġirūn</i>	‘they (m.) attack’

Here (1) represents a form in which Musil did record raising and (2) records forms in which he did not, and (1) is clearly the minority. The form *rićib* is interesting in that it would have been subject to the rule of /a/-raising above, and it is only on account of its membership of the I-CLASS (like *širib*) that we may infer by analogy that this is a case of /i/-spreading, since /šarib/ would not be subject to /a/-raising on account of the consonant /r/. If we consider the forms *‘išrīn* and *yġirūn* to be anomalous, however, we are left with /i/-spreading being a rule which is wholly restricted to one particular category, namely the FORM-I S-STEM I-CLASS. It is therefore doubtful whether a productive synchronic rule should be posited to explain it. Since Parkinson had only the verbal system on which to base his hypothesis, it must have explained the data very well, but taking into account the nominal forms there seems little evidence to argue that *nizal* exhibits /a/-raising and *gašīdah* exhibits /i/-spreading when the /a/-raising rule (without Parkinson’s additional stipulation that the following vowel should be /a/) would explain both. There are unfortunately too few examples in Ingham to verify whether other cases of raising before /ġ, ħ, ʕ, ħ, h, l, r, n/ could prove the necessity of /i/-spreading, and we have seen that raising is generally so

poorly attested in Musil that his much larger data set does not provide reliable evidence in this matter. A cross-dialectal response to Parkinson's theory is a matter for future research.

The major difficulty in assessing the extent to which Musil faithfully recorded low vowel raising (regardless of which of the two potential rules was at work) is his use of vowel symbols, and specifically the uncertainty over whether *e* could represent a raised variant of /a/ or merely a fronted one. There are clear cases where the transcription of *i* demonstrates raising: these include some of the I-CLASS verbs such as *šireb*, and there are also possible cases where syncope or backing can be assumed to have targeted a raised vowel, such as *ṭwīl* (/ṭawīl/ > ṭiwīl > ṭwīl) and *ḏuwābīḥ* 'killers' (/ḏawābīḥ/ > ḏiwābīḥ > ḏuwābīḥ, see next section), although this is much more tenuous evidence. There are also many cases where *e* occurs in a position where raising would be expected, such as *ḏebaḥ* (/ḏabaḥ/ > *ḏibaḥ?) and *yemīn* (/yamīn/ > *yimīn?), which might show the effects of raising. Finally, there are many cases where raising would be expected but has clearly not been recorded, such as *labes* (not *libes) and *nazal* (not *nizal).

It is interesting to note that the application of these two rules, if we accept them, were recorded by Musil with different levels of success: /i/-spreading was transcribed sporadically, while /a/-raising was almost never transcribed. Part of the reason for this may be that /i/-spreading as a phonological process occurs at a relatively deep level while /a/-raising may occur closer to the surface and correspondingly have a more gradient output. In summation, however, we may conclude that while other accounts have observed almost

exceptionless application of some rule of low vowel raising, Musil has not recorded it in the majority of cases.

A final example from Musil's data which appears to be a type of low vowel raising is the very frequent realisation of the P-STEM prefix *ya-* as *yi-*. In the form *yir'cb-ih* 'he rides it' (from Ingham's Rwala texts) the process appears to be one of /i/-spreading⁹⁶ since it occurs in a closed syllable; it could therefore be interpreted as the /i/ spreading from the suffix to the prefix after syncope has deleted the intervening vowel. It should be noted, however, that the evidence from Musil does not support this interpretation at all: when *yi-* occurs as the P-STEM prefix in Musil it is attested before *i* less than 5% of the time. Since there is usually no /i/ to trigger the process, it cannot be /i/-spreading, and since all of those forms are closed syllables (cases where the prefix was an open syllable were excluded) it cannot be /a/-raising either. It is arguable that since it is restricted to one morpheme, and no conditioning environments seem to exist to explain it, the matter of the *yi-* prefix represents a case of lexicalisation rather than an operant process (cf. §4.3.1).

⁹⁶ *yir'cb-ih* is derived from underlying /yar'cab/, as demonstrated by the forms *nar'cab* 'we ride' and *yarkibūn* 'they ride' (in the latter /a/ has been deleted via syncope and replaced via epenthesis). In these examples, Ingham's data can be seen to differ from Musil's in two regards: the first is the realisation of the 3MS object/possessive suffix as *-ih* (cf. Musil's *-eh*), and the second is the quality of the epenthetic vowel which is *i* in Ingham but *a*, *e* in Musil.

2.3.6.2 High vowel backing

There is a fairly clear phonological process which causes the backing of /i/ > u before /w/, and possibly also the raising and backing of /a/ > u as well. Examples include:

<i>suwād</i>	(< * <i>siwād</i>)	‘black (pl.)’
<i>ḏuwābīḥ</i>	(< * <i>ḏawābīḥ</i>)	‘killers’
<i>nuwar</i>	(< * <i>nawar</i>)	‘gypsies’
<i>ṭuwāyel</i>	(< * <i>ṭawāyel</i>)	‘high hills’
<i>suwa</i>	(< * <i>sawa</i>)	‘like’
<i>yistuwi</i>	(< * <i>yistawi</i>)	‘he is calm’

Since this process is only attested for open syllables, it seems likely that the cases of underlying /a/ are explained by a twofold process incorporating the rule of low vowel raising detailed in the previous section, namely:

UR	/nawar/
/a/-raising	niwar
backing	nuwar
output	<i>nuwar</i>

It seems, therefore, that in this case Musil faithfully recorded what is attested to be a fairly frequent phonological change in the north-Arabian dialects (and beyond, cf. Watson 2002: 265-266). When Cantineau (1936-37: 48) recorded that /a/ may be articulated as *u* his example (not from Rwala, however) was *šuwāreb* ‘moustaches’ which would appear to be exhibiting the same pattern. The representation is fairly systematic in Musil, since *iw* is only attested four times (one of which varies with *uw*), and *aw* (when it does not represent a diphthong) is mostly attested in situations where /a/-raising would be blocked by a

neighbouring consonant, lending support to the theory that raising is a prerequisite for subsequent backing.

There is some evidence of backing before other labial consonants as well. A clear case is that of *ṭrub-i* ‘my singing’, which is attested in its unsuffixed form in Kurpershoek (2005: 186) as *ṭarab* ‘joy, happiness’⁹⁷. Another example is *gṣuba* ‘hollow bamboo reed’, where the underlying vowel is unknown, but the derivation here shows that it could be /a/:

UR	/ṭarab + i/	/gaṣaba/
/a/-syncope	ṭrab-i	gṣaba
(/a/-raising)	(ṭrib-i)	(gṣiba)
backing	ṭrub-i	gṣuba
output	<i>ṭrub-i</i>	<i>gṣuba</i>

2.3.6.3 High vowel lowering

It is fairly frequently attested in the corpus that vowels which would be expected to be high on the basis of either internal or comparative evidence are recorded as low, *a*, in the contiguity of one of the guttural consonants /ḥ, ġ, ḥ, ʕ, h/. It has been seen in §2.3.6.1 that the guttural consonants block low vowel raising: in these cases, however, it appears that the opposite situation has occurred, since the vowels are usually high. Some of these forms have clear parallels in Kurpershoek’s (2005) glossary, where his syncopated forms attest to the deletion of high vowels:

⁹⁷ *ṭarab* is defined as ‘joy, rapture, inspiration’ in Holes (2001: 319), and is the emotion evoked in the listener by a *muṭrib* (or *muṭarrib*) ‘singer’ (Holes, *ibid.*). ‘Singing, vocal entertainment’ can therefore be seen as an extension of this basic meaning; compare *rāḥ iṭ-ṭarab* ‘I am no longer in the mood (to compose poetry)’ (Kurpershoek 1999: 60, cited in Holes, *ibid.*).

Musil	Kurpershoek	
<i>ma‘ādi</i>	<i>m‘ādi</i>	‘enemy’
<i>‘azz</i>	<i>‘izz</i>	‘power’
<i>ġana</i>	<i>ġna</i>	‘singing, songs’

In Musil these forms with *a* are not universally attested: *ma‘ādi* is consistent, but *‘azz* is usually attested as *‘ezz* and *ġana* is also attested as *ġna*. These examples are interesting because the forms where lowering has occurred are no longer contrastive with other words: *ma‘ādi* is usually a plural noun meaning ‘attacks’ (Kurpershoek 2005: 210); *‘azz* is a verb in some dialects, such as Bahraini ‘to love, cherish’ (Holes 2001: 348); and *ġana* is recorded as a singular noun meaning ‘wealth’ in Musil and in central Arabia (Kurpershoek, *ibid.*: 245). Another example from Musil may be *ša‘‘ār* ‘poets’, which does not look like a normal plural form but rather resembles the singular nominal template CaCCāC; *šu‘‘ār* would be an expected plural, as this template is regularly attested in words such as *šunnā‘* ‘blacksmiths’, also from this corpus. The word *ġadar* is an uncertain term which Musil claims to be a variant of *ġadīr* ‘pool’, although *ġadar* is not attested elsewhere with this meaning. It is possible that Musil could have meant the plural of *ġadīr*, which in this corpus is *ġudr*; it should be noted, however, that in this case it appears that there has been some corruption in the transmission of the poem, since the plural *ġud(u)r* would not make sense either.

2.4 Summary

The comparison of Musil’s Rwala material with that from other published sources means that it is possible to assess the extent to which his transcriptions are reliable. In the

first instance, his consonants should be interpreted in the light of the discussion in §2.1.1: in this work I have changed some of his symbols to a system of transliteration which is less misleading, and the reader of Musil should also assume that most if not all of his glottal stops should be deleted. Reading Musil's texts with the understanding that *e* may represent either /a/ or /i/ should suffice for interpretation, although the usage of *e* does render his material less useful for assessing the phonology of the dialect.

I have shown that in terms of syllable structure, epenthesis, assimilation, hiatus avoidance strategies, and word-final vowel shortening Musil's material is in line with what could be expected of the Rwala dialect. The one case of the non-application of word-final vowel shortening (i.e. the imperative suffixes) is not a serious one, since it is restricted to one context, and it seems possible that it was a mistake. Likewise, his representation of prothesis, while not agreeing entirely with that of other researchers, is not inconceivable and does not cast strong doubts upon his data. His material does, however, show surprising underattestation of several important phonological processes: these are syncope, the *ghawa*-syndrome, and low vowel raising. All other accounts of the Rwala dialect have demonstrated these processes to be thorough, but in Musil there are many exceptions to each of them.

Previous scholars have interpreted this situation as a failure on the part of Musil to adequately transcribe the dialectal material he heard: in this we may include Cantineau, who was Musil's contemporary and vocal critic, and also modern researchers such as Kurpershoek, who imposes these processes retroactively onto Musil's transcriptions when he

cites them in his glossary⁹⁸. Regarding the possibility that Musil's transcriptions contain errors, it has been argued here that the surprisingly high number of cases in which syncope failed to apply in Musil's material may, in some cases, be due to a lack of perception of the process; on this matter, Musil does not appear to be a unique case, since it was once remarked by Cantineau that "les enquêteurs débutants ont toujours tendance à noter, en arabe parlé, plus de voyelles qu'il n'y en a réellement" (Cantineau 1936-37: 155). It seems likely that there is some truth in this assessment, and it should not be dismissed lightly that Musil may also have been influenced by his conception of CLA phonology when he made his transcriptions. It has also been shown that many forms do not attest the *ghawa*-syndrome where it could have occurred, but in these cases it is harder to argue that Musil could have been insensitive to it, since this process resyllabifies the whole word and is therefore arguably a more observable phenomenon than syncope. The near-total absence of low vowel raising also appears to be at odds with what has been recorded in the dialectological literature; it may be significant here that most of the sources for this rule are considerably later than Musil, but earlier sources such as Cantineau do record the process even though the theory behind it was not elaborated until later.

There are other factors which must also be considered, however. We have seen that the structure of poetry may be an influencing factor on matters such as these. It is hard to

⁹⁸ An example of this which relates to both syncope and low vowel raising is Musil's *aḥīdaten* which Kurpershoek (2005: 4) cites as *iḥīdīn*, and for the *ghawa*-syndrome an example is Musil's *ya'zel* which Kurpershoek (*ibid.*: 216) cites as *y'azil*.

demonstrate clearly that poetry affected the rate of syncope, since, as noted above, there are many cases in which forms scan the same way whether syncope occurred or not: *fāhemīn* and *fāhmīn* have been given as an example, which both scan as – ʊ – (excluding the final consonant). It cannot be ruled out, however, that some performance styles may have led to short vowels in open syllables being retained in poetry more frequently than in normal speech. Moreover, it is certainly possible that the *ghawa*-syndrome was blocked in some poetic material, as non-*ghawa*-forms are attested plentifully in poetry in contexts where the *ghawa*-form would not fit the metrical requirements (i.e. the CaX syllable is analysed as a heavy syllable in scansion).

We must also take sociolinguistic considerations into account, since Musil's material, gathered as it was during a period of long-term observation, is more likely than that of other researchers to bear witness to the kinds of variability he recorded; variability, furthermore, which we should expect to find in a corpus this size in the light of sociolinguistic research. It seems possible that the surprising forms found in Musil's work—while not discounting the possibility that some were incorrectly recorded—reflect a measure of variability which concurs with with the sociolinguistic insight that speech behaviour is probabilistic and not categorical. In this sense, Musil may have succeeded in his stated aim of recording the dialect exactly as he heard it, and what appear to be “unsyncoated forms”, “non-*ghawa*-forms” and forms without low vowel raising in fact reflect the considerable variation he encountered while living among this large and diverse social group, which he himself

alluded to (albeit dismissively) when he described them as having “many inconsistencies in... pronunciation” (Musil 1928a: *xiii*).

3 Nominal morphology

This chapter deals with the morphologically defined categories of noun and adjective, which are an open class and inflected for gender (masculine and feminine), number (singular, dual and plural) and definiteness (definite or indefinite). The construct phrase, which is the juxtaposition of two nouns (or sometimes adjectives) is included here and includes a discussion of the allomorphs /a/ and /at/ of the feminine ending. The adnominal linker /en/ is also included because it is a bound morpheme which plays an important role in the morphology of nouns and adjectives. Finally, the numerals are included, although they constitute a very specialised semantic field and are not, at least in practice, an open class like the rest of the nominals.

3.1 Nouns

3.1.1 Gender

Masculine is the unmarked gender. Nouns are feminine if they have the feminine ending /a/, which has the allomorph /at/ in construct and upon suffixation (i.e. the *tā'* *marbūṭa* of CLA, cf. §2.2.3.3). Certain feminine nouns do not have this feminine ending from a historical perspective but do end in *-a* synchronically; an example is the substantivised usage of female colour adjectives (see below) which end in *-ā'* in CLA, e.g. *ḥamra* 'light

brown she-camel'. Compare also *ʿaša* 'stick', which occurs as *ʿašan* in CLA and is feminine in both varieties.

The list of feminine nouns which do not end in *-a* is finite. There are a number of nouns which are feminine by virtue of referring to female humans; these are described by Procházka (2004: 238) as "natural feminines". In the corpus these are: *oḥt* 'sister', *umm* 'mother', *bikr* 'virgin woman', *bint* 'daughter', *ḥāšī* 'adult girl with her face covered who performs with a sword during a circumcision ceremony', *ḥāmel* 'pregnant woman', *ṭamūḥ* 'married woman who desires another man', *ʿaḡūz* 'old woman', *ḡarw* 'beautiful woman', *ḡāzi* 'conquering maiden', *nāzel* 'pregnant woman'.

A similar number of terms refer to female animals: *ṭilb* 'toothless she-camel', *ḡill* 'adult she-camel', *ḥanić* 'riding she-camel which can eat her fill while on the march without stopping or deviating from her course', *ḥalūḡ* 'she-camel whose calf is killed immediately after birth and which notices it', *rub*⁶ 'she-camel in her sixth year which may be covered by the bull for the first time', *ruḥūl* 'female pack camel', *fāter* 'old she-camel', *mešer* 'pregnant camel up to the forty-fifth day', *mišwāl* 'fast mare or she-camel which raises her tail high when galloping', *mīlāf* 'she-camel which is loath to part from her young or herd and moans grievously for several days', *muḥres* 'pregnant mare in the last month before foaling', *muyasser* 'she-camel which is ready to be led in (to a rutting male which will cover her)', *mūṭi* 'camel in the last month of pregnancy', *nušūf* 'she-camel between her tenth and twentieth year', *heḡen* 'female riding camel'.

As in most varieties of Arabic, a number of other unmarked feminine nouns are attested in addition to the “natural feminines”. It has been noted (Procházka, *ibid.*) that a large number of unmarked feminine nouns tend to belong to certain semantic categories across the Arabic-speaking world, but that within these categories the attested unmarked feminines vary greatly between dialects, with the exception of a small number of words which are almost always feminine. One such category is parts of the body: Procházka lists ‘eye’, ‘ear’, ‘hand’, and ‘foot’ as terms which are almost universally feminine across all varieties of Arabic, this observation is upheld by Musil’s material since ‘*ayn* ‘eye’, *īden* ‘ear’, *yad* (~ *īd*) ‘hand’ and *riḡel* ‘foot’ are all demonstrably feminine. Other feminine body parts in the corpus are *ašbāʿ* ‘finger’, *ḡild* ‘skin’, and *drāʿ* ‘arm’. *rūḥ* and *nafs* (~ *nefes*), both meaning ‘soul’, are also feminine, though these are not body parts in an anatomical sense. An interesting case is that *galb* ‘heart’ is masculine in Musil’s material, but was recorded as feminine in the Najdi material of Socin (cited in Procházka, *ibid.*); *fawād* ‘heart’, on the other hand, is attested as feminine in Musil.

Another category is words relating to heaven, earth, and the elements. Procházka (2004: 241) gives ‘sun’, ‘earth, land’ and ‘fire’ as terms which are almost universally feminine, and in Musil we indeed find that *šams* ‘sun’, *arḍ* ‘ground’ and *nār* ‘fire’ are feminine. Other unmarked feminines in the earthy category from Musil are *gāʿ* ‘mud-flat’ (also listed as feminine for Najdi Arabic in Ingham 1994a: 62) and *bilād* ‘country,

countryside', and in the heavenly category we find *sheyl* 'Canopus'. It seems possible that other stars or constellations may also be feminine, but their gender is not attested.

Words for weather and seasons is another category of unmarked feminines, but one which is not well represented in Musil. It is interesting to note that *rīḥ* meaning 'wind' is masculine in Musil, although Procházka (*ibid.*: 243) claims that it is feminine in the dialects of the Arabian Peninsula; *habūb* 'breeze' is feminine, however. There is a great number of other specific terms for winds, but again their gender is not discernible from the corpus. All words for rain are masculine, although these are feminine in dialects of north Africa (*ibid.*). *sarāb* 'mirage' is an unmarked feminine, which may fit loosely into this category.

dār 'dwelling, camping ground' is attested as feminine, which fits into the category Procházka terms "inhabited places" (*ibid.*: 244). Other words which may be unmarked feminines in Arabic pertain to animals, but Musil's material provides no evidence for this: for example, the gender of *efā'i* 'short poisonous snake' in Musil is unknown, though the word is feminine in the majority of modern dialects (*ibid.*: 246). Likewise, the corpus does not reveal the gender of *arnab* 'hare' or *agrab* 'scorpion', though these are usually feminine in bedouin dialects (*ibid.*).

Unmarked feminine nouns for weapons are said to be attested only in the Arabian Peninsula (*ibid.*: 246), and indeed weapons such as *bārūd* and *bindug* 'rifle' are feminine in Musil (cf. feminine *bindig* in Johnstone 1961: 264). *ḥarb* 'war' (which is included in this category because the original meaning of the word was 'lance, spear') is masculine in Musil,

though feminine in Najdi Arabic (Ingham 1994a: 62) and in eastern Arabia (Holes 1990: 155).

Some nouns may be either masculine or feminine. Examples from Musil are: *ḥāl* ‘situation’, *rīg* ‘spittle’, *ḍabi* ‘(female) gazelle’, *anz* ‘(female) gazelle’, *faras* ‘horse, mare’, *lejl* ‘night, nighttime’.

By convention, terms of endearment are always masculine in form and agreement even when they refer to a female. These include *ḥabīb* and *ʿašīr*, both meaning ‘darling’.

3.1.2 Noun templates

The following lists contain the most productive noun templates in use in the corpus; other templates found in the glossary which are not listed here are either infrequent or there is some ambiguity over whether they represent a unique template or a variant of one already listed. The minimal trilateral templates are those forms which have no affixes but may have the feminine ending. Extended trilateral templates are all those which incorporate affixes, which may be internal vowel lengthening, gemination of the second radical, or prefixation; they may also exhibit suffixation, but this is dealt with separately. Finally, quadrilateral (and longer) templates are listed; since trilateral templates with a geminated radical pattern like quadrilaterals, however, these are grouped together. Forms with variation of vowels are shown (/a/ or /i/ varying with *e* and /u/ with *o*), as are syncopated forms; where the underlying form of either *e* or a deleted vowel cannot be

ascertained from internal evidence, the form in other dialects has been consulted and the form is given in brackets. Some of the following forms also show evidence of the *ghawa*-syndrome (cf. §2.3.1) or of backing or raising (cf. §2.3.6). These templates also form the basis of nouns from biliteral roots and other weak roots, but the outcome in these cases is predictable. For example, the root **ğdy** in the CaCC template yields *ğedi* ‘young goat’, while **byt** in the same template yields *beyt* ‘tent’, and **đmy** in the CuCaC template yields *đuma* ‘thirst’⁹⁹.

3.1.2.1 Minimal trilateral templates

CaCC	<i>bard</i> ‘chill, cold’, <i>tamr</i> ‘dates’ (collective noun), <i>telğ</i> ‘snow’
CaCCa	<i>tamra</i> ‘date’, <i>ğed’a</i> ‘camel in its fourth year’, <i>ghawa</i> ‘coffee’
CaCaC	<i>baħar</i> ‘basin’, <i>leben</i> ‘sour milk’, <i>ğedaħ</i> ‘milk pot’
CaCaCa	<i>raħama</i> ~ <i>rħama</i> ‘vulture’, <i>leħama</i> ‘part of a loom’, <i>ħağela</i> ‘white spot on a mare’s leg’, <i>ħeleme</i> ‘larva of the gadfly’, <i>bgara</i> ‘antelope’, (<i>gšuba</i> ‘hollow bamboo reed’)
CaCuC	<i>rağol</i> ‘man’
CiCC	<i>ğild</i> ‘skin’, <i>debs</i> ‘grape honey’
CiCCa	<i>ğibba</i> ‘jacket, coat’, <i>fetna</i> ‘revolt’
CiCaC	<i>widać</i> ‘suet’
CuCC	<i>ğurħ</i> ‘large wound’, <i>ong</i> ‘neck’
CuCCa	<i>ħurma</i> ‘woman’, <i>mohra</i> ‘young mare, filly’
CuCuC	<i>buřum</i> ‘terebinth’, <i>rořol</i> ‘rottol (a dry weight measure)’

As discussed in §2.2.2.1, for certain forms in this category which take the form CVCeC it appears that the *e* is an epenthetic vowel inserted into the templates CaCC, CiCC

⁹⁹ One less predictable exception is that nouns of the template mVCCa/āC (cf. §3.1.2.2) commonly end in *-āt* if the final radical is weak; examples are *marmāt* ‘objective, goal’, *margāt* ‘means of climbing up’, and *mařhāt* ‘desire’.

and CuCC. Many of the nouns which may be in this category have higher-sonority sounds as the final root consonant, /n/ and /r/ being the most frequent, and this may account for the epenthesis.

CaCC	<i>ṣaḥen</i> ‘shallow copper pan’, <i>maten</i> ‘waist’, <i>semen</i> ‘melted butter’
CiCC	<i>tiben</i> ‘straw’, <i>ḥeber</i> ‘ink’, <i>meleḥ</i> ‘salt’
CuCC	<i>ḥubez</i> ‘bread’, <i>goṣen</i> ‘twig’

3.1.2.2 Extended trilateral templates

CaCāC	<i>nahār</i> ~ <i>nhār</i> ‘day’
CaCāCa	<i>saḥāba</i> ‘dense grey cloud’, <i>rebāba</i> ‘rebāba, an instrument’, (<i>gḏāḏa</i> ‘white kerchief’)
CaCiC	<i>rabī</i> ‘abundance’, <i>ḡelid</i> ‘hoarfrost’, <i>wriḏ</i> ~ <i>warid</i> ‘jugular vein’
CaCiCa	<i>gaṭīfa</i> ‘carpet’, <i>temīma</i> ‘sacrifice’, <i>ḏbiḥa</i> ~ <i>ḏabiḥa</i> ‘sacrifice’
CaCūC	<i>habūb</i> ‘cold breeze’, <i>ḏelūl</i> ‘riding camel’, <i>š‘ūg</i> ~ <i>ša‘ūg</i> ‘dagger grip’
CaCūCa	<i>ḥalūba</i> ‘milking camel’
CāCūC	<i>bārūd</i> ‘rifle’
CiCāC	<i>bilād</i> ~ <i>blād</i> ‘country’, <i>helāl</i> ‘new moon’
CiCāCa	<i>ḥemāra</i> ‘she-ass’, (<i>bṭāne</i> ‘belly girth for saddle’)
CiCīCa	<i>si‘īme</i> ‘halter’
CiCāC	<i>dīwān</i> ‘drawing room of somebody wealthy’
CuCāC	<i>ḡurāb</i> ‘raven’, (<i>flān</i> ‘somebody, anybody’)
CuCūC	<i>ruḥūl</i> ‘she-camel for both riding and carrying light loads’, (<i>šrūd</i> ‘shying [camel]’, verbal noun)
CuCūCa	<i>rukūba</i> ‘female riding camel’
aCCaC	<i>abrag</i> ‘dark rock partially covered with reddish sand’
aCCaCa	<i>aṣba‘a</i> ‘finger’
aCCāC	<i>arḥāb</i> ‘inflammation of the hind leg of a camel’
maCCaC	<i>mag‘ad</i> ‘compartment’
maCCaCa	<i>mab‘ada</i> ‘far away place’, (<i>mardefa</i> ‘croup [camel]’), <i>mehkama</i> ‘court’, (<i>mezweḏe</i> ‘saddle bag for food’)
maCCiC	<i>manzel</i> ‘camping place’, <i>meḡles</i> ‘room for men to gather’
miCCaC	<i>miḡla‘</i> ‘sling’, <i>menga‘</i> ‘basin’, (<i>miḏneḏ</i> ‘edge of eyelid’)
miCCaCa	<i>mikhāle</i> ‘tiny tin box for keeping kohl’

maCCāC	<i>marhāš</i> ‘short downpour of rain’
maCCāCa	<i>markāba</i> ‘falcon stand’
maCCīC	<i>maskīn</i> ‘wretch’
maCCīCa	<i>mastika</i> ‘ordinary fabric’
maCCūC	<i>mašhūf</i> ‘small skin boat’ (cf. participles below)
maCCūCa	<i>madfūna</i> ‘small cushion for a camel saddle’ (cf. participles below)
miCCāC	<i>mihbāš</i> ‘wooden mortar’, (<i>meš‘āb</i> ‘short stick’)
miCCāCa	<i>miḥmāse</i> ‘shallow iron pan with a long handle’
muCCaC	<i>muḥmar</i> ‘vessel for kneading dough’, (<i>musned</i> ‘cushion’)
muCCaCa	<i>muḥtama</i> ‘[component part of a camel halter]’
muCCāC	<i>muṣlāḥ</i> ‘conscientious herdsman’

3.1.2.3 Nouns with initial consonant clusters

Nouns do not have initial clusters in their underlying representation, as can be seen in the above examples where variation occurs (cf. §2.2.3.5). Such clusters, where they appear, are formed by the syncope of a vowel. This list is a reference which shows which template from the above lists the nouns with initial clusters are derived from. Since the syncope process is fairly extensive in this dialect (even in Musil’s material), in many cases it is impossible to determine with certainty what the underlying vowel of a word may be, hence some of the following judgements are speculative. The supposed processes which produced the form has been given in each case:

CCaC	is from CiCaC or CuCaC	/i, u/-syncope
CCaCa	is from CaCaCa (or CaCCa)	/a/-syncope (g-syndrome)
CCāC	is from CaCāC, CiCāC or CuCāC	/i, u/-syncope*
CCāCa	is from CaCāCa or CiCāCa	/i, u/-syncope*
CCeCa	is from CaCaCa	/a/-syncope
CCiCa	is from CaCaCa	/a/-syncope, raising
CCīC	is from CaCīC	/i, u/-syncope*

CCiCa	is from CaCiCa	/i, u/-syncope*
CCuCa	is from CaCaCa	/i, u/-syncope*, backing
CCūC	is from CaCūC or CuCūC	/i, u/-syncope*
CCūCa	is from CaCūCa or CuCūCa	/i, u/-syncope*

Where syncope has targeted /a/ but the environment for /a/-syncope has not been met, it must be assumed that /i, u/-syncope has targeted /a/, either because it has a more extended application or because a two-step process has raised the underlying /a/ in an open syllable and subsequently the high high vowel has become the target of syncope (cf. §2.2.3.1). These cases are marked with an asterisk.

3.1.2.4 Participles and verbal nouns

The verbal nouns are abstract nouns expressing the action or state of the verb in question. Their forms, listed in §4.6 for FORM-I and individually in §4.8.1.1-9 for the derived forms, may be added to the templates already listed here.

The participle also regularly functions as both a noun and an adjective. Participles from all the derived forms may be substantivised, e.g. *mu‘azzeb* ‘host’ (FORM-II, active), *mefarrad* ‘weaned camel calf’ (FORM-II, passive), *mutlayyeše* ‘tired she-camel’ (FORM-V, f. active) and *mutalattam* ‘man with a kerchief tied below the chin’ (FORM-V, passive). Nouns having the form of the FORM-I participles are the most frequent, e.g.:

CāCiC	<i>ḥāmer</i> ‘yeast’
CāCiCa	<i>gāte‘a</i> ‘tent partition’, <i>ḥāfra</i> ‘head ulcer which breaks into the jaw’
maCCūC	<i>mafrūd</i> ‘weaned camel calf’
maCCūCa	<i>maṭbūḥa</i> ‘dates boiled in butter’

The question of agency is often transparent in these nouns. Nouns of the active participle forms express a form of agency, such as the yeast which leavens the bread (cf. *ḥamara* ‘to leaven’ in Wehr’s dictionary), the tent partition which divides the room and the ulcer that ‘digs’ its way into the mouth cavity. In contrast, nouns having the form of the passive participles usually express a patient, such as the weaned calf which is the patient of an action which separates it from its mother and the dates which have been boiled by another agent.

3.1.2.5 Quadriliteral and geminated templates

Forms with asterisks are attested only for quadriliteral roots and not for geminated trilateral roots; both true quadriliteral and geminated examples are given where attested.

CaCCaC	<i>ḍafḍa</i> ‘frog’, <i>ta‘leb</i> ‘fox’; <i>ṣallab</i> ‘cross’
*CaCCaCa	<i>gandara</i> ‘[a bird]’
CaCCiC	(<i>sayyed</i> ‘lord’)
CaCCāC	<i>ḍaḥḍāḥ</i> ‘vibration’, <i>dowlāb</i> ‘millwheel’; <i>nahhāb</i> ‘robber’, <i>šeyyāl</i> ‘somebody who burdens himself’
CaCCiC	<i>ḥanzīr</i> ‘wild pig’, (<i>derfīl</i> ‘waistcoat’); <i>ḥaḡḡīr</i> ‘relative who claims a girl exclusively for himself’
*CaCCiCa	<i>narkīle</i> ‘water pipe’
CaCCūC	<i>ḍarbūl</i> ‘badger’; <i>šabbūṭ</i> ‘male falcon’
*CaCCūCa	<i>ḡarbū</i> ‘a calf (of leg)’
*CiCCāC	<i>fiḡāl</i> ‘cup’, (<i>ḡerbāl</i> ‘sieve’)
CiCCiC	<i>šīršīb</i> ‘saddle tassel’, (<i>ḡifšer</i> ‘chest [camel]’)
*CiCCiCa	<i>sinsile</i> ‘chain’
CiCCiC	<i>mindīl</i> ‘kerchief’; <i>šiddīḡ</i> ‘friend’
*CiCCuC	<i>binduḡ</i> ‘rifle’
CuCCaC	<i>šumbar</i> ‘head covering’, (<i>ḥurneḡ</i> ‘young hare’); <i>sukkar</i> ‘sugar’, (<i>tummen</i> ‘rice’)
CuCCāC	<i>tumbāk</i> ‘loose tobacco’; <i>šubbāk</i> ‘sabre guard’, <i>tuffāḥ</i> ‘apples’

	(coll.)
*CuCCāCa	<i>şurmāye</i> ‘shoes of the kind worn in towns’
*CuCCuC	<i>kumkum</i> ‘coffeepot’, <i>gumfoḍ</i> ‘hedgehog’, <i>ḥoṣor</i> ‘[second or third] toe’
*CuCCūC	<i>ṭurmūd</i> ‘sorghum bread’

3.1.2.6 Nouns with more than four root letters

These nouns are not common at all. A few examples are: ‘*ankabūt* ‘spider’, *zamardağ* ‘emerald’, and *brīsem*, *briyyesem* ‘silk’ (although this last example may be based on the diminutive, see below). Other nouns with more than four root letters are denominals from non-native terms, such as *stambuliyye* ‘a *mindil* with variegated dots’ and *enklēsi* ‘a variety of Mauser revolver’, which are both nominalisations of adjectives.

3.1.2.7 Suffixed noun templates

Nouns may occur in the suffixed adjectival templates described below in section §3.2. In many cases these nouns are clearly substantivised adjectives and require no more explanation, e.g. *r(u)weyli* ‘of or pertaining to the Rwala’ (adjective) or ‘a man of the Rwala tribe’ (noun). In other cases, however, the relationship between the adjectives is less direct.

The suffixation of the relational suffix *-i* (see below, §3.2.4) creates a subset of suffixed noun forms, especially when the relational meaning it encodes is apparent. Many of them are clearly adjectival, such as *şeyfi* ‘summer rain’ < *şeyf* ‘the *şeyf* season (April-June)’. In other cases specific lexical meaning has developed, as in the case of *kubeysi* which means

both ‘a man from the settlement of al-Kubeysa (near Hīt)’ and ‘wandering merchant’, regardless of whether or not he is indeed from al-Kubeysa. Adjectives denoting relation to a place are commonly used as nouns for specific items associated with that place, such as *ḥomsi* ‘a *mindīl* with yellow or pink stripes’ (presumably a style associated with the city of Homs), and a great many weapons, such as *mogrebi* ‘a sword with three long grooves’, *hendi* ‘a sabre made of black steel’, and *enklēsi* ‘a variety of Mauser revolver’.

-i also occurs as a suffix on the CaCCāC form relating to professions, such as *garrāʿi* (~ *garrāy*) ‘snake charmer’, a form which does not appear to be a *nisba* but perhaps shows a semantic connection to the relational meaning of the morpheme. Forms such as this are widespread in Arabia and beyond: Holes (2016: 129) gives many examples for eastern Arabia, noting the use of CaCCāCi as an adjectival template for positions and shapes, e.g. *waggāfi* pl *waggāfiyya* ‘in a standing position’; *ga^{cc}ādi* ‘in a sitting position’¹⁰⁰, and *rawwāsi* ‘head-first’. Nouns having the same form may denote “specific styles of artefact or activity” (*ibid.*, note 56), e.g. *ḥaddāri* and *nawwāri*, which are both types of singing.

Final -i is particularly frequent in the names of animals, e.g. *rgēʿi* ‘swallow’, *slūgi* ‘saluki’ and *ḡerdi* ‘a large variety of jerboa’, and many names of plants also have it, such as *baḥatri* ‘*Erodium cicutarium*’. In all these cases its significance is unclear. I would tentatively suggest that the -i in these nouns may be a separate morpheme similar to one described for

¹⁰⁰ These terms have also been attested outside Arabia, e.g. in Baghdad (Blanc 1964: 85) where *waggāfi* (Muslim), *waqqāfi* (Jewish and Christian) and *ga^{cc}ādi* (Muslim), *qa^{cc}ādi* (Jewish and Christian) are attested with the same meanings. Holes also notes the attestation of CaCCāCi in Egyptian Arabic, e.g. *qaṭṭāʿi* ‘on and off, in fits and starts’ (Hinds and Badawi 1986: 708).

San'ani Arabic (Watson 2002: 192-3), which is *-ī* and is described as having a singulative and/or diminutive sense. The corpus does not explicitly provide evidence for this, but the categories in which it occurs (plants and animals) could very easily have developed collective and singular senses for the words in question. Apart from these, most nouns ending in *-i* are in fact weak roots in other templates, e.g. *rā'ī* 'herdsman' which has the root *r'y* and the template CāCiC.

Nouns having the feminine *nisba* suffix occur in similar distribution to those bearing the masculine suffix *-i*. Examples are *šerġiyye* 'east wind' < *šerg* 'the east, the inner desert', *ġamriyye* 'bread baked in hot ashes' < *ġamr* 'hot coals', *obṭiyye* 'a short *farwa* which only covers the breast' < *obṭa* 'armpit', and *šiniyye* 'decorated copper plate for serving coffee', so named because such dishes were originally brought from China. As with the *-i* forms, nouns ending in *-iyya* which do not seem to relate to the relational adjectival forms are usually weak roots in another template, e.g. *ra'iyye* 'herd' which has the root *r'y* and the template CaCiCa. The common use of *-iyya* as a suffix for forming abstract nouns attested in other dialects is not attested in this corpus.

Nouns ending in *-āwi* have the form of the adjectives discussed below; examples are *fedāwi* 'a free servant' < *feda* 'a ransom', and *ḥalāwi* 'a lone wolf' < *ḥala* 'barren desert'. These nouns are usually relational, and the list includes most of the seasonal rains, such as *sheylāwi* 'the rains of Canopus'. Likewise nouns derived from adjectives with the *-āni* suffix

are attested, such as *gowmāni* ‘hostile tribe’, which as an adjective means ‘from a tribe with which one is at war’ and is derived from *gowm* ‘enemy’.

Finally, adjectives having the form CaCCān are frequently substantivised as nouns denoting a person in a particular state; examples are *‘aṭṣān* ‘thirsty man’, *wağ‘ān* ‘sick person’ and *za‘lāne* ‘wife who, having suffered abuse, leaves her husband’s tent and seeks protection elsewhere’. The occurrence of *-ān* as a suffix in non-human nouns is rare but attested, such as *gurtān* ‘gout’, *burgān* ‘pot’ and *riḥāna* ‘breeze’ or ‘sweet basil’.

3.1.3 *Semantic classes*

Only a minority of the templates given above have a predictable meaning. The most important are given here (in addition to the use of participles discussed in §3.1.1.4 and the use of adjectival forms listed in the previous section).

CaCCāC, both with quadriliteral and geminated trilateral roots, usually indicates a type of occupation. Examples are *naḥḥāb* ‘robber’, *saggār* ‘falconer’, *ğallāb* ‘merchant’ and *gaṣṣād* ‘poet’, with *beyṭār* ‘blacksmith’ being a quadriliteral example. In cases such as *gaṭṭā* ‘that which eats away at something’ the usage is not professional as such but remains agentive. Other uses do not relate to humans but contain a similar sense of agency, such as *barrād* ‘the west wind’ which cools the heat of the day in summer. In other cases, such as *karrāt* ‘a type of garlic’ (compare ‘leek’ in Wehr’s dictionary), the derivation is unclear.

maCCaC(a) and maCCiC denote place, or possibly time. Examples are *magʿad* ‘tent compartment’ (place of sitting), *manzel* ‘camping ground’ (place of camping) and *mabʿada* ‘far off place’. Sometimes the term does not explicitly denote a physical location, such as *manḥar* ‘throat (camel)’, which means ‘place of slaughtering’ and refers to the practice of slitting a camel’s throat. Examples that denote time are less well attested, but could include words such as *mabāt* which means ‘the end of the night’, derived from the verb *bāt* ‘to sleep, to pass the night’. Wright states that in CLA the quality of the second vowel is dependent on the stem vowel of the P-STEM, with maCCaC(a) corresponding to verbs that have /a/ or /u/ in the P-STEM and maCCiC(a) corresponding to verbs that have /i/ in the P-STEM (Wright 1967 vol. 1: 124); this seems to be true of this dialect as well, since the root **nzl** yields *manzel* (CaCCiC) and has *e* in the P-STEM, while **qʿd** has *o* in the P-STEM and yields *magʿad*, and **ḍbh** has *a* in the P-STEM and yields *maḍbah* ‘spot where somebody (or a camel) was killed’ or ‘chest (camel)’.

miCCaC(a), maCCāC(a) and miCCāC(a) are used to describe tools or instruments in general. Examples are *mirwad* ‘small knife for applying kohl’, *mikḥale* ‘tiny tin box for keeping kohl’, *mabhāra* ‘pot in which coffee is boiled with ground spices’, *mihbāš* ‘wooden mortar’ and *miḥmāse* ‘shallow iron pan with a long handle, upon which coffee is roasted’.

The referent of many nouns of the form CaCiC(a) is the patient of an action, e.g. *ṭariḥ* ‘person who has been thrown from the saddle’ (cf. *ṭaraḥ* ‘throw, cast down’). Thus *daḥīl*

‘somebody to whom protection has been granted’ contrasts with *dāḥel* ‘somebody who seeks protection’ (and cf. *daḥal* ‘ask for protection’).

A particularly interesting case which is very productive in this dialect is the use of CVCCeyC(a) as names for plants. Examples are *ṭummeyr* ‘hemlock stork’s bill’ (*Erodium bryoniaefolium* or *Erodium ciconium*), *gurrēš* ‘a bean-like annual’ (*Trigonella hamosa*) and *ḥumbeyḍ* ~ *ḥammeyḍ* ‘a species of sorrel’ (*Emex spinosus*). Mandaville, studying the ethnobotany of central Arabia, found it so frequent that he dubbed it the *nomen plantarum* of the region (Mandaville 2011: 208), and Littmann devoted an article to the same template in which he collected 86 different plant names which have this pattern, although from a broader geographical area (Littmann 1926: 32-41).

3.1.4 Number

Plurals of the singular templates listed in the preceding section may be formed by internal or external plural forms, or both, and duals are also attested. In a minority of cases, partial or total suppletion has occurred, e.g. *ḍelūl* pl. *rċāb* (up to ten) or *ḡeyš* (ten or more) ‘riding camel’, *faras* pl. *ḥeyl* or *frās* ‘horse’ and *mara* pl. *nisa* ‘woman’. Some nouns have

irregular plural forms, e.g. *sana* pl. *sinīn* ‘year’, *eben* pl. *beni* ‘son’ and *umm* pl. *ummahāt* ‘mother’¹⁰¹.

It is not uncommon that nouns from the singular templates listed above behave as collective nouns, in which case the singular is marked by suffixation of the feminine morpheme /a/; examples are *tamr* ‘dates’, *tamra* ‘a date’ and *šaḡar* ‘perennials’, *šḡara* ‘a perennial’.

The matter of agreement with plural forms can be fairly complex. Human and non-human plurals are treated separately here, as this is one of several semantic variables which have been shown to play an important role in agreement patterns in Arabic (cf. Holes 2004a: 202, Holes 2016: 326). Agreement, both within the noun phrase and between the noun phrase and the adjectives, pronouns and verbs agreeing with it, may be ‘strict’ (i.e. plural agrees with plural) or ‘deflected’ (i.e. plural agrees with feminine singular, regardless of the gender of that plural noun in its singular form).

For human plurals, both strict and deflected agreement occurs; in the former case the gender applicable to the referent is used, e.g.:

¹⁰¹ In all of these cases, it seems likely that the plural form represents the survival of an ancient Semitic external plural which predates the broken plural system of Arabic. *sana* is one of a group of feminine nouns which end in *tāʾ marbūʿa* in the singular (especially having the third radical elided) which have the plural *-ūn/-īn* in CLA (Wright 1967 vol.1: 196); the plural form *beni* similarly represents the external plural form *banūn/banīn* with the final *n* removed in construct phrases (cf. *ibid.*: 195). The case of *ummahāt* is also an old external plural (cf. *ibid.*: 233).

<i>tā‘at l-i al-bedwān</i>	the bedouins obeyed (FS) me
<i>awlād murḥaṣa kill ḡāli</i>	boys considering cheap (FS) all that is expensive
<i>aš-ša‘‘ār mā yidākūrūn-eh</i>	the poets did not mention (MP) her ¹⁰²
<i>aš-šuyūḥ ... ḥarīb-hom yiḥāf</i>	chiefs whose (MP) adversary is afraid
<i>banāt al-bedw ḡann</i>	the daughters of the bedouins came (FP)

Non-human plurals also show both patterns of agreement, shown here for nouns which, in the singular, are feminine (‘*ayn* ‘eye’) and masculine (‘*elem* ‘news, report’):

<i>teṣḥar ‘ayūn-eh</i>	her ¹⁰³ eyes enchant (FS)
<i>‘ayūn yagtalen</i>	eyes [which] kill (FP)
<i>al-‘uyūn al-meḏālil</i>	the shaded eyes [i.e. by long lashes]
<i>‘elūm-ek te‘azzī-ni</i>	news of you consoles (FS) me
<i>al-‘elūm ar-radiyyāt</i>	the bad (FP) reports
<i>‘elūm all[i] farrāḡīnī</i>	reports which have split (MP) [us]

In these examples we see the internal plural form in deflected agreement with FS verbs, and in strict agreement with both FP and MP forms, and with an adjective which is itself an internal plural form. The occurrence of strict agreement for non-human plurals is considered to be a feature of the bedouin dialects of central Arabia and contiguous areas (Holes 2004a: 203, Holes 2011a: §5.2.1); it follows that it is also a notable feature of the language of *nabaṭī* poetry, in which regard it resembles the agreement patterns of CLA (cf. Fischer 2002: 71-2) and the pre-Islamic classical poetry.

¹⁰² As previously mentioned, it is a convention that “darlings” in poetry be referred to with masculine forms.

¹⁰³ See previous note.

Feminine external plurals (which may be non-human) may also take feminine singular agreement, such as *ċlēmāt abġi ebdi-ha* ‘aley-k ‘a few words [which] I should like to disclose to you’ where the pronoun referring to *ċlēmāt* is the singular *-ha*.

Strict and deflected agreement may be found even within the same phrase, as shown in the following examples:

<i>hāđi manāzel-hom ħelwāt</i>	this (FS) is their camping place, [which is]
	deserted (FP)
<i>dlāl^{en} mut^cabāt^{en} šwāmi</i>	tired (FP) Damascene (MS) coffeepots

Mixing agreement patterns can often be useful in poetry, since a form which fits the metre or rhyme of the poem may be employed; both of the above examples are poetic and in both the final adjective is the rhyme. Mixed agreement is not exclusively a feature of the poetic register, however (cf. Holes [2016: 334-7, 353-4] for examples of mixed agreement in regular speech from Bahrain), and the dialectological literature indicates that patterns of agreement are motivated by other factors. Notably, research has indicated that speakers’ perception of the noun as ‘generic’ as opposed to ‘specific’ is more likely to result in deflected agreement (cf. Holes 2004a: 203-4), but I have not identified clear examples of this in Musil.

The collectives may also have strict or deflected agreement¹⁰⁴, e.g.:

¹⁰⁴ Although the term *heġen* has been described as referring to any riding camel, Musil has at other times suggested that only female riding camels are referred to, which may be the reason why FP but not MP occurs in strict agreement in these particular examples.

<i>tefseg al-heğen</i>	the camels are ruttish (FS)
<i>al-heğen tasri b-i</i>	the camels will escape at night (FS) with me
<i>yigfen bi-na l-heğen</i>	the camels will leave (FP) with us
<i>al-heğen ... fowg-hen 'eyālen</i>	the camels [with] youths on them (FP)

3.1.4.1 Internal plural templates

The internal plural templates are listed according to their morphological form. As in other varieties of Arabic, the relationship between singular and plural templates is sometimes predictable, and in other cases a variety of plural templates may be used to pluralise any given template, the correspondence of which is purely lexical.

The following list shows variation where attested, and cases where syncope has occurred. As above, brackets indicate that for forms with *e* or syncope the underlying vowel is not certain from the corpus, and the form has been assigned to a particular template on the basis of comparative evidence, which is speculative. Forms from weak roots are not listed as separate templates, as the form is predictable according to regular phonological rules, e.g. *mezāwi* sg. *mezwi* 'light, black mantle' which has the root **zwy** and the preformative /m/ in the template CaCāCiC. Geminated consonants are represented by C_iC_i.

aCCāC	<i>aṭnāb</i> sg. <i>ṭoneb</i> 'tent rope', <i>enḏāl</i> ~ <i>anḏāl</i> sg. <i>naḏel</i> 'coward'
aCāCiC	<i>abā'er</i> sg. <i>ba'ir</i> 'camel'
CaCC	<i>waḏm</i> sg. <i>wḏime</i> 'strap used on a water trough'
CaCCān	<i>ra'yān</i> sg. <i>rā'i</i> 'herdsman'
CaCaC	<i>ğafar</i> sg. <i>ğufra</i> 'low ground in a plain where rain water gathers', (<i>derağ</i> sg. <i>derğe</i> 'piece of cloth')
CaCaCa	(<i>herara</i> sg. <i>herr</i> 'a river channel in clay soil with steep banks'), (<i>hereša</i> sg. <i>herš</i> 'male camel over twenty years old')
CaCāCa	<i>hayāya</i> sg. <i>hayye</i> 'snake', (<i>tenāya</i> sg. <i>teniyye</i> 'a passage between

	two rows of tents')
CaCāCi	<i>ḥabāri</i> sg. <i>ḥabāra</i> 'bustard', <i>hedāwi</i> sg. <i>ḥda</i> 'ditty'
CaCāCin ¹⁰⁵	<i>farāḡin</i> sg. <i>feriḡ</i> 'small camp', <i>b'ārīn</i> sg. <i>ba'īr</i> 'camel'
CaCiC	<i>ḥamīr</i> sg. <i>ḥmār</i> 'donkey'
CaCūC	<i>ra'ūn</i> sg. <i>ra'an</i> 'a line of rocky hillocks'
CiCC	<i>riḡl</i> sg. <i>riḡle</i> 'shallow valley', <i>zelm</i> ~ <i>zilm</i> sg. <i>zeleme</i> 'man'
CiCāC	<i>siyāḥ</i> sg. <i>sowḥ</i> 'shallow depression', (<i>ḡebāb</i> sg. <i>ḡibb</i> 'a deep rain well or cistern')
CiCāCa	<i>tiyāḥa</i> 'wanderers' (sg. n/a), (<i>refāga</i> sg. <i>rafiḡ</i> 'companion')
CiCūC	<i>ḍiyūl</i> sg. <i>ḍēl</i> 'tail'
CiC _i C _i āC	<i>ḡihhāl</i> sg. <i>ḡāhel</i> 'reckless youth'
CiC _i C _i āCa	<i>šikkāra</i> 'breeders of goats and donkeys' (sg. n/a)
CuCCān	<i>forsān</i> sg. <i>fāris</i> 'rider', <i>wudiyān</i> sg. <i>wādi</i> 'a long and wide major valley'
CuCaC	<i>rumak</i> sg. <i>rumaka</i> 'mature mare', <i>ḥoşaṣ</i> (sg. n/a) 'categories', (<i>šgag</i> sg. <i>šugga</i> 'coarse black tent fabric')
CuCaCa	<i>gušara</i> sg. <i>gašīr</i> 'neighbour'
CuCāC	<i>mohār</i> ~ <i>mhār</i> sg. <i>mohra</i> 'mare between her first and tenth year'
CuCāCa	<i>wuzāye</i> sg. <i>wiziy</i> 'cavity in the solid stony strata of a hillside'
CuCiC	<i>wuḡih</i> sg. <i>waḡh</i> 'face'
CuCuC	<i>ṭurug</i> (sg. n/a) 'low hills'
CuCūC	<i>ḡurūb</i> sg. <i>ḡarb</i> 'a reservoir from which water is distributed', <i>ḡoṣūn</i> sg. <i>ḡoṣon</i> 'twig', (<i>neḡūm</i> sg. <i>niḡm</i> 'star'), <i>ku'ūb</i> ~ <i>ek'ūb</i> (sg. n/a) 'different joints of a bamboo spear shaft', (<i>nhūd</i> ~ <i>nehūd</i> sg. <i>nahad</i> 'breast')
CuC _i C _i aC	<i>fuṭṭar</i> sg. <i>fāter</i> 'old she-camel'
CuC _i C _i āC	<i>šunnā'</i> sg. <i>šāne'</i> 'blacksmith'
CuC _i C _i āCa	<i>sukkāra</i> 'drunken men' (sg. n/a)

The geminated templates shown above only pertain to trilateral roots. The only quadrilateral forms are:

¹⁰⁵ Note that *ba'ārīn* is also attested as *ba'arīn*. While this is superficially similar to the external plural, it is in fact a form of internal plural. Not only are these forms attested in other dialects (e.g. Ingham 1994a: 33, Holes 2016: 125) but the internal structure of the stem is transformed, so it is not simply concatenation of an external plural suffix. The form also remains the same when the noun is the first element of a construct phrase, e.g. *ba'arīn telle* 'the camels of Telle', although the /n/ is not deleted in this context in external plurals either.

CaCāCiC	<i>maǧāli</i> ‘ sg. <i>miǧla</i> ‘ ‘sling’, <i>barāṭem</i> sg. <i>borṭom</i> ‘lip’, <i>ḥarāyeǧ</i> sg. <i>ḥarīǧe</i> ‘shallow well’, ‘ <i>awāref</i> sg. ‘ <i>ārefa</i> ‘hereditary native judge’ <i>tarāci</i> sg. <i>turǧiyye</i> ‘earring’
CaCāCiC	<i>ḥarābiš</i> sg. <i>ḥarbūš</i> ‘small tent’, <i>darāriǧ</i> sg. <i>darrāǧe</i> ‘spindle’, <i>maḥālil</i> sg. <i>maḥlūl</i> ‘calf in the process of weaning’

Singulars from trilateral roots which have prefixes in their singular templates also use these two templates, exemplified here by *maǧāli*‘, and the same for singulars from trilateral roots with suffixes, exemplified by *tarāci*. Singulars from trilateral roots having a long vowel in their singular template may also use these forms, which triggers glide insertion: where the long vowel is between C1 and C2, the glide is /w/ as in ‘*awāref*, and if the long vowel is between C2 and C3 the glide is /y/, as in *ḥarāyeǧ*. The form *darāriǧ* demonstrates a trilateral root which has been geminated in its singular form, and *maḥālil* represents a biliteral (“geminate”) root, **ḥll** with a preformative /m/ in the singular.

As with the singular templates, the syncope of short vowels creates internal plural forms with initial consonant clusters. These have been included where attested in the list above, but a reference to their derivation is here provided; naturally, in some cases the quality of the elided vowel is impossible to reconstruct. Since all of these examples are of /i, u/-syncope, it may well be that /i/ and even /a/ are targeted in the same forms:

CCaC	is from CuCaC	/i, u/-syncope
CCāC	is from CuCāC	/i, u/-syncope
CCāCīn	is from CaCāCīn	/i, u/-syncope*
CCūC	is from CuCūC	/i, u/-syncope

A particularly interesting phenomenon which pertains to the plural forms is that elision of the vowel in CVCVC forms yields the form CCVC, which is frequently resolved by

prothesis giving eCCV̄C. Where the long vowel is *ā*, the resulting form eCCāC is almost always from CVCāC, although eCCāC as a variant of the plural template aCCāC is also (though much less frequently) attested. As was discussed in §2.2.2.2 it seems unlikely that syncope would trigger prothesis, since there is no mora linked to the initial consonant; in this case, it probably occurs either on account of analogy with the productive aCCāC template, or where it is preceded by another unsyllabified consonant, triggering epenthesis between the two consonants (cf. §2.2.2.2).

3.1.4.2 External plural forms

The external plural masculine ending *-īn* is restricted in nouns to those describing human males (or groups of mixed gender). Examples abound among nouns which have the form of participles, such as *gādemīn* ‘chiefs’, *muslimīn* ‘Muslims’ and *meǧfīn* ‘clans who retrace their steps in the time of plenty’. Other examples come from the CaCCāC template, such as *fellāhīn* ‘inhabitants of cultivated territories’ and *sannāyīn* ‘men who pull the rope of a bucket up from a well’, although nouns of this pattern often also take a plural form CaCCāCa, such as *ḍabbāḥa*[*t^{en}*] ‘killers’ and *ṭallābt* ‘disciples of’ (construct form). As has already been noted, the plural of *sana* ‘year’ is *sinīn*, which is irregular.

The external feminine ending *-āt* has less restricted application. It applies to groups of women, e.g. *ḥabībāt* ‘darlings’ and *ḥaṭṭābāt* ‘women who gather fuel’, but it also forms the plural of non-human nouns having the feminine ending *-a*, such as *sāgāt* sg. *sāga* ‘dangerous

place to camp' and *ḥasārāt* sg. *ḥasāra* 'loss'. Other nouns may also form their plural with *-āt*, sometimes also showing internal transformation, e.g. *zebnāt* < *zebūn* 'outer shirt'. The ending is also productive as a plural suffix on foreign words, such as *riyālāt* 'reals (currency)' and *nērāt* 'pounds (currency)'.

The other external plurals attested are the *-a* suffix on nouns of the form CaCCāC mentioned above, and *-iyye* which is fairly productive for human groups, e.g. *fedawiyye* 'free servants', *ṭurgiyye* 'travellers', *ra'iyye* 'people who change from their permanent dwellings during the rainy season to movable tents' and *bindugiyye* 'men armed with *bindug* rifles'.

3.1.4.3 Dual

The dual is not very frequently attested in the corpus, with less than one percent of nouns having the dual number and even fewer adjectives. There are two dual forms, which are formed by the suffixation of either *-ēn* (or, rarely, *-eyn*) or *-ān*. Of these, *-ān* accounts for just under a fifth of occurrences, and the reason for this variation, if there is one, is unclear; it is certainly no marker of case, nor is it exclusively poetic. The combination of the dual suffixes with bound possessive pronouns is discussed in §5.2.1.3.

The evidence from Musil does not indicate that the use of the dual is especially productive, since its occurrence (outside poetry) is fairly predictable. A third of occurrences relate to body parts that come in pairs: *der'ān* 'arms', *riḡlēn* 'legs', *ḡenāhēn* 'wings'. Another common use of the dual is to denote objects or tools which also come in pairs: *delleṭēn* 'two

coffeepots' (which always come as a pair because one is used for boiling and the other for serving), *'amūdēn* 'two stout poles used to construct a water hoist', *negdēn* 'two marbles (used for a particular game)'. People who naturally come in pairs may also be attested in the dual, such as *wālideyn* 'both parents'. Similarly, certain technical terms employ the dual: *šerīcēn* is a legal term denoting the joint-owners of property, and the term *sofratēn* refers to the two small rests allowed a camel by its rider a little before sunrise and a little after sunset. The dual is found in fossilised expressions, such as in the greeting *halūtēn* 'welcome! (I welcome you twice)', and the name of a historical battle known as *manāḥ aḏ-ḏirēn* 'the Battle Camp of the Two Starvations'.

Apart from these rather formulaic uses, which all occur in conversation, the dual is used mostly in poetry¹⁰⁶. A great number of these instances relate to periods of time, e.g. *leylēn* '[for] two nights' and *šahrēn* '[for] two months', and places, e.g. *nazlēn* '[between] two camps' and *nahreyn* '[between] two rivers'. Although these two categories are the most common, it appears that the dual is generally productive in poetry. It is also used of animals, people, and miscellanea such as *rummāntēn* 'two pomegranates' and *beyramtēn* 'two strips of fabric'.

¹⁰⁶ There is only one exception to this: a reference in conversation to *fariġēn* 'two tribes', to describe a situation of enmity between the groups in question.

3.1.5 *Diminutives*

The use of the diminutive is not especially frequent in the corpus, but unlike the dual there do not appear to be any restrictions on its occurrence. The productivity of the diminutive in north-Arabian and other eastern bedouin varieties of Arabic was noted by Wallin¹⁰⁷ (1852: 212) and its absence from Cantineau, who elsewhere described it as a sedentary feature (Cantineau 1939: 82), is described as strange by Blanc (1970: 140 note 51) who records that it is used frequently by the Negev Bedouins.

There are two forms: one for the minimal trilateral templates (cf. §3.1.2.1) and one for the extended trilateral templates and geminated or quadrilateral templates (cf. §3.1.2.2, 4, 5). The former has the basic pattern CuCeyC (CuCēC), and the cluster-initial form CCeyC (CCēC) almost always occurs on account of high vowel syncope (cf. §2.2.3.1). Examples are *gleyb* ‘little heart’ < *galb* ‘heart’, *nheyd* ‘little breast’ < *nahad* ‘breast’, *wuseym* ‘camel brand’ < *wasm* (same meaning) and *‘uweyd* ‘little stalk’ < *‘ūd* ‘stalk’.

The longer template is CuCeyCeC (the *e* in the final syllable most probably indicates an underlying /i/), and again the syncopated form CCeyCeC predominates. For quadrilateral roots, each root consonant fills a slot, e.g. *ḥoneyṣer* ~ *ḥneyṣer* ‘fourth toe’ < *ḥonṣor* ‘second or third toe’. Nouns having the pattern CVCVC̄ exhibit glide insertion in the third consonant slot, and have the form CuCeyyeC: *ḡzeyyel* ‘little gazelle’ < *ḡzāl* ‘gazelle’, *ḥbeyyeb* ‘sweetheart’ < *ḥabīb* ‘darling’, *gu‘ayyed* ‘small young male camel’ < *ga‘ūd* ‘young male

¹⁰⁷ “Im Allgemeinen hört man in den Wüstendialekten sehr häufig Deminutivformen und gewöhnlich in ihrer ursprünglichen Bedeutung” (Wallin 1852: 212).

camel'. Nouns of the form $C\bar{V}CVC$ have glide insertion in the second consonantal slot yielding $CweyCeC$, e.g. *ṭwēreš* 'little pilgrim' < *ṭāreš* 'traveller'. It is possible that medially geminated forms have the alternative form CuC_iC_ieyC , as suggested by the terms *summeyr* 'night-time entertainment' and *kummeyde* 'grief', but the diminutive status of these nouns is not clear¹⁰⁸.

When concatenated morphemes occur on the basic forms, they are also present in the diminutive forms. The feminine ending is retained, e.g. *ša'eyle*¹⁰⁹ 'grey she-camel' < *ša'la* (same meaning), *ngeyra* 'hollow between the two fleshy parts of the back of the neck' < *nugra* 'hollow, basin', and *'uneyza* is the diminutive of *'anza* 'long, flat-topped hillock rising from a plain'. In some cases plural suffixes may also occur, e.g. *ḥmeyrāt* 'little donkeys'.

3.2 Adjectives

As in other varieties of Arabic, adjectival function is expressed by morphologically distinct categories. "True" adjectives having one of the adjectival templates given in the following section account for half (51%) of adjectives in the corpus, while the other half is a mixture of participles (37%) and adjectives formed with derivational suffixes (12%), which are predominantly the relational suffix *-i* (the so-called *nisba*).

¹⁰⁸ Whether the $CVCCeyC$ template for local names of plants, e.g. *ṭummeyr*, *ḥumbeyḏ*, *ḥammeyḏ* (cf. §3.1.3) is historically related to the diminutive is uncertain.

¹⁰⁹ The *a* in the first syllable of this word seems to be irregular.

3.2.1 Adjectival templates

The following singular noun templates are also attested as adjective templates, listed here in order of frequency with the most common first. The adjectives are inflected according to the following paradigm:

	singular	plural
m.	(zero)	-īn
f.	-a	-āt

Some templates, however, also have internal plural forms, listed in the next section.

Generally speaking, the external plurals apply to human referents, although they may also apply to other animate beings, e.g. *bel zeynāt* ‘fine she-camels’.

CaCiC	<i>ṣaġīr</i> ‘small’, <i>ġedid</i> ‘new’
CaCāC	<i>basāt</i> ‘thriving’, <i>defān</i> ‘buried’
CaCC	<i>ħarr</i> ‘hot’, <i>zeyn</i> ‘good, beautiful’
CaCCān	<i>za‘lān</i> ‘wrathful’
CaCūC	<i>raġūħ</i> ‘bent, inclining’, <i>dehūb</i> ‘golden’
CaCaC	<i>ba‘ad</i> ‘far off’, (<i>ħabel</i> ‘feeble minded’)
CaCCiC	<i>bayyen</i> ‘clear, visible’
CuCC	<i>ħorr</i> ‘free; noble’

3.2.1.1 Plural adjectival templates

CVCāC	<i>ħiyār</i> sg. <i>ħeyr</i> ‘good’, <i>kbār</i> sg. <i>ċebīr</i> ‘big’
CaCāCa	<i>‘arāya</i> sg. <i>‘ariyān</i> ‘naked’
CuCCaC	<i>ḍummar</i> sg. <i>ḍāmer</i> ‘slender’

3.2.2 *Colour adjectives*

A specific template exists for adjectives of colour, as in other varieties of Arabic. The template, with examples, is:

m.	f.	pl.	
aCCaC (CaCaC)	CaCCa	CiCC/CuCC	
<i>ašgar</i>	<i>šagra</i>	<i>šigr, šugr</i>	‘red-brown; blond’
<i>ašgaḥ</i>	<i>šagḥa</i>	<i>šigh</i>	‘grey; white-pink’
<i>aswad</i>	<i>sowda</i>	<i>sūd</i>	‘black’
<i>abyaḍ (bayaḍ)</i>	<i>bēḍa</i>	<i>bīḍ</i>	‘white’

The variation of aCCaC with CaCaC, which is the form attested for e.g. *ḥaḍar* ‘green’, is on account of the *ghawa*-syndrome; since *abyaḍ* does not have the guttural fricative which would trigger this process, the attestation of *bayaḍ* must be assumed to be one of analogy with the *ghawa*-forms in the same paradigm. As in other varieties of Arabic, the paradigm is also used for bodily defects, e.g. ‘*amya* ‘blind (f.)’ and ‘*arḡa* ‘lame (f.)’, but these are not well attested in the corpus.

3.2.3 *Elative*

The template is aCCaC, as in *ašgar* ‘smaller’, *aḡsa* ‘further’, *alaḍḍ* ‘more pleasant’. It is possible to form the elative from forms other than adjectives, as evidenced by the form *ašwa* in the phrase *aṭ-ṭowr ašwa* ‘the bull [becomes] easier to handle’, which is explained by Holes as being derived from the word *šway* ‘a little’ and having the sense of “something bad becoming less bad” (Holes 2016: 130, note 58). The phrase *aryaš al-‘ayn* ‘longest of lashes’

may be another example of this: the derivation is from *rīš* ‘feathers’ and the literal meaning ‘featheriest of eye’.

3.2.4 Derivational morphemes

The relational suffix *-i*, known elsewhere as the *nisba* or the gentilic suffix, is a productive suffix which derives adjectives from nominals. It inflects as follows:

	singular	plural
m.	<i>-i</i>	n/a
f.	<i>-iyya</i>	<i>-iyyāt</i>

Historically it denoted membership of a tribe or class (Badawi *et al.* 2016: 64), a usage which is encountered in the Rwala dialect, as in the tribal affiliations *sbēʿi* ‘of the Sbaʿa’, *ğandali* ‘of the Ibn Ğandal’ and *rweyli* ~ *ruweyli* ‘of the Rwala’. The suffix is not restricted to group affiliation, however, and we find adjectives being regularly derived from nouns, such as *flāni* ‘such and such’ < *flān* ‘so and so, such and such a person’, and possibly from other adjectives, such as *ğoreyri* ‘white’ < *ğarra* ‘white (f.)’.

As these examples demonstrate, certain morphological rules apply: feminine endings are dropped before suffixation (*sbēʿi*, *rweyli*), compound forms do not occur (*ğandali* rather than **ebenğandali*¹¹⁰), and the internal structure of the stem may be transformed. In Rwala

¹¹⁰ Compound forms do occur in MSA (as in other dialects), e.g. *šarq ʿawsaṭiyy* ‘Middle Eastern’ from *aš-šarq u l-ʿawsaṭu* ‘the Middle East’ (Ryding 2005: 268), but are not attested in Musil. Note also the definition from du Grandlauney (2011): “In phrases and sentences, the second element is deleted unless it is the more meaningful, so that the first part is deleted (*ħamsata* ‘ašar ‘fifteen’: *ħamsiyy*; *ʿabū bakr* ‘the father of Bakr’: *bakriyy*; *ibn az-zubayr* ‘the son of az-Zubayr’: *zubayriyy*; *ʿabd manāf* ‘the slave of Manaf’: *manāfiyy*; *taʿabbaṭa šarran* ‘he carried evil’: *taʿabbaṭiyy*).”

Arabic there is a restriction on the syllable structure of the stem, and the penultimate syllable (i.e. the pre-suffixed syllable) is necessarily heavy unless the antepenultimate syllable is heavy. The following table gives the suffixed adjectival forms divided into syllables, followed by a gloss and the nominal form they are derived from¹¹¹:

(1a)	šii.	n-i	‘Chinese’	< šiin	
	šaa.	m-i	‘Damascene’	< {aš.}šaam	
(1b)	sey.	f-i	‘of summer’	< seyf	
(1c)	maš.	r-i	‘Egyptian’	< mašr	
(2a)	f(v).	laa.	n-i	‘such and such’	< f(v).laan
	‘o.	maa.	n-i	‘Omani’	< ‘o.maan
(2b)	ku	bey	s-i	‘of al-Kubeysa’	< {al.}ku.bey.s{a}
(2c)	ye.	maa.	n-i	‘southern’	< *yV.man ¹¹²
(2d)	ru.	wey.	l-i	‘of the Rwala’	< r(u).wa.l{a}
	s(v).	bey.	‘-i	‘of the Sba‘a’	< s(v).ba.‘{a}
(2e)	ğ(v).	waa.	z-i	‘in a pair’	< ğawz
	ğ(v).	wey.	r-i	‘of Ġowr’	< ğawr
(2f)	ġo.	rey.	r-i	‘white’	< ġar.r{a}
(3a)	ħaz.	‘a.	l-i	‘of the Ĥazā‘el’	< ħa.zaa.‘el ¹¹³
	ġan.	da.	l-i	‘of the Ibn Ġandal’	< {e.ben.}ġan.dal
(3b)	haa.	še.	m-i	‘Hashemite’	< *haa.šim ¹¹⁴
(4)	‘a.	ğa.	m-i	‘Persian’	< ‘a.ğam

¹¹¹ In the following examples, elements enclosed in curly brackets {} are elements of the noun which are removed to form the *nisba* stem. Elements in brackets () may be elided and (v) represents an elided vowel the quality of which is not attested in the corpus (I assume that words do not have word-initial consonant clusters in their underlying representations, see §3.1.2.3 above).

¹¹² The word is not attested in Musil but we may compare the term *yiman* ‘south’ from Kurpershoek (2005: 359). V represents an unsyncopated vowel, the quality of which is unknown.

¹¹³ This form has clearly undergone a further transformation, but since it is the only example in the corpus of CVCVVCVC + i it is difficult to generalise what this would be. It is interesting, however, that forms such as *ħa.za‘.li, with CVC in the penultimate syllable, do not occur.

¹¹⁴ This is not attested in the corpus, but the name of the eponymous ancestor of the Hashemites is well known.

The forms in (1) have monosyllabic stems: in (1a), (1b) and (1c) the penultimate syllable is heavy, being CVV, CVy and CVC respectively. The forms in (2) have disyllabic stems, and it is here that templatic transformation occurs: in (2a) and (2b) the penultimate syllable is already heavy, having a long vowel or a diphthong respectively; in (2c) a light penultimate syllable is made heavy by vowel lengthening; in (2d) a light penultimate syllable is made heavy by diphthongisation. The forms in (2e) are derived from monosyllabic nouns having a diphthong: these are resyllabified, the glide of the diphthong being reanalysed as the onset of a syllable, and the penultimate vowel is either a long vowel or a diphthong to make a heavy syllable, like the rest of the forms in (2). Like (2e), the form in (2f) also seems to have been resyllabified, but on account of being a geminate rather than a hollow root. The forms in (3) have a heavy antepenultimate syllable, which is CVC in (3a) (by virtue of these words having quadrilateral roots) and CVV in (3b): (3) is the only category in which the penultimate syllable is not heavy. (4) represents the only example in the corpus of what we might consider a sub-minimal *nisba*, since both its penultimate and antepenultimate syllables are light; by analogy with (2), **ağāmi* or **ağeymi* would be expected.

Nominals ending in *-a* which is not the feminine ending and therefore not removed before suffixation take the suffix *-āwi* instead of *-i*. Examples are *gaṭāwi* ‘like a sand grouse’ < *gaṭa* ‘sand grouse’, *negāwi* ‘of war’ < *nega* ‘declaration of war’, and *hwāwi* ‘love stricken’ < *hawa* ‘love, passion’. These words may be analysed as having final long vowels which are

shortened word finally (cf. §2.3.3), which distinguish them from nouns having the feminine ending. The suffix is independently productive however and should not only be considered an allomorph of *-i* for nominals with final-weak roots. Thus we find that the relational adjective for *sheyl* ‘Canopus’ is *sheylāwi*, for *trayya* ‘the Pleiades’ it is *trayyāwi* and for *ḡawza* ‘Gemini’ it is *ḡōzāwi*. The last example provides a comparison with *ḡowz* ‘pair’ > *ḡwāzi* ‘in a pair’ as the former has not been resyllabified and the diphthong retains (albeit as the long monophthong *ō*). The suffix *-awi* is only attested once, in the form *dammawi* ‘bloodstained’ < *damm* ‘blood’. Other dialects have a productive *-awi* suffix¹¹⁵ so it could be that it is simply poorly attested in this corpus, or it may be that with the gemination of /m/, and the resulting heavy antepenultimate syllable, the penultimate syllable is light as in the case of (3) above.

The final productive suffix is *-āni*, which also creates adjectives out of other nominals. Examples are *gowmāni* ‘hostile’ < *gowm* ‘enemy [tribe]’ and *barrāni* ‘unrelated’ which is probably derived from the word *barr*, which is listed as ‘land outside the town, open desert’ in Kurpershoek (2005: 10) and is related to the word *barārīt* ‘desolate regions’ in Musil. It is common for adjectives to be modified with this suffix, e.g. *ḡamrāni* ‘golden’, compared with *ḡamra* ‘red’.

¹¹⁵ E.g. Cairene (Watson 2002: 197), San’ani (Watson 2002: 199), MSA (Badawi *et al.* 2004: 65).

3.3 The construct state

Construct phrases are the only productive way of expressing genitive relationship between two nominals in the dialect, since unlike other dialects of Arabia it has not developed an analytic genitive¹¹⁶.

The feminine suffix (the *tā' marbūṭa* of CLA) has two allomorphs: it is /a/ in any position unless it is the first element in construct in which case it is /at/ (note that the second element can be a bound pronoun as well as another nominal). Thus we find the following:

<i>ǧīra</i>	‘protection’	/a/ allomorph
<i>ǧīrat-i</i>	‘my protection’	/at/ allomorph
<i>ǧīrat allāh</i>	‘Allah’s protection’	/at/ allomorph

As discussed in §2.2.3.3, the vowel of this morpheme may be targeted for syncope before following vowels:

<i>dīre</i>	‘territory’	
<i>dīret aṣ-ṣawwān</i>	‘the territory of the Ṣawwān’	(no syncope)
<i>dīrt al-ǧenūb</i>	‘the southern territory’	(syncope)

The only other way in which the first element in a construct may be morphologically marked is when it is a dual, in which case the /n/ of the suffix is deleted:

<i>idēn</i>	‘(two) hands’
<i>idē-ha</i>	‘her (two) hands’

¹¹⁶ Compare the particles *ḥagg*, *māl* and *fēd* in Najdi (Ingham 1994a: 54 and 57-9).

Unlike in CLA, the /n/ is not deleted for the external masculine plural; rather, forms such as *nāfelīn an-nās* ‘most generous of all men’ (MP adjective) and ‘*āšegīn at-ṭamāmīh* ‘the heartthrobs of women whose husbands are cowards’ (MP noun) are regularly encountered.

The second element in the construct may take the adnominal linker (cf. §3.4) but this is optional. It is especially frequent when the second element is in some way modified, e.g.:

<i>bint ‘amm^{en} l-eh</i>	‘a girl of his kin’
<i>azwāl rebd^{en} muğaffal^{en}</i>	‘silhouettes of startled ostriches’

In other cases it occurs even though there is no modification, as in *ḥall b-dār gowm^{en}* ‘he penetrated the dwelling of [some of his] vassals’. In these cases there seems to be an element of specificity, as in the above translation which means ‘some particular vassals’ and not just ‘vassals’ (in general) or ‘the vassals’. Ingham has highlighted that in Najdi Arabic, when the adnominal linker appears on nouns in construct phrase, it signifies “one or a number of specific but undefined members of a class, whereas the presence of a noun without the marker signifies a member of the class in general” (Ingham 1994a: 50).

It should be noted that noun phrases comprised of adjective + noun¹¹⁷, e.g. *šugr^{an} ḍwāyeb* ‘fair of braid’ are in a different syntactic category to the regular construct phrase: while they are superficially similar, the adnominal linker often appears on the first element of these phrases, which never happens with true construct phrases. The adnominal linker is not compulsory, however, as in the phrase *yābsat ar-rīg* ‘at the end of her tether’ (lit. ‘dry of

¹¹⁷ Ingham (1994a: 60) terms these “attributive constructs”; the construction is also referred to as a ‘false *idāfa*’ or, in Arabic, *’idāfa gayr ḥaqīqiyya* (cf. Ryding 2011).

spit')¹¹⁸; Ingham's only cites phrases which do not have the linker, such as *wisī' aṣ-ṣadur* 'broad of chest' (*ibid.*: 60). In Musil's corpus this type of phrase is restricted to the poetic register, and it is recorded as a feature of *nabaṭī* poetry in Holes (2011a: §6.2) who notes that it occurs exclusively in relative clauses.

3.4 The adnominal linker

The adnominal linker, elsewhere described as dialectal *tanwīn* or nunation, has the form *-en* in the dialect. In a minority of instances (<5%) the same morpheme is transcribed as *-an*, although there does not seem to be any pattern in this variation. The form is given as *-in* for Najd (Ingham 1994a: 49¹¹⁹), Uzbekistan (Ingham 1994b), Qatar (Bettega 2014: 19) and Bahrain (Holes 2016: 131), although in the last of these the form *-an* does occur when it is suffixed to nouns to form adverbs, a distinction which is not found in this corpus¹²⁰. The form *-an* is attested in Sudanic Arabic (Owens 1993: 111) and Andalusī Arabic (Ferrando

¹¹⁸ For this idiom, cf. Bahraini *niṣaf riġ-i* 'I had a really hard time, got to the end of my tether' (Holes 2001: 520). *niṣaf* 'to dry out' is equivalent here to *yibes* in Musil (same meaning; *yābsat* is the feminine active participle in construct).

¹¹⁹ The one exception is *fōḍa-n* 'chaos', which is presumably parsed as underlying /fawḍaa/ with final long vowel shortening, so the /a/ is part of the noun, not the adnominal linker.

¹²⁰ Holes also gives one occurrence of *-un* in the phrase *aḥ^{un} 'azīz* 'dear brother' (a formula used when allowing somebody into your home), but this is described as exceptional (Holes 2016: 133 note 63). *-un* is also attested in Tihama Arabic in Yemen (Behnstedt 1985: 60, cit. Owens 2006: 102).

2018: 96)¹²¹. Both *-an* and *-in* are attested in Siculo-Arabic (La Rosa 2016: 362) and in Judaeo-Arabic (= Middle Arabic, Blau 1999: 173).

The reason for the variation between the two forms *-en* and *-an* in Musil is unclear. Owens (2006: 104) claims that the vowel of the adnominal linker has a single value (*-a*, *-i* or *-u*), but regarding this diversity admits that “the explanation for the appearance of a low vowel *-an* or high vowel *-in/-u(n)* is not self evident” (*ibid.*: 102 note 24). He also makes reference to the variation of *-in* and *-an* as the FP verbal suffix in Najdi Arabic as a parallel example of high and low vowels occurring in the same morpheme (*ibid.*), a situation which is recorded for the Rwala dialect by Prochazka (1988: 25ff.) but not encountered in Musil (cf. §4.2.1). Since *-en* could be a variant of either *-in* or *-an*, but *-in* is never attested, one could deduce that the underlying form is */-an/*¹²². From a purely comparative perspective, however, it is tempting to conjecture that Musil’s ambiguous *-en* form represents an underlying */-in/*, since the form found in Arabia is usually *-in* and the occurrence of *-an* is restricted to more peripheral dialects. If this is the case, then the minor variant *-an* in Musil could represent a transcription error, or variation such as is attested in Holes (2016: 131, although the occurrence of *-an* here may be functionally loaded) and perhaps in Johnstone (1967a: 63).

¹²¹ It is also mentioned as a variant alongside *-in* and *-ən* by Johnstone (1967a: 63) but for the individual dialects he describes he gives only the form *-in* (*ibid.*: 87, 117, 141).

¹²² This is the view of Cantineau, which may in fact be based on this same reasoning, since he uses Musil as one of his main sources for discussion on this point. It should be noted however that Cantineau’s assertion that the form encountered among the Rwala was in fact *-an* with *imāla* (Cantineau 1936-37: 102) is not based on any tokens he actually heard.

In all dialects where it is attested, the adnominal linker has the properties of being optional, and of being a single morpheme with no regard to case. It has been observed that while its main function is to mark an indefinite noun which has a modifier, the precise syntactic contexts in which the adnominal linker occurs vary between dialects (Holes 2004a: 17, Owens 2006: 105). In this regard the usage in Musil's corpus closely follows that described for Najdi by Ingham (1994a: 49) and for the north-Arabian bedouin dialects in Cantineau's *Deuxième article* (Cantineau 1936-37: 204-5), although since the latter heavily bases his description on Musil's evidence this similarity is unsurprising.

The adnominal linker frequently occurs where an indefinite nominal is followed by another nominal. This is commonly in a noun + adjective noun phrase, as illustrated by the following example:

fatt aš-šab'ān 'ala l-ǧē'ān fatt^{en} ġūwi
 it is terrible when well-fed people attack hungry people!
 [lit. the attack of the sated person upon the hungry person is a vile attack]

Here we see that the verbal noun *fatt* does not have the adnominal linker when it is in construct but it does have it when it is modified by an adjective. Similar phrases are *bint^{en} naḍīfe* 'a clean girl' and *fāl^{en} zēn* 'a good omen'. Less commonly, both the noun and the adjective have the morpheme, e.g. *'ādat^{en} sāniyyat^{en}* 'an old custom', a feature which may be particular to central (and, here, northern) Arabia but does not appear in the Gulf (Holes 2004b: 95). In poetry, noun phrases consisting of adjective + noun occur, in which the adjective often has the adnominal linker; these phrases may be interpreted as akin to

expressions such as “dark of hair” in poetic English. Examples from the corpus are *šugran* *ḍwāyeb rās-eha* ‘fair of braid’ [lit. fair in respect of the braids of her head] and *bakrat-i ḥelwen* *ḍemil-ah* ‘my camel has a gentle gait’ [lit. my camel is gentle in respect of her gait].

A very frequent variation on this usage, attested also in Najdi (Ingham, *ibid.*) but described as uncommon in Bahrain (Holes 2016: 131) is that of noun + modifying prepositional phrase. In Musil we find *lāgm^{en} min garṣat-eh* ‘a little piece of its [the moon’s] circle’, *fiṅḡāl^{en} ‘ala l-keyf* ‘a cup [of coffee prepared] just how I like it’, *ṣawib^{en} be-l-manām* ‘a wounded person on a couch’ and *dīret^{en} mā bi-ha ḡāze* ‘a land of ill fortune’ [lit. a land in which there is no luck].

As in Najdi, the adnominal linker also occurs when the noun is modified by any modifying clause. With verbal clauses we find *heḡeyniyyet^{en} galladaw-ha l-‘agal* ‘a *heḡeyni* [ditty] which they adorned with a head rope’ and *ḡazāl^{en} yeḡurr aṭ-ṭowb* ‘a gazelle which trails its dress’. We also find it with nominal clauses as in the example *walad^{an} ism-eh mašlān* ‘a boy named Mašlān’. It will be noted that these cases are relative clauses with an indefinite antecedent (cf. §5.2.3), and it appears that the adnominal linker in some ways functions as a relative pronoun; further evidence for this is that the adnominal linker is compulsory in these cases. Interestingly, Musil seems not to have fully understood this, as he often misses the relative function in these phrases. For example, he translates *zeml^{en} ahal-ha teḡill l-eh* as ‘she watches the pack camels of her kin’, an interpretation which ignores the fact that the adnominal linker never occurs on the first noun in a construct phrase (see above), and in

which the preposition *l-eh* is redundant¹²³. It should rather be interpreted as ‘pack camels whose kin [or perhaps owners] watch over them’, where the *l-eh* is an obligatory resumptive pronoun (since the subject of the relative clause is not *zeml*) and the adnominal linker is in a permitted syntactic position.

Ingham states that in Najdi the adnominal linker also occurs “regularly where an active participle is used with a verbal function when not followed by an object suffix” (Ingham, *ibid.*), and this is frequently seen in Musil. Examples are: *yā rāceb^{en} malḥa* ‘O you who are riding a black she-camel’, *hēg^{en} ḥāder^{en} ba‘d r-riyā‘* ‘a male ostrich descending through some defile’ and *ar-rabī‘ dāḥel^{en} be-l-geyḍ* ‘the season of plenty gives way to midsummer’.

There are cases in which the adnominal linker does not have an obviously linking function, as it may appear phrase finally. The first of these, as mentioned in the previous section, is that the second element of a construct noun phrase may take the adnominal linker if it is indefinite (and usually modified). Secondly, it is frequently used in Musil on indefinite nouns which have no adjunct. This is interestingly not attested in Ingham (1994a), although Cantineau (*ibid.*) cites examples from Musil and Montagne and Holes provides examples of this from Bahrain (Holes 2016: 132). Examples are *ḡābat walad^{en}* ‘she gave birth to a child’ and *yā leyt l-i ḡerw^{en}* ‘if only I had a hound!’, and it also occurs on indefinite nouns after prepositions such as *al-gamḥ alli ba‘ad ḥabb^{en}* ‘wheat which is still in

¹²³ Arguably both of these could be possible in poetry if the poet had supplied them simply in an attempt to satisfy the metrical requirements.

grain' and *b-šiniyyet^{en}* 'on a decorated plate'. Holes also includes the collective numericals in this category, e.g. his example *iḥna ḥamsatⁱⁿ* 'there are five of us' (*ibid.*, and see next paragraph). The use of the linker on proper names attested in central Arabia (Ingham 1986: 280) is not a feature of this dialect and seems to be restricted to the central dialects.

In most of the conditions described above the presence of the adnominal linker is optional, but there are several cases in which it is possible that its use is obligatory; these by definition may also occur phrase finally. Firstly, as in most varieties of Arabic, the adnominal linker combines with nouns to form denominal adverbs; examples in the corpus are *ba'īd^{en}* and *ġarīb^{en}* 'far off', *wukād^{en}* 'certainly' and *taw^{en}* 'just now' (although the last two examples are also attested without the linker in the same function). Secondly, a number of denominal conjunctions are attested, such as *gab^{en}* 'before' and *yowm^{en}* 'when' (again, the latter is also attested without the linker). Finally there are a number of denominal pronouns. The indefinite pronoun *šī^{en}* (~ *šay^{en}*) 'something, anything' always has this form, e.g. *mā 'end-ana šī^{en}* 'we have nothing', while its definite counterpart has the form *aš-ši*. The collective pronoun *kill^{en}* 'everyone' and its negative (*mā-*) *ḥad^{en}* 'anyone' always take the linker, as does the number 'one' when it has the meaning of 'someone' (*wāḥed^{en}*, *aḥad^{en}*), and the other numericals when they are collectives, such as *ḥams^{en}* 'five [things]', *elf^{en}* 'a thousand [people]'. A number of formulaic expressions may or may not have the adnominal linker, including *ahl^{an}* (*w-sahl^{an}*) 'hello', and *marhab^{an}* (= *marhaba*) 'welcome'.

3.5 Numerals

Not all the numerals are attested. The near complete list is:

<i>wāhed</i>	n/a	‘one’ (adjectival)
<i>aḥad</i>	f. <i>ḥada</i>	‘one’ (pronominal)
(e) <i>ṭnēn</i>	f. <i>ṭentēn</i>	‘two’
<i>ṭalāt</i>	f. <i>ṭalāta</i>	‘three’
<i>arba‘</i>	f. <i>arba‘a</i>	‘four’
<i>ḥams</i>	f. <i>ḥamse</i>	‘five’
n/a	n/a	‘six’
<i>saba‘ (sab‘)</i>	f. <i>sab‘a</i>	‘seven’
<i>ṭamān</i>	f. <i>ṭamāna</i> ¹²⁴	‘eight’
<i>tisa‘</i>	f. <i>tis‘a</i>	‘nine’
<i>‘ašar (‘ašr)</i>	f. <i>‘ašara (‘ašra)</i>	‘ten’
<i>eḥda‘aš</i>		‘eleven’
<i>eṭna‘aš</i>		‘twelve’
<i>ḥamsta‘aš (ḥansta‘aš)</i>		‘fifteen’
<i>‘ašrīn</i>		‘twenty’
<i>ṭalāṭīn</i>		‘thirty’
<i>arba‘īn</i>		‘forty’
<i>sittīn</i>		‘sixty’
<i>ṭamānīn</i>		‘eighty’
<i>tisa‘īn, tis‘īn</i>		‘ninety’
<i>miyye</i>		‘hundred’
<i>tisa‘ māyāt</i>		‘nine hundred’
<i>elf(en)</i>		‘thousand’
<i>elfēn</i>		‘two thousand’
<i>tis‘at alāf</i>		‘nine thousand’

It should be noted that the forms given here for the number seven are considered taboo and are usually replaced in conversation with the word *samḥ*; the feminine is not attested but is given as *samḥa* by Cantineau (1936-37: 206). Musil writes:

¹²⁴ This form is only attested in the construct state: *ṭamānt*.

“The number seven, *sab*^ʿ, is wholly under the domination of spirits; anyone pronouncing it irritates the evil spirits and repels the good. Colloquially *samḥ* is used instead of *sab*^ʿ, which is employed only in conjuring or swearing, and in cursing to perdition – that is, when calling on the spirits with some object in view.” (Musil 1928a: 390)

Euphemisms of this kind are attested in other dialects, such as *ʿodd yəddək* [lit. ‘count your hand’] for five among the Jewish women of Tunis, and *təsʿūd* (Morocco) or *təsʿad* (Lebanon) [lit. ‘you’ll be happy’] for nine, since *tisʿa* can mean ‘you will beg’ (Taine-Cheikh 2011).

It is interesting that *wahda* is not attested as the feminine form of *wāḥed*, although it is given specifically for Rwala by Cantineau (1936-37: 206). The existence of *tentēn*, however, is very characteristic of this group of dialects, being a feature of the ‘Anazi and Šammari dialects in Cantineau, in contrast to *les petits nomades* who do not use it (*ibid.*).

Ordinals from two onwards have the form CāCeC, e.g. *tāni*, *tālet*, *rābi*^ʿ ‘second, third, fourth’ and so on; *awwal* means ‘first’. Fractions are expressed as CiCC, e.g. *tilt*, *rib*^ʿ ‘a third, a quarter’; *nuṣṣ* or *nuṣf* means ‘a half’. The numbers also express some days of the week, e.g. *aṭ-tnēn* ‘Monday’, *aṭ-talāta* ‘Tuesday’, but Friday is called *ḡum*^{ʿa} and Thursday evening is *leylet al-ḡum*^{ʿa}.

The numerals have different syntactic relationships with the nouns they modify. Number one follows a singular noun, e.g. *ḡidd^{en} wāḥed* ‘one ancestor’, although its use is solely emphatic because singular is already established by the morphology. Two to ten precede a plural noun, e.g. *tnēn šhūd* ‘two witnesses’, *talāt sinīn* ‘three years’, *arba*^ʿ *raḥāyeb*

‘four riding camels’, *ḥamst arnāg* ‘five kinds’ and ‘*ašra leyāl* ‘ten nights’. As can be seen from these examples, there is considerable variability regarding the concord of gender, although the most frequent is the masculine form. The feminine allomorph /at/ rarely occurs, but when it does, it precedes a vowel, e.g. *ḥamst arnāg* above, and *ṭamānt ayyām* ‘eight days’. The phenomenon of the /at/ form only occurring before vowels is attested in other dialects: it is cited by Taine-Cheikh (2011) for Kfar ‘Abīda and Sudan and its absence from the dialect of the Bani Ṣaḥar is remarked on by Palva (1980: 131). There is also written evidence from as early as Middle Arabic of spellings such as *ḥams tiyyām* ‘five days’ (cf. Taine-Cheikh 2011) which indicates that phonologically the -t is felt to be part of the following noun. All higher numbers are invariable, and precede a singular noun, e.g. *eḥda‘ašš šahr* ‘eleven months’, *arba‘in yowm* ‘forty days’, *miyyet ḥāl* ‘a hundred uncles’ and *elf^{en} hala* ‘a thousand welcomes’. Unlike the numbers three to ten, *miyye* always has the form *miyyet* before a noun, and *elf^{en}* always has the adnominal linker.

3.6 Summary

Many of the topics discussed in this chapter are common in a broad sense across the Arabic-speaking world and as such do not help to evaluate the fidelity of Musil’s record of the Rwalā dialect. It is clear, nonetheless, that in terms of nominal morphology the dialect as Musil presented it is very much what would be expected of a Najdi dialect in general. In particular, the use of the adnominal linker is almost completely in line with that described

for Najdi dialects by Ingham (1994a and others). Furthermore, the productive use of the CuCCeyC(a) form for plant names attests a well-established usage which is specific to this area.

The chapter has considered noun and adjective forms with initial consonant clusters which are certainly a hallmark of dialects which have the syncope rules set out in §2.2.3, although the assessment of nominal templates also indicates, as has already been observed, that syncope does not always occur in Musil. The same treatment has shown the occurrence of low vowel raising and backing processes, although again not with great regularity. The absence of *ghawa*-forms is also conspicuous, although it has already been noted that this is more prominent in the verbal forms which will be discussed in the next chapter. This indicates that Musil's material is reliable in terms of the morphology and syntax. One notable absence from Musil's material is that of an analytic genitive, which is attested for other closely related dialects.

The use of different nominal templates is largely a lexical question which it is beyond the scope of this project to compare. The study of the *nisba* forms formed with the relational adjective ending *-i* has also raised an interesting question regarding restrictions on the syllabic structure of the base, for which a comparative study with a wider range of dialects would be desirable. Unfortunately, morphological studies on this matter are not forthcoming for the relevant dialects, meaning that this is another area in which a comparative approach to the lexica of different dialects is required. Nonetheless, there is

nothing in this assessment which would cause one to doubt the reliability of Musil's record of the Rwala dialect from this chapter.

4 Verbal morphology

The verbal systems of Rwala Arabic and closely related dialects have been extensively described by Cantineau (1936-37) and Prochazka (1988). This provides a considerable advantage in addressing the verbs in Musil's corpus, since these two dialectologists have furnished us with complete paradigms and scientific descriptions, meaning that gaps in Musil's material can be filled, and possible errors can be located. Cantineau's Rwala material, which was collected in 1936 (Cantineau 1936-37: 123) and is therefore very nearly contemporary to Musil's, is of particular importance in this regard, but Prochazka's Rwala data is much more complete and therefore many comparisons come from his record of the dialect. Since the two accounts have much in common, they have been collated here, so that the following presentation of the verbal system in Musil (1928a) can be easily compared. Initially, only FORM-I is dealt with, as it is the basis of the verbal system and occurs with the highest frequency: it accounts for just under half of all verbs in a large MSA corpus cited by Holes (2004a: 100-101), and this figure goes up to 70% in Musil. Discussion of the so-called weak verbs and of the biliteral, quadriliteral and derived forms follows the discussion of the FORM-I strong verb.

4.1 The Rwala verbal system in Cantineau and Prochazka

The inflectional suffixes of the suffix stem (S-STEM) and the prefixes and suffixes of the prefix stem (P-STEM) are recorded for the Rwala dialect by Cantineau and Prochazka. The following table lists the variants they recorded:

	S-STEM		P-STEM	
	suffix		prefix	suffix
3MS	(zero)		<i>ya-</i>	(zero)
3FS	<i>-at</i>		<i>ta-</i>	(zero)
2MS	<i>-t</i>		<i>ta-</i>	(zero)
2FS	<i>-ti(y)</i> or <i>-tay</i>		<i>ta-</i>	<i>-in</i>
1CS	<i>-t</i>		<i>a-</i>	(zero)
3MP	<i>-am</i> or <i>-aw</i> or <i>-ow</i>		<i>ya-</i>	<i>-in</i>
3FP	<i>-an</i> or <i>-in</i>		<i>ya-</i>	<i>-an</i> or <i>-in</i>
2MP	<i>-tam</i> or <i>-tow</i> or <i>-tuw</i>		<i>ta-</i>	<i>-in</i>
2FP	<i>-tan</i> or <i>-tin</i>		<i>ta-</i>	<i>-an</i> or <i>-in</i>
1CP	<i>-na</i>		<i>na-</i>	(zero)

Both accounts demonstrate that the S-STEM and the P-STEM have two opposing classes.

Within each stem, both classes have the same prefixes and suffixes but are differentiated by the different verbal stems used. This is illustrated here by the following paradigms from Prochazka, although Cantineau's account shows the same contrast:

S-STEM

	A-CLASS	I-CLASS
3MS	<i>kitab</i>	<i>simi^c</i>
3FS	<i>ktib-at</i>	<i>sam^c-at</i>
2MS	<i>kitab-t</i>	<i>simi^c-t</i>
2FS	<i>kitab-tiy / kitab-tay</i>	<i>simi^c-tiy / simi^c-tay</i>
1CS	<i>kitab-t</i>	<i>simi^c-t</i>
3MP	<i>ktib-am / ktib-aw</i>	<i>sam^c-aw</i>
3FP	<i>ktib-an / kitab-in</i>	<i>sam^c-an / simi^c-in</i>
2MP	<i>kitab-tam / kitab-tuw</i>	<i>simi^c-tam / simi^c-tuw</i>
2FP	<i>kitab-tan / kitab-tin</i>	<i>simi^c-tan / simi^c-tin</i>
1CP	<i>kitab-na</i> 'to read'	<i>simi^c-na</i> 'to hear'

P-STEM

	A-CLASS	I/U-CLASS
3MS	<i>ya-šrab</i>	<i>ya-ktib</i>
3FS	<i>ta-šrab</i>	<i>ta-ktib</i>
2MS	<i>ta-šrab</i>	<i>ta-ktib</i>
2FS	<i>ta-šrub-īn</i>	<i>ta-ktb-īn</i>
1CS	<i>ʔa-šrab</i>	<i>ʔa-ktib</i>
3MP	<i>ya-šrub-ūn</i>	<i>ya-ktb-ūn</i>
3FP	<i>ya-šrab-in</i>	<i>ya-ktib-in</i>
2MP	<i>ta-šrub-ūn</i>	<i>ta-ktb-ūn</i>
2FP	<i>ta-šrab-in</i>	<i>ta-ktib-in</i>
1CP	<i>na-šrab</i> 'to drink'	<i>na-ktib</i> 'to write'

The verbal stems differ between the classes not only in the quality of the stem vowels but also in the syllable structure. Furthermore, in all classes except for the P-STEM A-CLASS there appear to be two different stems recorded, one which (primarily) occurs before a vowel-initial suffix and another before a consonant-initial suffix or a zero suffix. Even if one ignores vowel quality it is possible to differentiate the classes: in the S-STEM both classes have the stem CVCVC before zero and consonantal suffixes but the A-CLASS has CCVC before

a vocalic suffix in contrast to the I-CLASS which has CVCC. Likewise, in the P-STEM the A-CLASS has the stem CCVC throughout the paradigm while the I/U-CLASS has CCVC when unsuffixed (all P-STEM suffixes are vowel-initial) but CCC when suffixed¹²⁵. The following sections will discuss these distinctions and assess the extent to which the same system is attested in Musil's material.

While the four paradigms above demonstrate the regular syllable structure in all four classes, the vowel quality of the stem vowels shows some variation. Table 4.1 shows all the stems listed in Cantineau and Prochazka¹²⁶.

TABLE 4.1: Verbal stems in Cantineau and Prochazka

	S-STEM	P-STEM
	A-CLASS	I/U-CLASS
zero/consonantal:	CaCaC (Cantineau), CVCaC (Prochazka)	CCoC (Cantineau), CCiC, CCuC (Prochazka)
vocalic:	CCöC (Cantineau), CCVC (Prochazka)	CCC
	I-CLASS	A-CLASS
zero/consonantal:	C(v)CeC (Cantineau), CiCiC (Prochazka)	CCaC
vocalic:	CeCC (Cantineau), CaCC (Prochazka)	CCaC (Cantineau), CCuC (Prochazka)

¹²⁵ The only exception to this is that the *-in* ending of the 3FP (which is only attested in Prochazka) takes the same stem as consonantal suffixes, while the *-an* ending (which is attested in both) takes the stem for vocalic suffixes (Prochazka 1988: 28).

¹²⁶ The upper case V in Prochazka's forms represents a vowel that is phonologically conditioned: V is realised as /a/ in the environment of a back fricative /ğ, ħ, ʕ, ʁ, h/ or if followed by /l, r, n/, otherwise it is realised as /i/. I consider Cantineau's CCoC form in the P-STEM I/U-CLASS to be a possible underrepresentation, suspecting that CCiC forms are also possible (in line with Prochazka). Note that his <ö> is a central vowel and (here) an allophone of /a/ rather than /u/.

It may be noted that there is no direct correspondence between the S-STEM classes and the P-STEM classes, although it is predictable to an extent, given that the relationship between *fa'ila* and *yaf'alu* known from CLA is true of verbs in this dialect too, and *fa'ala* verbs with C2 or C3 guttural also tend to be *yaf'alu*.

4.2 S-STEM in Musil

As in the analysis of Cantineau and Prochazka just described, the S-STEM in the Rwala dialect as recorded by Musil consists of a verbal stem which is concatenated with a full set of suffixes. The suffixes are invariable but the verbal stem varies on account of different stem vowels in the underlying representation which yields two classes, as in Cantineau and Prochazka. The stem also varies depending on the suffix, owing to syncope which occurs after suffixation, e.g. *ḍakar*, *ḍkarat* 'he/she remembered' and *širib*, *šarbet*¹²⁷ 'he/she drank'.

4.2.1 Suffixes of the S-STEM

The inflectional endings of the S-STEM are listed, and the forms from Cantineau and Prochazka have been repeated for convenience:

¹²⁷ The vowel alternation between *a* and *i* is addressed below.

	Cantineau/Prochazka	Musil
3MS	(zero)	(zero)
3FS	-at	-at
2MS	-t	-t
2FS	-ti(y) or -tay	-ti
1CS	-t	-t
3MP	-am or -aw or -ow	-ow or -aw or -ū
3FP	-an or -in	-en
2MP	-tam or -tow or -tuw	-tom or -tu
2FP	-tan or -tin	n/a
1CP	-na	-na

There are some disparities between these two columns. Regarding the 3MP *-am* suffix, Cantineau remarked that the masculine plural suffixes ending in *-m* were rare: he specifically states that he didn't hear it for 3MP and only heard it twice for 2MP among the Bani Şaḥar¹²⁸ (Cantineau 1936-37: 186). Prochazka does claim to have encountered the 3MP form *-am* in the Rwala dialect, but here Musil aligns with Cantineau. Commenting on Najdi dialects in general, Ingham remarks that "in all dialects this is an optional variant" (Ingham 1979: 55). Musil's data for the 2MP form show the same distinction between a suffix ending in *m* and a suffix ending in a vowel/diphthong. In Musil, the 2MP *-tu* suffix occurs twice as often as the *-tom* suffix¹²⁹. Cantineau also notes that the diphthongal 3MP suffix *-ow* (also *-aw* in Musil) is an important characteristic of the bedouin dialect group; likewise in Musil the monophthongal variant *-ū* is very rare (8% of forms) except when there is an object pronoun suffixed, in which case it almost always occurs.

¹²⁸ Note however that the *-am* suffix is absent from Palva's description of the dialect of the Bani Şaḥar (Palva 1980).

¹²⁹ Note that in following paradigms, only the more frequent *-tu* form is given.

The 2FS form *-tay* does not occur in Musil, and the 2FP is unattested; we may infer that the form would be **-ten*, by analogy with the 3FP suffix. 3FP forms are not particularly well attested, but since they always take the disyllabic verbal stems which precede a zero or consonantal suffix, rather than the monosyllabic verbal stems which precede a vocalic suffix, we may infer by analogy with Cantineau and Prochazka that Musil's *-en* corresponds to their *-in* suffix rather than their *-an* suffix, and that their *-an* is unattested (more on this below). It is difficult to speculate as to why only one of each two variants is attested in Musil in the cases where Prochazka or Cantineau gave two. Unfortunately these other descriptions do not provide the type of sociolinguistic data which could shed light on this variation.

There are two variations on this paradigm attested (albeit very infrequently) in the corpus. The first is a variant 1CS suffix *-eyt* which is attested once in the corpus, in the phrase *nagleyt-eh* 'a-l-īdi 'I shifted it into [my] other hand'. Variants such as this are common in eastern Arabia and southern Iraq¹³⁰ but have not been recorded, as far as I know, for northern or central Arabia before. Where it occurs, it appears that *-ē-* may be inserted between the verbal stem and any consonant initial suffix, in any S-STEM verb, as opposed to just the final weak and doubled verbs where this type of suffix is widespread (cf. §4.10.2-3). Furthermore, the verbal stem is routinely that of the 3MS (Holes 2016: 185), producing forms such as *nāmēt* 'I slept' rather than *nimt* (Ingham 1982a: 39). The attestation of this

¹³⁰ Cf. Holes (2016: 186) on Bahrain; Johnstone (1967a: 92, 110) on Bahrain and Qatar, Ingham (1982a: 39) on southern Iraq, Khuzistan, and Bahrain), as well as Holes (2016: 33-35) for full references and general discussion.

verbal form in Musil is surprising, given that it is considered to be restricted to the areas already mentioned, and it has been proposed that it is a “survival of a coastal Arabic dialect” (Holes 2016: 36). From just one example in the corpus, however, it is difficult to speculate further on the significance of this occurrence.

The second type of variation is that in eleven instances the S-STEM 3MP suffix appears to be *-ūn/-own* rather than the much more common *-ow/-aw/-ū* ending (this also occurs twice on plural imperatives). Examples are *sayyalū-n-eh* ‘they (m.) caused it to flow’, *za^calū-n-eh* ‘they (m.) made him angry’ and *tahāwašow-n* ‘they (m.) fought each other’. In the first two examples (and nine examples in total) this occurs before a vowel-initial suffix pronoun, which differs from the regular form in the corpus which would be **sayyalū-h*, **za^calū-h*, while the third example (one of two in total) suggests that this can also occur without suffixation. We may speculate that the origin of this formation could have been analogy with the *-ūn* suffix of the 2MP and 3MP P-STEM, although the data do not provide enough evidence to allow for further speculation as to whether this arose internally or via contact, or as to the contexts in which such variation occurred. It may be noted that similar variation is attested in other varieties of Arabic (although no historical link between these occurrences is claimed here). For example, in Jewish and Christian Baghdadi Arabic an intervening *-n-* is recorded between vocalic inflectional endings and vowel-initial suffix pronouns: this is attested for the imperative in Jewish *ketbō-n-u* and Christian *ketbū-n-u* ‘write (pl.) it!’ (Blanc 1964: 101). Furthermore, *-nu* is recorded as the post-vocalic 3MS direct object pronoun for

suffixation on verbs (as opposed to *-u* following a consonant) in Christian Baghdadi (Abu-Haidar 1991: 80). These cases support the hypothesis that this process is analogical, since in Baghdadi the *-n-*, which may have originally been phonologically motivated, has been morphologised and may also occur in nominal and prepositional constructions such as Jewish and Christian *abū-n-u* ‘his father’ and *bī-n-u* ‘with him’ (Blanc *ibid.*: 66). That forms such as these were recorded by Musil on at least two separate occasions shows that this was not an isolated occurrence, but these examples are insufficient to shed further light on the variational context of this phenomenon.

4.2.2 *Classes of the S-STEM*

The opposition of the A-CLASS and I-CLASS as described above can be seen to apply in Musil’s Rwala data. These classes are rather difficult to extract from data such as Musil’s, where some forms are completely unattested, and it is very rare that a single verb is attested in more than a few inflected forms. As such, there are a number of verbs which cannot be clearly assigned to either class. The morphosemantic and structural differences between the classes are discussed in the sections on vowel incidence and stem syllabification below.

4.2.2.1 The A-CLASS

The stem is CaCaC before zero or consonantal endings. When a vocalic ending produces the sequence CaCaC + V, /a/-syncope may occur (cf. §2.2.3.1), since this syncope

rule targets sequences of low vowels in open syllables. This produces alternation between the unsyncopated stem CaCaC (e.g. *ḍakar/ḍakart* ‘he/I remembered’) and the syncopated stem CCaC (e.g. *ḍkarat* ‘she remembered’), for which reason it is proposed that the underlying representation of the stem is /CaCaC/. In Musil’s data, syncope occurs in less than half of attested forms, and the unsyncopated stem CaCaC also occurs before vocalic endings; this differs from the reports of Cantineau and Prochazka, who imply that /a/-syncope always occurs (see below, §4.2.4). This is in line with the rest of Musil’s material, however, in which syncope is less frequently attested than in these other accounts. The paradigm for *ḍakar* ‘he remembered, spoke of sth.’ is therefore:

3MS	<i>ḍakar</i>	3MP	<i>ḍakarow, ḍkarow</i>
3FS	<i>ḍakarat, ḍkarat</i>	3FP	<i>ḍakaren (ḍkaren)</i>
2MS	<i>ḍakart</i>	2MP	<i>ḍakartu</i>
2FS	<i>(ḍakarti)</i>	2FP	<i>(ḍakarten)</i>
1CS	<i>ḍakart</i>	1CP	<i>ḍakarna</i>

Forms in brackets are predicted to occur based on the other forms in the paradigm, but are not actually attested in the corpus. Examples of each of these forms as attested in the corpus are given:

3MS	<i>ḍakar</i> ‘he remembered’	3MP	<i>ḍabaḥow</i> ‘they killed’ <i>ksabow</i> ‘they captured’
3FS	<i>sanadat</i> ‘she ascended’ <i>bradat</i> ‘it became cool’	3FP	<i>ḍakaren</i> ‘they (f.) remembered’
2MS	<i>ḍakart</i> ‘you remembered’	2MP	<i>nagaltom</i> ‘you (pl.) rendered’
2FS	n/a	2FP	n/a
1CS	<i>naḍart</i> ‘I made a vow’	1CP	<i>sanadna</i> ‘we ascended’

Variations of the stem are only attested with *e*: CeCaC, CaCeC and CeCeC for CaCaC, i.e. the shift $a > e$ may occur in either or both positions in CaCaC. This produces forms which are sometimes difficult to distinguish from the I-CLASS.

It can be seen from the above paradigm that syncope of CaCaC > CCaC before vocalic suffixes is attested for the 3FS and 3MP forms but not for the 3FP form. This could simply be a coincidence, since syncope does not always occur in Musil, and 3FP forms are generally not well attested in the corpus. It is interesting to note, however, that Prochazka (1988: 28-29) recorded two 3FP suffixes for the Rwala dialect, one (-*in*) which does not trigger syncope and another (-*an*) which does: compare *ḍibaḥin/ḍbaḥan* ‘they (FP) killed’. It is therefore also possible that Musil’s records parallel the -*in* form in Prochazka rather than the -*an* form. The case of the 3FP form in Prochazka is curious, since the variation he recorded is unique to the Rwala dialect: the other northern dialects he recorded only have the *ktiban* form, as do those in Cantineau (1936-37: 79) and Ingham (1994a: 23), and the southern dialects in Prochazka have a different form altogether, *katabnah* (Prochazka 1988: 27). Unfortunately Prochazka does not offer comment on this variation, and it is difficult to speculate as to why he came to hear two variants from his Rweyli informants only.

4.2.2.2 The I-CLASS

The stem is CiCeC before zero or consonantal endings and CaCC before vocalic endings: compare *šireb/širebt* ‘he/I drank’ and *šarbat* ‘she drank’. It is proposed that the

underlying representation is /CaCiC/, and that the alternation between *a* and *i* in the first stem vowel is explained by raising (cf. §2.3.6.1) while the alternation between *e* and zero in the second stem vowel is explained by /i, u/-syncope (cf. §2.2.3.1)—this derivation is demonstrated below in §4.2.4. CCaC may also occur before vocalic suffixes through levelling with the A-CLASS: this variation was also recorded for the Rwala dialect by Prochazka (1988: 30-31) who gives the same explanation. The forms are not all attested. The paradigm for *šireb* ‘he drank’ is therefore:

3MS	<i>šireb</i>	3MP	<i>šarbow, šrabow</i>
3FS	<i>šarbat, (šrabat)</i>	3FP	<i>(šireben)</i>
2MS	<i>(širebt)</i>	2MP	<i>(širebtu)</i>
2FS	<i>(širebti)</i>	2FP	<i>(širebten)</i>
1CS	<i>širebt</i>	1CP	<i>(širebna)</i>

Examples attested in the corpus:

3MS	<i>šireb</i> ‘he drank’	3MP	<i>labsow</i> ‘they wore’ <i>gbalaw</i> ‘they accepted’
3FS	<i>šarbet</i> ‘she drank’, n/a	3FP	n/a
2MS	n/a	2MP	n/a
2FS	n/a	2FP	n/a
1CS	<i>širebt</i> ‘I drank’	1CP	n/a

Variation in the stem is encountered in the first vowel, where sometimes raising does not occur and the resulting stem is CaCeC/CaCiC, e.g. *labes* ‘he wore’, *wariden* ‘they (f.) came to water’ and *račebna* ‘we mounted’. Sometimes the indeterminate vowel *e* occurs, where it is not clear whether this vowel has been raised or not; this produces the stem CeCeC and forms such as *selem* ‘he saved himself’. In theory, the discussion of the 3FP form

above applies to the I-CLASS too: Prochazka records *simiʿin* and *samʿan* ‘they (f.) heard’ so we would predict **šireben* rather than **šarben* in line with *ḍakaren* in the A-CLASS, but it is unattested in Musil.

4.2.3 *Stem vowel incidence of the S-STEM*

The incidence of different stem vowels is a matter of interest in the typology of Arabic dialects. Modern vernaculars have developed different systems of verb classification, which tend to be simplified *vis-à-vis* the more specified CLA system. These changes have been motivated by phonological developments, e.g. the tendency for /i/ and /u/ to become nondistinctive, and the tendency for vowel quality to be affected by the surrounding consonants. They have also been morphosemantic, e.g. the loss of stem types and derived forms (Holes 2004a: 126).

Examples of both motivations can be observed in Cairene Arabic. In terms of phonological development, the S-STEM exhibits vowel harmony resulting in the three forms CaCaC, CiCiC and CuCuC, and furthermore the forms CiCiC and CuCuC have (with some exceptions) become interchangeable, creating a height contrast between /a/ stems and /i, u/ stems. In terms of morphosemantic development, the /a/ stems are mainly transitive and the /i, u/ stems are mainly intransitive, and roots are differently distributed compared to the CLA system, resulting in a system which marks transitivity more or less predictably. By contrast, the ʿArab dialect of Bahrain has no correspondence between S-STEM theme vowel

and verb type, since the theme vowel is entirely phonologically conditioned: the form is CVCaC, where V is /a/ in the environment of a guttural consonant, or when followed by /l, r, n/, and otherwise /i/ or /u/, which have partially merged so that /i/ occurs unless in the environment of a labial consonant (examples from Holes 2004a: 126-130).

Prochazka's account of the Rwala system shows phonological considerations in the A-CLASS only. He describes the verbal stems as CVCaC and CCVC, where V is phonologically conditioned: V is realised as /a/ in the environment of a back fricative /ħ, ǧ, ḥ, ʕ, h/ or if followed by /l, r, n/, otherwise it is realised as /i/ (cf. the /a/-raising rule in §2.3.6.1). This effectively produces four paradigms for the A-CLASS, depending on the position of back fricatives, liquids or /n/:

	3MS	3FS	gloss
i-i	<i>kitab</i>	<i>ktibat</i>	'he/she wrote'
i-a	<i>ḏibaḥ</i>	<i>ḏbaḥat</i>	'he/she killed'
a-i	<i>ḥalab</i>	<i>ḥlibat</i>	'he/she milked'
a-a	<i>gaʕad</i>	<i>gʕadat</i>	'he/she sat'

On account of /a/-raising not occurring regularly in Musil, the high vowels do not feature in the paradigm according to his Rwala data. Accordingly, /a/ is always the stem vowel for the A-CLASS; even if the occurrence of *e* in the A-CLASS were analysed as evidence of raising, the phonological distribution described by Prochazka is not applicable.

In terms of morphosemantic adjustments, Ingham asserts that in Najdi dialects the A-CLASS is more or less correlated to transitive verbs and the I-CLASS to intransitive verbs, rather like the Cairene example above. This terminology is perhaps misleading here,

however, as Ingham is referring to something much broader and admits that “it may be considered odd to treat [multiple] different types of feature under the heading of ‘transitivity’” (Ingham 1994a: 70). Additionally, the difficulty encountered in separating out the A-CLASS and the I-CLASS from Musil’s data makes it hard to say whether his material agrees with this statement. For the purposes of discussion here, therefore, it will have to suffice to acknowledge that the structures exist in the form of two distinct classes.

4.2.4 *Stem syllabification*

The morphology of the S-STEM has been presented as requiring two differently syllabified verbal stems, one before zero and consonantal suffixes and another before vocalic suffixes, in both the A-CLASS and the I-CLASS. Such a situation is characteristic of many Arabian dialects, and the above paradigms concur with the data from Cantineau, Prochazka, and Ingham for example, in contradistinction to other dialects of Arabic (including CLA) in which the verbal stem remains unchanged throughout the paradigm: we may compare Rwala *barak*, *brakat* ‘he/she knelt’ with CLA *baraka*, *barakat*. The disparity between the resyllabification of CaCaC to CCaC in the A-CLASS but CiCiC to CaCC in the I-CLASS is explained in terms of the syncope and raising rules of the language.

The case of the A-CLASS is simply one of the application of /a/-syncope as described in §2.2.3.2: the stem has the underlying form /CaCaC/ and syncope renders it CCaC before vocalic suffixes. Its partial absence from Musil (shown by forms such as *ḍakarow* instead of

dkarow) is simply due to the fact that /a/-syncope is not represented as frequently in Musil as it is in other sources of the dialect.

The I-CLASS is a more complex and in some ways more interesting case. It was proposed in §4.2.2.2 that the stem has the underlying form /CaCiC/, which accounts for the vowel alternation in the stem before zero/consonantal suffixes (*širib-*) and before vocalic suffixes (*šarb-*)¹³¹. The first stem vowel alternates between *a* and *i*, and the second vowel alternates between *i* and zero. These alternations are accounted for through the application of both /i, u/-syncope and raising¹³², which are the only productive rules in the dialect which produce these types of alternation. This is shown in the following derivation:

UR	/šarib/	/šarib + at/	/šarib + t/
/i, u/-syncope	————	šarbat	————
raising	širib	————	širibt
output	<i>širib</i>	<i>šarbat</i>	<i>širibt</i>
	‘he drank’	‘she drank’	‘I drank’

This derivation applies, even though the usual surface form of these verbs in Musil is

šireb, šarbat, širebt. The suffixation of object pronouns produce the same results:

UR	/šarib + at/	/lamis + ah/
/i, u/-syncope	šarbat	lamsah
raising	————	————
output	<i>šarbat</i>	<i>lams-ah</i>
	‘she drank’	‘he touched it’

¹³¹ This proposed underlying representation appears to coincide with the CLA stem /CaCiC/, but in the analysis here it is based on the alternations shown.

¹³² Here the neutral term raising is used since it is not clear whether the raising exhibited in the I-CLASS, which in some accounts is not blocked by neighbouring gutturals or following /l, r, n/, is sufficiently different from what we have termed /a/-raising (which is blocked in those consonantal environments) to be considered a separate rule.

In Cantineau’s data, /i/-syncope may target the first syllable as well. Since syncope needs to be ordered before raising to produce the form *šarbat*, this may be indicative of a cycle producing the following variants:

UR	/rakib/	/rakib/
/i, u/-syncope 1	————	————
raising 1	rikib	rikib
/i, u/-syncope 2	rkib	*****
raising 2	————	————
fronting	rícib	rićib
output	rícib	rićib ¹³³
	‘he mounted’	‘he mounted’

Here the notation ***** is indicative of the non-application of a rule. An alternative analysis would be that this is a case where /i, u/-syncope targets the low vowel /a/ which, as we have seen (cf. §2.2.3.1), is attested in Musil (but not in Cantineau). Cantineau claims that “la première voyelle ... peut se maintenir si elle est encore sentie comme une voyelle a ... elle peut tomber si, par assimilation, elle est passée au timbre *i*” (Cantineau 1936-37: 50-51). Since his assumption was that the dialectal forms he describes are ultimately derived from something resembling the CLA *qatila* forms, which happens to correspond with my proposed underlying representation, this interpretation is rather close to the proposed raising and /i, u/-syncope rules set out in this work.

Palva interprets the I-CLASS forms as representing a historical rather than a synchronic process. He argues that in the dialects of the Bani Şaḥar (Palva 1980: 124) and of

¹³³ As with Musil’s examples in the derivation above, Cantineau’s surface forms actually have *e*: *récib* and *rećeb* (Cantineau 1936-37: 79).

the Ḥwēṭāt (Palva 1984-6: 299) a historical development /šarib/ > /širib/ accounts for the variation between *širib* and *šarbat* and that the different forms exist on account of the historical ordering of this change and a hypothetical historical point at which syncope was fixed. He explains the different forms attested in other dialects as various degrees of levelling, and compares hybrid systems (exemplified by the dialects he writes about, which have paradigms similar to the Rwala one) with levelled ones, pointing out that the sedentary dialects of the area have all levelled the paradigm by analogy to the CiCiC stem, with the exception of the dialect of al-Karak which has CaCiC throughout. I have preferred to interpret this data as indicating a set of phonological rules rather than a historical change, since the case for a dialect simultaneously accepting anachronistic verbal stems in the same paradigm has not yet been made convincingly. I agree, however, with his assessment that, in the sedentary dialects, *širbat* (rather than *šarbat*) probably occurs through levelling and I would suggest (not having looked closely at his data, however) that in al-Karak, rather than having bypassed his hypothesised historical development, it is more likely that raising is not a feature of that dialect (as discussed below).

There is huge variation in the I-CLASS paradigm throughout north Arabia. While it is beyond the scope of this work to attempt to address this for multiple dialects, the following adjustments to the system proposed here may indicate how the situation came to be so

varied across this part of the Peninsula¹³⁴. The following example paradigms may be compared¹³⁵:

	Rwala	Najrān	Sudayr	al-Karak
3MS	<i>šireb</i>	<i>šmi^ʕ</i> (sic)	<i>sam^ʕ</i>	<i>šarib</i>
3FS	<i>šarbat</i>	<i>sim^ʕat</i>	<i>sam^ʕat</i>	<i>šarbat</i>
2MS	<i>širebt</i>	<i>šmi^ʕt</i> (sic)	<i>simi^ʕta</i>	<i>šaribt</i>

The large scale variation in the I-CLASS shown here represents some of the diversity among north- and central-Arabian dialects¹³⁶. It appears that this variation can be explained in terms of dialect-specific rule ordering:

(1) 'he drank/heard'	Rwala	Najrān ¹³⁷	Sudayr	al-Karak
UR	/šarib/	/sami ^ʕ /	/sami ^ʕ + Ø/	/šarib/
/i, u/-syncope	————	*****	sam ^ʕ Ø	————
raising	širib	simi ^ʕ	————	*****
/i, u/-syncope	*****	smi ^ʕ	*****	*****
output	širib	smi ^ʕ	sam ^ʕ	šarib

¹³⁴ This analysis is presented particularly in response to Parkinson (1992), where a similar system is proposed. It is unfortunate that Parkinson did not engage with any of the contradictory examples from other dialects, since he derived all of his Rwala data from Prochazka (1988), a work which provides a huge corpus of evidence of different systems in other north- and central-Arabian dialects.

¹³⁵ Najrān examples from Prochazka (1988: 31) (these are similar to those discussed by Cantineau); Sudayri examples from Ingham (1994a: 24); and al-Karak examples from the description in Palva (1980: 124 note 42).

¹³⁶ Najrān, while being geographically in the south of the Peninsula, is analysed typologically as a Najdi dialect by Prochazka (1988: 11).

¹³⁷ Here the š has been removed from the paradigm for clarity.

(2) 'she drank/heard'

UR	/šarib + at/	/sami ^c + at/	/sami ^c + at/	/šarib + at/
/i, u/-syncope	šarbat	*****	sam ^c at	šarbat
raising	————	simi ^c at	————	*****
/i, u/-syncope	*****	sim ^c at	*****	*****
output	šarbat	sim ^c at	sam ^c at	šarbat

(3) 'I drank/heard'

UR	/šarib + t/	/sami ^c + t/	/sami ^c + ta/	/šarib + t/
/i, u/-syncope	————	*****	————	————
raising	širibt	simi ^c t	simi ^c ta	*****
/i, u/-syncope	*****	smi ^c t	*****	*****
output	širibt	smi ^c t	simi ^c ta	šaribt

In these examples, it can be seen that the dialect of al-Karak only has the rule of /i, u/-syncope, while that of Najrān orders raising before /i-u/-syncope in order to produce the correct output. Rwala and Sudayri have the same rules in the same order, and I have preliminarily analysed the difference in output in the first example as being on account of a rule in Sudayri which considers the zero suffix to pattern like the vocalic suffixes rather than like the consonantal suffixes as in the Rwala dialect; further speculation on this matter is beyond the scope of the present work, however. Although this is no more than a preliminary sketch, I am persuaded that the above derivations demonstrate that the variation found in this dialect group may be satisfactorily explained through different applications of the phonological processes already discussed here.

In Musil, the contrast between /a/-syncope and /i, u/-syncope sufficiently distinguishes the two classes. The lack of regular raising in Musil, however, means that 1-

CLASS forms such as *šireb* (which shows raising) occur alongside forms without raising, such as *wariden* ‘they (f.) came to water’ and *raćebna* ‘we mounted’.

There is one further type of variation encountered in Musil’s I-CLASS, which is that verbs in this class often seem to pattern like A-CLASS verbs. This is not unprecedented, however, since Prochazka observed that “the stem C1aC2C3- in [Rweyli, Ḥāyil, Al-Qašim and Riyadh] freely varies with the stem C1C2VC3-, borrowed from the C1VC2aC3 conjugation [i.e. the A-CLASS]” (Prochazka 1988: 30-31). This may be explained more simply in terms of analogical levelling, by which the I-CLASS, as the less productive of the two classes, is levelled with the more productive A-CLASS. Such a hypothesis is supported statistically, since the I-CLASS is much less attested in Musil than the A-CLASS. Examples in Musil include the variation of *labsow* ~ *lbasow* ‘they wore’ and the occurrence of *gbalow* ‘they accepted’ rather than **gablow*.

4.3 P-STEM in Musil

Like the S-STEM, the P-STEM is formed from a verbal stem and a fixed set of affixes, which in this case are a full set of prefixes and a partial set of suffixes. As in the S-STEM, in the P-STEM the affixes are invariable but the verbal stem varies depending on the nature of the suffix attached and because of lexically encoded vowels. The evidence for the P-STEM in Musil is outwardly different to the systems as set out by Cantineau and Prochazka, however, in that the syncope rules do not create two different classes; while they both record the

existence of two classes, Musil's data is most easily expressed in one, and the reasons for this are outlined below.

4.3.1 *Prefixes of the P-STEM*

In Prochazka (1988: 32-37) the vowel of the P-STEM prefix is /a/. Cantineau (1936-37: 187) also gives /a/, and notes that *a* is found in all the dialects studied for his 1937 article¹³⁸, at which point he specifically disagrees with Musil's claim that *a* in the prefix is unique to the Šammari dialects (cf. Musil 1928b: 19).

That Musil considered the *a* vowel in the P-STEM prefix to be a Šammari feature, to the exclusion of the ʿAnazi dialects including that of the Rwala, is not borne out in his Rwala data, in which *a* features prominently. Table 4.2 lists the occurrences of each possible prefix.

TABLE 4.2: Frequency of P-STEM prefix vowels

	y		n		t		zero		total
a	ya-	54	na-	23	ta-	42	a-	34	153
e	ye-	94	ne-	7	te-	74	—		175
i	yi-	60	ni-	1	ti	5	—		66
zero	y-	2	n-	1	t-	3	—		6

There are two difficulties here. The first is that some of these forms may belong in FORM-IV and not FORM-I, since these are outwardly indistinct in the P-STEM (cf. §4.8.1.3) and verbs that are only attested in the P-STEM are very difficult to assign to one form or the

¹³⁸ These are: Bani Ḥāled, Weld ʿAli, Rwala, Sardiye, Bani Şaḥar, Ḥsane, Sbaʿa, Mawāli, Šammar, and the sedentary population of Ğōf.

other, except on the basis of comparative evidence. Prochazka claims that the prefix *yi-* is exclusively a FORM-IV marker; this generalisation does not hold for the data from Musil, however, and the forms here are those which are either certainly or probably FORM-I, with some possible errors in the latter category where neither internal nor external evidence is able to illuminate the matter.

The second difficulty is in discerning whether the *e* variants represent an underlying /a/ or /i/. Given the evidence from both Cantineau and Prochazka of an /a/ prefix vowel in Rwala, it seems likely that the *e* forms are variants of /a/. This is supported by the data for *i* variants, which show that forms with *i* are virtually unattested except after *y*; that is, they almost only occur in the 3MS, 3MP and 3FP forms. Although the literature makes no account of this for Rwala, such a situation is attested in other dialects, e.g. in the dialect of N^čēm, which has *yirgab*, *yirgabūn* in the third person but *tergab*, *tergabūn* in the second person (Cantineau 1936-37: 82). Such a situation is phonologically unsurprising since /y/ and /i/ are featurally similar. The *yi-* form in Rwala is not morphologically motivated, since it does not correlate with the verbal stem, and does not appear to be phonologically motivated either, since it occurs in the same consonantal environments as the *ya-* and *ye-* forms. The exception is that it does not seem to be attested before gutturals, though it is before emphatics. It also occurs regularly in closed syllables (since the syllabic structure of the P-STEM is CVC.CVC), so does not seem to be a case of /a/-raising, and (if /i/-spreading is taken to have Parkinson's [1992] definition) it occurs very infrequently before a following

/i/, so is unlikely to be /i/-spreading either. We may posit therefore that *yi-* varies freely with *ya-* and *ye-* unless it is blocked by a guttural.

The table above also demonstrates that while the *e* variant is almost twice as common as the *a* variant after *y* and *t*, this trend is not evidenced after *n*. 1CP forms provide little evidence, but it is a general feature of the dialect that *na* surfaces twice as often as *ne* in the entire corpus, whereas *te* is slightly more frequent than *ta*, and *ye* is much more frequent than *ya*. It is difficult to speculate on the reasons for this since, as has already been discussed in §2, the variation between *a* and *e* is very uncertain. The non-vocalic prefixes *y-*, *t-* and *n-* only occur as a result of resyllabification when C1 is a guttural (cf. §4.9.1).

The inflectional prefix is therefore:

3MS	<i>ya-, ye-, yi-</i>	3MP	<i>ya-, ye-, yi-</i>
3FS	<i>ta-, te-</i>	3FP	<i>ya-, ye-, yi-</i>
2MS	<i>ta-, te-</i>	2MP	<i>ta-, te-</i>
2FS	<i>ta-, te-</i>	2FP	n/a
1CS	<i>a-</i>	1CP	<i>na-, (ne-)</i>

The 2FP is not attested in Musil but evidence from other dialects shows that it should have the same prefix as 2MP.

The fact of *yi-* occurring in FORM-I is a matter of interest, not only because it does not concur with previous descriptions of the dialect, but because of the implications for the verbal system in general. We have already seen that the contrast of *ya-* and *yi-* in the P-STEM has been recorded as the fundamental differentiator between FORM-I and FORM-IV; furthermore it serves to differentiate the active from the passive in FORM-I, according to

Prochazka and supported by the evidence from Ingham (1994a: 27) and Abboud (1964: 18). Since, however, Musil elsewhere commented (albeit incorrectly) on the P-STEM prefix vowel in the Šammari dialect (Musil 1928b: 19) being different to that of the Rwala, it would be surprising indeed if he should have recorded it incorrectly in his Rwala material. Although the Rwala dialect is recorded as having a fixed /a/ in the active FORM-I prefix, some dialects related to Rwala do have /i/ as their fixed prefix vowel, such as the dialects of Sudayr and Riyadh as recorded by Prochazka, and the bedouin dialect of the Ghāmid much further south: the form is not problematic in itself, however, but because of its variation with /a/. There are enough examples, however, to force us to confront the fact that Musil's data presents us with a picture of the dialect in which *yigtel*, *yegtel* and *yagtel* 'he kills' are all attested variants of the FORM-I P-STEM.

4.3.2 *Suffixes of the P-STEM*

The inflectional endings of the P-STEM are:

3MS	(zero)	3MP	- <i>ūn</i>
3FS	(zero)	3FP	- <i>en</i> or - <i>an</i>
2MS	(zero)	2MP	- <i>ūn</i>
2FS	- <i>i</i> or - <i>īn</i>	2FP	n/a
1CS	(zero)	1CP	(zero)

In Musil's corpus 2FS forms are found, in equal numbers, with the suffixes -*i* and -*īn*, e.g. *ta'nezi* 'you (f.) go astray' but *taz'alīn* 'you (f.) are angry'. Cantineau and Prochazka both record only the -*īn* variant. Rosenhouse distinguishes three groups among bedouin dialects

based on their 3MP and 2FS suffixes in the P-STEM: those with *-ūn* and *-īn*, those with *-um* and *-i*, and those with *-u(w)* and *-i(y)* (Rosenhouse 1984: 27-28). Since *Rwala* has *-ūn*, *-īn* and *-i*, it seems that it belongs in the first group, which spreads over north Arabia, eastern Arabia and some dialects of the north of Israel. This analysis makes the *-i* variant even more curious, but the corpus indicates that both forms are in completely free variation. The 2FP form is not attested in *Musil* but evidence from other dialects shows that it should have the same suffix as 3FP.

Interestingly, Cantineau and Prochazka differ with regards to the feminine plural suffixes, the former having only *-an* and the latter having only *-in*. The data from *Musil* shows nine tokens of *-an* and twenty-six tokens of *-en*, but it is unclear whether the more common *-en* variant represents a *-in* form like that in Prochazka or is merely an variant form of *-an*. Similarly to the situation with the 3FP suffix in the S-STEM, Prochazka's suffix *-in* does not trigger syncope of the stem (*yaktbūn* but *yaktibin*) while Cantineau's suffix *-an* does trigger syncope (*yik^ōtbūn* and *yik^ōtban*¹³⁹). As we will see, however, in *Musil* it is not clear whether the P-STEM resyllabifies at all, so it is impossible to tell which of the forms he recorded.

¹³⁹ While Cantineau asserts that this paradigm (examples given from N^ēm) is adequate to represent all the dialects in his study, it seems clear from the position of epenthesis here that the N^ēm dialect is in fact typologically different, since it apparently represents a rhyme dialect. I assume that his claim of similarity between these dialects is based on similar stem-altering syncope patterns and that the position of epenthesis was not relevant to him in this particular matter.

4.3.3 *Classes of the P-STEM*

Cantineau characterises the Rwala dialect as having two P-STEM “conjugations”, one for verbs with *a* in the stem and one for verbs with *e* or *o*. The reason for the grouping is the different syllable structure that occurs in forms which have P-STEM suffixes, namely, 2FS, 3MP, 2MP, 3FP, and 2FP. His A-CLASS follows the pattern *yašrab*, *yašrabūn*, with no change in the syllable structure of the verbal stem after suffixation (since /a/-syncope cannot apply and /i, u/-syncope is not predicted to target /a/), while his E/O-CLASS follows the pattern *yaktob*, *yaktbūn*, which shows /i, u/-syncope. Prochazka is in agreement, although he records an I/U-CLASS rather than an E/O-CLASS, he excludes the feminine plurals from the differently syllabified group, and he records raising and backing in his examples (in accordance with all of his examples), thus giving *yašrubūn*.

It is difficult to say whether the data from Musil supports this division. His paradigm is identical in all “classes”, which all have *e* before a suffix:

/a/	/i/	/u/
<i>yagbal</i> , <i>yagbelūn</i>	<i>yanšed</i> , <i>yanšedūn</i>	<i>yaskon</i> , <i>yaskenūn</i>
‘he/they accept(s)’	‘he/they ask(s)’	‘he/they live(s)’

The syncopated form in the E/O-CLASS (i.e. *yaktbūn*) is therefore entirely absent from the transcription, but there are reasons why this should be so. Firstly, out of the strong verbs in which no root consonant is guttural (any other verb would block the resyllabification in various ways), there are only ten 2FS, 3MP, and 2MP forms attested, of which at least five are in the A-CLASS anyway. There are therefore hardly enough data to prove the presence or

absence of such a class. Secondly, as has been addressed in §2.2.3.4, Musil does not represent syncope in the majority of cases where it might be expected. In §2.2.3.4, three interpretations of this paradigm were proposed. Firstly, the *e* could represent a super-short vowel (*yagbelūn*, *yanš^edūn*, *yaskenūn*), as was alluded to by Cantineau with his reference to these forms having a “point vocalique” (Cantineau 1936-37: 186). Secondly, the vowel could be deleted by syncope and then reinserted by epenthesis; this could seem to apply to *yagbelūn* (< **yagbalūn*) and *yaskenūn* (< **yaskonūn*) since there is a change in vowel quality, while the same process in *yanš^edūn* would not be observable. And lastly, Musil could simply have not perceived, or otherwise not recorded, the syncope in these forms. Comparing them with the data from Cantineau and Prochazka, it is also possible that two of these interpretations could apply separately; for example, *yaskenūn* (< **yaskonūn*) could have the super-short vowel because of /i, u/-syncope while the *e* in *yagbelūn* (< **yagbalūn*) represents a full undeleted vowel.

The arrangement of the data below in a single paradigm, albeit with columns for the lexically governed vowel which appears in some forms, belies the fact that in Musil’s data the forms are transcribed the same way. If a contrast exists between the full *e* vowel of *yagbal*, *yagbelūn* and the reduced *e* vowel in *yaskon*, *yask^enūn*, the transcriptions do not show it. Likewise, it may be that the unattested feminine plural forms for the “A-CLASS” should retain an *a* rather than *e*; this would provide enough evidence for a two-way class distinction as in Cantineau and Prochazka, but it is also unattested.

4.3.3.1 The paradigm

The stem is CCVC, where V is either *a*, *e* or *o*. The nature of this vowel is lexically governed by a vowel phoneme /a/, /i/ or /u/ which is associated with the stem at an underlying level; on the surface, their representations are the following:

/a/	>	<i>a</i> or <i>e</i>
/i/	>	<i>e</i>
/u/	>	<i>o</i>

When suffixes are added the stem is always CCeC. The paradigm is therefore:

3MS	yaCCVC	3MP	yaCCeCūn
3FS	taCCVC	3FP	yaCCeCen
2MS	taCCVC	2MP	taCCeCūn
2FS	(taCCeCīn)	2FP	(taCCeCen)
1CS	aCCVC	1CP	naCCVC

Example paradigm:

	/a/	/i/	/u/
3MS	<i>yagbal</i>	<i>yanšed</i>	<i>yaskon</i>
3FS	<i>tagbal</i>	<i>tanšed</i>	<i>taskon</i>
2MS	n/a	<i>tanšed</i>	<i>taskon</i>
2FS	n/a	n/a	n/a
1CS	<i>agbal</i>	<i>anšed</i>	<i>askon</i>
3MP	<i>yagbelūn</i>	<i>yanšedūn</i>	<i>yaskenūn</i>
3FP	n/a	<i>yanšeden</i>	<i>yaskenen</i>
2MP	<i>tagbelūn</i>	<i>tanšedūn</i>	n/a
2FP	n/a	n/a	n/a
1CP	<i>nagbal</i>	<i>nanšed</i>	<i>naskon</i>
	‘receive’	‘ask’	‘live’

Examples attested in the corpus¹⁴⁰:

3MS	<i>yašbaṭ</i> ‘he climbs’	<i>yaʿseb</i> ‘he brings booty’	<i>yaktob</i> ‘he writes’
3FS	<i>tagbal</i> ‘she lets in’	<i>tasned</i> ‘she goes up’	<i>tatmor</i> ‘she rubs dough [into leather for tanning]’
2MS	n/a	<i>taḍreb</i> ‘you beat’	<i>tarzog</i> ‘you reward’
2FS	n/a	n/a	n/a
1CS	<i>arkab</i> ‘I mount’	<i>anšed</i> ‘I ask’	<i>atrok</i> ‘I desert’
3MP	<i>yarʿebūn</i>	<i>yenšedūn</i>	<i>yeġtelūn</i> ‘they kill’
3FP	n/a	<i>yenšeden</i>	<i>yagtelen</i>
2MP	<i>terʿebūn</i>	<i>tanšedūn</i>	n/a
2FP	n/a	n/a	n/a
1CP	<i>narkab</i>	<i>nanzel</i> ‘we set up camp’	<i>nasgof</i> ‘we pitch a tent’

Variation of the stem is limited to verbs with underlying /a/, i.e. the first column, between CCaC and CCeC. This produces forms such as *yalbes* and *albes* for **yalbas* and **albas*.

As described above, variation of the prefix allows for *ya-*, *ta-* and *na-* to become *ye-*, *te-*, and *ne-*, and *ya-* often becomes *yi-* under the influence of the /y/. This produces forms such as *yislam* and *yislam* ‘he saves himself’, *yigbel* ‘he receives’, *telbes* ‘she wears’, *yilbesūn* ‘they wear’ and *terʿebūn* ‘you mount’.

4.3.4 *Vowel incidence of the P-STEM*

As with the S-STEM, the incidence of different vowels in the P-STEM is a matter of interest in the typology of Arabic dialects. Firstly, Arabic dialects can be differentiated in

¹⁴⁰ Some of these forms are not attested with the *ya-* prefix, but this must be purely coincidence, and does not concern the presentation of paradigms here.

regard to the relationship between the vowel of the P-STEM prefix and the vowel of the verbal stem. Two contrastive systems are described in Holes (2016: 142-147), who shows that the ‘Arab dialect of Bahrain exhibits vowel dissimilation, whereby the prefix vowel dissimilates from that of the verbal stem (e.g. *yīšrab*, *yaskin* ‘he drinks, he lives’) while the Baḥarna dialect maintains the *yi-* prefix irrespective of the stem vowel (e.g. *yīšrab*, *yiskin*). As a third category, the Arabian dialects of the Southern Hijaz and the Tihāma exhibit vowel harmony, whereby the prefix vowel is governed by the stem vowel (e.g. *yuktub*, *yiksir*, *yarčab* ‘he writes, he breaks, he rides’) (Prochazka 1988: 32-34). As shown above, the Rwala dialect is of the fixed prefix vowel type, since the prefix vowel is always /a/ except for variation in forms that have *yi-*. Forms such as *yislam* ‘he saves himself’ show that vowel harmony is not a factor, and forms such as *yigtel* ‘he kills’ (in which the *e* seems to be a realisation of underlying /i/) suggest that dissimilation is not a factor either.

The incidence of vowels in the verbal stem is more complicated. As with the stem vowels of the S-STEM described above, Arabic vernaculars fall on a spectrum with regards to the stem vowels of the P-STEM. At one end are dialects in which the choice of stem vowel is fully morphosemantic, and serves as, for instance, a transitivity marker; this is shown to be the case in many central Arabian dialects in which the S-STEM and the P-STEM are correlated on the basis of transitivity (Ingham 1994a: 23-24 and chapter 7). At the other end, there are dialects where the choice of stem vowel is entirely influenced by the phonological environment. Many dialects lie somewhere in between, which the same examples from

Holes (2016) illustrate. In the Bahraini ‘Arab dialect, the stem vowel is governed firstly by the presence and position of guttural consonants in the root, and then by a division into /a/ stems and /i, u/ stems, in which latter case the vowel is governed phonologically depending on the consonantal environment. In the dialect of the Baḥarna, the height contrast of CLA is mirrored very closely.

The Rwala system is also a hybrid: the lexical stem vowel, which in many cases is the same as its CLA counterpart, maintains, but certain phonological redistributions are made. It is useful to consider the verbs in groups according to how they relate to their CLA counterparts¹⁴¹, since there is a high level of overlap. In this matter, it must always be remembered that CLA is not to be taken as an ancestor dialect of the data presented here, since, although its structures are similar, we cannot assume that it resembles the ancestor dialect of any modern vernacular. The comparison of CLA stem vowels with Rwala stem vowels produces seven groups:

Group 1: CLA /a/, Rwala /a/ 67 roots (+ 13 roots with *e* in Rwala*)

e.g. *yaslamu, yaslam* ‘he is safe’
 yaqbalu, yaqbal ‘he accepts’

Group 2: CLA /i/, Rwala /i/ 15 roots*

e.g. *yaʿrifu, yaʿref* ‘he knows’
 yanzilu, yanzel ‘he sets up camp’

Group 3: CLA /u/, Rwala /u/ 23 roots

e.g. *yatruku, yatrok* ‘he deserts’
 yaṭlubu, yaṭlob ‘he asks for’

¹⁴¹ Here I follow Holes (2016: chapter 4).

Group 4: CLA /i/, Rwala /u/	4 roots
e.g. <i>yaqlibu, yaglob</i> ‘he turns around’	
Group 5: CLA /u/, Rwala /i/	19 roots*
e.g. <i>yaqtulu, yagtel</i> ‘he kills’	
Group 6: CLA /i/, Rwala /a/	4 roots
e.g. <i>yargi‘u, yargā‘</i> ‘he returns’	
Group 7: CLA /u/, Rwala /a/	8 roots
e.g. <i>yab‘udu, yab‘ad</i> ‘he is far away’	

What can be seen from the data¹⁴² is that CLA /a/ and /i/ are fairly stable in corresponding to /a/ and /i/ in the Rwala dialect. The biggest difference is seen in the case of CLA /u/ which is split in Rwala, either three ways, between /a/, /i/, and /u/, or two ways, between just /a/ and /u/.

A three-step process for the determination of stem vowels may be posited. Firstly, the data show that there are no guttural consonants in C2 or C3 position in any of the Rwala /i/ or /u/ groups, and furthermore, roots with C2 or C3 gutturals make up half the roots in Groups 6 and 7. Therefore, the first step is that roots with C2 or C3 gutturals are naturally assigned to the /a/ class, which accounts for Groups 6 and 7. Secondly, it can be seen that CLA /a/ roots are also /a/ roots in the Rwala dialect, so the second step is that /a/ corresponds to /a/. There is a possibility that the 13 roots only attested with an *e* stem

¹⁴² The above figures are not fully valid on account of the neutralisation of both /a/ and /i/ as *e* in Musil’s representation of the Rwala dialect: it must therefore be noted that members of Group 2 may, in fact, be Group 6; members of Group 5 may, in fact, be Group 7; and the extra thirteen members of Group 1 may, in fact, belong to a new group of CA /a/, Rwala /i/ (these are marked with an asterisk).

vowel represent a shift /a/ > /i/, but since 8 of them contain gutturals, such a move would be in contravention of the first step, which suggests that the /a/ class as a whole does not change, leaving us with Group 1. Finally, CLA /i/ and /u/ roots are differently distributed in Rwala, which suggests an overall vowel height distinction in the P-STEM rather than a functionally three way system. The phonological criteria, which mirrors that documented by Holes for the Baḥarna (Holes 2016: 146), is that /u/ only occurs when both C2 and C3 of the root are labials, liquids, velar stops, or emphatic consonants. This group, consisting of /b, m, l, r, k, g, ʔ, ʕ, ɖ/, forms a natural class since the labials are marked for roundness and the velar stops and emphatics are marked for backness. Roots with CLA /u/ correspond to Rwala /i/ if these criteria are not met (Group 5). The process seems to happen in reverse as well, although there are not many data to show this. Of the CLA /i/ roots, where the conditions for /u/ are met, these correspond to /u/ in 4 out of 6 cases (Group 4). This accounts for at least 52 of the 61 roots in Groups 2-5. The nine exceptions, in addition to three guttural roots which did not move to the /a/ class, account for all 12 exceptions out of 153 verbal roots attested for the strong verb.

4.4 Participles

FORM-I participles are formed of a base form, which demonstrates a templatic transformation of the finite verb stem, concatenated with feminine and plural suffixes.

4.4.1 Active Participle

The form is CāCeC. As established by McCarthy and Prince (1990: 28) and Watson (2002: 153-4) the form is derived from the template of the finite verb (CVCVC) by initial mora reduplication, at which point the active participle vocalic melody *a-i* (which is always *a-e* in Musil) maps onto the moraic slots, while the consonants of the finite verb are associated with the non-moraic slots.

When inflected, the forms are:

MS	CāCeC	<i>šāreb</i>	
FS	CāCeCa	<i>šāreba</i>	
MP	CāCeCīn	<i>šārebīn</i>	
FP	CāCeCāt	<i>šārebāt</i>	‘drinking’

As discussed in §2.2.3.4, it is curious that the suffixed forms should retain the vowel in all but a few instances, since forms with vowel deletion are predicted to occur on account of /i, u/-syncope and furthermore are recorded in the majority of relevant dialectal literature. Cantineau and Ingham only give the syncopated forms in their works, and in Prochazka it is only a small minority of dialects, such as that of Tanūmah, which have forms such as *rākībīn* ‘riding_{MP}’ as opposed to the majority which have *rākībīn*. The Šammari texts recorded by Montagne contain a mixture: he does include forms such as *ṭālb-ek* ‘asking_{MS} you’ (MŠa V, 5) but also *rācebīn* (MŠr 27) and *ḏābeḥ-o* ‘killing_{MS} him’ (MŠa IX, 11, 13; all references from Cantineau 1936-37). Cantineau, who was possibly thinking of Musil’s transcriptions as well as Montagne’s, asserts that sometimes “points vocaliques” (which, as above, I understand as super-short vowels) could be mistaken for short vowels (Cantineau

1936-37: 189), leading to erroneous transcription. Such an argument seems to tally with Musil's data, since we have already seen that in his material the deletion of high vowels in open syllables is more of a tendency than a rule. It is likely therefore that Musil's transcriptions have obscured the operation of this process, since the transcriptions would show no difference whether vowel reduction occurred (rather than a total deletion) or syncope induced epenthesis of an super-short vowel.

4.4.2 Passive Participle

The form is maCCūC. It is morphologically complex, derived from the finite verb by both the prefixation of *mμ-* and by final mora reduplication. The vocalic melody of the passive participle is *a-u*.

When inflected, the forms are:

MS	maCCūC	<i>maḍbūḥ</i>	
FS	maCCūCa	<i>maḍbūḥa</i>	
MP	(maCCūCīn)	n/a	
FP	maCCūCāt	<i>maḍbūḥāt</i>	'(having been) killed'

I have seen no record of these forms in other dialectological literature for north-Arabian dialects. Being formed of a heavy syllable followed by a superheavy syllable however, they are phonologically stable and variation of the kind that we have discussed elsewhere in this chapter is unlikely to occur. With the exception of two forms, *mḡalūt* 'mad (dog)' and *mḥazūga* 'ragged', *ghawa*-forms are not attested for the passive participle in Musil.

4.5 Imperative

The imperative is formed from the P-STEM verbal stem. Correspondingly, there are three classes: the stem is CCVC where V may be *a*, *e*, or *o* representing an underlying vowel /a/, /i/, or /u/ which is encoded lexically within the root. It appears that only the stem of the imperative is derived from the P-STEM because, as will be seen from the paradigms, the imperative has its own set of suffixes which differs from that of the P-STEM. The FP imperative is unattested. The paradigms are:

	/a/	/i/	/u/
MS	<i>erkab</i>	<i>ebšer</i>	<i>ogʻod</i>
FS	<i>erkabī</i>	n/a	n/a
MP	<i>erkabū</i>	<i>ebšerū</i>	<i>ogʻodū</i>
FP	n/a	n/a	n/a
	‘mount!’	‘rejoice!’	‘sit!’

Prochazka records the forms *ʾiktib* ‘write!’ and *ʾišrab* ‘drink!’ for the Rwala, which correlate with *ebšer* and *erkab* here. As in the P-STEM, his suffixed forms show syncope and vowel alternation which is not present in Musil’s data, e.g. *ʾiktbay* for the FS in the I-CLASS, which shows /i, u/-syncope, and *ʾišrubay* in the A-CLASS, which shows backing. Musil’s unsyncoated forms are consistent with what is recorded for the P-STEM, however. The *-ū* suffix, which differs from the *-ūn* suffix of the P-STEM, is found in all bedouin dialects of Arabia and is considered to be an archaic trait of the dialect group by Cantineau (1936-37: 187), although it is not always recorded as long.

The interesting kind of vowel harmony of the prothetic vowel demonstrated by the imperatives has been described in §2.2.2.2, and the matter of the long vowel of the *-ī* and *-ū*

suffixes has been addressed in §2.3.3. In §4.2.1 it was noted that *-n-* occurs infrequently between P-STEM vocalic suffixes and vowel-initial object pronouns: this is attested twice for imperatives as well, e.g. *ḥadderū-n-eh* ‘lower her’¹⁴³.

4.6 Verbal noun

The FORM-I verbal noun is unpredictable in form. Here the most frequently attested forms are listed with an example of each. Of those listed, the form CaCC is by far the most common, accounting for over half the attested verbal nouns. All forms are attested outside of the verbal system as nominal templates (cf. §3.1.2).

CaCC	<i>dabḥ</i> ‘killing’
CiCC	<i>diḥr</i> ‘remembering’
CuCC	<i>ogd</i> ‘tying’
CaCaC	<i>ba‘ad</i> ‘being far from’
CaCāC	<i>ṭalāg</i> ‘cutting off’, <i>ḍrāb</i> ‘leading in [camel]’
CuCūC	<i>sukūn</i> ‘staying in a place’, <i>rkūb</i> ‘riding’
CaCCa	<i>hebde</i> ‘running until one’s head shakes’

Here it may be mentioned that verbal nouns are attested in the corpus in what is termed an ‘absolute accusative’ construction (cf. the CLA *maf‘ūl muṭlaq*, Wright 1967 vol. 2: 53). This refers to a verb (or participle, although this is unattested in Musil) taking its own verbal noun as a complement, and the meaning is described by Holes for Bahrain, which seems similar to the usage in Musil:

“The rhetorical effect is to add emphasis, often with an element of extensiveness or intensiveness to the action described.” (Holes 2016: 454)

¹⁴³ The use of the 3MS pronoun for female referents is common in the poetry.

An example from Musil is *galb-i yūgef wġāf* ‘my heart stops dead’ (literally ‘my heart stops a stopping’), which illustrates this emphatic quality.

4.7 Passive forms

The presumed internal passive of Old Arabic has only survived in Arabian dialects, and even then only in some of them (Holes 2004a: 135). Based on the low number of attestations it does not seem to be particularly productive in the Rwala dialect, but it is attested nonetheless.

4.7.1 Internal passive of the S-STEM

Prochazka (1988: 116) gives the forms *ħmil*, *ħimlat* ‘he/she was carried’ and Ingham (1994a: 27) gives *sriġ*, *sirġat* ‘he/she was robbed’. It is confirmed that the passive S-STEM has the underlying representation /CiCiC/ in e.g. Al-Mozainy (1981: 50) and Parkinson (1992: 103). Musil’s data is inconsistent: his form *q̄cer* ‘it was remembered’ seems to have the same form as Ingham’s *sriġ*, but anomalous disyllabic forms such as *ġasel* ‘it was washed’ (not **ġisel* or **ġsel*) are attested. Best attested are /u/ forms which are more outwardly similar to CLA than to the recorded north Arabian forms, such as *q̄ubaḥ* ‘he was killed’, *gudrat* ‘it (f.) was effected by’ and *ħoṭemt* ‘you were rapped on the nose’. I do not consider the data to be sufficient for a full appraisal of this paradigm. There is, however, some support for the

continued existence of /CuCiC/ in north-Arabian dialects in Al-Mozainy (1981: 59) for the Ḥarb dialect.

Prochazka has only one /u/ form in his entire study, which does happen to come from Rwala: *rubtat* ‘she was tied’, but even for this form he gives the 3MS as *rbut*, so the forms are probably /ribiṭat/ and /ribiṭ/ with syncope of /i/ and backing in the environment of the labial /b/. Johnstone (1967a: 109) gives the form *w-uhidū* ‘and they were captured’ for Qaṭari, where he specifically points out that the Qaṭari dialect makes use of the internal passive due to its stronger connection to the ‘Anazi dialects; this form corresponds to one found in Musil, *wuḥedū* ‘they were robbed’¹⁴⁴. In Johnstone the *u-i* melody is only found with this C1 = /ʔ/ verb, and it will be seen below (§4.10) that the internal passive is particularly productive in so-called “weak” verbs, perhaps because the resulting forms have particularly salient features. The *u* in *wuḥedū* could be explained phonologically as backing of /i/ in the environment of /w/, but in Musil the *u-i* melody is used for strong verbs as well in a way that does not seem to be paralleled in the dialectal literature. It seems unlikely that Musil imposed this structure onto the words, since he records other passive forms with /i/ as the passive morpheme, which fit better with descriptions of the passive in other dialects. In general, the forms that are attested do not produce identifiable patterns; in this regard, we may consider Rosenhouse’s observation that

¹⁴⁴ Note that Musil correctly parses the /w/ not as the conjunction but as the reflex of /ʔ/ in the passive voice. It appears that Johnstone may have missed this, since he inserts a hyphen and translates “and”, but with the quotation being given out of context one cannot know for sure.

“in persons other than 3rd m. sg. [the] distinction [between active and passive] is often blurred” (Rosenhouse 1984: 36). It is a shame that Musil’s records do not contain more passive verbs for a detailed study, but it is certainly interesting that he has recorded some forms with the melody *u-i* outside the C1 = /ʔ/ group.

The use of the internal passive in the S-STEM appears to be much more productive in the biliteral verbs, the paradigms of which are described below. For the deficient (C1 = /ʔ/) verbs, they are formed with an initial /w/, so the forms are *wuḥed*, *wuḥidat* ‘it was taken (MS, FS)’ (Prochazka records the syncopated form *wuḥdat*). For the hollow verbs there is a medial /ī/, e.g. *ḡil*, *ḡilat* ‘it was said (MS, FS)’. Doubled verbs also have /u/, e.g. *ḥull* ‘it was refused’. In the final-weak verbs we see forms which agree with Prochazka, such as *bini*, *binyat* ‘it was built (MS, FS)’ alongside /u/ forms such as *ḡufa* ‘it was snatched’. It seems likely that the particularly distinctive form of these passives has increased their productivity in the dialect.

4.7.2 *Internal passive of the P-STEM*

Prochazka gives the form *yihmal* ‘he is carried’ which matches the most common form in Musil, which is *yidkar* ‘it is said [lit. it is remembered]’. The passive is not attested in the P-STEM for the biliteral verbs, except for the hollow verbs where there is a medial /ā/ vowel, e.g. *yeḏām* ‘he will be injured’. Apart from this form, the passives rarely occur in the P-STEM; this seems to conform to a pattern whereby the internal passive is most productive

where it has clear morphological markers, such as the initial /w/ in the s-STEM for the deficient verb and the characteristic /ī/ and /ā/ vowels for the hollow verb. All other P-STEM forms would be indistinguishable from active forms, which may explain why they are not frequently attested.

4.8 Derived forms

Nine derived verb forms are attested in Musil for Rwala Arabic. These augmented verbal patterns are sometimes called morphosemantic (cf. Holes 2004a: 100) and are derived from the basic verb stem by the affixation of one or two moras, as will be seen below. Following the orientalist tradition, these forms are referred to by Roman numerals. However, the use of this terminology is not intended to suggest that the system described here is identical to that of CLA. In the following table, the use of the consonants *f*, *ʕ*, and *l* is a traditional representation of the three root consonants:

FORM-I	<i>faʕal ~ fiʕil</i>
FORM-II	<i>faʕʕal</i>
FORM-III	<i>fāʕal</i>
FORM-IV	<i>aʕʕal</i>
FORM-V	<i>tafaʕʕal</i>
FORM-VI	<i>tafāʕal</i>
FORM-VII	<i>ɲfaʕal</i>
FORM-VIII	<i>ftaʕal</i>
FORM-IX	<i>fʕall</i>
FORM-X	<i>stafʕal</i>

There is also one instance in the corpus of a doubled verb in the local form *ntafaʕʕ* which is dealt with in the section on FORM-VII and is provisionally referred to as FORM-VII-t.

All forms are derived from FORM-I, whether directly or indirectly: the following are synchronic derivations proposed by McCarthy (1981). Forms II-IV are derived directly by infixation of a mora, which respectively associates with the medial root consonant (FORM-II), with the leftmost vowel (FORM-III) or with the vowel of the causative prefix (FORM-IV). Forms V-VI are derived indirectly, being synchronically derived from forms II-III by prefixation of /ta/, or /t/ in some cases in the P-STEM and imperative. Forms VII-IX are derived directly through prefixation of a mora, which is associated with /n/ (FORM-VII), with /t/ (FORM-VIII), and with the initial root consonant (FORM-IX). As shown in §2.2.2.2, this prefixation produces an incomplete initial syllable in these forms which is syllabified by prothesis (cf. Figure 2.17). In the case of FORM-VIII, it has been proposed that a second rule metathesises the prefixed /t/ with the initial root consonant (cf. Watson 2002: 140; McCarthy 1981: 389-90), and it is this two-step process of prefixation and metathesis which ensures that the additional mora is ultimately associated with the initial (incomplete) syllable, as necessary to account for prothesis¹⁴⁵. In the case of FORM-IX, the association of the prefixal mora with the initial root consonant results in gemination of the final root consonant, as moras cannot be lost from the original template. FORM-X is historically derived from FORM-IV, through the infixation of /t/ (cf. Watson, *ibid.*) but synchronically seems to be derived from FORM-I, having been reanalysed as a prefix /sta-/.

¹⁴⁵ According to the syllabification algorithm established in §2.2.2, a mora associated with /t/ in *fta'al* would cause it to be syllabified as **fatta'al* rather than (*e*)*fta'al*.

4.8.1 *Morphology of the derived forms*

Some features are common to all the derived forms. Most have the vocalic melody *a-* *a* in the S-STEM and *a-e* in the P-STEM, imperative and (active) participle, and most have only one participle; where there is a differentiation between active and passive participle, the latter has the vocalic melody *a-a*. Forms VII-X have initial consonant clusters which are resolved by prothesis. In §2.2.2.2 it was argued that the initial consonant is linked to an incomplete syllable, explaining why the cluster is syllabified through prothesis rather than medial epenthesis, resulting in *vCC* rather than *CvC*. The prothetic vowel is *e* (or, rarely in FORM-X, *i*) from an underlying /*i*/, as according to the prothesis rule established in §2.2.2.2. The only exception is the poorly attested FORM-IX, where in one example it seems that the initial /*a*/ of the *a-a* and *a-e* vocalic melodies extends into the prothetic vowel as well. This may be an anomaly; alternatively, since FORM-IX is unique among the derived forms in only having one vowel in its template, it is possible that the prothetic /*a*/ surfaces through analogical extension which, being an irregular process, overrides the phonological rule.

4.8.1.1 FORM-II

The FORM-II finite verb stem is derived from the FORM-I finite verb stem through medial gemination. The stem is $CaC_iC_i aC$ in the S-STEM and passive participle, and $CaC_iC_i eC$

in all other forms¹⁴⁶. Examples given are for the verb *baddal* ‘to change one thing for another’.

In the S-STEM:

3MS	<i>baddal</i>
3FS	<i>baddalat</i>
2MS	<i>baddalt</i>
3MP	<i>baddalow</i>

A minority of verbs (14%) have the stem $CaC_iC_i eC$ in the S-STEM as well. This situation defies easy explanation: of the nine FORM-II verbs which are attested also in FORM-I, six are in the I-CLASS, e.g. *‘allemt* (FORM-I *‘alemt*), but it would be unprecedented for different classes to exist in FORM-II, and furthermore this is not a strong correlation on account of examples such as *farregat* (FORM-I *farag*). There is also some correlation between the occurrence of *e* and the proximity of /l/ or /r/ as C2 or C3; this accounts for 81% of $CaC_iC_i eC$ stems, but cannot ultimately be the motivating factor since $CaC_iC_i eC$ only occurs in 28% of verbs with /l/ and 10% of verbs with /r/. Since there is no clear phonological reason to explain these data, and since *a-a* is the vowel melody of the S-STEM in all the derived forms, it is likely that the *e* here is simply a variant of /a/.

In the P-STEM:

¹⁴⁶ I assume that the underlying representation of these stems is / $CaC_iC_i aC$ / and / $CaC_iC_i iC$ /, and this pattern of *a-a* and *a-i* melodies is a feature of most of the derived forms of the verb. I do not give the underlying representations, however, on account of the fact that *i* is never attested in these paradigms and this could potentially be another case for phonemic /e/ (cf. §2.1.2.1).

3MS	<i>yebaddel</i>
2MS	<i>t(e)baddel</i>
1CS	<i>abaddel</i>
3MP	<i>yebaddelūn</i>

Here it is interesting that Musil's forms, though highly consistent in themselves, contrast with the description given by Prochazka. The latter states that the Rwala dialect exhibits syncope of the stem-final vowel when suffixes are added, i.e. *yebaddelūn* above would be *yebaddlūn*. This is completely unattested in Musil's data, although the reason for the discrepancy is clearly the less frequent occurrence of syncope in Musil. Prochazka also states that the prefix has no vowel in FORM-II, giving forms such as *ykassir* 'he smashes'. This is attested in Musil only with the *t-* prefix (*tbaddel*), with the syncopated form occurring in a third of cases. As discussed in §2.2.3.5, Musil may not have perceived this on account of the differing sonority profiles of /y/ and /t/.

The imperative form is *baddel*. Active and passive participles are differentiated, with the active forms having the vowel melody *a-e* and the passive forms having *a-a*. The prefix is *mu-* (often syncopated to *m-*) which gives the active participle form *m(u)baddel* and the passive form *m(u)baddal*. Again, Prochazka records only the syncopated form. The verbal noun is *tabdīl*.

Prochazka records internal passive forms for both the S-STEM and the P-STEM, which are respectively *kissir* 'it got smashed' and *ykassar* 'it gets smashed'. Here, as in the derived forms of all the dialects he records, the passive morpheme is /i/ in the S-STEM and /a/ in the P-STEM. The same internal passive of the P-STEM is attested in Musil, such as in *tbanna*

‘[the tent] is pitched’ which opposes *tbanni* ‘she pitches [the tent]’. The *kissir* type in the s-STEM is not attested, however. Rather, there are some instances of a *u-i* vocalic melody for the s-STEM passive, such as *futtiḥen* ‘they (f.) were opened’. These are all found in poetry which may explain their absence from Prochazka’s data, although it has already been seen that while Prochazka did not record the *u-i* passive melody, the melody is not unattested in north- and central-Arabian dialects (cf. §4.7).

4.8.1.2 FORM-III

The FORM-III finite verb stem is derived from that of FORM-I through lengthening of the first stem vowel. The stem is CāCaC in the s-STEM and CāCeC in all other forms.

Examples given are for the verb *rādaf*, ‘to sit behind somebody on a camel’.

In the s-STEM:

3MS	<i>rādaf</i>
3FS	<i>rādafat</i>
2MS	<i>rādaft</i>
3MP	<i>rādafow</i>

These concur with Prochazka’s forms, although he also records some raising and backing of the second vowel when vocalic suffixes are added. Thus his form *ḥārab* ‘he fought’ becomes *ḥārubaw* and *ṭārad* ‘he pursued’ becomes *ṭāridat*. This variation may be somewhat reflected in Musil’s s-STEM CāCeC forms, which pattern like the s-STEM CaC_iC_ieC forms in FORM-II. /a/-raising is found throughout Prochazka’s account, and seems to parallel /a/ > e in Musil in the same position, but backing is never attested in this paradigm.

In the P-STEM:

3MS	<i>yirādef</i>
2MS	<i>t(e)rādef</i>
1CS	<i>arādef</i>
3MP	<i>yirādefūn</i>

Here the data concur with Prochazka's, except for the lack of syncope of the prefix vowel, which once again only occurs after /t/ (and once after /n/, in *nḥāreb* 'we fight'). The /y/ prefixes (3MS, 3MP and 3FP) usually have *i* as their vowel, although *e* is also attested.

The imperative form is *rādef*. The active participle is formed with the prefix *me-* (which is syncopated to *m-*), which differs from the *mu-* form in FORM-II, although the underlying representation of the morpheme is uncertain, but presumably the same. The *mu-* prefix is attested for the passive participle, however, e.g. *mubārak* 'blessed'. The verbal noun has the form m(u)CāCaC, e.g. *murāga* 'climbing' and *muṭā'an* 'stabbing'; this form of the verbal noun is not recorded in Prochazka but has been recorded in other bedouin-descended dialects (cf. Holes [2016: 207] for eastern Arabia). As Holes (*ibid.*) notes, this form of the verbal noun is identical with the passive participle. There are no S-STEM or P-STEM passive forms recorded for FORM-III in Musil, though Prochazka records internal passive forms *ġibīl* 'he was met' and *ygābal* 'he is met'.

4.8.1.3 FORM-IV

The FORM-IV finite verb stem is derived from that of FORM-I through the prefixation of *a*. The stem is aCCaC in the S-STEM, and CCeC in the P-STEM and other forms. Examples given are for the verb *aḏhar*, ‘to show’.

In the S-STEM:

3MS	<i>aḏhar</i>
3FS	<i>aḏharat</i>
2MS	<i>aḏhart</i>
3MP	<i>aḏharow</i>

Musil’s forms concur with the forms in Prochazka, except that the latter gives the raised form *aḏhirow* instead of *aḏharow*. This reflects the usual pattern and, as would be expected, there is some variation in Musil between *aḏharow* ~ *aḏherow*, but no *i* is recorded.

In the P-STEM:

3MS	<i>yīḏher</i>
2MS	<i>teḏher</i>
1CS	n/a
3MP	n/a

These forms are poorly attested, the reason for which may be, in large part, the fact that they are indistinguishable from FORM-I, which makes it very difficult to say with any confidence what form a verb belongs in if it is only attested in the P-STEM. Prochazka records /i/ throughout this paradigm, e.g. *yīb‘id* ‘he moves away’, and thus establishes a neat contrast between FORM-I verbs which have /a/ as the prefix vowel and FORM-IV verbs which have /i/: the forms are therefore distinguishable regardless of what the vowel of the verbal

stem may be in FORM-I. In Musil, as we have seen, *yi-* prefixes are common in FORM-I, although /i/ only rarely occurs in the prefix after /t/ and does not occur in the first person.

The imperative form is *aḏher*, also attested in the plural as *eḏherow*. This concurs with Prochazka, who records *ʿabʿid* ‘move away!’ which retains the /a/ vowel, as opposed to his central and southern dialects which often have /i/ as the first vowel. The most common active participle form is *meḏher*, and there seems to be a passive participle form *maḏhar*. This reflects the situation in other derived forms in which *e* (</i/) expresses the active participle and contrasts with /a/ in the passive; there is a fair amount of variation in the forms of participles in this derived form, however. Prochazka does not record FORM-IV participles for the Rwala, but his form for the closely related dialect of Ḥāyil is *miCCiC* which concurs with Musil, and he does not record any passive participles. The verbal noun is *eḏhār*. Prochazka records no verbal noun for Rwala, but it appears to be productive in Musil.

It should be noted that a number of FORM-IV verbs have lost the formal distinction with FORM-I verbs. A common example is the variation *rād* ~ *arād* ‘he wanted’, where the more common S-STEM form *rād* has the shape of a FORM-I verb. Other verbs show this partial coalescence in different parts of the verb, such as *nšerow* ‘they set out to make a distant raid’ which, as with other attested forms from this root, should probably be interpreted as FORM-I (cf. Kurpershoek 2005: 318), but the verbal noun *enšār* appears to be that of FORM-IV. The verb *ašbaḥ* ‘he came/left in the morning’ usually has FORM-IV forms in the S-STEM but in the

P-STEM varies between ambiguous forms (*yeşbeh*) and those which seem much more likely to be FORM-I (*yaşbah*). In all of these cases there is no semantic difference between these forms.

4.8.1.4 FORM-V

The FORM-V finite verb stem is derived from that of FORM-II through the prefixation of *ta-*. The stem is $taCaC_iC_i aC$ in the S-STEM, and $t(a)CaC_iC_i eC$ in all other forms. Examples given are for the verb *tağassal*, ‘to scrub oneself’. In the S-STEM:

3MS	<i>tağassal</i>
3FS	<i>tağassalat</i>
2MS	n/a
3MP	<i>tağassalow</i>

As above, this concurs with Prochazka’s description except for the forms with vocalic suffixes, where the contrast is the same as for FORM-IV.

In the P-STEM:

3MS	<i>yetağassel</i>
2MS	<i>tetğassel</i>
1CS	<i>atğassel</i>
3MP	<i>yetağasselün</i>

Here there are three departures from Prochazka’s account. Firstly, Prochazka records syncope of the prefix vowel, as we have seen elsewhere. Secondly, his P-STEM verbal stem is $tiCaCCaC$ on the surface rather than Musil’s $t(a)CaCCeC$, which differs from Musil’s data both in the first and last vowels, although the first of these is a case of /a/-raising. Finally, his account of the 3FS and 2FS forms (presumably all forms where the prefix has /t/) shows

forms such as *tićallam* ‘she speaks’; such forms are attested at least once in Musil, in the 3FS form *tawaṭṭa* ‘she steps forward’. This occurs as a result of the following rule applications:

UR	/tatakallam/	/tataḥammalīn/
/a/-syncope	ttakallam	ttahammalīn
/a/-raising	ttikallam	————
fronting	ttićallam	————
backing	————	ttaḥammulīn
degemination	tićallam	taḥammulīn
output	<i>tićallam</i>	<i>taḥammulīn</i>
	‘she speaks’	‘you (f.) bear’

The rule marked degemination is a simplification process presumed to occur when derivation has produced a word-initial geminate. These derivations present some problems for Musil’s data, where *tetgassel* is hard to account for. The explanation may be found in a comparison with Prochazka’s 1CS form:

	Prochazka	Musil
UR	/atakallam/	/ataḡassil/
/a/-syncope	*****	*****
/a/-raising	atikallam	atiḡassil
/i, u/-syncope	*****	atḡassil
fronting	atićallam	————
output	<i>atićallam</i>	<i>atḡassel</i>
	‘I speak’	‘I scrub myself’

Here /a/-syncope fails to apply, and I suggest that the 1CS prefix resists syncope on account of a restriction on the deletion of entire morphemes. As can be seen, Prochazka’s form also resists the /i, u/-syncope rule, although we may compare his form from Hofūf where it does apply: *atkallam* (this form also indicates that fronting is ordered after syncope). It is possible that Musil’s form *tetgassel* is derived by levelling, whereby the non-

application of /a/-syncope in the 1CS form has extended to other inflections. Since analogical processes such as this tend to be random, the variation of forms in Musil's data (i.e. *tetgassel* but not **yetgassel*) are not particularly surprising.

The final /a/ in Prochazka's paradigm represents the norm in bedouin-descended dialects of north Arabia and the Gulf coast; Johnstone notes that "the retention of the second vowel of the imperfect is strange, but such forms are well attested" (Johnstone 1961: 258). The vocalic melodies of the derived forms are given for the Rwala dialect by Prochazka (my emphasis):

FORM-II	<i>ykassir</i>	'he smashes'	<i>a-i</i>
FORM-III	<i>ysāfir</i>	'he travels'	<i>a-i</i>
FORM-V	<i>ytićallam</i>	'he speaks'	<i>a-a</i>
FORM-VI	<i>ytaḥārab</i>	'he fights (with sb.)'	<i>a-a</i>
FORM-VII	<i>yinǧilib</i>	'he turns over'	<i>a-i</i>
FORM-VIII	<i>yištiǧil</i>	'he works'	<i>a-i</i>

The vocalic melody is also *a-i* in forms VII-VIII, but this is obscured by /a/-raising. In Musil, the P-STEM conforms to a general pattern whereby all the derived forms have the vocalic melody *a-e*, but given the dialectal evidence it is possible that the apparently systematic nature of Musil's transcriptions could have been imposed onto them by the transcriber himself.

The imperative is poorly attested, but seems to be (predictably) **taǧassel*, as evidenced by the MP form *tasamme'ū*. 'hear!'. Prochazka predictably records the imperative with /a/-raising: *tićallam* 'speak!'. The participle is *mut(a)ǧassel*, which is an interesting case of syncope, since /i, u/-syncope has targeted /a/ rather than an available high vowel,

producing *mutġassel* rather than **mtaġassel*. The motivation of maximising bimoraicity could be argued in this case, especially since (in contrast to *tetġassel* above) /a/-syncope cannot apply (cf. the bimoraicity constraint, §2.2, and examples in §2.2.3). One form with /a/ in the final syllable is attested, *mutšallah*, ‘having one’s sleeves thrown back and tied’, which could arguably be a passive participle, but the usual semantic function of FORM-V, i.e. reflexivity (see below) makes it very difficult to discern active and passive meaning in this form, and it is possible that there is no passive participle. The verbal noun is poorly attested: it occurs once as *taħadđor*, ‘living as a sedentary’, in which the final *o* is difficult to explain, and in another instance as *twaħmet* ‘craving’, which appears to be a shortened form of **tawahħimat*. Prochazka gives the form *tikissir* ‘breaking’, where it is unclear why the second *i* is raised, as it would be expected to be **tikassir*.

4.8.1.5 FORM-VI

The FORM-VI finite verb stem is derived from that of FORM-III through the prefixation of *ta-*. The stem is taCāCaC in the S-STEM, and t(a)CāCeC in all other forms. Examples given are for the verb *taħāšar*, ‘to share out’.

In the S-STEM:

3MS	<i>taħāšar</i>
3FS	n/a
2MS	n/a
3MP	<i>taħāšarow</i>

Once again, apart from the difference in the stem-final vowel before vocalic endings, these forms concur with those found in Prochazka.

In the P-STEM:

3MS	<i>yitaḥāšer</i>
2MS	<i>tet(a)ḥāšer</i>
1CS	<i>atḥāšer</i>
3MP	<i>yetaḥāšarūn</i>

All the disparities between the data from Musil and Prochazka regarding the P-STEM in FORM-V apply here also. The stem-final *a* vowel in the 3MP form is difficult to explain and may simply represent an anomaly in the data, since 3MP is not well attested and all other persons have *e*; alternatively it could be seen as an alignment with the *a-a* melody found in Prochazka.

The imperative and verbal noun are unattested, but Prochazka gives *taḥānag* ‘have a quarrel!’ for the imperative, which would concur with the predicted (but hypothetical) **taḥāšer* form in Musil. The participle is also unattested, except for one form, *muṭṭālibīn*, ‘quarrelling MP’, which exhibits assimilation of the morphemic /t/ of the verbal stem to the following emphatic, and also syncope of the penultimate vowel, representing a possible underlying form /mutaṭālibīn/.

4.8.1.6 FORM-VII

The FORM-VII finite verb stem is derived from that of FORM-I through the prefixation of /n/. The stem is nCaCaC in the S-STEM and nCaCeC in all other forms. Examples given are for the verb *(e)nḏabaḥ*, ‘to be killed’.

In the S-STEM:

3MS	<i>(e)nḏabaḥ</i>
3FS	n/a
2MS	n/a
3MP	<i>(e)nḏabaḥow</i>

Prochazka’s forms are of the type *’inkisar* ‘it got broken’ or *’inhárag* ‘it burned down’, where /a/-raising occurs predictably in the former, but is blocked in the latter by the preceding guttural and following /r/ (cf. §2.3.6.1). Musil’s examples do not show /a/-raising, but this is in line with the rest of his data. The reason for the differently stressed forms in Prochazka is uncertain.

In the P-STEM:

3MS	<i>yinḏabeḥ</i>
2MS	<i>tendēbeḥ</i>
1CS	n/a
2MP	<i>tinḏabeḥūn</i>

Here Prochazka only records initial stress, as in *yínḡilib* ‘he turns over’, and the /a/ surfaces when syncope closes the syllable, as in *yinḡalbūn*.

The imperative is not attested; based on Prochazka’s *’inḡilib* it is not unreasonable to predict that the form in Musil would be **(e)nḏabeḥ*. Participles are formed by the prefixation

of /mi/ to the P-STEM verbal stem, hence *mīndabeḥ*. This concurs with Prochazka who records *mūngīlib*, f. *mīngālbah*. Syncope in the suffixed forms is attested in Musil for FS, e.g. *mīntabra*, ‘having been caught’, but not for MP, e.g. *mīnšāḥenīn*, ‘being anxious’. The verbal noun is *nḍebāḥ*, which agrees with Prochazka who records both the syncopated form *ʾinglāb* and the unsyncopated form *ʾingīlāb*.

The only example of FORM-VII-t in the corpus is *yintaṣeḥḥ* ‘be cured’ from the root **ṣḥḥ**. The combination of /n/ and /t/ as a prefix is well documented for Najdi Arabic dialects, examples being *yintalabbas* ‘he can/should be dressed, is dressable’ (Abboud 1979: 474) and *yintuwaggaf* ‘it is stoppable’ (Ingham 1994a: 74). These forms are analysed by both researchers as combining /n/ and /t/ prefixes on the FORM-II stem, rather than the /n/ prefix on the FORM-V stem: this is argued by Ingham (*ibid.*: 75) who demonstrates that *yintifahham* ‘he is capable of being enlightened’ is semantically related to FORM-II *fahham* ‘he enlightened’ more closely than it is to FORM-V *tifahham* ‘he came to understand’. Both note that the FORM-III stem may also be modified in this way. It is interesting that in the one example from Musil the stem is neither FORM-II nor FORM-III, and the example in this corpus thus differs from what has been previously recorded for Najdi dialects. The formation is claimed to occur only or mostly in the P-STEM in Najdi (Ingham 1994a: 199 note 48; Abboud *ibid.*). Holes (2016: 165) records the /n/ prefix being flexibly applied to other derived stems in a similar way in eastern Arabia, examples being *ingāṭṭa-* ‘to be covered up’ and *intifarrag-* ‘to be split up, to be dispersed’ (Bahrain), and *in‘awwar-* ‘to get hurt, injured’ (Oman).

4.8.1.7 FORM-VIII

The FORM-VIII finite verb stem is synchronically derived from that of FORM-I through the prefixation of /t/ and subsequent metathesis with C1 (cf. §4.8). As with FORM-VII, this results in an initial consonantal cluster which is resolved, where needed, by prothesis. The stem is CtaCaC in the S-STEM and CtaCeC in all other forms. Examples given are for the verb (e)*ħtalaf*, ‘to renege’.

In the S-STEM:

3MS	(e) <i>ħtalaf</i>
3FS	<i>eħtalafat</i>
1CS	<i>eħtalaft</i>
3MP	n/a

The forms are generally the same as those recorded by Prochazka, except that Prochazka records some syncopated forms. Thus his *‘iħtarag* ‘it burned down’ and *ħtaragin* concur with Musil’s data but forms such as *‘iħtrígat* do not. This syncope leads to some remarkable forms such as *šġalow* ‘they worked’ which has an initial CCCV syllable; forms in Musil are not attested with syncope of the stem-initial vowel.

In the P-STEM:

3MS	<i>yíħtalef</i>
2MS	<i>tíħtalef</i>
1CS	<i>aħtalef</i>
3MP	n/a

As in FORM-VII, Prochazka's forms have initial stress and /a/-raising, as in *yīštiḡil* 'he works', but the suffixed forms such as *yīštaḡlūn* show that the /a/ is present in the underlying representation.

The imperative is interestingly attested more as *eḡtelef* than as the predicted *eḡtalef*, although both do occur; they therefore more closely resemble Prochazka's *ʾīštiḡil* with its raised vowel. The participle prefixes /mi/, resulting in *miḡtalef*. As with FORM-VII, this concurs with the pattern of Prochazka's FS form *miḡtarḡah* where the vowel is not raised, but his MS form is *mīḡtiriḡ*. The verbal noun is variously *(i)ḡtilāf* and *(e)ḡtelāf*, and syncopated forms to match Prochazka's *ʾiḡtrāḡ* are not attested. The verbal noun is particularly frequent in poetry, where it is often found at the end of a line and has the poetic short vowel suffix *-i*; the uncommonly frequent occurrence of this form is presumably because of its syllable structure CVC.CV.CVV.CV, which fits the requirements of some poetic metres.

4.8.1.8 FORM-IX

FORM-IX is very infrequently attested in the corpus. Its verb stem is derived from that of FORM-I through prefixation of an initial mora, which is filled by the first root consonant; as a result, the final root consonant is geminated and the stem is trimoraic. As with FORM-VII and FORM-VIII, its initial consonantal cluster is resolved, where needed, by prothesis. The

stem is CCaCC in the S-STEM and CCeCC in the P-STEM. Examples given are for the verb (v)*ṣfarr*¹⁴⁷, ‘to become yellow’.

In the S-STEM:

3MS *ṣfarr*
3FS *(a)ṣfarrat*

These forms precisely match those recorded by Prochazka, who also gives the 3MS variants *ʔḥmar* ~ *ḥmarr* ‘he became red’, the first of which has initial stress and subsequent final degemination.

In the P-STEM:

3MS *yṣferr*

The form *yḥmar*, corresponding to the *ʔḥmar* form in the S-STEM, is the only form recorded for 3MS in the P-STEM by Prochazka, but he gives the suffixed forms as *yḥmarrūn* which, without initial stress, have not undergone degemination. Since the P-STEM is only attested once in Musil, it is difficult to compare the situation with that recorded by Prochazka, but the one attested form suggests either that initial stress did not occur in Musil or that, if it did occur, it did not trigger degemination.

The imperative, participle and verbal noun are not attested. They are also absent from Prochazka, except for the verbal noun which he gives as *ʔḥmrār*.

¹⁴⁷ Following the usual pattern, it is predicted that prothesis will be of *e* (< /i/). The one example in FORM-IX, however, is of *(a)ṣfarrat* which is unique in the corpus. I assume that it is an anomaly; an alternative explanation could be that the vocalic melody of the S-STEM *a-a* overrides this, since FORM-IX is the only derived form which has only one vowel in its template.

4.8.1.9 FORM-X

FORM-X is also very infrequently attested in the corpus. Synchronically, its verb stem is derived from that of FORM-I through the prefixation of /sta/. As with other forms that have an initial consonantal cluster, this is resolved, where needed, by prothesis. The stem is staCCaC in the S-STEM and, based on one attestation alone, the same in the P-STEM. This may be unreliable, as the participle has the stem staCCeC, and it would be predicted from the other derived forms both that the P-STEM should have *e* and that it should share its verbal stem with the imperative. Examples given are for the verb (*i*)*stafza*ʿ, ‘to ask for the help of (somebody)’.

In the S-STEM:

3MS	n/a
3FS	<i>istafzeʿat</i>
3MP	<i>istafzaʿū</i> ¹⁴⁸

These forms concur with those found in Prochazka, with the significant difference that in Prochazka’s account prothesis does not occur. In Musil, however, forms without prothesis are only attested contextually following a vowel, e.g. *tumma-starāḥ* ‘and then he was appeased’ as opposed to *gad istabšarū* ‘they rejoiced’. The prothetic vowel is almost always transcribed as *i*, unlike forms VII-VIII where it is *e*, and FORM-IX where it is once *a*. In FORM-X it is attested only once with *e*, in the 3FP form *estašraben*, ‘they drank their fill’.

¹⁴⁸ Note that the more frequent suffix *-ow*, as given in all other S-STEM paradigms, is not attested for FORM-X (but the evidence is scant, as there is only one example of a 3MP S-STEM for FORM-X).

In the P-STEM:

3MS *yistafza*[˘] (**yistafze*[˘]?)

Prochazka's form *yista*[˘]*mil* 'he uses' has final *i* which mirrors forms VII-VIII, perhaps strengthening the case that the usual form in Musil would have *e*. It is possible that the final *a* in this example surfaced because of the influence of the guttural /[˘]/.

The imperative and verbal noun are unattested, but Prochazka gives the imperative form *sta*[˘]*mil*. The participle prefixes /*mu*/, giving *mustafze*[˘]. Prochazka also gives passive forms for FORM-X, *sti*[˘]*mil* 'it was used' and *yista*[˘]*mal* 'it is used', which show the same morphemic vocalic melodies as in forms II-III, namely *i-i* for the passive S-STEM and *a-a* for the passive P-STEM. These forms are absent from Musil.

4.8.2 *Semantic correspondences of the derived stems*

In this section comparisons are made between verbs of the same root in different derived stems. For clarity, all forms have been expressed in the 3MS S-STEM form, but where the form is only speculative in terms of being A-CLASS or I-CLASS for FORM-I an asterisk has been added to make this clear.

4.8.2.1 FORM-II

FORM-II has the highest number of occurrences of all the derived forms in the dialect and is very productive. In a number of cases, FORM-II directly modifies the meaning of FORM-

I: out of all cases where a root is attested in both FORM-I and FORM-II, the latter modifies the FORM-I meaning in 84% of cases (the exceptions are mostly denominative, see below)¹⁴⁹.

Firstly, FORM-II often expresses a causative or factitive meaning:

<i>barad</i>	‘cool off’	<i>barrad</i>	‘make sth. cool off’
<i>zi‘el*</i>	‘be angry’	<i>za‘‘al</i>	‘make sb. angry’
<i>ħadar</i>	‘go down’	<i>ħaddar</i>	‘lay sth. down’

In the related, factitive sense:

<i>ħabar*</i>	‘know’	<i>ħabbar</i>	‘tell’ [i.e. make known]
<i>kašaf*</i>	‘come out’	<i>kaššaf</i>	‘uncover’ [i.e. make come out]

FORM-II frequently expresses an intensive meaning of FORM-I:

<i>yenšed</i> ¹⁵⁰	‘ask for’	<i>naššad</i>	‘ask a lot of questions’
<i>šafag</i>	‘clap sth.’	<i>šaffag</i>	‘stamp noisily’

In some cases the passive of FORM-I is expressed by FORM-II:

<i>ħasab</i>	‘count, think of’	<i>ħassab</i>	‘be reckoned’
--------------	-------------------	---------------	---------------

FORM-II often expresses further modifications *vis-à-vis* FORM-I which do not fall into these categories:

<i>šireb</i>	‘drink’	<i>šarrab</i>	‘rinse out’
--------------	---------	---------------	-------------

¹⁴⁹ Note that this only applies to the 54% of FORM-II verbs which are attested in the corpus in both FORM-I and FORM-II. 46% of FORM-II verbs are not attested in FORM-I at all: many probably do not have FORM-I counterparts, but in many cases it cannot be ruled out that the corresponding FORM-I verb is simply not attested.

¹⁵⁰ This verb is only attested in the P-STEM, and it seems possible that it is only used in the P-STEM since *sāl* (s’l) is only attested in the S-STEM and this looks like a possible case of suppletion.

Similarly, FORM-II frequently expresses technical meanings which are derived from their FORM-I counterparts but which only have specific areas of application:

<i>‘abar</i>	‘cross’	<i>‘abbar</i>	‘offer money’ [financial]
<i>‘araḏ</i>	‘show’	<i>‘arraḏ</i>	‘lead a mare to a stud’ [agricultural], ‘make an appeal’ [legal]
<i>ṣabaḥ*</i>	‘do in the morning’	<i>ṣabbaḥ</i>	‘attack at dawn’ [military]
<i>ṣa‘aṭ</i>	‘put reins on’	<i>ṣa‘‘aṭ</i>	‘put drugs in the nostril’ [veterinary]
<i>ṭalaġ</i>	‘let go’	<i>ṭallaġ</i>	‘divorce’ [family, legal]
<i>ṭaman</i>	‘pay for’	<i>ṭamman</i>	‘pay a fair price for’ [business]

The meaning expressed by FORM-II is occasionally *petal*; this term, borrowed from Ingham (1994a: 86), denotes an action in which the subject is asking for something:

<i>faza‘*</i>	‘help’	<i>fazza‘</i>	‘ask for help’
---------------	--------	---------------	----------------

Some of the major functions of FORM-II, however, do not encode a modification of a FORM-I meaning. These include denominative transitive verbs, in which verbs with various semantic values are formed from nominals:

<i>zēn</i>	‘lovely’	<i>zayyen</i>	‘adorn’
<i>ṣiḥ</i>	‘ <i>ṣiḥ</i> , a plant’	<i>ṣayyaḥ</i>	‘supply [a horse] with <i>ṣiḥ</i> ’

Some denominative verbs are also intransitive:

<i>ṣta</i>	‘winter, rainy season’	<i>ṣatta</i>	‘graze in winter’
<i>ġelāl</i>	‘blanket’	<i>ġallal</i>	‘wrap oneself in a <i>ġelāl</i> ’

And some are inchoative:

<i>rabi‘</i>	‘abundance’	<i>rabba‘</i>	‘be/become abundant’
<i>ṭa‘ām</i>	‘food’	<i>ṭa‘‘am</i>	‘be/become bitter’

Finally, FORM-II lends itself well to new forms such as those which are onomatopoeic:

<i>ṭha</i>	‘[noise made to call a sheep]’	<i>taḥḥa</i>	‘call a sheep with the sound <i>ṭha</i> ’
------------	--------------------------------	--------------	---

This last function of FORM-II is related to its use in other dialects to create verbs from foreign loan words (cf. *fannaš* ‘to dismiss, fire’ from English *finish*, Holes 2016: 150), but this is not attested in Musil’s corpus. It should be noted, additionally, that in some cases FORM-I and FORM-II of the same root may be used interchangeably, e.g. *ʿagal* and *ʿaggal* which both mean ‘fetter’.

4.8.2.2 FORM-III

The main use of FORM-III is to express an action which takes place with a patient. In many cases there is a FORM-I equivalent:

<i>šabak*</i>	‘mingle’	<i>šābak</i>	‘come into the midst of’
<i>lawā</i>	‘twist’	<i>lāwā</i>	‘twist together’

In many cases, however, the FORM-I either does not exist or is unattested:

<i>ḥārab</i>	‘fight with’
<i>ḥāšar</i>	‘share out’
<i>ʿāyaf</i>	‘enter into an agreement with’
<i>ʿāšar</i>	‘live with’

It is very frequent that the co-operative nature of this semantic association causes FORM-III verbs to take on a sexual nuance:

<i>laʿab</i>	‘play’	<i>lāʿab</i>	‘play [sexually]’
<i>ḡammaʿ</i>	‘gather’	<i>ḡāmaʿ</i>	‘have sexual intercourse’
<i>ḡaḏaʿ</i>	‘throw down’	<i>ḡāḏaʿ</i>	‘lie next to’
<i>nahaḏ*</i>	‘lift’	<i>nāhaḏ</i>	‘lift up the part of a dress covering the breasts’

FORM-III may also express action acted out on a patient rather than with them:

<i>faraḡ</i>	‘separate’	<i>fāraḡ</i>	‘leave [sb.]’
<i>naḏar</i>	‘see’	<i>nāḏar</i>	‘keep watch on, guard’

The petal sense is attested:

<i>manaʿ</i>	‘pardon’	<i>mānaʿ</i>	‘ask for pardon’
--------------	----------	--------------	------------------

Some verbs seem to be denominative:

<i>ʿaliḡ</i>	‘food bag for horse’	<i>ʿālaḡ</i>	‘chomp [horse]’
	(cf. <i>ʿallaḡ</i> ‘attach [a food bag]’)		
<i>ṣāḥeb</i>	‘friend’	<i>ṣāḥab</i>	‘become friends with’

A number of verbs have /aw/ or /ay/ instead of /ā/, and these are considered a subset of FORM-III in some grammars (e.g. FORM-IIIa in Johnstone 1967a: 45). I analyse them as quadrilaterals, and they are explained in the section below.

4.8.2.3 FORM-IV

While many modern dialects have lost FORM-IV (Kaye and Rosenhouse 1997: 293; Fischer and Jastrow 1980: 116), it is in common use in Najdi dialects (Ingham 1994a: 72) and related dialects, and indeed in the dialect of the Rwala. Its most common use is causative or factitive, especially *vis-à-vis* FORM-I. Among the causatives are:

<i>baʿad</i>	‘be far from’	<i>abʿad</i>	‘make distant’
<i>siher*</i>	‘wake up [intrans.]’	<i>ashar</i>	‘wake up [trans.]’
<i>ḥalaf</i>	‘swear an oath’	<i>aḥlaf</i>	‘make sb. swear an oath’

In some cases, both FORM-II and FORM-IV express the causative of FORM-I in different ways:

<i>ḥaḍar</i>	‘descend’	<i>ḥaddar</i>	‘lay down’
		<i>aḥdar</i>	‘throw down’
<i>gaʿad</i>	‘sit; wake up’	<i>gaʿʿad</i>	‘set down’
		<i>agʿad</i>	‘cause to sit up, rouse’

In the first example, FORM-IV expresses a more intensive meaning than FORM-II, while in the second example each form picks up on a different valency of the FORM-I meaning.

Intensive meaning is associated with FORM-IV more generally as well:

<i>ṭalab</i>	‘seek, ask for’	<i>aṭlab</i>	‘pursue’
<i>wadaʿ*</i>	‘put down’	<i>awdaʿ</i>	‘drive away’

Certain other meanings are sometimes expressed by FORM-IV. They include permissive:

<i>sanad</i>	‘go’	<i>asnad</i>	‘let sb. go’
--------------	------	--------------	--------------

And petal:

<i>fazaʿ*</i>	‘help’	<i>afzaʿ</i>	‘call for help’
---------------	--------	--------------	-----------------

And possibly ascriptive (no FORM-I is attested):

<i>aḡaz</i>	‘declare sb. incapable of doing sth.’
-------------	---------------------------------------

In other cases, FORM-IV simply imbues the verb with additional nuance:

<i>nakaf*</i>	‘return’	<i>ankaf</i>	‘return home without booty’
---------------	----------	--------------	--------------------------------

4.8.2.4 FORM-V

The predominant function of FORM-V is as a passive or reflexive of FORM-II, from which it is derived:

<i>ḍakkar</i>	‘remind’	<i>taḍakkar</i>	‘think of’ (be reminded of)
<i>ḡassal</i>	‘pour water for washing’	<i>taḡassal</i>	‘scrub oneself’
<i>zabban</i>	‘give protection to’	<i>tazabban</i>	‘be protected by’

In some cases the passivisation or reflexivisation is rather more obscure but the semantic development is marked by detransitivisation:

<i>sannad</i>	‘go up’	<i>tasannad</i>	‘lie on one’s back’
<i>‘assar</i>	‘trouble sb.’	<i>ta‘assar</i>	‘have a painful (troubling?) labour’

This usage is not restricted to modifying FORM-II:

<i>galab</i>	‘turn sth. over’	<i>tagallab</i>	‘turn from side to side’
--------------	------------------	-----------------	--------------------------

As will be apparent from the above examples, the semantic equivalence is sometimes hard to justify and may be verging on folk etymologies.

4.8.2.5 FORM-VI

Similarly to FORM-V and its relationship with FORM-II, the main function of FORM-VI is to modify FORM-III. Since FORM-III often expresses action with or on a patient, the result is that FORM-VI often expresses reciprocity in that action:

<i>ḥārab</i>	‘fight’	<i>taḥārab</i>	‘fight each other’
<i>ḥāwaš</i>	‘fight’	<i>tahāwaš</i>	‘have an altercation’
<i>lāga</i>	‘meet sb., have a duel’ ¹⁵¹	<i>talāga</i>	‘meet each other’

As an extension of this, FORM-VI frequently expresses communal activity:

<i>ḥafar*</i>	‘dig’	<i>taḥāfar</i>	‘bury sb. as a group’
<i>ga‘ad</i>	‘sit’	<i>tagā‘ad</i>	‘meet (sit) together’

Ingham (1994a: 83) notes that in Najdi dialects FORM-VI is commonly used to express gradual change. This is perhaps seen somewhat in Musil’s *Rwala* data, e.g.:

<i>hammal</i>	‘cry, well up’	<i>tahāmal</i>	‘become hazy with tears [eyes]’
---------------	----------------	----------------	------------------------------------

4.8.2.6 FORM-VII

One of the major functions of FORM-VII is as a passive for FORM-I, and functions alongside the internal passive:

<i>ḍabaḥ</i>	‘kill’	<i>(e)nḍabaḥ</i>	‘be killed’
<i>gatal</i>	‘kill’	<i>(e)ngatal</i>	‘be killed’

¹⁵¹ N.B. This verb is only attested in Musil as the verbal noun, *mulāga* ‘duel’.

These passives are frequently used in constructions which Holes (2016: 155) refers to as the potential passive, whereby *arđ tinwasem* refers to ‘ground which is able to be soaked by the rain’. FORM-VII does not always create true passives, however, and often it simply denotes an intransitive verb:

<i>ğada</i> ^c	‘throw down’	(<i>e</i>) <i>nğada</i> ^c	‘lie down’
<i>hazzam</i>	‘help sb. escape’	(<i>e</i>) <i>nhazzam</i>	‘escape’

It occasionally happens that the /n/ prefix is prefixed to an already intransitive verb, and the result is no change in meaning, e.g. *gaṭan* and (*e*)*ngaṭan* which both mean ‘camp in one place for a long time’.

Finally, it is interesting to note that in many related dialects the /n/ prefix may be added to FORM-II and FORM-III as well to form passives. This does not occur in Musil, and notably does not occur in the following example:

<i>kašaf</i> [*]	‘come out’		
<i>kaššaf</i>	‘uncover’	(<i>e</i>) <i>nkašaf</i>	‘be uncovered’

It would appear that in this case FORM-VII expresses the passive of FORM-II much more precisely than it detransitivises FORM-I, and yet the medial gemination of FORM-II does not occur.

The meaning of FORM-VII-t, as shown by the Najdi forms *yintalabbas* ‘he can/should be dressed, is dressable’ (Abboud 1979: 474) and *yintuwaggaf* ‘it is stoppable’ (Ingham 1994a: 74), is usually interpreted for Najdi dialects as a potential passive. It has therefore been suggested by Palva (1992: 25 note 32) that the meaning of *yintaṣeḥḥ*, translated by

Musil as ‘he was cured’, should be interpreted as ‘he could be cured’ in light of recorded Najdi usage.

4.8.2.7 FORM-VIII

FORM-VIII principally has a generally reflexive sense:

<i>mala</i> *	‘fill’	(<i>e</i>) <i>mtala</i>	‘be full’
<i>naha</i>	‘forbid’	(<i>e</i>) <i>ntaha</i>	‘cease from doing sth.’
<i>ra</i> ‘ <i>ad</i> *	‘thunder’	(<i>e</i>) <i>rta</i> ‘ <i>ad</i>	‘tremble’
<i>walad</i>	‘give birth to’	(<i>e</i>) <i>ttalad</i>	‘hail from’

This is also sometimes co-operative:

<i>farag</i>	‘part [sth. from sth.]’	(<i>e</i>) <i>ftarag</i>	‘part [from one another]’
--------------	-------------------------	----------------------------	---------------------------

FORM-VIII also frequently expresses a change of state:

<i>liheb</i> *	‘be burning’	(<i>e</i>) <i>ltahab</i>	‘become red hot’
<i>rafa</i> ‘*	‘lift up’	(<i>e</i>) <i>rtafa</i> ‘	‘reach a height’

For some roots, a particular nuance of meaning is expressed by FORM-VIII that is not predictable:

<i>ṣabaḥ</i> *	‘do sth. in the morning’	(<i>e</i>) <i>ṣtabaḥ</i>	‘have breakfast’
----------------	--------------------------	----------------------------	------------------

4.8.2.8 FORM-IX

FORM-IX is not well attested in the corpus. Holes says of Bahraini Arabic that in those dialects it does ‘occasionally occur in dialect poetry’ (Holes 2016: 158), and while it is true that most attestations in Musil are also poetic, the conversational use of *aṣḡarrat* ‘it turned

brown' to describe the process of making coffee shows that some forms are still in use, at least in certain technical capacities. It is unclear whether the functions of FORM-IX, predominantly expressing the subject being or becoming a certain colour, physical defect, or other characteristic¹⁵², have been taken over by other forms, since other verbs expressing e.g. colour are not attested.

4.8.2.9 FORM-X

Several semantic categories are expressed by FORM-X. The reflexive is frequently attested:

<i>ġafar</i> [*]	'reopen a wound'	(i) <i>staġfar</i>	'worsen [wound]'
---------------------------	------------------	--------------------	------------------

The reflexive sense may be benefactive, which is to say that the subject does something for their own benefit, or petal:

<i>šireb</i>	'drink'	(i) <i>stašrab</i>	'demand water' or 'drink one's fill'
<i>faza</i> ^{˘*}	'help'	(i) <i>stafza</i> [˘]	'ask for help'

There is a transformative sense expressed by FORM-X, in which the subject experiences a change of state as a result of the actions of another agent:

<i>baššar</i>	'bring good news'	(i) <i>stabšar</i>	'rejoice' (i.e. after receiving good news)
---------------	-------------------	--------------------	--

¹⁵² Although the categories of colour and physical defect are traditionally held to be the remit of FORM-IX verbs, it has been demonstrated that in Standard Arabic colour makes up half of attested verbs and physical defects only 7%; see Eisenstein (1990), cited in Watson (2002: 140).

Finally, as in other verb forms, sometimes FORM-X is used to express a particular or technical meaning:

<i>ağğal</i>	‘urge sb. to speed’	(i) <i>stağal</i>	‘induce labour’ (i.e. hasten the birth)
<i>dana</i>	‘touch’	(i) <i>stadna</i>	‘reach out for’

4.9 Phonological adjustments

The dialect has a number of biliteral verbs which are described in the next section. Some trilateral verbs, however, show adjustments on the basis of their root consonants in both FORM-I and the derived forms, but do not produce a biliteral root. These include verbs that have guttural fricatives in the root, verbs that have either /w/ or /y/ as C1, and verbs in which C2 was historically /ʔ/. Verbs that historically had /ʔ/ as C1 may be either biliteral or trilateral in the dialect, and these are described in the next section. When C3 was historically /ʔ/, they pattern like the final-weak verbs.

4.9.1 C1 = *guttural*

If C1 is guttural (/ğ, ħ, ʕ, h/) then resyllabification may occur in the P-STEM on account of the *ghawa*-syndrome (cf. §2.3.1); interestingly, while Ingham (1994a: 24) and Prochazka (1988: 36) describe this as obligatory, it is only a minority of verbs in Musil’s corpus which exhibit this. The resyllabification is CaX to CXa (where X is a guttural consonant). In Musil, this leads to P-STEM forms such as *yʕarfū* ‘they know’ and *nħadar* ‘we go down’, but this process does not occur in the majority of cases.

4.9.2 C1 = /w/ or /y/

In FORM-I there is no change in the S-STEM, since the /w/ or /y/ behaves as any other consonant word initially. In the P-STEM the situation is inconsistent. Compared with a hypothetical *yawCVC, derived from the strong verb pattern yaCCVC, three options are available. Firstly, the form may retain, as in *yowda* ‘it becomes’ (*a* > *o* under the influence of the /w/), which in turn may be monophthongised to *yūCVC* in a minority of cases, such as *yūgef* ‘he ceases’ (but interestingly never to *yōCVC*, although *ō* is the usual monophthong which varies with *aw/ow*). Alternatively the /w/ may be deleted, with *yawCVC* > *yaCVC* as in *yaret* ‘he inherits’. In the third case, which is least represented in the corpus, *aw* > *ā* produces *yāCVC* as in *yāšel* ‘he reaches’. The imperatives demonstrate aphaeresis e.g. *wašal*, *yāšel* ‘reach’ which has the MP imperative *šelū*. It is frequent for verbs in this category that the FORM-I imperative is substituted with a FORM-II imperative, e.g. *waššelū*.

The situation in the derived forms is much the same as for FORM-I, with most forms unaffected. FORM-IV undergoes similar changes to those in FORM-I, with a diphthong formed in the S-STEM, e.g. *awda* ‘he prepared’, and either diphthongs or long vowels in the P-STEM, e.g. *yowšel* ‘he carries’ or *yūgef* ‘he stops’ (note that this form is identical to the FORM-I *yugef* above). In FORM-VIII the glide is eliminated and gemination is triggered, hence *ettaga* ‘he went into hiding’ (< **ewtaga*) and *ettelād* ‘being from somewhere’ (< **ewtelād*). In one instance *tifāg* is attested for **ettefāg* ‘thoroughly, harmoniously’. FORM-X is only attested once

for this group, and in this instance the hypothetical diphthong is simplified and shortened:

testafi ‘you recover [a debt]’ rather than **testāfi* or **testawfi*.

4.9.3 C2 = /ʔ/

These verbs are very poorly attested in the corpus; in fact, many verbs of this type have been lost from the dialects altogether (cf. Sawayan 1992: 79). For example, the verb *sāl* (from root **sʔl**) ‘he asked’ is rare, and *yensəd* ‘he asks’ (from root **nšd**) is common. In the S-STEM of FORM-I all examples pattern like verbs where C2 is /w/ or /y/ (i.e. the hollow verb, cf. §4.10.2), e.g. *sāl*, *selt* ‘he asked, I asked’. The form *selt* is presumably the surface form of an underlying /silt/, which again represents the convergence of the C2 = /ʔ/ paradigm and the hollow verbs paradigm; it is possible that /salt/ < /saʔalt/ is the underlying form, although since /aʔa/ sequences invariably surface as /ā/ in the dialect this is rather unlikely. The P-STEM, however, is completely unattested, and a form such as **yasāl/yasūl/yasīl* following the hollow verb pattern is completely hypothetical. Medial glottal stops are expressed as long vowels (e.g. *bīr* ‘well’ < /biʔr/) but the sequence CʔVC, as would be analogous to the P-STEM morphology, is also completely absent from the corpus. The only comparable word is *malak* ‘angel’, where *malak* < /malʔak/ (at least historically), with no vowel lengthening; by extension, the P-STEM of *sāl* could also have a short vowel. The verb ‘to see’ from root **rʔy** (it is not used in either the P-STEM the S-STEM) is used twice as an imperative, once as *erʔow* ‘see! 2MP’ and once as *erʔ-eh* ‘see him! 2MS’; in the first instance

the glottal stop is heard (at least, according to the transcription), and in the latter it is expressed as /^ʕ/, a phenomenon found also in the Šammari dialects (Sowayan 1992: 80). One could therefore speculate as to whether the P-STEM of *sāl* would be *yasāl* like a hollow verb, *yasal* like *malak*, *yasʕal* like *erʕow* or *yas^ʕal* like *er^ʕ-eh*. Practically, it is more likely that the form does not exist, and that the verb exists only as a suppleted form, occurring in the S-STEM but having **nšd** in the P-STEM (**nšd** does not occur in the S-STEM). Suppletion of this kind is remarkably seen in every attested C2 = /^ʕ/ verb, such as **rʕy** which is only used as an imperative, and has **šwf** in the P-STEM and the S-STEM, and **ṭr** which occurs as 3MS but which occurs in FORM-V for other persons, e.g. *teṭārayt* ‘I avenged’, or is expressed periphrastically, e.g. *aḥaḍt ṭār* ‘I took vengeance’.

Where these verbs are attested in the derived forms, the /^ʕ/ surfaces as a glide. Thus in FORM-II we have *layyemna* ‘we assembled’ < **laʕemna*, and in FORM-III *masāyel* ‘dreading’ < **masāʕel*. In FORM-IV the form is *arwa* ‘he revealed’ < **arʕa*.

4.10 Biliteral verbs

The dialect has four types of biliteral verbs, which differ from the trilateral verbs so far described in a number of ways. The verbal stems are different, and accordingly they have their own templates, as will be seen. The “deficient verbs”, such as *ḥaḍa* ‘he took’ are those in which C1 was historically a glottal stop; these surface as both trilateral and biliteral variants in Musil’s corpus. “Hollow verbs”, such as *zāl* ‘he vanished’, lack a medial

consonant, while “final-weak verbs”, such as *rama* ‘he threw, shot’, lack a final consonant.

“Doubled verbs”, such as *hazz* ‘he beckoned’, have a geminate C2.

4.10.1 *Deficient verbs*

These verbs historically had a glottal stop as C1. In the S-STEM, not all forms are attested, but from those that are it can be seen that there are trilateral and biliteral variants for these verbs. The trilateral variants pattern like the strong verb, with the exception that there is no resyllabification when vowel-initial suffixes are added (*akalat* ‘she ate’, not **kalat*). The biliteral variant can be seen to pattern similarly to other biliteral verbs, specifically the final-weak and doubled verbs, in the diphthongal form *kaleyna* ‘we ate’. There are no attested verbs in the I-CLASS, although Prochazka does record the form *amin* ‘he was safe’ for Rwala. The paradigm given is for *akal* ~ *kala* ‘to eat’.

In the S-STEM:

3MS	<i>akal</i> or <i>kala</i>	3MP	<i>akalow</i>
3FS	<i>akalat</i>	3FP	<i>kalen</i>
2MS	n/a	2MP	n/a
2FS	n/a	2FP	n/a
1CS	<i>akalt</i>	1CP	<i>akalna</i> or <i>kaleyna</i>

In the P-STEM the deficient C1 causes vowel lengthening of the prefix vowel, which is always /ā/. The verbal stem is CVC, and as with the strong verb, V is lexical. Only the o variant is given here but forms with *a* and *e* occur as well. The paradigm is for *yāhoḍ* ‘he takes’:

3MS	<i>yāḥoḍ</i>	3MP	<i>yāḥeḍūn</i>
3FS	<i>tāḥoḍ</i>	3FP	n/a
2MS	<i>tāḥoḍ</i>	2MP	n/a
2FS	<i>tāḥeḍi</i>	2FP	n/a
1CS	<i>āḥoḍ</i>	1CP	<i>nāḥoḍ</i>

Prochazka records the Rwala as only having the form *ikal* rather than *akal*, but *i* doesn't appear a single time in Musil, and Cantineau gives *ekal* for Rwala which is indeterminate. Cantineau also gives a variant form *hakal* which he describes as unique to the Rwala, although Prochazka noted this and writes that his speakers did not recognise it (Prochazka 1988: 67). A great many different surface forms of this verb are attested even in very closely related dialects: we find *kala* in Ḥāyil, *'akal* in al-Qaṣīm, and *kal* in Riyadh (forms from Prochazka). Furthermore Prochazka himself records *'aḥaḍ* for Rwala where Musil has *aḥaḍ*, showing that the form aCaC is not unattested elsewhere for C1 = /ʔ/ verbs. It may therefore be assumed that the irregularity of /ʔ/ has produced a range of realisations of C1 = /ʔ/ verbs, and discrepancies between Musil's and others' accounts should not raise concern about the accuracy of his data.

Both Prochazka and Cantineau list the biliteral forms for neighbouring dialects but not for Rwala; interestingly, however, the biliteral forms they do record do not have the diphthongised suffixes as seen in *kaleyna*. I have only found these forms in Montagne, who records *ḥaḍeyna* 'we took' (MŠa, VII, 3 cited by Cantineau), but since this type of levelling occurs with final-weak and doubled verbs as well, it is not unusual for the levelling to have extended across the biliteral verbs in general. Cantineau also writes that there is not an A-

CLASS in the P-STEM, and Musil's one instance of *āṭar* 'I vex' may not be sufficient evidence to contradict this.

The participles are *āḥed* 'taking MS' and *māḥūd* 'taken MS'; *ma'ḥūd* is also attested but it seems very unlikely that the dialect would support the pronunciation of a glottal stop in this position. The imperative is *ḥoḍ* 'take! MS' which matches Cantineau's and Prochazka's data. Cantineau records that these pattern like the imperatives of hollow verbs, i.e. *kel*, *kūlū* 'eat! MS, MP', but suffixed imperatives are not attested in Musil. Prochazka records short vowels however, i.e. *kil*, *kulū*. Internal passives are well recorded in the S-STEM, where /ʔ/ surfaces as /w/, e.g. *wuḥid*, *wuḥidat* 'he/she was captured' (the latter shows lack of syncope, since **wuḥdat* would be expected). This is possibly part of a general trend in which internal passive forms which contrast more clearly with their active counterparts are better attested than those which do not; it can thus be theorised that the /w/, which is present only in the passive forms in FORM-I, distinguishes passive verbs from active ones in this category enough to render them productive.

In the derived forms the situation is rather diverse. In FORM-II, in the S-STEM, C1 surfaces as /w/, e.g. *wahḥar* 'he abandoned', with only one exception (*addaw* 'they returned sth.');

this /w/ features also in FORM-V, e.g. *twaḥḥar* 'he slid backwards in the saddle'. In the P-STEM, however, the /ʔ/ resurfaces in forms such as *t'awwel* 'you lead back', which is not attested as **twawwel*. This occurrence appears to be conditioned by the C2 /w/, and may be an effect of the Obligatory Contour Principle:

OBLIGATORY CONTOUR PRINCIPLE:

In a given autosegmental tier, adjacent identical autosegments are prohibited
(McCarthy 1979a: 238)

This hypothesis was first proposed to account for distributional regularities in tonal systems (Leben 1973) and was reformulated by McCarthy (1979a) in an analysis of Arabic roots. It proposes that successive identical elements are banned, which is demonstrated by McCarthy (1986: 209) for the Arabic verbal stem *samam* (cf. CLA *samamtu* ‘I poisoned’). *samam* is analysed as having a consonantal melody which is not /smm/ but rather /sm/, which is associated with a template on a left-to-right basis with the rightmost element spreading to unfilled positions:

CaCaC → *samam*

This proscribes any association which would produce a stem **sasam*, a result which concurs with the evidence from Arabic in which **sasam* is not a permitted stem in any context (except possibly loan words). McCarthy (*ibid.*) shows that this is restricted to single morphemes, as shown by the fact that **tatak* is not permitted as a verbal stem but occurs in *tatakallam* ‘you converse’, where each /t/ belongs to a different morpheme. The application of the Obligatory Contour Principle in the case of *t'awwel* would suggest that a consonantal melody /wl/ would produce a form **twallel* rather than **twawwel*, since only the rightmost consonant may spread to unfilled positions in the template. It would therefore seem that the change /ʔ/→/w/ usually seen in these verbs is blocked by the following /w/, since /wwl/ is not a permitted consonantal melody.

In one case in the corpus, coalescence occurs and **neʔaddi* ‘we are returning sth.’ surfaces as *neddi*. The only attested participle is *mʔawwel*, ‘leading back’, where again the change /ʔ/→/w/ has probably been blocked by the Obligatory Contour Principle. The /w/ also surfaces in the FORM-IV S-STEM, e.g. *awda* ‘he caused a nuisance’; the only example of a P-STEM in this form is passive, *tūteh* (for **tūta*) ‘you are offered [sth.]’. FORM-X is attested only in the P-STEM and again the situation is mixed: the forms are variously *yistaysar* ‘he seizes sb. as a captive’, *yistaʔhel* ‘he deserves’ or *yistāhel* ‘he deserves’.

4.10.2 *Hollow verbs*

The so-called “hollow” verbs are traditionally understood as having a medial glide. In the P-STEM the verbal stem is $C\bar{V}C$, where \bar{V} is lexically governed, either \bar{a} , \bar{i} or \bar{u} . In the S-STEM the verbal stem is $C\bar{a}C$ before zero and vocalic suffixes, and CVC before a consonantal suffix, where the nature of V is rather complex. It appears that at some point V represented the quality of the underlying glide, whereby a verb with the P-STEM verbal stem $C\bar{u}C$ would have the verbal stem CuC in the S-STEM, giving the following three principal parts:

gāl, gult, yagūl
 he said, I said, he says

The dialect as recorded by Musil shows a partial merger of the CuC stem into the CiC stem, however, and his data predominantly show:

gāl, gilt, yagūl
 he said, I said, he says

As such, while CūC stems account for 48% of all hollow verbs in the P-STEM, CuC stems account for only 25% of hollow verbs in the appropriate forms of the S-STEM (i.e. before consonantal suffixes). Cantineau (1936-37: 196) does not provide Rwala data on the hollow verb, but the dialects for which he does give data fit into the former category (*ibid*: 91). Prochazka, on the other hand, writes that Rwala has merged CuC entirely into CiC (Prochazka 1988: 77). On this basis, it seems possible that Musil created a record of the dialect at a time when the merger was only partially complete: while *gāl* ‘he said’ shows a marked preference for *gilt* (86%), the common verb *šāf* ‘he saw’ shows a very even split with *šuft* in 56% of cases and *šift* in 44%. The merger is evidenced in all parts of the verb except 2MP (2FS, 2FP and 3FP are, however, unattested), which may simply be because the 2MP is poorly attested, or may possibly indicate that this form was more resistant to the change, perhaps influenced by the /u/ of the suffix.

Some verbs have the verbal stem CāC in the P-STEM. In these cases, the underlying root consonant may be either /w/ or /y/, according to the traditional analysis. In CLA, these verbs have the stem CiC before consonantal suffixes in the S-STEM, a feature which is shared by the Rwala dialect. The paradigms, divided on the basis of the vowel in the P-STEM verbal stem, are given for *nām* ‘to sleep’, *šār* ‘to become’ and *gāl* ‘to say’.

In the S-STEM:

	/a/	/i/	/u/
3MS	<i>nām</i>	<i>ṣār</i>	<i>gāl</i>
3FS	<i>nāmat</i>	<i>ṣārat</i>	<i>gālat</i>
2MS	<i>nimt</i>	<i>ṣirt</i>	<i>gult/gilt</i>
2FS	n/a	n/a	n/a
1CS	<i>nimt</i>	<i>ṣirt</i>	<i>gult/gilt</i>
3MP	<i>nāmow</i>	<i>ṣārow</i>	<i>gālow</i>
3FP	n/a	<i>ṣaren</i>	n/a
2MP	<i>nimtu</i>	n/a	<i>gultu</i>
2FP	n/a	n/a	n/a
1CP	n/a	n/a	<i>(gulna)/gilna</i>

In the P-STEM:

	/a/	/i/	/u/
3MS	<i>yanām</i>	<i>yaṣīr</i>	<i>yagūl</i>
3FS	<i>tanām</i>	<i>taṣīr</i>	<i>tagūl</i>
2MS	<i>tanām</i>	<i>taṣīr</i>	<i>tagūl</i>
2FS	<i>(tanāmi)</i>	<i>(taṣīri)</i>	<i>t(a)gūli</i>
1CS	<i>anām</i>	<i>aṣīr</i>	<i>agūl</i>
3MP	<i>(yanāmūn)</i>	<i>yaṣīrūn</i>	<i>yagūlūn</i>
3FP	<i>(yanāmen)</i>	<i>yaṣīren</i>	<i>(yagūlen)</i>
2MP	<i>t(a)nāmū(n)</i>	<i>(taṣīrūn)</i>	<i>t(a)gūlūn</i>
2FP	<i>(tanāmen)</i>	<i>(taṣīren)</i>	<i>(tagūlen)</i>
1CP	<i>(nanām)</i>	<i>naṣīr</i>	<i>nagūl</i>

Where whole forms are enclosed in brackets, the form is not attested and what is given is a reconstruction. Where individual letters are in brackets, they are missing from the attestations of that particular form but present across the paradigm. It can be seen that in several forms the *a* of the prefix is not attested. This reflects the fact that the prefix (*ya*, *ye*, *yi*, etc.) is only targeted for syncope when its syllable is open, hence syncope of the prefix is attested for the hollow verbs.

The participles are CāyeC and maCyūC, and the imperatives are CVC but CVC̄ with suffixes, e.g. *gum* ‘stand! MS’, *gil* ‘say! MS’, *gūlū* ‘say! MP’, *hātū* ‘bring! MP’ and *šīrū* ‘become! MP’. These forms are the same as in Cantineau and Prochazka. The only exception is the imperative *hāt* ‘bring! MS’ which always has a long vowel, although this lexeme may have been reanalysed as an interjection. Internal passive forms are well attested in the S-STEM third person where the long vowel is /ī/, e.g. *gīl* ‘it was said’, *gīdat* ‘she was led. As with the deficient verb, it seems possible that the clearly differentiating nature of this /ī/ in the passive voice has led to its high productivity, especially since second and third person forms, which would be less distinctive, are not attested. P-STEM passives such as *yīḏām* ‘he will be injured’ are considerably rarer, probably because they are not easily distinguishable from the active paradigm.

In principle, the same phonological changes occur in the derived forms. Forms II-III are unaffected, although in one (anomalous) case *zānat* ‘she adorned’ occurs alongside *zayyanat*, as though it were a FORM-I verb. In FORM-IV, as in FORM-I, the stem vowel is long in the third person and short elsewhere, hence *arād* ‘he wanted’ but *aradna* ‘we wanted’. The vowel is long /ī/ in the P-STEM, e.g. *yerīd*, *yerīdūn*, in the imperative, e.g. *ewīgū* ‘examine well! MP’, and in the participles, e.g. *mḡīmīn* ‘staying in a place MP’. FORM-VI shows the same variation as FORM-III, although it is rare: in most cases the /w/ or /y/ is retained, as in *tahāwašow* ‘they fought’ and *tahāyag* ‘he ascended’, but forms such as *thāb* ‘he feared’ (< **tahāyab*) and *tagārow* ‘they rode at full gallop’ (< **tagāwarow*) do occur. It is possible that

these forms are not, in fact, FORM-VI, but rather represent a novel class for the hollow verbs of detransitivisation through the /ta/ prefix on FORM-I. FORM-VII has /ā/ throughout, although the first and second persons of the S-STEM are not attested and this could be short: *enlāh* ‘it waned’, *yinhāš* ‘he puts himself under protection’ and *men‘āğ* ‘swaying’. FORM-VIII and FORM-X pattern like FORM-VII.

4.10.3 *Final-weak verbs*

The final-weak verbs have a vowel-final stem, and as such have a different set of paradigms to the strong verb; verbs which etymologically have C3= /ʔ/ are also in this group. In Rwala the stem-final vowel is short in unsuffixed forms on account of word-final vowel shortening (cf. §2.3.3). In the S-STEM there is only one class, with the following paradigm given for *laga* ‘to meet’:

3MS	<i>laga</i>	3MP	<i>lagow</i>
3FS	<i>lagat</i>	3FP	<i>lagen</i>
2MS	<i>lageyt</i>	2MP	<i>lageytu</i>
2FS	<i>lageyti</i>	2FP	n/a
1CS	<i>lageyt</i>	1CP	<i>lageyna</i>

This in itself is remarkable. From Prochazka and other sources one would expect to see two classes, one resembling the one shown above and another with a stem vowel *i*, which would manifest as CVCi in the bare form and e.g. CVCit, rather than CaCeyt, for the 1CS. It may be that such forms are simply poorly attested in Musil, but some verbs seem to indicate that in fact they have merged with the A-CLASS, e.g. ‘to meet’ which is attested once

as *lagīna* in the 1CP, which would be I-CLASS, but otherwise appears as *lageyt* in other forms, which is clearly A-CLASS.

The apparent absence of an I-CLASS for the final-weak verbs is very unusual. The closest parallel to this may be Palva's description of the dialect of the Bani Şaḥar:

“the C₃y verbs follow the well known pattern *ḥaća, ḥaćat, ḥaćeť, etc., ruma, rumat, rumêť, etc., miša, mišat, mišêť, etc., gara, garat, garêť, etc., but nisi, nisyat/nisat, nisêť*” (Palva 1980: 128).

From the phrasing, we might deduce that the final category (the I-CLASS) is rare, which may give precedent for its absence from Musil's data. Palva also supplies *nisat* as a variant of *nisyat*, which is identical to the A-CLASS form and absent from Prochazka, and his form *nisêť* similarly conforms to the A-CLASS pattern in contrast with the latter's *nisit*, meaning that the I-CLASS is identical to the A-CLASS in most forms. This interestingly suggests that some forms attested in Musil could belong to an I-CLASS but not be definitively marked as such, meaning that Musil's system simply attests a similarly underspecified morphology of final-weak verbs, rather than having an unprecedented single-class paradigm in the S-STEM. This analysis is supported by the evidence from the Dōsiri dialect, where the I-CLASS is described as comprising “alternate forms additional to” the A-CLASS (Johnstone 1961: 261).

In the P-STEM, the verbal stem is CCV without suffixes, where V is a lexically governed vowel /a/ or /i/, and simply CC before suffixes. The paradigms are therefore:

	/a/	/i/
3MS	<i>yalga</i>	<i>yamši</i>
3FS	<i>talga</i>	<i>tamši</i>
2MS	<i>talga</i>	<i>tamši</i>
2FS	<i>talgi</i>	<i>tamši(n)</i>
1CS	<i>alga</i>	<i>amši</i>
3MP	<i>yalgūn</i>	<i>yamšūn</i>
3FP	<i>yalgen</i>	<i>yamšen</i>
2MP	<i>talgūn</i>	<i>tamšūn</i>
2FP	<i>(talgen)</i>	<i>(tamšen)</i>
1CP	<i>nalga</i>	<i>namši</i>

The verbs with stem vowel /a/ are substantially lower in occurrence than those with /i/, accounting for only 27% of unsuffixed forms (forms with suffixes appear the same in both classes). Nonetheless, they are certainly a class of their own, since verbs in the A-CLASS surface only in that class, and the same for the I-CLASS. Ingham claims that in Najdi dialects the S-STEM A-CLASS corresponds to the P-STEM I-CLASS, and the S-STEM I-CLASS to the P-STEM A-CLASS (Ingham 1994a: 26). This does not appear to be the case for Musil's corpus, where the S-STEM I-CLASS is almost unattested; this is also unsurprising since Ingham makes the same claim about the strong verb and that does not apply to Musil's Rwala data either. Furthermore, the concept of strong correspondence between the stem vowels is challenged by the attestation in Musil of *dara, yadri* 'to know' and *dara, yidra* 'to care about', where the two verbs are identical in the S-STEM. Moreover, Kurpershoek (2005: 90) records the verbs as *diri, yadri* 'to know' and *dara, yadra* 'to care about', where the correspondence is of A-CLASS to A-CLASS and I-CLASS to I-CLASS.

The active participle ends in /i/, e.g. *bāgi* ‘wanting MS’; in this case there is no final consonant to associate with the final slot in the template, and an interesting effect of this in Rwala is that the vowel surfaces as *i*, rather than as *e* which surfaces in the strong verb. Suffixed forms have a glide, e.g. *nābiye* ‘bulging FS’. Passive participles are unattested, and the imperative is not attested enough to provide a clear picture. Prochazka records only monosyllabic forms of the imperative, e.g. *ʔins* ‘forget! MS’, which is attested in Musil too, e.g. *erʕ* ‘behold! MS’: this loss of the final vowel of the imperative is considered typical of eastern bedouin dialects (cf. Rosenhouse 2011). In Musil the final vowel is often retained, however, and these forms may have either /a/ as in *erʕa* ‘behold! MS’ or /i/ as in *emši* ‘go! MS’. Suffixed forms are of the type *emšū* ‘go! MP’, where the CC stem can shed no light on the classes. Passive forms are attested in the S-STEM with both /i/ and /u/ melodies: *bini*, *binyat* ‘it (m./f.) was built’, *ǧufa* ‘it was snatched’. Here the /u/ melody matches the rest of Musil’s data; since the form *ǧufa* is unique it is possible that it also represents an /i/ melody which has been backed before the labial /f/, but such a change is not recorded well enough in Musil’s data to support this interpretation, and it does not account for the final *a*.

In the derived forms the final vowel is /a/ in the S-STEM, with diphthongs in the second and third persons, giving forms such as *lagga* ‘he met’ and *laggeyt* ‘you met’, *agfat* ‘she went back’ and *agfeyt* ‘you went back’. In the P-STEM and participles, FORM-II shows the same active/passive contrast as it does in the strong verb, namely, between the active *a-i* melody and the passive *a-a* melody, e.g. *yeʕaddi* ‘he pursues’ and *yeʕabba* ‘it is filled’: as with the

FORM-I participle, this paradigm sees *i* surface in the active forms where the strong verb always has *e* (compare *ye‘addi* and *yebaddel*). Imperatives in the derived forms attest the loss of the final vowel more regularly than the FORM-I examples in this corpus, e.g. FORM-II *semm* ‘name [it]!’ and FORM-VIII (*e*)*štar* ‘buy!’.

4.10.4 *Doubled verbs*

These verbs have a stem which ends in a geminate consonant. The verbal stem is monosyllabic, and is always CaC_iC_i in the S-STEM and CVC_iC_i in the P-STEM, where the quality of *V* is encoded in the root, as in the strong verb, and C_iC_i is a geminated consonant. The paradigm given is for *radd* ‘to return sth.’.

In the S-STEM:

3MS	<i>radd</i>	3MP	<i>raddow</i>
3FS	<i>raddat</i>	3FP	<i>radden</i>
2MS	<i>raddeyt</i>	2MP	<i>raddeytu</i>
2FS	(<i>raddeyti</i>)	2FP	(<i>raddeyten</i>)
1CS	<i>raddeyt</i>	1CP	<i>raddeyna</i>

The P-STEM effectively has three classes, which are the same but for the underlying vowel. Some variation of this lexical vowel is seen in the doubled verbs, with around 20% of roots showing variation compared with less than 4% for the strong verb. Interestingly, in the strong verb underlying /i/ and /u/ never surface as *i* or *u*, but only as their allophones *e* and *o*; this is not true of the doubled verb where the more close variants co-occur with their more open counterparts. This may be because the second syllable, where the stem vowel

occurs, is stressed in this paradigm. The unstressed nature of the first syllable can be seen also by the frequent occurrence of *e* in the prefix to the exclusion of *a*, if it is possible that this represents reduction of the vowel. Syncope is attested in the second person (but not in 3FS). The paradigms for ‘*add* ‘to tell’, ‘*šadd* ‘to saddle’ and *radd* are:

	/a/	/i/	/u/
3MS	<i>ye‘add, ye‘edd</i>	<i>yešedd, yešidd</i>	<i>yerudd, yerodd</i>
3FS	<i>te‘add, te‘edd</i>	<i>tešedd, tešidd</i>	<i>terudd, terodd</i>
2MS	<i>t‘add, t‘edd</i>	<i>tšedd, tšidd</i>	n/a, <i>trodd</i>
2FS	n/a	n/a	n/a
1CS	<i>a‘add, a‘edd</i>	<i>ašedd, n/a</i>	<i>arudd, arodd</i>
3MP	<i>ye‘addūn, ye‘eddūn</i>	<i>yešeddūn, yešiddūn</i>	n/a, <i>yeroddūn</i>
3FP	<i>ye‘adden, n/a</i>	n/a, <i>yešidden</i>	<i>yerudden, n/a</i>
2MP	<i>t‘addūn, t‘eddūn</i>	<i>tšeddūn, n/a</i>	<i>truddūn, n/a</i>
2FP	n/a	n/a	n/a
1CP	<i>ne‘add, ne‘edd</i>	<i>nešedd, nešidd</i>	<i>nerudd, nerodd</i>

Note that the forms shown here are illustrative and do not attest to the coincidence of prefix-vowels and stem-vowels, which are variable.

It is curious that these data do not conform to the expected Najdi system in the P-STEM. Ingham’s Najdi data (Ingham 1994a: 25) concur with Prochazka’s Rwala data (Prochazka 1988: 57), which give *yírid* ‘he returns sth.’ in place of Musil’s *yerudd*. The form *yírid* has the stress on the first syllable, which triggers degemination of the final consonant; this contrasts with suffixed forms in which, according to their data, the gemination is maintained, as in *yiriddūn* ‘they return sth.’. Although the stress placement is not known for Musil’s corpus, the syncope of the first syllable which frequently occurs throughout the paradigm, and the total absence of degemination of the final consonant, suggest that the

stress falls on the second syllable in, e.g. *yerúdd*. Prochazka also records the 1cs prefix with /a/-raising as *i-* while Musil has only recorded *a-*. The Rwala system presented here closely resembles Cantineau's 'Omur data (Cantineau 1936-37: 86) and eastern dialects such as that of Hofūf (Prochazka 1988: 57) and those recorded by Johnstone (1967a: 46). It is also interesting that these writers note only the existence of an I-CLASS and a U-CLASS, but not an A-CLASS as is found in Musil.

Historically, these verbs were mapped onto the same template as the strong verb, and were disyllabic before consonantal suffixes (Watson 2002: 146); compare CLA *radadna* 'we drove [them] back' with Rwala *raddeyna*. A 'Double Rule' described by Moore (1990: 81) produces a monosyllabic base with a geminate consonant before vocalic suffixes; compare CLA *radda* 'he drove [them] back' with Rwala *radd*. Here the Rwala dialect, and indeed all modern dialects of Arabic, preserve only the monosyllabic base in all parts of the verb. That the S-STEM shows only one vocalic melody, CaCC, as opposed to the two classes of the strong verb, is evidence that this process has been lexicalised and has produced a unique template for the doubled verbs.

Participles have the shape CāCC in the active and maCCūC in the passive, e.g. *bārr* 'feeling [for somebody] (MS)' and *mašdūd* 'saddled (MS)'. In this, the dialect resembles CLA *mādd^{un}* 'stretching (MS)' rather than dialects where the Double Rule does not apply, e.g. Cairene *mādid*. I have not found any descriptions of this participle for other north-Arabian dialects, but Johnstone gives active *mādd* and passive *mamdūd* for the east-Arabian dialect

group so the form is not unattested. There is one example where Musil appears to have recorded shortening of the final geminate consonant, in *mā šāḥ le-ha ḥatta* ‘a second drink of water has not taken place for them’. It seems possible that this kind of shortening may occur regularly when CāCC is not followed by a vowel. The imperative is CVCC where V is the same as in the P-STEM: ‘*add* ‘tell! MS’, *šidd* ‘saddle! MS’, *rodd* ‘drive [camels] together! MS’. Prochazka only records imperatives in /i/ for Rwala, but this matches his data for the P-STEM which, as we saw, are very different to what is described in Musil, since all three classes are very well attested in Musil. The passive is attested in the S-STEM with a /u/ melody, e.g. *ḥull* ‘it was refused’. Doubled verbs do not show unpredictable behaviour in the derived forms.

4.11 Quadriliteral verbs

There are a number of verbs in all varieties of Arabic which have four consonants in the lexical root, rather than two or three. These verbs have the template CVCCVC which is shared with the FORM-II trilateral verbs. The paradigm for the verb *darham* ‘to trot (camel)’ is as follows:

S-STEM	<i>darham</i>
P-STEM	<i>yedarhem</i>
imperative	<i>darhem</i>
participle	<i>mudarhem</i>
verbal noun	<i>derhām</i>

All of these forms concur with the forms given by Prochazka, although he does not provide participles or verbal nouns, and it should be noted that the verbal noun is very poorly attested in Musil as well. It goes without saying, however, that the quadrilateral verbs, while identical in structure to the FORM-II trilateral verbs, cannot have the same verbal noun (taCCiC) since this only accommodates three root consonants; it is also unsurprising that the verbal noun is formed, as it is in many other derived forms, simply through lengthening of the final vowel and application of the vocalic melody *e-a*. A second qualification is that the form of the P-STEM (and the imperative and participle forms which are synchronically derived from it) is less clear cut than FORM-II in having the vocalic melody *a-e* in the stem: a sizeable minority of verbs have either *a-a*, e.g. *tahalhal* ‘she waves her hair’, or *e-e*, e.g. *tederhem* ‘she trots’.

Quadrilateral roots arise for a number of reasons. Firstly, some are derived from quadrilateral nominals, for example *‘ašrag* ‘to be/become blue’, which is known from CLA and is traditionally held to be a denominal verb from the name of a plant, *‘išriq*. The verb *mar‘ad* ‘he thundered’ may be derived from the attested word *mar‘id* ‘thunderer’, although this may represent a rare case where a word has been invented by the poet, as both *mar‘ad* and *mar‘id* are obscure. Some more direct denominal derivations include:

<i>meršeḥa</i>	‘horse saddle’	<i>maršaḥ</i>	‘saddle [a horse]’
<i>ma‘rega</i>	‘light horse saddle’	<i>ma‘rag</i>	‘saddle [a horse]’
<i>zağarīt</i>	‘cry of joy’	<i>zağrat</i>	‘cry for joy’
<i>nowmas</i>	‘glory’	<i>nowmas</i>	‘lead to glory’

The first two of these verbs are derived from nouns which would be analysed as having trilateral roots, since the /m/ is a prefix; the verbs, however, are truly quadrilateral.

Other quadrilateral verbs are derived directly from both biliteral and trilateral verbs.

For the biliteral roots, this often involves reduplication:

<i>farr</i>	‘fly away’	<i>farfar</i>	‘lift one’s wings’
<i>hall</i>	‘loosen one’s hair’	<i>halhal</i>	‘wave one’s hair’
<i>hazz</i>	‘shake’	<i>hazhaz</i>	‘shake hands’

For the hollow verbs especially, /w/ or /y/ becomes C2 of the root and C1 is reduplicated, as in the case of *lawlah* ‘flutter’, which appears to be related to the noun *melwāh* ‘a cloth waved about to coax a falcon’, to which we may compare CLA *lawwaḥa* ‘to wave a thing about’ and *milwāh^{un}* ‘decoy bird’ (Hava 1951: 700). Another example is *zeyzam* ‘shield sth./sb.’ which, unless it is a denominative verb from *zeyzūm* ‘protector’, appears to be related to the hollow root noun *zūm* ‘endurance’.

The reduplicated verbs also lend themselves well to onomatopoeic verbs, such as *‘ay’a* ‘to crow (cock)’ and *dawda* ‘coax (camels) by calling “daw”’.

When the roots are trilateral, the root is augmented by insertion of a new consonant. The consonants used tend to be /w, y, l, n, r/ (i.e. the high sonority semi-vowels and the most sonorous coronals) in C2 position, or occasionally in C3 position, and /š/ in C1 position. In these cases, there is a semantic relationship between the trilateral root and the augmented quadrilateral root, with the latter usually being more nuanced or specific:

<i>ḍamal</i>	‘trot’	<i>ḍowmal</i>	‘trot fast’
<i>ḥadar</i>	‘go down’	<i>ḥawdar</i>	‘go down to pasture’
<i>ḥazam</i>	‘draw together’	<i>ḥeyzam</i>	‘gird’
<i>lahab</i>	‘be aflame’	<i>šalhab</i>	‘flare up’

Alternatively, quadrilaterals may be derived from trilateral verbs by the reduplication of C1 as C3, as in the case of *gahgar* ‘drive away (a camel)’, which is derived from *gahar* ‘stop (a camel)’. This derivation resembles that of the quadrilaterals derived from hollow verbs such as *lawlah* above, which are derived in the same way if the /w/ or /y/ is analysed as a second root consonant.

There is a reduced inventory of derived forms for the quadrilateral verbs. They are¹⁵³:

FORM-QI	CaCCaC
FORM-QII	taCaCCaC
FORM-QIII	(e)nCaCCaC
FORM-QIV	CaCCaC _i C _i

No verb is attested with more than one derived form after FORM-QI. It should be noted that verbs which have /w/ or /y/ as C2 are treated in some grammars as a variation of FORM-III of the trilateral verb (see, e.g., Johnstone 1967a: 45), but I treat them here simply

¹⁵³ I have used these labels since they best represent what is attested in the Rwala corpus. To avoid confusion, it should be made clear that they are not synonymous with the labels Q1, Q2, Q3 and Q4 in Lane’s lexicon, although the first two are the same, and the last one also has final gemination. Lane does not record a quadrilateral with /n/ prefixed, and his Q3 *if’anlala* is not found in the Rwala dialect (cf. Lane 1984: xxviii).

as FORM-QI, preferring to see the augmenting consonant as a consonant proper rather than as a diphthongal variant of the /ā/ of FORM-III¹⁵⁴.

Unfortunately too few examples are attested in the corpus to accurately appreciate the semantic correspondences between the derived forms of the quadrilaterals. As with the derived forms of the trilateral verb, the derived forms may add a sense of cooperation:

ḥawdar ‘go down to pasture’ *enḥawdar* ‘go down together’

Others simply express a nuance that is hard to capture with precision:

ḥanšal ‘raid on foot’ *enḥanšal* ‘start out a raid on foot’

Like the relationship between FORM-II and FORM-V of the trilateral verb, FORM-QII may express a medio-passive of FORM-QI, although this is hard to ascertain with confidence from the data:

nowmas ‘lead to glory’ *tanowmas* ‘win glory’ (i.e. be led to glory?)

There are very few phonological adjustments to be considered concerning the quadrilateral verb. When C2 is /w/ the first syllable is Caw or Cow and when C2 is /y/ the first syllable is Cey. If C4 is /w/ or /y/ the verb ends in /a/, e.g. *dawda* ‘he coaxed [the camels] by calling “daw”’ and *yetamaṭna* ‘he marches through the territory of sb.’.

¹⁵⁴ Holes (2016: 176) cites Fleisch (1944): “At the conclusion of a lengthy analysis, [Fleisch] comments that qautala and qaitala (= our CōCaC and CēCaC) [CawCaC and CeyCaC in Musil] ‘developed in all the Arabic dialects completely outside qātala, and, in general, in their own way as intensives’.”

4.12 Summary

The following summary of this chapter's appraisal of the verbal system of Rwala, as presented in Musil's corpus, focusses on the characteristics of the dialect which differ from the published descriptions of Cantineau and Prochazka.

The phonological processes described in Chapter 2 are routinely absent from the verbal forms. Firstly, Musil's system exhibits an almost total lack of phonologically conditioned vowels. Prochazka's description of the A-CLASS has four possible conjugations divided on the basis of the presence and position of guttural consonants or sonorous coronals, i.e. *kitab* (*ktibat*), *dibaḥ* (*ḍbaḥat*), *ḥalab* (*ḥlibat*) and *ga'ad* (*g'adat*), and phonologically conditioned vowels occur elsewhere in the system, where high vowels replace low vowels in unstressed open syllables: *ḥāribow/ḥārubow* for Musil's *ḥārabow*, *aḍhirow* for Musil's *aḍharow*, etc. The almost total absence of this from Musil's data demonstrates the lack of both /a/-raising and (/i/-)backing. At the same time, the verbal system has demonstrated that some raising (or /i/-spreading) is regularly attested, since it is formative in the S-STEM I-CLASS.

The verbal system also illustrates the lack of syncope in Musil. It has been demonstrated that both the A-CLASS and the I-CLASS of the S-STEM have regular syncope in other descriptions, but these forms are only partially attested in Musil. The behaviour of the P-STEM has been explained in terms of transcription: Musil's *yaktebūn* in place of Cantineau and Prochazka's *yaktbūn* 'they write' most probably belies an attempt to transcribe some sort

of audible release, and the form of the active participle, where we find *šārebīn* rather than **šārbīn* ‘drinking’, may also attest to this. Other examples include the syncope of P-STEM prefixes in some of the derived and biliteral forms, where Musil’s record of syncope after /t/ but not after /y/ or /n/ (e.g. *tkasser* ‘you smash’ but *yekasser* and not **ykasser*) possibly indicates that the sonority profile of the resulting cluster may have affected Musil’s perception (cf. §2.2.3.5). In all of these examples, /i/-syncope did not occur, according to Musil’s transcriptions. Unusually, /a/-syncope is much less attested for verbals than it is for nominals. In the S-STEM, as we have described, the syncopated forms (*ḍkarow* rather than *ḍakarow* ‘they mentioned’) do not occur in half of all cases, and the unsyncopated FORM-VIII verbs (cf. Prochazka’s *štǧalow* ‘they worked’) also come in this category. While the low application of syncope fits into the general phonological pattern attested to by Musil’s transcriptions, it seems to occur even less frequently in verbal forms than in other parts of speech. One possible explanation for this discrepancy is that Musil could have been more strongly influenced by his preconceptions of the verbal system of the language (from his presumed knowledge of CLA) than he was with nominals. For instance, if he assumed that verbs should have a disyllabic verbal stem in all persons, this would have interfered with his perception of these phonological processes if they did occur.

A similar situation is found with the *ghawa*-syndrome, which would be expected (at least in some examples) to render *yaʿref* as *yʿaref* or *yaʿaref* ‘he knows’. Although this does occur in Musil, it is restricted to a minority of cases and forms such as *yaʿref* occur most of

the time. This process is a feature of the whole dialect group to which the Rwala dialect belongs, and furthermore the phonological rule which governs it is (at least usually) active in the nominal forms, but absent from most of the verbal forms. We have seen that Musil's records most probably shed light on variability in the case of these processes, but the discrepancy between the regular variation (between *ghawa*-forms and non-*ghawa*-forms) in the rest of the corpus and its more conspicuous absence from verbal forms raises questions. It seems to be a possibility that Musil did fail to record it correctly in some cases in the case of verbs, perhaps by analogy to the verbs without guttural consonants.

There are a number of features specific to the verbal system where Musil's records are at odds with the existing descriptions of Rwala and other related dialects. In general terms, however, these are not particularly divergent. Here the major loci of discrepancies are summarised.

The *e* in the final syllable of the P-STEM of forms V-VI is certainly unusual when other evidence shows that it is more likely to be an *a*, and this seems to be a place where Musil could have systematised his data to conform to the rest of the derived classes. FORM-V also shows unusual syncope in the forms like *tetgassel* 'she scrubs herself' (cf. Prochazka's *teḡassel*), but this is explained as /i, u/-syncope targeting the vowel which will produce a bimoraic syllable as is attested elsewhere, e.g. in Prochazka's *ḥimlat* (< **ḥimilat*) 'she was carried'; it could also be analysed as the preformative /ta/ being analysed as /t/ in many of the P-STEM and participle forms. Musil's form *akal* 'he ate' contrasts with the forms in some

other published accounts but is not unattested, and this paradigm shows variation in many dialects. The two differently stressed forms on FORM-VII in Prochazka, *ínhirag* and *inhárag*, are interesting, although Musil's data cannot shed light on this variation. Prochazka does not record prothesis in FORM-X, although the prothesis in Musil is consistent with the phonology of the dialect. The fact that Musil's doubled verbs in the P-STEM do not show degemination, i.e. **yírid* < *yerúdd*, seems surprising from the dialectological literature, but it is difficult to say more about this with no indication of stress placement.

There are some more significant departures as well. Some S-STEM suffixes attested in other descriptions are missing from Musil, although this may be mitigated by the fact that Cantineau described them as rare. Similarly, the existence of both *-i* and *-in* as 2FS suffixes in the P-STEM is typologically unusual and hard to account for. The apparent absence of an I-CLASS for final weak verbs is startling, but it has been seen that it is possible that the I-CLASS may be identical to the A-CLASS in many forms. The attestation of the P-STEM prefix as *yi-* in addition to *ya-/ye-* is particularly problematic since it blurs the contrastive nature of FORM-I active forms with passive forms and FORM-IV forms. I have analysed this as possible assimilation of /y/ features, since *ti-* and *ni-* are hardly attested, but this does not really mitigate the potential effects of this change on the verbal system. The vowel harmony of the prothetic vowel, as observable from the imperative, is not recorded for Rwala and therefore unexpected, although it is not out of place for north-Arabian dialects in general. Finally the attestation of a *u-i* passive is also not recorded for this group generally, but is very common

in Musil. Its frequency suggests that it was not imagined, although it is possible that it represents backing (due to the surrounding consonants) of the expected /i/ of the paradigm; this would have to be considered overrepresented by an unfortunate coincidence in the corpus. Since it is similar to the passive forms of CLA, however, it cannot be discounted that Musil could have assumed the passive to have this form by analogy with the classical language.

With these reservations aside, it seems evident that the most surprising forms in Musil's presentation of the verbal system are the same as those given in Chapter 2 for the phonology. The fact that rules such as syncope and the *ghawa*-syndrome are represented more frequently in the nominal system than in the verbal system suggests that Musil may have been a less faithful transcriber of verbs than of other parts of speech, perhaps because he had a preconception about the verbal system from CLA; this could be supported by some of the apparently classicising features of his interpretation of verbal forms, such as the *u-i* passive, although this hypothesis is tenuous. In general Musil seems to have recorded the verbs faithfully, especially in his presentation of the complex developments regarding the two classes of the S-STEM in FORM-I and the distribution of stem vowels in the FORM-I P-STEM which match those of other related dialects.

5 Closed classes

Chapters 3 and 4 have dealt with the open classes of nouns, including adjectives, and verbs. These are so called because the classes may still acquire new members; this chapter deals with the closed classes which are unlikely to acquire new members at all. The exception to this is some adverbials and prepositionals which may still be productively derived, but they have also been included here in the relevant sections on adverbs and prepositions, which are at least partially closed classes. It will be seen that, unlike the treatment of nominals and verbs, the items in the following sections cannot be arranged solely according to their morphological structure, and it has made sense to define them rather by their function. For this reason the relevant syntax is discussed inasmuch that it helps to clarify the presentation of the data.

§5.1 deals with particles (such as prepositions, conjunctions and interrogatives), which differ from the open classes in the previous chapters in that they are in some cases bound, and in that they are uninflected and therefore cannot be marked for definiteness, gender, or number. The majority of these are not derived from consonantal roots, although it will be seen that some are nominal forms which do have consonantal roots but which are distinguished from nouns by their lack of inflection. §5.2 deals with partially inflected classes such as demonstratives and pronouns which are nominal in status but which do not have nominal form: these differ from the nominals of §3 because they are closed classes

which may also be bound and because none of them have consonantal roots, and as such they have been included with the closed class of particles in this chapter. At the end of the chapter, §5.3 collects some nominal forms which are included here by virtue of having also become closed classes.

In all sections of this chapter, relevant comparable forms from related dialects have been highlighted, although in many cases the words in the classes included here have been documented sporadically by researchers and a treatment comparable to that of the previous chapters has not always been possible. This is partly on account of the fact that many of the categories treated here are not relevant for the purpose of dialect classification, with many terms being more or less common to perhaps most varieties of Arabic. As such, typically north-Arabian, bedouin, or Najdi features have been highlighted but it has not seemed necessary to elaborate on more generic elements. It will also be seen that some terms appear in more than one section: this is due to the fact that some lexical classes are not clear-cut, and the focus has been on presenting the data in the most accessible way, even when this requires some reiteration. Some interrogative adverbs may therefore be listed twice, both with “interrogatives” and with “adverbs”, while other terms which function both as adverbs and as exceptives will also appear in more than one category.

5.1 Particles

5.1.1 *The vocative*

The vocative particle is *yā*, which is always transcribed as long; as addressed in §2.3.3, it seems likely that word-final vowel shortening does not apply to monosyllabic words. In Ingham's (1995) Rwala material it is transcribed both as *ya* and as *yā*, indicating that the vowel may or may not be shortened. It is used before personal names, titles or terms of address:

yā fheyd tarā-ni
O Fheyd, look at me!

yā 'amm-i
O uncle

ent min ēn yā walad
where are you from, boy?

It also occurs with descriptive adjectives, phrases or participles (including relative phrases), in which case they complement an inferred 'you':

yā ḥaḍīb al-eydeyn
O you with the dyed hands

yā šāreb at-tumbāk
O [you] smoker of tobacco! / O [you who are] smoking tobacco!

yā-lli 'elūm-ek te'azzī-ni
O you, the news of whom consoles me!

Frequently *yā* occurs before non-human nouns, in poetry or exclamations:

yā dār wudd-i
O camping place of my beloved!

Some nouns have forms which seem to be markedly vocative, namely, *yā yumma* ‘mother, little mother’, *yā ḥay* ‘brother, little brother’ and *yā ḥayye* ‘sister, little sister’, which never appear without *yā*. Forms such as *yumma*, *yā yumma* have been recorded in related dialects as a specific form of so-called bi-polar address, which encodes the relationship between speakers¹⁵⁵. This is not explicitly recorded by Musil, where these terms seem to function as diminutives (or, by extension, as terms of endearment) and vocatives, but since bi-polar address forms are widely attested in other dialects it is possible that this usage is not apparent from the corpus on account of the type of material collected.

5.1.2 *Exclamations*

Most of the following exclamations are particles, and most are free. The only bound exclamation is *w(a)-* which is used in oaths. Most oaths are by God, *w-allāh* ‘I swear by Allāh’, although other constructions do occur:

w-allāh l-aḡṭaʿ rās-ak

I swear I will cut your head off

w-alli mdir an-nahāra

I swear by him who measures the days

In the second example the oath is still by God, but by means of a relative clause rather than a name. A compound variant, *lā w(a)-*, exists with the same meaning. As well as

¹⁵⁵ Cf. Holes (2016: 437-9) for Bahrain and Yassin (1977: 297ff) for Kuwait; it has also been recorded in the Negev (Henkin 2010: 187ff) and in bedouin poetry from northern Jordan (Holes and Abu Athera 2009: 149 n.110, both cited in Holes 2016: 437, 442).

being used in oaths, *w-allāh* also functions as an exclamation of surprise or admiration, as in the first example below, or as a general intensifier, as in the second example:

w-allāh ramī-ha zēn
she shoots well! [lit. her shooting is good]

w-allāh yā 'omr-i ḥasāra
surely my life is destroyed

The rest of the exclamations are free. The particles for agreement and dissent are *na'am* 'yes' and *lā* 'no'. Particularly strong denial is expressed with *lā w-allāh* or *li-llāh*. The former, meaning 'absolutely not', may be understood as an extension of the oath formula described above, 'by God, no!'.

Exclamations of warning are compounds of a particle or noun and a bound pronoun (cf. §5.2.1.3). The pronoun is usually second person, since these exclamations are often imperative in nature, but this is not always the case. The attested particle is *ayyā-k* 'beware!' or 'be careful (of)!'. Other terms with the same meaning are *ḥaḍr-ak* and *bāl-ak*; and a dubious form, *ḥass-ekom*, also seems to have this meaning.

Exclamations of 'how...!' are variously made with *mā*. The fossilised classical form *mā 'af'ala* exists exclusively in the formulaic expression *mā ḥala* (often *yā mā ḥala*), 'how sweet!, how beautiful!'; the form *ḥala* is derived from **aḥla*, having undergone *ghawa*-syndrome resyllabification (**aḥla* > **aḥala* > *ḥala*). The construction (yā) *mā* + verb means variously 'how often...! how many...! how much...!', as in the following examples:

mā 'aḍḍalnā-hom
how we warned them!

mā gaṭa'na leylet^{en} ḥarmešiyye
how many dark nights we have ridden through!

yā mā gaḍeyna
how often we have been on a raid! / how many raids we have been on!

Expressions of this type are also made with *kam* 'how many...!'.

By far the most frequent category of exclamations in this corpus is exclamations of sorrow, and these are also very formulaic, although a lot of variation occurs. The most common expression uses *weyl*, apparently a fossilised noun, often combined with a bound pronoun, which has the meaning of 'woe to...!', and frequently occurs with the particles *yā* or *wa*¹⁵⁶. *lā* also features in these exclamations, as do the nouns *ḥeyf* 'shame', *waḡd* 'sorrow' and *ḥasāyef* 'losses'; expressions of grief can therefore be expressed in the following ways:

<i>weyl 'ala</i>	'woe to...'
<i>yā weyl-i, wa' weyl-i</i>	'woe is me!'
<i>lā wa'</i>	'oh woe!'
<i>lā wa' b-i</i>	'woe is me!'
<i>yā ḥeyf</i>	'oh shame!'
<i>yā waḡd-i</i>	'woe is me!'
<i>lā wa' ḥasāyef</i>	'what a shame!', 'woe is me!'

Wishes are expressed with the exclamative particle *leyt* 'if only, would that', which occurs either with a noun or with bound pronouns (the negative *mā* may intervene, as in the third example), or may be followed by a clause as in the fourth example:

yā leyt ḥelm-i dāyem
may my dream last forever!

¹⁵⁶ It seems possible that the glottal stop may be genuine in this form, since the exclamative will presumably always be uttered with a particularly intense prosody, but otherwise I would suggest that the term is *wā*.

leyt-aha tenšād al-yowm
may she be caught today

yā leyt mā ‘omr-i ṭawīl
may my life not be long!

yā leyt mā l-eh gwāni
may he have no kin!

The presentative *hā* ‘here is’ (cf. §5.2.2.1) and the grammaticalised forms *tara* and *er/er’/er’* ‘behold’, both with or without bound pronouns, function as exclamations; these latter terms are fossilised verbals from the root *r’y*, which in the dialect only survives in imperatives (all other parts of the verb are expressed by the root *šwf*). There are also exclamations of greeting, including:

<i>hala, ahla, ahl^{an}, yā-hla, yā hala</i>	‘hello, welcome’
<i>hala bi-</i>	‘welcome to...’
<i>marḥaba, marḥab^{an}, ahl^{an} w-sahl^{an}</i>	‘welcome’
<i>as-salām ‘aley-kom</i>	‘greetings’

Certain miscellaneous expressions have become functionally exclamative, e.g. *(lā) yā ba‘ad* ‘may [he] live long’.

5.1.3 *Coordinating conjunctions*

These conjunctions coordinate items of the same syntactic importance. The only bound conjunction is *w(a)-* ‘and’. The form is *w-* in most cases, which forms an initial cluster when followed by a single consonant, as initial consonant clusters are permitted in the dialect. *wa-* exists exclusively before consonantal clusters, where a CCC cluster would be

illegal. This is interestingly one of the only cases in the dialect where initial consonant clusters are not derived through syncope. The precise distribution of this distinction is not clear cut, however, as *w-CC* does occur in over a third of cases, and the difference between *w-CC* and *wa-CC* does not seem to be phonologically conditioned. A further complication is that words with prothetic vowels sometimes retain the quality of the vowel and sometimes not, so that *(e)štarayt* ‘I purchased’ may be rendered *wa-štarayt* or *w-eštarayt*. When *w(a)-* is bound to the definite article *(a)l-*, the resulting form is */wal/* (or */waC/* in the case of assimilation to “sun letters”, cf. §5.1.14); it should be noted that the transcription *w-al-* is used here to divide the morphemes, but strictly speaking there is no reason to prefer this to the transcription *wa-l-*.

A possible second bound conjunction appears in only one example in the corpus, and is the conjunction *fa-*. This record is rather dubious for two reasons. Firstly, *fa-* is considered to be lost in the modern Arabic dialects (Brockelmann 1908–1913 II, 488, cited in Waltisberg 2011), although it is listed in Werbeck (2001: 249, cited in Waltisberg 2011) for Manāḩa; to the best of my knowledge, it is not a feature of any of the north-Arabian bedouin dialects, however. Secondly, the usage in this example is not the same as that which is usually associated with *fa-*, which is ‘and’ or ‘and so’:

lā yagdar yu‘enn-ihā illa fa-yarċi-ha ‘ala šenḩūba
 he can’t bridle [the horse] unless he makes her lean against a huge rock

It is difficult to speculate from only one attestation just what syntactic function *fa-* has in this sentence. It may be noted, however, that ‘unless’ may be expressed by *illā wa-* in

CLA (Fischer 2002: 169, Lane 1984: 78) and it is possible that what Musil heard (or thought he heard) was a dialectal variant of this.

The following free coordinating conjunctions are attested:

<i>amma</i> (<i>yumma</i>)	‘either... or...’
<i>aw, w-, w-lā, w-illa</i>	‘or’
<i>ayy/low... illa...</i>	‘whether... or...’
<i>bal</i>	‘in fact’
<i>tumm(a)</i>	‘so’
<i>mār</i>	‘but’
<i>w-lācen</i>	‘but’

Of the particle *mār*, Blanc writes that “no particle is more characteristic of the confines of northern Arabia and the Syrian Desert” (Blanc 1970: 146)¹⁵⁷.

5.1.4 *Subordinating conjunctions*

The bound conjunction *w(a)-* may also have a subordinating function. An example is:

tenḥa fheyd w-hw šahr zemān^{en} ‘an-na
 she calls on Fheyd, even though he is a month away from us!

The attested free subordinating conjunctions are:

<i>an, anna, ann</i>	‘that’
<i>en, elya¹⁵⁸</i>	‘when, after, if’
<i>elyamā</i>	‘until’
<i>amma</i>	‘whether’
<i>bann</i>	‘because’
<i>ḥatta, ḥatt-en</i>	‘so that’

¹⁵⁷ Here it may be noted that the form *mār* ‘then; good; okay’ is attested in the Modern South Arabian language Šherēt, cf. Johnstone (1981: 169) *māǧǧr* E[astern dialect] *mór* ‘well then; well; indeed’.

¹⁵⁸ The particle which is *elya* in the Rwala corpus seems to be *ila* in other bedouin dialects from the region (cf. Ingham 1991: 42) but is attested as *ilya/elya* in Cantineau (1936-37: 210).

<i>swāt, swāk</i>	‘like, just as’	
<i>gabl^{en}, min gabl</i>	‘before’	
<i>gablmā</i>	‘before’	
<i>killmā</i>	‘whenever’	
<i>ćann</i>	‘like, as if’	(cf. CLA <i>ka-’inna</i>)
<i>lann</i>	‘because’	(cf. CLA <i>li-’anna</i>)
<i>low</i>	‘even if’	
<i>yowm</i>	‘when’	

yowm meaning ‘when’ is considered by Blanc (1970: 146) to be as characteristic of this dialect group as *mār*. Unlike some other dialects (e.g. Negev, *ibid.*, and Ḥwēṭāt, Palva 1984-86: 305, although both of these are of the northwestern Arabian group) *nhār* as a synonym of *yowm* in the sense of ‘when’ (both mean ‘day’ as nominals) is unattested for this dialect.

5.1.5 Conditionals

The following conditional particles are attested:

<i>illa</i>	‘unless’
<i>en</i>	‘if’
<i>elya</i>	‘if’
<i>ćān</i>	‘if’
<i>law (low)</i>	‘if’

ćān is never realised as *kān*. It is frequently found in compounds such as *ćān en(na)* (cf. *inna*, §5.1.7), while it more clearly preserves the verbal meaning of *kān* in the compound *en ćān...* ‘if it were the case that...’. This list matches that described by Ingham (1991: 45) for bedouin-type dialects except for the absence of *lamman* signifying ‘when’.

5.1.6 *Exceptives*

The following exceptive particles are attested:

<i>illa</i>	‘except’
<i>ġeyr</i>	‘other than’
<i>kūd, akūd</i>	‘except, unless’

kūd is recorded by Holes (2011a) alongside *kūd in* and *kūdīn* as occurring with a range of meanings in many Arabian and Jordanian bedouin dialects; in addition to the meaning of ‘except’ it may also express ‘perhaps’ (cf. Ingham 1994a: 125–126) and ‘with difficulty’ (cf. Kurpershoek 2005: 289), although these meanings are not attested in Musil. The particle is attested with the meaning ‘unless’ as *yakūd* in some of the Šammari texts of Montagne, according to Cantineau (1936-37: 210).

5.1.7 *Emphasisers*

The first type of emphasiser is that of sentence-introducing particles:

<i>inn, inna</i>	[introduces noun- or pronoun-initial main clause]
<i>‘ād</i>	[introduces a discursively new subject]
<i>leyt</i>	‘if only’

Some of these introductory particles also introduce subordinate clauses:

<i>ann, anna</i>	[introduces noun- or pronoun-initial subordinate clause]
<i>‘énn</i>	‘as if’ (cf. CLA <i>ka-’inna</i>)
<i>lann</i>	‘because’ (cf. CLA <i>li-’anna</i>)
<i>w-lā‘én</i>	‘but’

There are some particles which reinforce the verb in various ways:

<i>la-</i>	[future reinforcer]
<i>gad, ġid</i>	[past reinforcer]
<i>gatt</i>	[negative reinforcer]

la- is used before a P-STEM verb with the sense of ‘surely, most certainly’. This usage may also be extended to account for oaths (cf. §5.1.2 and Holes 2011a), so *l-agta‘ rās-ak* may be interpreted as emphatic ‘I will cut off your head straight away’ (as in Musil) or as ‘I swear I will cut off your head’. Similar usages are found elsewhere in Najdi (cf. Ingham 1994a: 129) and in poetry from Sinai (cf. Holes and Abu Athera 2009: 75, 78).

A reflex of CLA *qad* is occurs in dialects across the Arabian Peninsula, especially in bedouin-descended dialects (Watson 2011a: 898). The form in Musil is *gad* or *ġid*; the forms *gid* and *ġid* have been recorded for Najdi Arabic generally (Ingham 1994a: 104-7), alongside a variant *ćid* which is claimed to occur in northern-Najdi dialects and is not attested in this corpus. In Musil it precedes the S-STEM only and has several functions.

One usage is resultative (Holes 2016: 268), described as “a presently relevant result of a past action or a state of being realised as a result of a past action” (Holes, *ibid.*). An example from Musil is:

*šār baḥīṭ be-l-maġāri... w-gad ištahar šīt-eh*¹⁵⁹
 he had great success on raids... and therefore he was famous

¹⁵⁹ Note that this example is not actually presented as a continuous utterance in Musil, although the context indicates that the first phrase leads into the second. The phrasing (commonly found in Musil, as he illustrates narrative sections with interspersed Arabic) is, “[he] had great success on his frequent raids, *šār baḥīṭ be-l-maġāri*. Therefore he was famous all over the desert, *w-gad ištahar šīt-eh*” (Musil 1928a: 593).

The meaning here is that the spread of his fame occurred as a result of his success in raids. Another function may be described as ‘anteriorising’ (Holes, *ibid.*: 269). Ingham notes that this may be translatable with ‘already’ in English when the reference point is the present (Ingham 1994a : 104):

telgā-h ġid māt
you find that she¹⁶⁰ has died/is already dead

When the reference point is in the past, this usually expresses the pluperfect; this is not explicitly recorded in Musil, but an example would be:

**lageyt-eh ġid māt*
you found that she had died

Ingham (*ibid.*) describes the usage with a present reference point as emphatic, while the usage with a past reference point does not have the same element of emphasis. He also notes that with a past reference point this carries the implication that the state of affairs has ceased, although this is not applicable to the above example. His example of this usage is:

abū-ha ġid aḥaḍ umm-ih
her father had married his mother

This statement implies that the couple had married at some point in the past, but that they were no longer married at the time of speaking. The particle only occurs in Musil with the functions shown here, and some functions recorded for other Arabian dialects are not attested: *gad/ġid* does not occur before the P-STEM, a usage which has been recorded in eastern Arabia with the modal meaning ‘perhaps oftentimes’ (Holes 2016: 272), and its

¹⁶⁰ Here, as elsewhere, the female referent is described with grammatically masculine terms.

usage in southern Arabia as an introductory particle before a nominal sentence is not found (cf. Ingham 1986: 279 for southern Najd; Watson 1993: 413 for San’ani). It is also not attested with bound suffixes (cf. Ingham, *ibid.*, Watson, *ibid.*, Holes 2016: 269).

lā gatṭ or *mā gatṭ* occur before a verb in either the S-STEM or P-STEM with the meaning ‘never’. It is interesting to note that in Ingham’s Najdi data, this kind of ‘experiential’ negative is expressed by *gad/ǧid* rather than *gatṭ* (Ingham 1994a: 104-5). This usage of *gad/ǧid* is not attested in Musil, but it is difficult to discern the significance of these occurrences as there are only two examples of *gatṭ* in the corpus.

The following exclamatives may also be included in this group:

<i>kam</i>	‘how many...!’
<i>mā</i>	‘how...!’

Both of these are listed as typical of *nabaṭī* poetry by Holes (2011a: §7.3).

5.1.8 *Restrictives*

The restrictive particle *bass* is attested with the meaning ‘only’. It does not occur in verbal clauses in the corpus, where a negative construction using one of the exceptive particles (§5.1.6) is found instead:

ṭalbet-ak bass ḥorra
your demand is for a thoroughbred only

lā yišūf-ni kūd [ar-]rā’i
only the herdsman sees me

5.1.9 *Negatives*

The following negative particles are attested:

ḡeyr
(w-)lā
lam
mā
mā ‘ād

lā and *mā* are transcribed by Ingham (1994a: 44) and Cantineau (1936-37: 207) as *la* and *ma*, although Ingham’s Rwala transcriptions have both with long and short vowels (Ingham 1995: 125-126); other words in the same texts which do not constitute minimal phonological words (cf. §2.2.2) show the same variation (e.g. *ya* ~ *yā* [vocative particle], *ḡa* ~ *ḡā* ‘he came’). Given this evidence, it seems likely that Musil’s record of these particles not exhibiting word-final vowel shortening in any instance should be questioned, although Ingham’s material shows that there is variation in this regard.

The particles *lā* and *mā* negate verbal sentences. It would be expected of dialects in this group that *lā* would occur exclusively with imperatives and jussives and *mā* in other contexts (Ingham 1994a: 44), but this usage is more mixed in Musil’s corpus. Both particles may negate a past action:

lā radd as-salām
he did not return the greeting

mā ḡalla ḡawiy-ye
he did not leave my brother

This somewhat unconventional use of *lā* is not unprecedented, however, having also been observed among the Rwala by Cantineau (1936-37: 207), who recorded with surprise the phrase *dawwar u la laga* ‘he searched but he did not find’ rather than **u ma laga*. The fact that this construction occurs fairly frequently in Musil’s material suggests that it may be a feature of this particular dialect. Ingham (*ibid.*), however, records the usage in Najdi dialects in general of *lā* negating plain verbal statements as an emphatic negative, offering the following examples:

<i>ma ġā-na muṭar</i>	<i>w-lā ġā-na muṭar</i>
no rain came to us	no rain came to us at all

It may therefore be the case that this usage of *lā* is not a Rwala peculiarity but a general use of *lā* as an emphatic negative, and perhaps that Cantineau’s phrase should be translated as ‘he searched but he found nothing at all’. As in other Najdi dialects, only *lā* negates past actions in “neither... nor...” constructions (Ingham 1994a: 45):

waṣf-k lā zāġ w-lā rāġ
 your description did not deviate [from the facts] nor did it miss [the target]

Other uses of the negative particles in verbal sentences are as expected. *lā* is used with both the P-STEM and the S-STEM in the imperative and optative:

lā tebċīn amr^{en} maḍa
 do not mourn something which has passed away

lā bārak allāh be-l-ğaṭa
 would that Allāh had not blessed the sand grouse!

An extension of this usage is that *lā* is used with the P-STEM in negative purpose clauses and in unreal situations in general:

*afekk-aha lā tḏūg-aha*¹⁶¹

I will protect it so that you will never get a taste of it

ebḥaš as-sebīl lā yinsedd

clean the pipe so that it will not clog

Contrastively, *mā* with the P-STEM usually indicates a real future situation:

mā yarwa w-mā yešba^c

he will never find enough to eat or drink

A characteristically Najdi usage is *mā... b*, with incorporation of the relevant personal pronoun, in the negation of nominal sentences (Ingham 1994a: 44). In Musil this construction is also attested with *lā... b*:

mā-ni b-malyūm

I should not be blamed¹⁶²

lā-ni b-rā'i

I am not a herdsman

The occurrence of *b* in these phrases seems to be more a feature of central-Arabian dialects so its fairly low attestation in Musil's material probably reflects the fact that Rwala

¹⁶¹ Musil transcribed *affek-aha* (Musil 1928a: 21) which I can only assume is a misprint, since *afekk* 'I protect' is frequently attested.

¹⁶² It appears that Musil did not grasp the generic meaning of this construction, since his translations of phrases of this type usually make some visible attempt to interpret *b* with its usual prepositional meaning, such as in this case, 'I am not among those who are abused' (and note that here the misconstrual is also lexical, since *lowm* means 'blame', not 'abuse', cf. Glossary).

is both typologically and geographically on the periphery of where this occurs. Both *mā* and *lā* frequently occur with personal pronouns without *b* in negating nominal sentences:

mā-ni bāgi
I don't want [to]

lā hw ḥemāra
he is not an ass

mā hu eben al-‘amm
he is not one of our close relations

mā ḥenna ši^{en}
there's nothing we can do [lit. we are nothing]

It will be noted that *mā-ni* in the first example differs from the others in that *lā/mā hw/hu* and *mā ḥenna* use the free personal pronoun (1CS *ana*) rather than the bound pronoun (1CS *-ni*, cf. §5.2.1). This is attested in other Najdi dialects (Ingham 1994a: 45) where it is reported that *mā-ni* may vary with *mā-na*, which is a contraction of *mā ana*. The latter does not occur in Musil's corpus, and nor does the contracted form *mu* (<*mā hu*) known from central Arabia (Ingham, *ibid.*). Both *lā* and *mā* may be used in categorical negation without subject pronouns if prepositions are involved:

lā l-ak
you (MS) do not have

mā b-ak ḥemāra
you do not have the flame of courage within you

As with the verbs, only *lā* expresses “neither... nor...” in phrases of categorical negation:

lā goṭn manfūs lā mer^{ezen} leyān baṭn-eh
neither fine cloth nor cotton is as soft as her [lit. his] belly

lā ḥayyen w-lā māt
neither living nor dead

The negative particle *lam* is only found occasionally in poetry and may be a feature of the poetic register. It occurs with the P-STEM and only indicates an unreal future situation:

lam tağed rūḥ sawā-ha
you won't find a soul like that one

Sometimes the exceptive *ğeyr* is used in categorical negation:

mā yarwi l-bel ġeyr muserbel
he shall not water his camels who does not pull up [the rope] steadily

In this example, it seems appropriate to consider *ğeyr* as a negative particle although it most probably represents a contraction of the more periphrastic relative “other than he who...”.

It should be noted that the negative construction *mā ‘ād* means ‘no longer’; this differs from the more frequent usage of *‘ād mā* where the particle *‘ād* is not part of the negative phrase but simply emphasises a new subject, or in the phrase *elya ‘ād mā* ‘should it never again happen that’:

mā ‘ād l-eh zebde
‘he no longer has (any) butter’

elya ‘ād mā yeḏḥak
‘if he does not smile’

5.1.10 *Interrogatives*

This set is not a lexical class, but rather consists of pronouns, adjectives and adverbs united by their function as question words, or *wh*-words, used in the forming of *wh*-questions (as opposed to yes-no questions, cf. §5.1.11). The following is the complete attested set, and each is discussed below:

<i>ayy</i> , f. <i>ayye</i>	‘which?’
<i>ćam</i> , <i>kam</i>	‘how many?’
<i>ćeyf</i>	‘how?’
<i>lēh</i> , <i>lēš</i>	‘why?’
<i>limin</i>	‘whose?’
<i>mata</i>	‘when?’
<i>min</i>	‘who?’
<i>min ēn</i>	‘from where?’
<i>waš</i>	‘what?’
<i>wēn</i> , <i>ēn</i> , <i>weyn</i>	‘where?’

ayy, *ayye*, ‘which?’, is a particle which can function as an adjective or as a pronoun;

here the usage is pronominal:

ayy al-ğedīd
which [one] is the new one?

and here adjectival:

ayye ġebīle
which tribe?

It is inflected for gender, but the particle does not have a plural in any dialect (or in CLA). Gender agreement is strictly adhered to in the singular (as in the above examples) but with plurals the MS is attested, rather than ‘deflected’ agreement with the FS, (cf. §3.1.4), as shown by *ayy al-‘orbān* and *ayy ‘arab* ‘which Arabs?’. Interestingly the particle is not

recorded for any dialect in Cantineau (1936-37), and Palva (1980: 123) also notes its absence from the Najdi texts of Abboud (1964), but it is recorded for the Bani Şaḥar (Palva, *ibid.*) and described as widespread across the Gulf (Johnstone 1967a: 68).

It should be noted that the words *ćam* and *kam* ‘how many?’ are not contrastive, and *ć* and *k* do not distinguish meaning in the same way as *ćeyf* and *keyf*, which are distinct words. Both forms occur, although the *kam* variant is less frequent; this is to be expected since the sequence *ka* is half as frequent as *ća* across the whole corpus. ‘How many [people]?’ is expressed by *ćam wāhed^{en}*. *ćam/kam* is always followed by the adnominal linker (cf. §3.4), e.g. *ćam leylet^{en}* ‘how many nights?’; there is only one exception in the corpus (*ćam ġamal* ‘end-akom ‘how many camels do you have?’) which, being found in the context of a dialogue, possibly suggests that the linker is not always used in informal conversation.

ćeyf has a monophthongized variant, *ćēf*, attested once in conversation. As an interrogative adverb it is found always with *ć*, as opposed to the word *keyf* which is a substantive meaning ‘habit’, ‘pleasure’ or ‘rest’; as noted elsewhere, this distinction, as well as the distinction in the 2_{MS} and 2_{FS} bound pronouns (cf. §5.2.1.3), establishes /ć/ as a marginal phoneme, although this status is liminal given its restriction to these words. Variants of *ćeyf* are given by Cantineau for the whole ‘Anazi group (*ćēf*) as well as for the Syro-Mesopotamian group (*ćēf*) (Cantineau 1936-7: 111, 208); this contrasts with western dialects of Greater Syria which have unaffricated *k*. The corpus does not attest the adverb

šlōn (or a variant thereof) for ‘how?’; *šlōn* occurs alongside *čēf*, *kēf* and similar variants throughout the eastern areas, including the Gulf and Šammari dialects (Bergsträsser 1915: map 29).

lēh or *lēš* is by far the most frequently attested term for ‘why?’; an alternative form, *warāk*, does not appear to be a native term among the Rwala and therefore is not listed above. Musil must have found the form obscure, since he provides an account of having investigated its usage: “according to the generally accepted explanation, [*warāk*] is equivalent to *leh* or *lēš*” (Musil 1928a: 216). The term is found as *wara* in the Gulf dialects recorded by Johnstone (1967a: 157), who records *wara dhart* ‘why have you come?’, indicating that the /k/ of this *warāk* is, in fact, a bound pronoun, and that this use of the term probably developed from that of the preposition *wara* ‘behind’. Certainly Kurpershoek (2005: 348) takes this view, and his glossary shows different pronouns agreeing with the referent of the question: *warā-č hniyye* ‘what are you (f.) doing here?’ and *warā-h šēn* ‘why is he so ugly?’ (*ibid.*). Intriguingly, the pronoun in Musil does not show agreement:

warāk tezhed yeryeš al-‘ayn bi-na
why does she furrow her brow at us?

Clearly we would expect **warā-ha* in line with Kurpershoek’s evidence, rather than the apparently 2MS *warā-k*. This indicates that in the dialect of the poet in question, *warāk* has undergone lexicalisation and the suffix is no longer semantically active. The poem in which the term occurs was composed by Feyṣal al-‘Angari, a sedentary villager from al-Baṣra who came among the Beni Ḥāled by the Persian Gulf during the summer (Musil 1928a: 214)

and his use of *warāk*, though apparently recognised by the Rwala reciters, seems to be from his dialect rather than native in the Rwala dialect¹⁶³.

min ‘who?’ occurs with an independent pronoun when its referent is definite (e.g. to verify the identity of somebody present), but does not take a pronoun when the referent is unspecified:

min hom ha-l-‘arab
who are those Arabs?

min hw melfā-h
who should I give it to? [lit. who is the receiver of it?]

min yebaššer aš-šammar
who will tell the good news to the Šammar?

In the first example the referent is clearly specified and being pointed at; the second example shows that, even when the referent is unknown, it remains a specified participant, since only one person can receive the letter. This is in contrast to the third example in which any unspecified person may respond to the call. Palva (1980: 122) describes the form without clitic pronoun as being infrequent in the dialect of the Bani Šaḥar, but does not specify whether it follows the same distribution as described here. *min* without a pronoun is certainly not infrequent in the Rwala corpus, in which the referent is unspecified in the majority of cases.

¹⁶³ The reciters also argued about the phrase *yeryeš al-‘ayn* ‘forehead’ in the same poem, preferring their native *rīš al-‘ayn*, again showing that the dialect of this particular poem was not native to the Rwala.

There are many forms of ‘what?’, all of which are developments of the OA expression *w-ayy-šay*¹⁶⁴. *waš* is included in the above list because it is the major variant (75% of all tags) but there are a number of variants: *wēš*, *wušu*, *wuš* and *wšu*, in descending order of frequency. This matches the forms in Cantineau’s ‘Anazi group (Cantineau 1936-7: 108, 207); the longer forms such as *weššāh* and *wešnōh* attested in the Šammari dialects are not found in this corpus, and nor are the short forms typical of other groups from the region such as *šī*, *š-* (Mawāli).

The distribution of the variants *wēn* ~ *weyn* is 2:1, much higher than the equivalent variation of *čēf* ~ *čeyf* which almost always has a diphthong. *wēn* means both ‘where?’ and ‘to where?’, meaning that *wēn ent* means both ‘where are you?’ and also ‘where are you going?’.

Some dialects of Arabic make a lexical distinction between ‘how much?’ and ‘how many?’ (Watson 2011b) including many bedouin ones, e.g. Negev Arabic *kam* ‘how much?’ and *gaddayš*, *gaddēš* ‘how many?’ (Blanc 1970: 144), and such an occurrence cannot be ruled out for this dialect where it is simply unattested.

5.1.11 *Question particle*

The only instance of a question particle found in the corpus is the interesting form in the following line from a love ditty:

¹⁶⁴ Forms containing the word *šay* ‘thing’ are found as the pronoun ‘what?’ in almost all modern varieties of Arabic, cf. Versteegh (2011).

a-taḥseb en mā-ni yamm-ak
did you (MS) think that I wasn't with you?
[lit. do you think that I am not with you]

Given the infrequent occurrence of this particle, and its apparent absence from related dialects, we may wonder whether a simple interjection is intended, either the sound “ah” or the attested *ā* ‘alas!’. Musil’s transcription <ataḥseb> implies that he did not consider it to be an interjection (which he always separates with a space); he may have inferred from CLA that a question particle could exist, although his translation “oh, thinkest thou...” leaves the possibility that he did consider there to be an interjection and left the transcription ambiguous.

Yes-no questions are generally not marked morphologically or syntactically, as seen in the following examples, but we may infer that there are prosodic markers:

ent ‘amyān
you are blind / are you blind?

mā laḡeyti ḡeyr ḡowz-i
you (FS) couldn’t find anyone other than my man / couldn’t you (FS) find
anyone other than my man?

5.1.12 **Adverbs**

Nouns and adjectives are frequently used adverbially, therefore words with adverbial function do not form a closed class. There are, however, words which are exclusively adverbs, such as the local and manner demonstratives, which are listed in this section, and

the ways in which adverbs may be derived in the dialect are discussed. Interrogative adverbs are dealt with in §5.1.10.

As in most modern dialects of Arabic, there is a closed class of demonstrative adverbials. A number of dialects have three distance levels for the local demonstratives (Watson 2011b), and Musil's corpus suggests that there are three in his material, although the closest is surprisingly unattested:

here	n/a
there	<i>hunāk</i>
over there	<i>ġād, ġādi</i>

'Here' was recorded for the Rwala as *hnaya* by Cantineau, a form which he considered to be an isogloss of the 'Anazi and Šammari dialects, as opposed to *hin*, *han* of the petits nomades. The form is probably a diminutive, and is attested in some Yemeni dialects as *hanayyāh*, *hinayyāh*, *hiniyyih* and others (Piamenta 1990-91: 512). *hnaya* is also lacking from Palva's account of the Bani Ṣaḡar, where only *hān* is attested as a cognate, alongside the form *b-hāḡa* which has the same meaning (Palva 1980: 132). *hunāk* was recorded as the syncopated form *hnāk* for the Rwala by Cantineau.

The manner demonstrative is *hič* 'like this'. The manner demonstrative is held to be a cognate of CLA *kaḡā* or *hākaḡā* in all modern dialects (Fischer and Jastrow 1980: 83). *hič* or *hiči* have been recorded for north-Arabian bedouin elsewhere, and *hēč* in nearby Ḥōrān (Fischer and Jastrow, *ibid.*).

Other adverbs are derived through grammaticalisation of nouns or adjectives. Examples of this for temporal adverbs are *(h)alyowm* ‘today’, derived from *al-yowm* ‘the day’ or *ha-l-yowm* ‘this day’, and *halleyle* ‘tonight’, derived from *ha-l-leyle* ‘this night’. *al-‘ām* ‘the year’ has the sense of ‘last year’ in adverbial contexts, and *yowm* without the definite article has been grammaticalised as the adverbial conjunction ‘when’. Other temporal adverbs are *ams* ‘yesterday’, *bukra* or *bā‘er* ‘tomorrow’, and *taw^(en)* or *tawwa* ‘just now’. ‘Yesterday’ may also be expressed by *al-bāreħa*, a term which Cantineau considered to occur among the bedouin only due to contact with the sedentary population.

Some dialects have derived adverbs from the grammaticalisation of prepositional phrases. For example, Cantineau recorded the local adverb *balaħla* ‘outside’ for the Rwala: *be-l-ħala* is translated as ‘in the desert’ throughout Musil, although there are instances where ‘outside’ would seem to be a more apposite translation. These grammaticalised forms have sometimes undergone phonological change which renders their origins opaque, e.g. *balāš* ‘free, for nothing’, which, historically, is derived from the phrase **b-lā šī* (cf. Holes 2001: 52). Some interrogative adverbs (listed in §5.1.10) fall into this category, such as *lēš* ‘why?’ (< **l-ayy šī*) and *waš* ‘what?’ (< **w-ayy šī*).

Another process by which the modern vernaculars have enriched the inventory of adverbs is by borrowing (Watson 2011b). There are not many examples of this in the Rwala dialect as recorded by Musil, but the usage of *bass* ‘only’ is a historical example (it is a

borrowing from Persian); this is an adverb which is described as a pan-Arabic feature by Watson (*ibid.*). *bālci* ‘perhaps’ also appears to be a borrowing (cf. Holes 2001: 53).

A number of adverbs are formed with the suffix *-an*. This is a matter of interest in the Rwala dialect since the adnominal linker (cf. §3.4) occurs frequently and these forms are indistinguishable. Notably, some dialects which make use of the adnominal linker nonetheless seem to differentiate between this and the adverbial suffix *-an*: one example may be Bahraini where in the majority of cases *-an* is an adverbial suffix and *-in* is the adnominal linker (cf. Holes 2016: 134), but in Najdi dialects there tends to be no differentiation (cf. Ingham 1994a: 49). In the Rwala dialect as recorded by Musil, we have seen that the adnominal linker is transcribed as both *-en* and *-an* with no apparent difference in meaning; adverbs are also derived with both, but the variant *-an* is in fact more commonly attested on adverbs (40% of cases) than with the other uses of the adnominal linker (20% of cases). Since *-an* is always the minority variant, it is hard to tell whether the more frequent attestation of *-an* as an adverbial suffix is a coincidence or whether it is an indication that there is some differentiation between the use of the adnominal linker in the functions established in §3.4 and its use in forming adverbs from nominals. It is also hard to tell whether this suffix is a productive means of deriving adverbs or whether the adverbs that are attested with it are fossilised as such. Examples of adverbs with this suffix are:

<i>eyḏ^{an} w-</i>	‘and, also’
<i>dāym^{en}</i>	‘always’
<i>wukād^{en}</i>	‘certainly’
<i>dowm^{en}</i>	‘still’
<i>taw^{en}</i>	‘just now’

Most of these have exclusively adverbial function, and all except *dāym^{en}* are also attested without the ending; ‘always’ is, however, also expressed as *dūm*. Watson (2011b) gives a similar contrast of *dāyman* and *dīma* ‘always’ for Khartoum, and writes that adverbs which take the *-an* ending are “particularly frequent in higher registers of speech, often where more obviously colloquial alternatives exist” (Watson *ibid.*). It is likely that *wukād* may also be used as a noun, although this is not attested in Musil: *wakād* ‘certainty, fact, corroboration’ has been recorded in bedouin usage (Holes and Abu Athera 2009: 237), and the adverbial use ‘certainly’ may also be expressed as *bi-l-wikād* (Kurpershoek 2005: 354). *be-* is highly productive at forming adverbials from nominals, e.g. *be-luṭf^{en}* ‘gracefully’ and *be-ḡmīl* ‘nicely, kindly’.

5.1.13 *Prepositions*

Arabic historically has two word categories which may be considered prepositions, namely the “pure” prepositions (such as *bi-* and *min*) and the prepositionals, which are semi-inflected nominal forms with prepositional function (these have the inflection *-a* in CLA, such as *bayna* and *taḥta*) (cf. Badawi *et al.* 2016: 66). Since this inflection is not a feature of the dialect in question, this division has not been drawn here. Compound prepositionals

comprising preposition + preposition and preposition + nominal have been collected separately and again it is unclear whether these represent closed classes. Many preposition(al)s have broad semantic valency and these may be consulted in full in the glossary. The following list gives the main preposition(al)s used in the corpus:

<i>ila</i>	‘(in)to’
<i>b(e)</i>	‘in, at, to, with’ etc.
<i>ba‘ad</i>	‘after’
<i>bēn</i>	‘among’
<i>taḥat</i>	‘under’
<i>tegel</i>	‘like’
<i>ḥadar</i>	‘below’
<i>ḥa‘a</i>	‘like’
<i>ḥlāf</i>	‘behind’
<i>sawa, swāt</i>	‘like’
<i>šibh, šebīh</i>	‘like’
<i>šufi</i>	‘like’
<i>‘ogb</i>	‘after’
<i>‘ala, ‘a</i>	‘for, unto, on, against’ etc.
<i>‘an</i>	‘from; about’
<i>‘end(a)</i>	‘among’
<i>ḡabb</i>	‘after’
<i>ḡeyr</i>	‘except for’
<i>fowg</i>	‘on, over’
<i>fi</i>	= <i>b(e)</i> (poetic)
<i>gabl</i>	‘before’ (time)
<i>gubāl, ḡebāl</i>	‘before’ (place), ‘in front of’
<i>ḡebīl</i>	‘opposite’
<i>gafa, gafw</i>	‘behind’
<i>‘a</i>	‘like’
<i>‘ema</i>	‘like’
<i>‘ann</i>	‘like’
<i>kūd, kowd</i>	‘except for’
<i>la, li</i>	‘to, for’ etc.
<i>miṭl</i>	‘like’
<i>ma‘</i>	‘with’
<i>min</i>	‘from’

<i>wara</i>	‘beyond, behind’
<i>yamm</i>	‘at; to, towards’

There are a number of complex prepositions. The following are formed from the combination of two distinct prepositions (they are divided according to the first element):

<i>bi-ġeyr</i>	‘without’
‘ <i>ala miṭl</i>	‘in the same way as’
<i>min bēn</i>	‘between’
<i>min dūn</i>	‘between; without’
<i>min fowg</i>	‘above’
<i>min wara</i>	‘from behind’
<i>min ‘end</i>	‘from’
<i>min ‘ogb</i>	‘after’
<i>min gabl</i>	‘before’
<i>min ġeyr</i>	‘without’

The following complex prepositions are formed from the combination of a preposition and a noun:

<i>b-ġenāb</i>	‘by, beside’
<i>b-waġh</i>	‘in front of’
‘ <i>enda waġh</i>	‘in front of’
<i>ma‘ waġh</i>	‘in the direction of’
<i>ma‘ gafa</i>	‘from behind’
‘ <i>ala ‘ayn</i>	‘before, in front of’
<i>min wasi‘</i>	‘distant from’
<i>min šān</i>	‘for the sake of’

Cantineau (1936-37: 209) also records the preposition *gafa* ‘behind’ for the Rwala dialect, though it should be noted that its basic sense is as a noun meaning ‘back’ and that the sense of ‘behind’ is an extension of this meaning.

tegel ‘like’ has the literal meaning ‘you say’, being derived from the 2MS P-STEM of the root **qwl** and originating akin to ‘you’d think’ in English. For example:

ṣāfiyen tegel ġarnūg
pure white, like a crane / you’d think it were a crane

Musil was aware of the connection between *tegel* and **qwl** but was dismissive of the form, glossing it as “*tegel*, instead of the correct *tegūl*, ‘thou mightest say’” (Musil 1928a: 320). In Musil the form is always *tegel* with the meaning ‘like’, while *tegūl* is the straightforward verb. *tegūl* or variations thereof are found with this usage in other dialects, however: *itkūl* is attested in ordinary speech in eastern Arabia (Holes 2005: 197), *tugūl*, and *tigūl* are found in poetry from Sinai (Holes and Abu Athera 2009: 72, 79), and *tugul* in poetry from northern Jordan (*ibid.*: 142). It is interesting to note that other particles with the same meaning are derived from roots the basic meaning of which has to do with speaking: *ḥāca* ‘like’ (literal meaning ‘he said’) is attested in Musil, and the particle *ʿidd* ‘like, as if’ (literal meaning ‘tell!’) is attested in Jordanian poetry recorded by Palva (1992: 173).

5.1.14 *Definite article*

The definite article is /**(a)l**/. As in CLA, /**l**/ of the definite article is assimilated to /**t, ṭ, d, ḏ, r, z, s, š, ṣ, ṭ, ḏ, n**/ and (theoretically) /**l**/, i.e. the so-called “sun letters”:

<i>al-‘arab</i>	‘the Arabs’
<i>ar-rwala</i>	‘the Rwala’

As in many dialects, this list comprises the coronal consonants, with the exception of [ɖʒ] (although, as we have seen in §2.1.1.2, the articulation of /ǧ/ is often palatal in this dialect); this is considered to be because /ǧ/ was probably palatal or velar in early CLA and probably in Old Arabic as well. The affricated allophones *ć* and *ǧ* are also both coronal but do not assimilate the definite article in this dialect:

<i>al-ǧebel</i>	‘the mountain’	(not * <i>aǧ-ǧebel</i>)
<i>al-ǧebīle</i>	‘the tribe’	(not * <i>aǧ-ǧebīle</i>)
<i>al-ćalb</i>	‘the dog’	(not * <i>ać-ćalb</i>)

As indicated by the brackets in the form (a)l-, the vowel of the definite article is usually elided after vowels, although see below, e.g.:

<i>al-leyl</i>	the night
<i>be-l-leyl</i>	at night, in the night

It is presumed that glottal stop epenthesis occurs when the definite article is phrase initial (cf. §2.3.5.3):

<i>al-leyl</i>	>	ʾal. leyl
----------------	---	-----------

There are no instances in the corpus of the unelided vowel being anything other than *a* (even as *e*) and this is supported by other Rwala texts such as Ingham (1995).

Prochazka (1988: 21) provides evidence that in the Rwala dialect (and presumably in all closely related dialects, although he does not provide much detail on this point) the

definite article forms part of the phonological word with the noun it precedes, as shown by the fact that it is stressable, e.g.:

ba-l-ḥala > *bálḥala*
'in the desert'

This is attested without a preposition for the dialect of Ḥāyil which illustrates the point more clearly (Prochazka, *ibid.*):

al-ǧimal > *álǧimal*
'the camel'

al-barr > *álbar*
'the open desert'

In the second example, the stress placement has caused concomitant degemination of the coda of the following (unstressed) syllable, hence *al-bar* < **al-barr*. This situation contrasts with other dialects of Arabic in which the definite article forms a syntactic unit with the following nominal but a phonological unit with what precedes it (cf. Watson 2002: 61-2): here the definite article appears to be both phonologically and syntactically associated with the following nominal.

This situation is described by Palva (2011) as a general feature of bedouin dialects not shared by the sedentary populations of the same areas. Palva notes that the /a/ of the definite article is stable enough to maintain after V deletion even following /ī/, as shown in the following example:

rā' ál-ḡanam < *rā'ī*
'the shepherd of the sheep and goats'

Although stress patterns are not recorded by Musil, his data do suggest that the situation he heard was no different. Compared to Palva's example, we find two different situations in Musil:

rā' al-hawa
 'the beloved, the sweetheart'

rā'i l-beyt
 'the master of the tent'

Here, *al-hawa* has a syllabic structure that could potentially have antepenultimate stress (álhawa) while *al-beyt* certainly has final stress (albéyt). In this instance it seems likely that the antepenultimate stress on *al-hawa* has caused the elision of /i/, whereas if *al-hawa* had penultimate stress (as it does without the definite article: háwa) then it seems likely that the /a/ of the definite article would have been elided as in the example of *rā'i l-beyt*.

Palva (*ibid.*) also provides examples of the /a/ of the definite article being maintained when it is not stressed, and also following /ū/ as well as /ī/, but here the situation in Musil is varied. We may compare:

Palva	Musil
<i>f-al-bēt</i> (< fī)	<i>rā'i l-beyt</i>
'in the tent'	'master of the tent'
<i>ab al-bint</i> (< abū)	<i>ab al-ḥṣeyn</i>
'the father of the girl'	'fox'

The definite article is almost certainly unstressed in all of these examples, but in Palva it is maintained both times whereas in Musil only in the second example. This appears to represent a different situation. Dialects in which /a/ of the definite article is maintained

in all contexts all belong to the northwest-Arabian group (e.g. Ḥwēṭāt, Palva 1984-86: 306; Negev, Blanc 1970: 143; Bduur, Bani Yasin and Owens 1984: 219) but in Musil, whose material comes from the north-Arabian group, this is not the case, and the same /a/ is frequently elided. It seems that in Musil the /a/ of the definite article is maintained in a hiatus situation only when it is stressed, and the exceptions represented by irregular words such as *abu* ‘father of’ and *aḥu* ‘brother of’ probably represent a separate development.

5.1.15 *Indefinite article*

There is no morphological marker of indefiniteness, except for the adnominal linker (cf. §3.4). Nominals are indefinite unless marked for definiteness, which may be by means of the definite article, by being annexed to a definite noun, or by suffixation.

5.1.16 *Existential particle*

The majority of north-Arabian bedouin dialects use *fī-h* as an existential particle, ‘there is, there are’¹⁶⁵ (Palva 1980: 134, Cantineau 1936-37: 109), but this particle is absent from Musil’s Rwala material. Ingham describes the situation more explicitly, stating that in “dialects with /fi/ [as a preposition] the word for ‘there is’ is /fih/” (Ingham 1982a: 90). He assigns an existential particle derived from *b* (*b-ah*, *b-uh*, *b-aha*, *b-ih* depending on the

¹⁶⁵ **fī* and **bi* have merged in the north-Arabian bedouin dialect area where only *bi* is retained, as opposed to the bedouin dialects further south (approximately south of Ḥāyil) where a distinction is retained, i.e. *fī* designates ‘in’ and *bi* designates ‘at, by’ (cf. Ingham 1982a: 89-90). In those dialects where *fī* is not found, it may exist only as the existential form *fī-h* as here.

dialect) to the northern dialects (those in which *fī* is nonexistent as a preposition) and the existential particle *fī-h* to the central dialects (those in which *fī* also exists as a functioning preposition). This seems to be a slight oversimplification, since some dialects which do not have *fī* as a preposition retain *fī-h* as an existential particle (cf. Palva, *ibid.*). A clear example of *b-eh* as an existential particle is found only once in the Rwala material, in the following exchange:

- A: *mā-dri-n cān aḥad be-ha-l-gaṣr* (*< *mā adri en*)
 B: *b-eh*
 A: I don't know if there's anybody in this [stone built] house.
 B: There is.

It will be observed from this example that *b-eh* is a development of the regular use of *b* as a preposition¹⁶⁶, and there are therefore many examples in which the distinction between prepositional use (often as a relative clause) and use as an existential particle is blurred. The following example from a Euphrates Bedouin dialect is given by Ingham (1982a: 99) as existential and is closely paralleled in the subsequent Rwala example:

- ad-dinya sāḥi u lā b-ah ġēm*
 the weather is clear and there is no cloud

dirt ar-rwala maḥal mā b-aha rabī^c
 the Rwala territory is droughtland, [in which] there is no abundance

The development from **dire(t^{en}) b-aha rabī^c* 'land in which there is abundance' to a non-specific **b-aha rabī^c* 'there is abundance' is extant in related dialects, e.g. Qaṣīm,

¹⁶⁶ As Ingham notes, "the interior type continues the system of Classical Arabic insofar as there is no existential verb 'to be' ... In the outer [Mesopotamian, Gulf] type on the other hand actual existential verbal particles have developed" (Ingham 1982a: 100).

Shammar, ‘Aniza, Ḍafir and Euphrates Bedouin (Ingham, *ibid.*) and the one example above shows that it is found in the Rwala material too, although the prepositional use is much more frequent.

5.2 Pronouns, demonstratives and relatives

These categories are nouns in status but not form (Badawi *et al.* 2016: 30). They are unlike particles in that, with the exception of the relative pronoun *alli* (see below), they are inflected for number and gender, but they are unlike nouns in that they are not inflected for definiteness (being inherently definite) and in that they are closed classes, the members of which do not have consonantal roots.

5.2.1 Pronouns

The pronouns comprise a bound set and a free set. The free pronouns are the independent personal pronouns, and the bound pronouns are the possessive and object suffixes; the inflectional “verb endings” may also be considered bound subject pronouns but these are listed in §4.

5.2.1.1 Free independent pronouns

The system is as follows:

	basic forms	poetic forms
1CS	<i>ana</i>	<i>ani</i>
2MS	<i>ente, ent</i>	
2FS	<i>enti, ente</i>	
3MS	<i>hw, hu</i>	<i>hwa</i>
3FS	<i>hi</i>	<i>hiye</i>
1CP	<i>ḥenna, eḥna</i>	
2MP	<i>entom</i>	
2FP	<i>(e)nten</i>	
3MP	<i>hom</i>	
3FP	<i>hen</i>	<i>henna</i>

Where the basic form has two variants, these are in free variation. The forms listed as poetic are rare, even in poetry where the basic forms still predominate. The form *hu* is also very rare; I take Musil's transcription to indicate that in *hw* the vowel has been reduced to a central vowel (e.g. [ə]) in an unstressed position, but it is very difficult to speculate as to the articulatory nature of this clitic from the transcription given (Prochazka transcribes it as *huw*). Ingham (1995) also transcribes both *hu* and *hw* for *Rwala*, where the variation seems to match that of his two transcriptions of the conjunction /w(a)/ as *u* and *w-*, usually with the semivowel surfacing as a glide if it occurs in pre-vocalic position; in Musil, this distinction does not apply, as shown by the phrase *hu eben al-ʿamm* 'he is one of our close relatives' (not **hw eben*). It is unclear why Musil occasionally transcribed *hu*, when his data suggest that he simply generalised *hw*, whatever its articulatory value, across the system.

It is likely that the *ente* forms also show centralisation of two different phonemes, /a/ > e for the 2MS pronoun and /i/ > e for the 2FS pronoun (cf. §2.1.2.1), as *enti* is the variant form for the feminine and **enta*, while not recorded in this dialect, is the usual form

in related dialects, e.g. *anta* for the Hijazi dialects (but only *int*, *ant* for the north-Arabian dialects, Prochazka 1988: 125). There is no 1FS form in the dialect, which is usual in north-Arabia although it is found in some peripheral Arabian dialects such as in Bahrain (Holes 2016: 81) and in some Yemeni dialects spoken in the southwestern mountain range (Watson 2011c).

The 1S, 2S, 1P (*eḥna* form) and 2P forms do not have an underlying or phonemic glottal stop, as evidenced by the fact that they elide following vowels. A glottal stop will thus only be articulated when the pronoun is phrase-initial, as illustrated in the following syllabifications:

ana-rīd

I want...

/ana ariid/ > ?a. na. riid

mā-rīd ana

I don't want...

/maa ariid ana/ > maa. rii. da. na

illa-na

If I were not...

/illa ana/ > ?il. la. na

These pronouns are normally independent and function as subject:

ana-ḥu guṭne

I am Guṭne's brother

They are also frequently employed resumptively as a pseudo-copula:

rā'ī l-beyt mā hw ḥāḍer

the master of the tent is not at home [lit. *he* is not at home]

They are not always the subject, however, and can also be used in apposition to a participant in any case, e.g.:

al-walad alli ġā-k hāda li-ye ana
the boy that has come to you belongs to me

In this example, the pronoun *ana* is in apposition to the possessive pronoun *-ye* (an allomorph of *-i*, see below) and serves to emphasise it.

5.2.1.2 *iyya-*

In many dialects, including most Najdi dialects (cf. Ingham 1994a: 30), the word *’iyya* combined with the bound pronouns listed below creates an object pronoun described as a “compound free pronoun” (Badawi *et al.* 2016: 54), which could be considered a subset of the free pronouns. In the dialects of Syria and Iraq and some of the Gulf dialects this pronoun is frequently employed as a direct object pronoun (Ingham 1994a: 30), e.g. *ġilt-li-hum-iyyā-h* ‘I said it to them’ (*ibid.*). This usage is unattested in Musil, although evidence from the Najdi dialects indicates that it may well exist in Rwala too. In the corpus the form is only attested following *w-*; the resulting *wiyya*, described by Procházka (2011) as a reflex of CLA *wa-’iyyā*, is a functioning preposition ‘with’ in many dialects, such as in the dialects of Iraq in which *ma‘a* is not productive except in fossilised expressions (cf. Maamouri 2013: 562). In Rwala it is used exclusively after a subject pronoun in the sense of ‘with’, which pragmatically functions as a second subject, as indicated by verbal and adjectival agreement with both parties and not just the first:

mutaḥāwīn hw w-īyyā-ha ṭalāṭ sinīn

they [lit. he and she, he with her] were in hopeless love for three years

5.2.1.3 Bound objective and possessive pronouns

These two sets of pronouns are identical except for the 1CS form, for which *-i* is possessive and *-ni* is objective. The former is attached to nominals and prepositions and the latter to verbs.

	basic forms	poetic forms
1CS	<i>-i / -ni</i>	<i>-n</i>
2MS	<i>-ak (-k)</i>	
2FS	<i>-eċ (-ċ)</i>	
3MS	<i>-eh</i>	
3FS	<i>-ha (-ah)</i>	
1CP	<i>-na</i>	
2MP	<i>-kom</i>	
2FP	n/a	<i>-ken</i>
3MP	<i>-hom</i>	
3FP	<i>-hen</i>	<i>-henna</i>

There are a number of phonotactic consequences when these pronouns are suffixed which are outlined in this section.

The 1CS possessive suffix is unique in being the only vocalic suffix and this effects a change when it follows a vowel, but it is always *-i* after a consonant. Following a vowel the form is *-y* or *-yi/-ya/-ye*, e.g. *šakwa* > *šakwā-y* ‘my sorrow’, *mawla* > *mawlā-yi* ‘my Lord’ (here the preceding vowels surface as long because word-final vowel shortening does not apply). These transcriptions imply that two strategies are employed: in the case of *-y* the *i* is reanalysed as the semi-vowel /y/ (glide-formation) and in the case of *-yi* the hiatus is

resolved with the epenthesis of /y/ (glide-insertion), but the allomorphs *-ya* and *-ye* are harder to account for. The geminate forms *-yyi/-yye/-yya* are attested following *i*, e.g. *ḥawi* > *ḥawiy-yi*, *ḥawiy-ye* ‘my travelling companion’; this seems to be due to a convention of Musil’s transcriptions whereby /īy/ is represented as /iyy/, so there is no real discrepancy between *mawlā-yi* and *ḥawiy-yi* (= *ḥawī-yi*). The status of the variation (*-y ~ -yi ~ -ye ~ -ya*) is unclear but from the evidence available it does not seem to be phonologically conditioned. These forms are also used after diphthongs, e.g. *idēn* (> *idē*) > *idey-ye* ‘my (two) hands’, *bunay* > *bunay-ya* ‘my son [diminutive form]’.

The plural pronouns, the 3FS pronoun *-ha* and the 1CS object pronoun *-ni* are consonant-initial and the behaviour of this group is more or less homogeneous. They follow all vowels including diphthongs, and preceding vowels appear lengthened when word-final vowel shortening does not apply. They may follow a consonant but where this results in a medial superheavy syllable this regularly triggers epenthesis (cf. §2.2.2.1): epenthetic vowels are inserted in almost all cases after CC (*zaml-ana* ‘our pack camels’) to avoid a CCC sequence, including after geminate consonants (*rabb-ana* ‘our Lord’), but optionally after VVC (*‘eyāl-ana ~ ‘eyāl-na* ‘our children’), including when VV is one of the diphthongs, VwC or VyC (*ḥeyl-ana ~ ḥeyl-na* ‘our horses’).

The 3FS pronoun *-ha* behaves identically to the rest of this group, but in cases where epenthesis is triggered it freely varies with *-ah*, e.g. *‘ayūn-ah ~ ‘ayūn-aha* ‘her eyes’. *-ah* occurs in complementary distribution with the 3MS pronoun *-eh*, e.g. *‘ayūn-eh* ‘his eyes’.

‘*ayūn-ah* ‘her eyes’, and the vowel /a/ is unchangeable. This demonstrates that this is a case of allomorphy, and not of apocoptation (of *-aha* > *-ah*), because /ha/ may have either *a* or *e* as the epenthetic like the other consonant-initial pronouns¹⁶⁷, while the allomorph always has the form /ah/. There does not seem to be a distributional context for these allomorphs so it appears that they occur in free variation. /ah/ also occurs in a minority of cases (<8%) when epenthesis would not be triggered before /ha/, e.g. *ḥāber-ah* ‘knowing her [active participle]’ (not **ḥāber-ha*).

The 3MS pronoun, 2MS and 2FS allomorphs *-ak* and *-eċ*, and the 3FS allomorph *-ah* are vowel-initial; they may follow any C and their initial vowel is deleted if they are suffixed to a vowel-final word (cf. §2.3.5) or one which ends in a diphthong. Suffixation to C triggers a process of resyllabification, as the final C of the preceding syllable is resyllabified as the onset of the suffix syllable across the morpheme boundary, as with epenthesis.

It is apparent that the underlying representation (as well as, in most cases, the surface form) of these morphemes incorporates a vowel because the vowel is found after any C, and is not only induced by epenthesis in order to avoid unacceptable syllables. The only variation in this system is that after CVC *-ak* may occur as *-k* and *-eċ* as *-ċ*; the variation is roughly equal, with *-k* occurring after CVC in 60% of cases and *-ċ* occurring in 50% of cases

¹⁶⁷ There are also a very small minority of cases where the epenthetic vowel was recorded as *i*. As discussed in §2.2.2.1 these do not seem frequent enough to be significant. There may be some evidence that epenthetic *i* is somewhat motivated by vowel harmony, since it occurs after /i/ or /i/ (*ḥarīb-ina* ‘our enemy in war’, *ḡit-ini* ‘you came to me’) but the evidence is too scant to verify this with any certainty.

(-*eh* occurring as -*h* after CVC is unattested). This variation produces forms such as *šāreb-k* ‘your moustache’ (rather than **šāreb-ak*). Since it is unlikely that syncope would target a vowel in a closed syllable, we may posit that these are allomorphs: /k/ with /ak/ and /ć/ with /eć/. The distributional context of the allomorphs is not clear: there seems to be preference for /ak/ and /eć/ after guttural and emphatic consonants, but this is not without exception. After all other consonants the forms seem to be in free variation, which is particularly noticeable after the feminine suffix /at/ where both allomorphs are frequently attested. /k/ and /ć/ almost always occur after CVC but do occasionally surface after CVCC or CVVC, although these cases (e.g. *galb-k* ~ *galb-ak* ‘your heart’, *ḍelūl-k* ~ *ḍelūl-ak* ‘your riding camel’) are rare and seem anomalous¹⁶⁸. Prochazka records both /k/ and /ć/ allomorphs for Rwala, and Cantineau considers these allomorphs to be a defining characteristic of the ‘Anazi group. Their descriptions unfortunately do not give details about the variation, so it is impossible to speculate as to whether the forms are underrepresented in Musil, although they do not specifically state that /ak/ and /eć/ never occur in the same environment.

This system mostly parallels that described by Ingham for Najdi dialects in general (Ingham 1994a: 30, illustrated by Sudayri) in which the same three pronouns (3MS, 2MS and

¹⁶⁸ Final CVCCC syllables are allowed in San’ani Arabic according to Watson (2007: 348), e.g. *mā libis-t-š* ‘I did not wear’, where concatenation (1CS suffix *t*, negative marker *š*) has produced a final CVCCC; she explains this rare case however as being permissible only because *tš* is considered an affricate, [tʃ], which takes up only one mora, effectively making it CVCC. No such explanation exists here, however, and there is no concrete evidence for the allowance of final CVCCC syllables.

2FS) are vowel-initial and the rest are consonant-initial and may trigger epenthesis. In Sudayri the 3S pronouns are also differentiated at a phonemic level, with /ih/ being the 3MS pronoun and /aha/ the 3FS, although in Sudayri, unlike in Rwala, both of the 2S pronouns have /i/. We might postulate that in Rwala the 2FS *-eć* and 3MS *-eh* represent /ić/ and /ih/ respectively, but /i/ is only realised as *e* in this context and never as *i*. It is therefore perhaps more useful to view *e* here as a marginal phoneme /e/, which is the approach taken by Al-Wer (2011) for the dialect of Amman. The 2S pronouns are also the basis¹⁶⁹ for the marginal phonemic status of /ć/, on account of the post-vocalic forms, e.g. *hawā-k* ‘your (MS) love’ contrasts with *hawā-ć* ‘your (FS) love’.

The objective pronouns express direct objects on verbs:

asharat-ni
she woke me up

as well as indirect objects:

e‘tī-ni ‘elūm-ak
give me your news

In this example the indirect object could also be prepositional if the direct object were suffixed, **e‘tī-ha l-i* ‘give it (FS) to me’. It is also possible to use two pronouns:

tefaṭṭen-ni-h
she reminds me of it

e‘tī-na-h
give it to us

¹⁶⁹ This, in addition to the contrast established in highly grammaticalised contexts, e.g. *keyf* ‘habit, pleasure, rest’ and *ćeyf* ‘how?’ (cf. §5.1.10).

The resulting compounds can be rather unpredictable in form, however. When the first pronoun ends in a vowel (as in the above examples), they combine easily and the fact that they are not long indicates that these pronouns have short vowels in their underlying representation (cf. §2.3.3). When the first pronoun ends in a consonant and the following pronoun is consonant initial, epenthesis is induced; from the few examples available, there seems to be no rule to determine the nature of the epenthetic vowel, e.g. the long vowel in ‘*atā-k-ī-ha*’ ‘he gave it to you’, which is hard to account for. The case of ‘*aṭeytī-n-eh*’ ‘you FS gave it to me’ is unusual since the vowel of *-eh* has been preserved over the vowel of *-ni* producing *-n-eh* (cf. *-ni-h* in the above example).

The possessive pronouns on nouns simply indicate the possessor of the referent of the noun:

ṣāḥeb-i
my friend

Their use on prepositions is detailed separately below. Nouns, because they are in a state of annexation with the pronoun, lose all markers of definiteness and indefiniteness (cf. §5.1.14 and §5.1.15). The possessive pronouns are also attached to verbal nouns, and care should be taken to distinguish in this case the agent function from the patient function, as illustrated by the following examples:

nawa ḡatl-i
he resolved to kill me (lit. my killing)

ḥagg eḏḥār-i l-eh min ha-l-ḥufra
my reward for bringing him out of that hole (lit. the reward of my

bringing)¹⁷⁰

5.2.1.4 Bound pronouns on prepositions

The behaviour of the prepositions with attached pronouns is somewhat unpredictable, as Table 5.1—which lists *all* attested forms (the *p* indicates that it is only found in poetry)—demonstrates. In many cases the variation is predictable from what has already been described: the optional epenthetic vowel exemplified in *bēn-ana* ~ *bēn-na* and *bēn-akom* ~ *bēn-kom* has already been noted, as has the free variation seen in *‘aley-yi* ~ *‘aley-ye* for example. Other cases are specific to these prepositions, such as the situation of *l* and *b* which have no syllable nucleus of their own and are thus constrained to form a complete syllable with the attached pronoun, which has resulted in multiple variants. A complete list of prepositions is found in §5.1.13.

It will be seen that *min* and *‘an* have geminate forms *minn-* and *‘ann-* (cf. §2.3.4) which occur before the vocalic suffixes (*-i*, *-ak*, *-eċ*, *-eh* and *-ah*) but non-geminate forms occur before the non-vowel-initial allomorphs *-k*, *-ċ*, and *-ha.*. This is a feature found in many dialects of Arabic, and which occurs in CLA before the 1CS suffix giving *minn-ī*, *‘ann-ī* (cf. Wright 1967 vol. 1: 281 *Rem. b*). In §2.2.2 it was seen that monomoraic words such as *min* and *‘an* are sub-minimal; the geminate forms *minn* and *‘ann* however are bimoraic and meet the minimal word requirement. The gemination of /n/ is therefore explained by

¹⁷⁰ The preposition *l* here introduces the direct object.

TABLE 5.1: Bound pronouns with prepositions (Musil)

	l	b	min	‘end	ma‘	‘ala	miṭl	‘an	bēn
1CS	<i>li</i> <i>liye</i> <i>liyye (p)</i> <i>leyya (p)</i>	<i>bi</i> <i>beyye (p)</i>	<i>minni</i>	<i>‘endi</i>	<i>ma‘i</i>	<i>‘aley</i> <i>‘aleyye</i> <i>‘aleyyi</i> <i>‘aley (p)</i>	<i>miṭli</i>	<i>‘anni</i>	-
2MS	<i>lak</i> <i>lāk</i>	<i>bak</i>	<i>minak</i> <i>mink</i>	<i>‘endak</i>	<i>ma‘ak</i>	<i>‘aleyk</i>	<i>miṭlak</i>	<i>‘ank</i>	<i>bēnak</i>
2FS	<i>le‘</i> <i>lā‘ (p)</i>	<i>bi‘</i>	-	<i>‘ende‘ (p)</i>	-	<i>‘aley‘</i>	-	<i>‘anne‘ (p)</i>	-
3MS	<i>leh</i>	<i>beh</i>	<i>mineh</i> <i>minneh (p)</i>	<i>‘endeh</i>	<i>ma‘eh</i>	<i>‘aleyh</i>	<i>miṭleh (p)</i>	<i>‘anneh</i>	-
3FS	<i>laha</i> <i>lha (p)</i> <i>leha (p)</i> <i>lah (p)</i>	<i>bha</i> <i>bah (p)</i> <i>bāh (p)</i> <i>beha (p)</i> <i>biha (p)</i> <i>biha (p)</i> <i>baha (p)</i>	<i>minha</i>	<i>‘endah (p)</i>	<i>ma‘aha</i>	<i>‘aleyha</i>	-	<i>‘anha</i>	-
1CP	<i>lena</i> <i>lana (p)</i>	<i>bīna (p)</i> <i>bena (p)</i>	<i>minna</i>	<i>‘endena</i> <i>‘endana</i>	<i>ma‘ana</i> <i>ma‘na (p)</i>	<i>‘aleyna</i>	-	<i>‘anna</i>	<i>bēnana</i> <i>bēnna</i>
2MP	<i>lekom (p)</i> <i>lakom (p)</i>	<i>bikom (p)</i>	<i>minkom</i>	<i>‘endakom</i>	<i>ma‘kom</i>	<i>‘aleykom</i>	<i>miṭlekom</i>	-	<i>bēnakom</i> <i>bēnkom</i>
2FP	-	-	-	-	-	-	-	-	-
3MP	<i>lahom</i> <i>lehom</i> <i>lhom</i>	<i>bihom (p)</i> <i>bhom (p)</i> <i>behom (p)</i>	<i>minhom (p)</i>	-	<i>ma‘hom</i>	<i>‘aleyhom</i> <i>‘alēhom</i>	-	<i>‘anhom</i>	<i>bēnhom</i> <i>bēnahom</i> <i>(p)</i>
3FP	<i>lehen (p)</i> <i>lahen (p)</i>	<i>bihen (p)</i> <i>behen (p)</i>	-	-	-	<i>‘aleyhen</i> <i>(p)</i>	<i>miṭlehen</i>	-	-

Watson (2002: 205-206) for San’ani and Cairene Arabic as fulfilling a requirement that suffixes “are affixed to a form at least the size of the minimal phonological word”. She goes on to say that “in fast, casual speech the gemination of /n/ may fail to apply before consonantal suffixes” (*ibid.*), which may account for the non-geminate forms, although the situation as attested in Rwala is not directly comparable. Firstly, the Musil corpus attests

non-gemination also before vocalic suffixes (e.g. *min-eh*), and, secondly, the dialects Watson describes attest geminate forms such as *minn-ina* and *minn-anā* (Cairene and San’ani respectively) which are not attested at all in Musil, and variation cannot be proven to occur. It is interesting that *min-k* and *‘an-k* are attested in the corpus to the exclusion of *minn-ak* or *‘ann-ak* (although a larger sample may have shown different results); indeed, the dialect seems to show a preference for the non-geminate forms, perhaps over and above what could be attributed to “fast, casual speech”. At first glance, therefore, the Rwala dialect seems not to have the requirement that the suffixes must only be affixed to a bimoraic syllable, since forms such as *min-k* are frequent. These forms ultimately fulfil (in their concatenated [suffixed] state) the minimum word requirements (*ma*‘ ‘with’ is also a degenerate foot prior to suffixation); it may also be argued, as Watson does, that the additional mora of *minn* and *‘ann* is maintained pre-consonantly even when gemination does not occur. This would have to apply pre-vocalically as well in the Rwala dialect, unless forms *min-ak* and *min-eh* are considered anomalous (on the basis that each only occurs once). Otherwise, the data imply that the dialect permits any outcome in which the preposition *after* suffixation fulfils the minimum word requirement, which would in turn suggest a process of lexicalisation of the suffixed forms of these prepositions. All the other forms shown in the table (apart from under *l* and *b*) show only predictable variation, since the prepositions already constitute minimal words.

These forms may be compared with those given by Prochazka for the Rwala dialect in Table 5.2.

Table 5.2: Bound pronouns with prepositions (Prochazka)

	l	b	min	ʿan	maʿ
1CS	<i>li</i>	<i>bi</i>	<i>minni</i>	<i>ʿanni</i>	<i>miʿi</i>
2MS	<i>lak</i>	<i>bak</i>	<i>mink</i>	<i>ʿank</i>	<i>maʿak</i>
2FS	<i>lić</i>	<i>bić</i>	<i>minć</i>	<i>ʿanć</i>	<i>miʿić</i>
3MS	<i>lih</i>	<i>bih</i>	<i>minih</i>	<i>ʿnih</i>	<i>miʿih</i>
3FS	<i>laha</i>	<i>baha</i>	<i>minha</i>	<i>ʿanha</i>	<i>maʿha</i>
1CP	<i>lina</i>	<i>bana</i>	—	—	<i>miʿna</i>
2MP	<i>ləkum</i>	<i>bəkum</i>	—	—	<i>miʿkum</i>
2FP	<i>lićin</i>	<i>bićin</i>	—	—	<i>miʿkin</i>
3MP	<i>ləhum</i>	<i>bəhum</i>	<i>minhum</i>	<i>ʿanhum</i>	<i>mi/aʿhum</i>
3FP	<i>lihin</i>	<i>bihin</i>	—	—	<i>miʿhin</i>

These data are not generally at odds with Musil’s forms. Prochazka’s examples possibly suggest that *l* and *b* have different vowels in their underlying representation, although this only surfaces in the forms *lina* and *bana*; in Musil, the ambiguous quality of the epenthetic vowel *e* makes it difficult to make a comparison, however. *min* and *ʿan* show even less /n/-gemination in Prochazka, although his forms do show the expected /i, u/-syncope in the 3MS forms where Musil’s *ʿanneh* would not be targeted for syncope on account of gemination rendering the first syllable closed, but his *mineh* theoretically would be targeted by syncope. It is interesting to note that, as in Musil, the paradigm of *maʿ* does not show *ghawa*-syndrome resyllabification, as described in §2.3.1.

5.2.1.5 Interrogative pronouns

Some of the interrogatives are pronouns. Although the “question words” do not belong in the same morphological or even syntactic category, it has seemed sensible nonetheless to group them together; indeed, this is standard practice in many descriptions of Arabic varieties, even to the point that the whole group is often referred to as the interrogative pronouns despite the fact that only a minority of them are technically pronouns. They are described individually in §5.1.10, and further reference to the function of the pronouns as relatives can be found in §5.2.3.3.

5.2.2 *Demonstratives*

Demonstratives in all dialects of Arabic are morphological composites in which up to three basic morphemes are combined: a presentative morpheme, a deictic element which denotes number and gender, and a morpheme which denotes distance, the absence of which denotes proximity (Holes 2004a: 184-5). Table 5.3 shows all the demonstratives attested in Musil arranged by their combination of morphemes.

Table 5.3: Morpheme combinations of attested demonstratives

	h	ḍ	h + ḍ	h + k	ḍ + k	h + ḍ + k	ḍ + l (+ k)
MS	<i>ha</i>	<i>ḍa</i>	<i>hāḍa</i>	<i>hāk</i>	<i>ḍāk</i>	<i>haḍāk</i>	[<i>ḍāl</i>], [<i>ḍālek</i>]
FS	<i>ha</i>	–	<i>hāḍi</i>	<i>hāk</i>	<i>ḍiċ</i>	–	–
CP	<i>ha</i>	<i>ḍowla</i>	<i>haḍōlen</i> , <i>haḍōla</i>	<i>hāk</i>	<i>ḍowlāk</i>	<i>haḍowk</i>	–

In all the above examples¹⁷¹, /ha/ (or /haa/) is the presentative element, the deictic morpheme is /ḍa/ (m.), /ḍi/ (f.), and /ḍoo/ or /ḍool/ (pl.), and the /k/ (=ĉ) morpheme denotes distance from the speaker. Some of these elements are optional (e.g. *ḍa* is functionally the same as *hāḍa*), so the following table presents the full set of possible demonstratives, with brackets marking those elements which may be omitted:

	proximal	distal
MS	(hā)ḍa	(ha)ḍāk
FS	(hā)ḍi	(ha)ḍiĉ
CP	(ha)ḍōlen/ḍōla	(ha)ḍowk/(ha)ḍowlāk
indecl.	ha, ḍa	ha, hāk

The Rwala evidence concurs with Ingham’s observation that in Najdi Arabic *ha*, *ḍa* and *hāk* can occur with nouns of any number or gender (Ingham 1994a: 55). The material also provides evidence of the trend in Arabian dialects for the demonstratives to be employed as substitutes for the free personal pronouns in the third person (Ingham 1994a: 56, Holes 2004a: 185), e.g.:

min nām kill al-leyl haḍāk habīl
 he who sleeps all night [he] is an idiot

al-walad alli ḡā-k hāḍa liy-ye
 the boy who has come to you [he] belongs to me

¹⁷¹ *ḍālek* and *ḍāl* have been discounted on the basis of what appears to be a misunderstanding on the part of Musil: he uses *ḍālek*, which he presumably borrows from Classical Arabic, to gloss what he interpreted as a demonstrative “*ḍāl*” but the word probably doesn’t exist. The two forms he records as *ḍāl ḍa* and *ḍa’an ḍa* should rather be parsed as *ḍa l-ḍa* and *ḍa ‘an-ḍa*, i.e. ‘that to that’ and ‘that from that’.

In both cases the sentence would have been perfectly grammatical with *hw* ‘he’. The distinction between proximal and distal is conveniently illustrated in the following phrase:

be-l-gurba haḏōla agrab le-na min haḏōlak

in terms of proximity, these ones are closer to us than those ones

The demonstrative can function both as a determiner and as a pronoun, and the differentiation of these functions is explained in the following paragraphs. As a determiner, the demonstrative + noun relationship is one of apposition, so the demonstrative and noun must therefore agree in number, gender and definiteness; since the demonstrative is inherently definite, the noun must always be definite. If the definiteness of the noun is marked by means of the definite article, the demonstrative precedes and the noun is in apposition to it:

(m.) *hāḏa r-raḡḡāl*
this man

(f.) *hāḏi s-sene*
this year

(pl.) *hāḏi l-aḑāl*
these deeds

haḏōla l-‘arab
these tent-dwellers

Note that in the first plural example, the agreement rule for non-human plurals requires the feminine singular demonstrative, and there is no instance of a non-human plural agreeing with a plural demonstrative, unlike the variation with adjectives, verbs and

pronouns described in §3.1.4. Human plurals, however, may occur in strict agreement with a plural demonstrative, as illustrated by the second plural example.

The indeclinable demonstratives also precede the noun:

(m.) *ha r-raġġāl*
this man

(f.) *ha s-sene*
this year

(pl.) *ha l-‘arab*
these tent-dwellers

Badawi *et al.* (2016: 149) observe that “the preposed demonstrative can be seen as an extension of [the definite article], creating a cluster of quasi-prefixed deictic elements”, and this seems to hold especially true for the indeclinable demonstrative, which must precede a noun with the definite article and cannot occur in postposition with a determining function.

When the demonstrative is a determiner qualifying an annexation structure, including nouns suffixed with bound pronouns, it follows the unit and is therefore in apposition to it:

(m.) *walad-i hāḍa*
this boy of mine

(f.) *dīnyā-k hāḍi*
this world of yours

(pl.) *‘elūm-i hāḍi*
these tidings of mine

The demonstrative in postposition may also be used for human plurals with the definite article, and *al-‘arab haḏōla* is attested as a variant of *haḏōla l-‘arab* ‘these tent-dwellers’, perhaps with an emphasis on the demonstrative (‘*these* tent-dwellers, as opposed to some others’).

When the demonstrative is used as a pronoun, it may stand alone and can be used in all the functions that a noun can. If it is the head of a nominal sentence, the syntax indicates that it is not a determiner in the following ways. It may precede indefinite nouns, in which case it clearly cannot be a determiner:

hāḏa bawwāg
this man is a traitor

It may precede annexation structures, which is a reversal of the word order when it has a determining function, compare:

hāḏa walad-i
this is my boy

walad-i hāḏa
this boy of mine

When the predicate of the nominal sentence is a noun with the definite article, a separating pronoun must be inserted in order to distinguish this construction from that of the demonstrative as determiner. This is known in traditional grammars as *ḏamīr al-faṣl*, the “distinguishing pronoun”.

hāḏa r-raḡḡāl
this man

hāḍa hw ar-raġġāl
this is the man

Note that when the declinable demonstratives are used in nominal sentences, they still agree with the predicate in number and gender, though not in definiteness, and the rule for non-human plurals still applies:

hāḍi manāzel-hom
this (FS) is their camping place (pl.)

The indeclinable demonstratives can also be used as pronouns:

ha liy-ye
this is mine

The nominal demonstrative may perform any of the other syntactic functions of the noun, which do not all need to be demonstrated here. When the construction is not subject + predicate, there are no restrictions on word order because these constructions are generally not confusable. Thus as the object of a verb the demonstrative is free to be in pre- or postposition (although the resumptive pronoun is required when the object is fronted), e.g.:

mā gilt hāḍa
I didn't say this

hāḍi teġib-ah
you are bringing this

The demonstratives, especially the indeclinable set, are used very frequently as heads of relative structures:

b-ha lli zell

in this month just passed [lit. in this which passed]

It should be noted that although all of the examples in this section have shown the use of the proximal set of demonstratives, the distal set is used in exactly the same way.

5.2.2.1 Presentatives

The combination of *ha* with personal pronouns produces a presentative, of which there are two types recorded in the data set.

The first type is only for the second person and is formed of *hā* + bound pronoun, which is attested only for 2MS: *hā-k*. It is used when the speaker is physically pointing at or handing over something:

hā-k gḏāḏt-i

here, take my handkerchief

The other type of *ha*-presentative is formed of *hā* + free personal pronoun. Its function is described by Holes as “a means of foregrounding a fact or an event as pragmatically or discoursally new, especially if unexpected or contrary to expectation” (Holes 2016: 89). It is attested in the following persons, although it is not unreasonable to assume that the whole set is feasible:

1CS	-
2MS	-
2FS	-
3MS	-
3FS	<i>hā hi</i>
1CP	<i>hā ḥenna</i>
2CP	-
3MP	<i>hā hom</i>
3FP	-

Its usage is very contextual and is best illustrated with examples:

hā hi l-‘aša

this here is the supper (physically pointing at it)

w-hā hom ḥalgen‘ cētīr

they were much stronger numerically [lit. they were a numerous group]

In this example *hā hom* creates a comparative element, cf. Bahraini *hāk-hum yīṭla‘ūn*

‘they would go out of the house [contrasting with today’s generation]’ (*ibid.*: 90).

hā ḥenna ḡīnā-k

behold, here we are [lit. we have come to you] (unexpected event)

The usage of grammaticalised imperatives from the verb of seeing, namely *tara* and *er/er’/er‘*, appear to be functionally presentatives. Musil routinely translates these with their original lexical sense, e.g.

er‘-aha ‘end ahal-ha

but see, she stays with her kinsfolk

Here it seems likely that a more plausible interpretation would be emphatic: ‘*she* stays with her kinsfolk’ or ‘she is actually staying with her kinsfolk’. Similar usage of this

particle is recorded for the Negev dialect in Blanc (1970: 144-145). Both particles frequently take pronouns, as the *ha*-presentatives do.

5.2.3 *Relatives*

The main relative pronoun is *alli*, which is usually elided following a vowel to *-lli*. Forms of this pronoun are found across the dialects of Arabic: Cohen (1962: 140) hypothesises that a form *illi* was prevalent in Old Arabic and has survived in virtually all the dialects, but that it was not standardised in CLA, which adopted the less used form *allaḍi* (*allaḍi* and the feminine form *allati* are attested infrequently in Musil in poetry). Despite being uninflected, *alli* is included in §5.2 on account of being, at least in some cases, functionally pronominal.

5.2.3.1 With a definite antecedent

The relative pronoun is employed in relative clauses which qualify a definite nominal antecedent. In the simplest cases the subject of the relative clause and the antecedent are coreferential, e.g.:

naḥart al-ḥaṭṭāṭa alli ʿend-ana

I went to the soothsayer who was staying with us

In cases where the subject of the relative clause is a new participant, and therefore is not coreferential with the antecedent, a resumptive pronoun must be employed. This applies to all positions on the accessibility hierarchy below the position of subject, namely: direct or

indirect object, prepositional object, and genitive¹⁷². For example, as direct (here) or indirect object¹⁷³:

al-aḍm alli yeṭhan-eh
the bone he is grinding

as prepositional object:

eṭla^{ci} ma^c makān-e^c min alli ḡṭi minn-eh
crawl back to the place you came from!¹⁷⁴

and as genitive:

allāh alli ^cazīz^{en} šān-eh
God, whose power is eminent

There are, however, cases in which resumption does not occur. Variation is mostly found when the antecedent has the role of direct object in the relative clause, allowing phrases such as:

nehd-i alli tnašša
my breast, which he smelled

¹⁷² According to Comrie (1989: 155), languages place restrictions on the syntactic function of the antecedent (i.e. the noun or noun phrase being modified) within the relative clause. The accessibility of the positions is ordered cross-linguistically from most to least accessible in the following order: subject, direct object, indirect object, oblique, genitive, and object of comparative (Keenan and Comrie 1977: 66). This dialect is able to relativise all positions by means of *alli*, although all positions other than subject require resumption (and relativisation of the object of a comparative is not attested in the corpus).

¹⁷³ No actual examples of the antecedent becoming an indirect object in a relative clause are attested, although an example (reworded from the attested *‘aṭeyti-n-eh* ‘you FS gave it to us’) would be **ḥenna alli ‘aṭeyti-n-eh*, **ḥenna alli ‘aṭeyti-h la-na* ‘we to whom you gave it’.

¹⁷⁴ Note that the first *min* is redundant here.

Here we would expect to see **tnaššā-h*. The suppression of the resumptive pronoun occurs in 30% of attested cases, and usually in poetry, where it may be omitted for metrical reasons. Further down the accessibility hierarchy the phenomenon becomes rarer; specifically, there are only two cases for prepositional object and one for genitive.

5.2.3.2 With an indefinite antecedent

As is the case in most dialects of Arabic, the relative pronoun as detailed above is only used when the antecedent is definite, whether by means of the article, or the construct state, or if it is definite by nature (Ángeles 2011). When the antecedent is indefinite, *alli* is not used and the relative clause is juxtaposed to the main clause, as in:

lā tebċīn amr^{en} maḍa
do not mourn something [which] has passed away

Instead, relative clauses of this type are marked with the adnominal linker (see §3.4). Again, resumption will occur if the indefinite antecedent has any role in the relative clause other than that of the subject; for example, as a direct object:

asrār^{en} ketam-hen bi-ṣandūg
secrets [which] he had locked up [as if] in a box

5.2.3.3 With a zero antecedent

As well as being used to modify a defined antecedent, there are cases where *alli* is used pronominally and the antecedent is assumed. The relative pronoun (and assumed antecedent) may have any syntactic role in the main clause, for example, subject:

alli yākol al-ḡedi yaḥma umm-eh
he who eats a young goat must protect its mother

or object:

hātī lli warā-ć
bring what I have paid for you!¹⁷⁵

As above, resumption occurs when the subject of the relative clause is not coreferential with the (zero) antecedent, whether the antecedent is the direct or indirect object of the relative clause:

rāćeb alli mā laheḡ-ha l-ḡenīni
the rider of [a she-camel] which calves have not yet oppressed¹⁷⁶

or a prepositional object:

alli mandūr^{en} ʿaley-h
the one for for whose benefit a vow was made

or a genitive:

ten^{ʿa} lli b-ḥadd-ah rešūmi
you mourn [a girl] on whose cheek [there are] tattoos

¹⁷⁵ Lit. ‘bring [that] which [is] behind you’, where ‘behind you’ alludes to the bride-price that the groom has paid for his wife.

¹⁷⁶ This refers to a camel which has never given birth.

The pronouns *mā* and *min* can also be used to form relative clauses, although they are used less frequently than *alli*. These usually function as relatives with an assumed antecedent (as in the first example below), *mā* meaning ‘that which’ and *min* meaning ‘he/she who’; in the second example below *min* does have an antecedent and just means ‘who’:

allāh ya‘mal mā yerīd
God will do what he wants

al-‘eyn min ġā-h al-ġurāb whīd
the eye [of him] whom the single raven has approached¹⁷⁷

5.3 Nominal forms

The remaining sections in this chapter concern items which are morphologically nominal but which have acquired specific grammatical function. These all appear to be closed classes; this contrasts with the adverbials and prepositionals given above (§5.1.12-13) which also have nominal form but which may represent classes which are still open.

¹⁷⁷ In this example, I interpret the eye (*‘eyn*) to be a synecdoche for the person whose eye it is, since *min* refers to a person and not just an eye, and furthermore the resumptive pronoun *-h* is masculine whereas *‘eyn* is feminine. The supplied translation ‘[of him]’ does not imply that *min* is pronominal here, since the definite *al-‘eyn* cannot be the first element in a construct phrase with *min*.

5.3.1 *Quantifiers*

5.3.1.1 *kill*

kill is a nominal determiner meaning ‘each, every, all’. It occurs in annexation to the noun it qualifies. With an indefinite singular, the meaning is ‘each, every’, with or without the adnominal linker on the noun:

kill ʿūd^{en} yibes
every stalk dries up

kill yowm
every day

With a definite singular or plural the meaning is ‘all, the whole of’:

kill al-ğezīre
the whole of the Ğezīrah (the land between the Tigris and the Euphrates)

ğammaʿū kill al-aḏāra
they gathered all the maidens

It is also annexed to relative clauses, meaning ‘everyone who’ or ‘everything that’:

kill min ġā-h
every passer by [lit. everyone who came¹⁷⁸]

arḥašt l-eh kill mā-rād
I permitted him whatever [lit. everything that] he wished

It should be noted that *killma* is also a conjunction meaning ‘whenever’, which does not represent the annexation of *kill* to a relative clause, e.g.:

¹⁷⁸ *ğā*, an intransitive verb, commonly takes an impersonal object pronoun *-h*. This seems to express the destination, i.e. *ğā-h* ‘he came [here]’; it has become lexicalised as a by-form of the verb (*ğā* = *ğā-h*) and does not require translation. It is unclear to what extent the two are acoustically dissimilar.

killmā-gūl ana-nsā-h

(< *killma agūl ana ansā-h*)

whenever I say that I will forget him

In addition to occurring in annexation, *kill* also occurs in apposition to the noun it determines and has an attached pronoun, although this is less frequent:

ḥālow ʿala l-ḡeyš kill-eh

they surrounded all the camel riders

Such appositional usage is most common when the noun it determines has already been expressed as a pronoun, e.g.:

min nām ṭilt al-leyl nām-ha kill-aha

he who slept a third of the night [it is as though he] slept all of it

ar-reḡāl mā hi kill-eha raḡḡāl

not all men are heroes [lit. men, they are not all heroes]

In the first example the speakers inclusion of *-ha* ‘it’ after *nām-ha* makes *kill-aha* an appositional annexation; the same meaning could have been expressed **nām kill-aha* ‘slept all of it’ although in this case the *kill-aha* would be pronominal (see below). In the second example *hi* is resumptive (see §5.2.1.1) and triggers the annexation of *kill-eha* in apposition to *hi*, which is by definition in apposition to the subject *ar-reḡāl*. The other structure in which apposition commonly occurs is in relative clauses (cf. §5.2.3) where a resumptive pronoun is required; *kill* may be employed here:

al-māl alli ḥešet kill-eh

all the property you stored up [lit. the property all of which you stored up]

In this case it is clear that *kill* could be removed, leaving **al-māl allī ḥešet-eh* ‘the property you stored up’, and also that the resumptive pronoun could be duplicated as in the first example above, **al-māl allī ḥešet-eh kill-eh* ‘all the property you stored up’.

In addition to meaning ‘each, every, all’, *kill* routinely has the meaning ‘any’. This is especially frequent in negative phrases, although the third example shows that this occurs in affirmative phrases as well:

mā talḥag-ni kill ‘obeyye
no ‘obeyye mare could overtake me

zahar mā lāg bi-kill wādi
such blossoms as are not found in any other valley

aṭ-ṭarš kill^{en} yehāb-eh
any herd would be afraid of him¹⁷⁹

kill^{en}, always with the adnominal linker, is also used as a free form, meaning ‘each one’ or ‘all of them, everyone’:

kill^{en} tāḥoḍ ‘ašīr-ah
each [woman] will take her beloved

Such pronominal usage is frequently resumptive, as seen in §5.2.1.1 in the case of the personal pronouns:

al-bedw w-al-ḥaḍrān kill-hom swiyye
bedouins and sedentaries are all equal

¹⁷⁹ This phrase is poetic and therefore the language is less clear than in the other examples from this section. There are cases of *kill* occurring in postposition, hence this reading. It is possible, of course, that the *kill^{en}* in this instance is functioning pronominally in apposition to the subject, i.e. “[in respect of] herd[s], any would be afraid of him”. In order to provide the modal sense of “would” I would expect *kill ṭarš yehāb-eh* to be the simplest strategy, with the definite form employed here poetically.

In this example, *kill-hom* could be understood to be in apposition to *al-bedw* and to *al-ḥaḍrān*, meaning ‘all the bedouins and sedentaries are equal’. Since a clause of this type would usually trigger a pronoun, however (e.g. **al-bedw w-al-ḥaḍrān hom swiyye* ‘bedouins and sedentaries [they] are equal’), we may understand *kill-hom* to be acting as a resumptive pronoun in this case; i.e. ‘bedouins and sedentaries, they are all equal’.

5.3.1.2 *ḡamīʿ*

Like *kill*, *ḡamīʿ* is annexed to a definite noun with the sense of ‘all’:

tmaḍḥalat ʿan-na ḡamīʿ al-ḥaċāya
all the talk about us has come to nought

ḡamīʿ has a more restricted usage than *kill*, however, and does not occur with indefinite nouns in the sense of ‘each’; this is the same situation as in MSA, for example (Badawi *et al.* 2016: 258), as could be expected from the semantic connotation of togetherness expressed by the root **gmʿ**. Appositional usage is not attested in this corpus, but its occurrence in other dialects suggests that it may be possible. As a free form, *ḡamīʿ* or *al-ḡamīʿ* means ‘everyone’, ‘all’:

ḥāl al-leyl ʿala ḡamīʿ
night covered all of them

gowm ʿala l-ḡamīʿ
war against all of them

5.3.1.3 *ba‘aḏ*

ba‘aḏ is annexed to definite plural nouns and means ‘some, any’:

hēḡen ḥāderen ba‘aḏ ar-riyā‘

a male ostrich descending through some defile

kuff aš-šerr ‘an ba‘aḏ al-mela

prevent evil from (being committed by) certain people

5.3.2 *Reflexive pronouns*

The derivational system permits that many reflexive constructions are covered synthetically by verbal patterns, in which case the use of reflexive pronouns is not required.

Compare:

ḡasel al-‘abd be-š-šābūn

he washed [his] stomach with soap

taḡassal be-š-šābūn

he washed himself with soap

Here, the FORM-V pattern distinguishes the reflexive verb, ‘to wash [oneself]’, from the transitive FORM-I, ‘to wash [+ obj.]’. These distinctions do not represent a productive reflexivising strategy but rather a (fairly extensive) set of lexicalised reflexive verbs. The use of certain grammaticalised nominals as reflexive pronouns allows for analytic constructions to exist as well, however, and this strategy is completely productive. The nominals in

question are *ḥāl*, *nafs* and *nesem*¹⁸⁰, with attached possessive pronouns. The words literally mean ‘state’ (*ḥāl*) or ‘soul’ (*nafs*, *nesem*) but have lost their semantic meaning in the process of grammaticalisation. The forms of ‘myself’ are therefore:

	<i>ḥāl</i>	<i>nafs</i>	<i>nesem</i>
1CS	<i>ḥāl-i</i>	<i>nafs-i</i>	<i>nesm-i</i>

Regrettably, the construction is not attested in the corpus as a plural, and it is therefore uncertain whether the form of the plural pronoun should attach plural pronouns to the singular base form, e.g., **nafs-ahom* ‘themselves’, or whether the noun itself should also be plural, e.g. **anfōs-hom* ‘themselves’. The plural form is found in CLA and possibly in the Šammari dialects, e.g. *ḥaḡīnⁱⁿ b-anfis-hum* ‘[they were] self confident’ (Sowayan 1992: 301), but the singular form is used in other dialects, e.g. Bahraini *nafis-na* ‘ourselves’ (Holes 2016: 99).

The reflexive functions as direct object in a clause:

ašūf ḥāl-i
I see myself

yirmi ḥāl-eh
he throws himself (into battle)

It should be noted that the gender of the nouns holds, even though they have lost their semantic connotations, and therefore feminine agreement should be expected with *nafs*:

¹⁸⁰ The use of *nesem* + pronoun as a reflexive pronoun is described by Musil (1928a: 665) in parallel with *nafs*, but it is never actually attested in the corpus. Note that *rās* + pronoun may also be used as a reflexive pronoun but this is infrequent and in uncertain contexts.

yēšūf nafs-eh ḡāyede

he boasts of his greatness [lit. he sees himself as great]

The reflexive may also function as prepositional object:

yebaddel-ha l-nafs-eh

he is selling her for for himself

When the reflexive is governed by a preposition in this way, however, syntax permits the preposition to provide the reflexive element solely by means of a bound pronoun, and therefore the use of the nominal reflexive is optional. The above example can be compared with:

yāḥod-eh l-eh

he is taking it for himself [lit. for him]

In the second example, the fact of the reflexive being introduced by the preposition *l* allows *l-eh* to carry the reflexive meaning of *l-nafs-eh*. There is no evidence of the bound prepositions expressing reflexive meaning when they are not governed by a preposition; i.e. **ašūf-(a)ni* does not replace *ašūf ḥāl-i* for 'I see myself'.

5.4 Summary

The above treatment of the closed classes demonstrates that Musil's Rwala material falls in line with what would be expected of a north-Arabian dialect. This is perhaps unsurprising since it is Musil's system of transcription which predominantly casts doubt on the reliability of his data and there is no reason why the lexical information he provides should not be reliable. Thus the main discrepancies in the data presented in this chapter fall

into the category of forms obscured by his transcriptions, such as the absence of *-ih* as the 3^{MS} bound suffix, which is always recorded as *-eh* by Musil.

It may well be that certain lexical items of the type included in this chapter are useful for typological classification of dialects. The difficulty of making a thorough lexical examination of much of this data has already been highlighted by Palva, who wrote that “the comparative material is too scattered to allow a typological comparison” (Palva 1980: 134). This situation has changed little since Palva wrote those words, with some notable exceptions such as Kurpershoek’s (2005) glossary of Najdi vocabulary, and the wealth of data (including that of Musil) remains unsystematised in compilations of texts. From the comparison with available material, it can be seen that there are a few noticeable absences of particular forms, such as the diminutive form of the adverb *hnaya* ‘here’ instead of *ha* or *hin*, but this is probably a predictable result of working with data from a corpus rather than targeted fieldwork.

The main classificatory criteria for the ‘Anazi dialects recorded by Cantineau which are relevant to the material in this chapter are all partially attested in Musil. These are: the maintenance of ^{FP} pronouns; the 1^{CS} pronouns *ħenna* rather than *eħna*; and the existence of the /k/ and /ć/ allomorphs of the bound suffixes /ak/ and /eć/. ^{FP} pronouns are attested, and although the 2^{FP} pronoun is admittedly rare, there are also few contexts in which it could occur in the material. *ħenna* and *eħna* are both attested in Musil, but the former is much more frequent, occurring in 75% of cases. The /k/ and /ć/ allomorphs of the

pronouns /ak/ and /eć/ are well attested in Musil, although their distributional context remains rather uncertain. Other characteristics relevant for classification are the absence of *šlōn* as an interrogative (as would be found in the speech of the *petits nomades*) and the 3MS suffix, where there is a distinction between the Šammari dialects, which have -o, and all other forms such as -ih, or -eh in this dialect. Many other terms Musil recorded have been shown to be prototypical of dialects of the group to which Rwala belongs, such as *mār* and *ćann*, and I find nothing recorded in this chapter to be at odds with what can be expected from the dialectological literature.

6 Conclusion

This section summarises the analysis of the previous four chapters in order to assess the ways in which Musil's Rwala material conforms to the expected norms of a central-Najdi dialect according to the literature and, specifically, to an 'Anazi dialect. §6.1 revisits the classificatory criteria for bedouin, north-Arabian and central-Najdi dialects as set out in Chapter 1 and evaluates the extent to which Musil's record of the Rwala dialect attests these features. §6.2 briefly summarises the features already mentioned which diverge significantly from other records of the Rwala dialect and related dialects. In §6.3 there is a preliminary discussion of the topic of shared vocabulary, which has not been addressed in this thesis but which nonetheless is often discussed in published dialect descriptions.

6.1 Attestation of classificatory characteristics

6.1.1 Attestation of bedouin-type characteristics

As set out in §1.3.1, the following list from Palva (2011) is used as a basic selection of distinguishing features attested in bedouin dialects:

- (i) voiced reflex of **q*
- (ii) the *ghawa*-syndrome
- (iii) gender distinction in 2nd and 3rd person plural forms
- (iv) productivity of FORM-IV
- (v) stressable definite article *al-* and relative pronoun *alli*
- (vi) shared bedouin lexical items
- (vii) productive use of diminutives

For feature (i) I have argued in §2.1.1.4 that the sound transcribed by Musil as <ḵ>, (the reflex of OA *q) must be voiced. Although Musil's records do not give direct evidence for this, indirect evidence comes from two sources: firstly, his records attest an affricated allophone ḡ, which he transcribes as <ž>, and it is well established that this affrication can only have developed from a voiced sound ([g] or possibly [G])¹⁸¹; secondly, the substitution of /ḡ/ in place names from dialects which have unvoiced reflexes of *q (e.g. *ḡāder* 'Qāder') demonstrates that the speakers were unsure how to interpret an unvoiced sound such as [q] or [ḵ] (= [kʰ]) according to their inventory of consonant sounds. The fact that all other sources of the Rwala dialect also record a voiced reflex of *q strengthens the argument that Musil's <ḵ> was an etymological rather than acoustic representation.

Concerning (ii), it has been demonstrated in §2.3.1 that Musil's record of the Rwala dialect does show evidence of the *ghawa*-syndrome, and indeed his examples account for many of the examples of the phenomenon in Cantineau's early cross-dialectal treatment of the subject (Cantineau 1935-36: 168-169). It has also been observed, however, that the *ghawa*-syndrome is only partially recorded by Musil, and in particular that he regularly transcribes its occurrence in certain categories (e.g. CaXCa or aXCaC nominal forms) but almost never in others (e.g. yaXCVC and maXCūC verbal forms). This situation is difficult to account for, and I have suggested that it may be indicative of a more prescriptive attitude towards the verbal system than the rest of the grammar on the part of Musil. This is also a

¹⁸¹ The possibility of the sound becoming devoiced at a historical point after affrication had already occurred is dismissed since this has not been recorded for any dialect that I know of.

case in which it would be helpful to have some information on how Musil kept his records while in the field, since any notes he recorded in Arabic script would not show a difference (unless he specifically vocalised it as such) between *ghawa*- and *gahwa*-forms¹⁸².

The gender distinction described in (iii) refers to both verbal and pronominal forms. It has been seen that 3FP forms are attested in all categories, while 2FP forms are mostly unattested: they are not given at all in either the S-STEM (cf. §4.2.1) or the P-STEM (cf. §4.3.2) of the verb, and only in poetry for the bound objective/possessive pronouns (cf. 5.2.1.3). It is difficult to argue *ex silentio* that 2FP forms do not exist in the dialect, since the nature of the material means that situations where such forms would be used are not frequent, if they occur at all. The dialectological literature distinguishes between dialects which have gender distinction in plural forms and those which have lost it, on which basis the fairly plentiful attestation of 3FP forms in the corpus indicates that the Rwala dialect is in the former group.

Feature (iv), the productivity of FORM-IV, has been demonstrated in §4.8.1.3 and §4.8.2.3. In fact, as has been noted previously, it is possible that FORM-IV is even more frequent in the corpus than this description shows, since the P-STEM forms are (unlike in the paradigms recorded by Prochazka) indistinguishable from many FORM-I forms, and in the S-STEM some forms also resemble FORM-I A-CLASS forms when the characteristic *a*- prefix has

¹⁸² I have kindly been given scans of some of Musil's notes by the Academic Society of Alois Musil which show that his records were kept in a mixture of Arabic and Latin script, although it is unclear whether these were field notes or written up later while he was organising his data. It seems unlikely that the majority of his material was recorded in Arabic script, however, since this would imply that he transcribed his five-way vowel system at a later stage.

been elided in hiatus, e.g. *gbalat* ‘she accepted’ (FORM-I) and *gbalat* (= [a]gbalat) ‘she stepped forward’ (FORM-IV)¹⁸³. While Musil’s corpus aligns with the bedouin-type in respect of having a productive FORM-IV, therefore, the paradigm of FORM-IV itself nonetheless raises some questions which are addressed below.

Concerning (v), the relative pronoun *alli* is frequent in the corpus (cf. §5.2.3). This contrasts with the form *illi* which is found in many Arabic vernaculars outside the bedouin-type group. His observation of the stressable definite article *al-* is discussed in §5.1.14 where I observed that Musil’s material does not align with that described by Palva and others. There is some evidence that the definite article is probably stressed in some instances, which therefore aligns it with the bedouin type in this regard, but it is less stable than that recorded by Palva, since I have shown that the vowel of the definite article is deleted following another vowel unless it is in stressable position, while evidence from Palva and others shows a stable definite article in which the vowel is never deleted. In spite of this difference, it appears that it is the stressability of the definite article rather than the stability of its /a/ vowel that is the defining bedouin characteristic, as other central Najdi sources also show elision of the vowel.

The concept of shared vocabulary given as criterion (vi) is difficult to assess since no definitive list of what this includes exists, although many authors allude to it. In Palva’s summary he gives only four examples: “*gōṭar* ‘to go’, *sōlaf* ‘to tell, narrate’, *ṭabb* ‘to arrive’,

¹⁸³ In this case, I have distinguished the form on the basis of the interpretation, but other cases are ambiguous.

and *nišad* ~ *nišád* ‘to ask’, etc.” (Palva 2011). All of these are attested in Musil. It is interesting to note that two out of four are examples of the CōCaC (~ CawCaC) verb form (which also has a /y/ variant, CēCaC ~ CeyCaC), analysed by some scholars as an innovative Najdi form¹⁸⁴ but described in this work as a regular quadriliteral derived from the semantic extension of the root (fʿl > fwʿl/fyʿl). This phenomenon, however we choose to analyse it, is clearly a feature of the dialect group in question and amply attested in Musil. The matter of shared lexical items is discussed in more detail below (cf. §6.3).

The productive use of diminutives outlined in (vii) was discussed in §3.1.5 where it was observed that diminutive forms are perhaps not quite as frequently attested as would be expected. This situation is not unique, since Cantineau also recorded fairly few diminutive forms in the dialects he studied, a fact which was considered strange by Blanc (1970: 140 note 51). The reason for this relative paucity may be partly sociolinguistic, since the use of diminutives has been observed to be especially common in women’s speech both cross-linguistically (cf. Jurafsky 1996: 545) and for other Arabian dialects of Arabic (cf. Holes [2016: 127] on Bahrain and Watson [2006: 192] on diminutive verbs in San’ani). This tendency extends also to speech addressing or about women, illustrated in Musil by *ḥbayyeb* and *šweyḥeb* ‘little darling’, and elsewhere in forms such as *umayma* ‘little mother’ (Holes, *ibid.*). Musil did not record much speech by women or addressed to them, which may explain the lower attestation of diminutives in this corpus. It may be added that the nominal

¹⁸⁴ Ingham (2011) gives FORM-XI CōCaC, FORM-XII tiCōCaC, FORM-XIII CēCaC, and FORM-XIV tiCēCaC; Johnstone (1967a: 45) refers to FORM-IIIa CooCaC and FORM-VIA tCooCaC.

template CVCCeyC(a), which is especially productive for plant names (cf. §3.1.3), is very frequent (both in this corpus and in Najdi dialects in general), and it is possible that previous observations of the productive use of the diminutive in this dialect group included this form in that assessment. Without relevant statistical data from any recorded dialect it is difficult to assess the extent to which Musil's Rwala material does or does not conform to the expected level of productive use of diminutives in bedouin-type dialects, but we have observed that there are no restrictions on its use and that it occurs in all types of speech (not just in poetry). The existence of CVCCeyC(a) as a common nominal template probably gives an observer the casual impression that diminutive forms are more common than in contiguous sedentary dialects, so there is reason to think that Musil's material does conform to the bedouin norm in this regard.

6.1.2 *Attestation of north-Arabian characteristics*

Of the following characteristics, (i)-(iv) are taken from Johnstone (1967a: 2) and (v)-(ix) are taken from Cantineau (1936-37: 222-224)¹⁸⁵.

- (i) the affrication of /k/ and /g/
- (ii) the *ghawa*-syndrome
- (iii) /a/-syncope
- (iv) /a/-raising
- (v) “differentiation” between /i, u/-syncope and /a/-syncope
- (vi) plural s-STEM suffix is diphthongal (-ow, -aw)
- (vii) P-STEM suffixes are -*in* and -*ūn*

¹⁸⁵ (iii) and (iv) were recorded as a single category in Chapter 1, following Johnstone, but here have been separated since they are two distinct phonological processes.

- (viii) lack of *b*- imperfect
- (ix) usage of adnominal linker

In §2.1.1.4 we observed that the affrication of /k/ and /g/ described in (i) is recorded as a regular phenomenon throughout Musil. The major issue in this regard is his transcriptions which suggest that the affricated allophones are [tʃ] and [ʒ] rather than the sounds which they almost certainly must have been, namely [ts] and [dʒ]. Nonetheless, in terms of distributional context the attestation of these allophones is regular and in line with what could be expected of a north-Arabian dialect.

Feature (ii), the *ghawa*-syndrome, is given by Johnstone (1967a: 2) as a characteristic of the north-Arabian group, but this was dealt with in the previous section since it has also been suggested by Palva (2011) that it is a feature of bedouin-type dialects in general.

The rule of /a/-syncope outlined in (iii) was discussed in §2.2.3.2, where it was noted that the rule does apply in Musil's data but not with the regularity that would be expected. In §2.2.3.4 we noted that the sequence Ca.Ca. was not targeted for syncope (or at least not recorded as such by Musil) at least 92 times word initially, which is fairly high. In general, it seems that Musil's material does attest the /a/-syncope rule that is characteristic of the whole north-Arabian group, but with less regularity than other descriptions have observed.

(iv) refers to the regular process of the raising of /a/ to *i* in open syllables unless it is blocked in the environment of a guttural fricative /ħ, ǧ, ḥ, ʕ, h/ or if the following consonant is /l, r, n/. In §2.3.6.1 we found that this process is hardly recorded by Musil,

except in some fairly dubious contexts: namely, forms in which /a/ became a target for either /i, u/-syncope or backing, which may suggest raising as a middle stage, e.g. /ṭawīl/ > ṭiwīl > ṭwīl ‘long’ and /ḍawābīḥ/ > ḍiwābīḥ > ḍuwābīḥ ‘killers’. There is also fairly frequent attestation of forms in which /a/ was transcribed as *e* where raising would be expected to have occurred, e.g. *ḍebaḥ* ‘he killed’ (/ḍabaḥ/ > *ḍibaḥ) and *yemīn* ‘right hand’ (/yamīn/ > *yimīn), which potentially shows that raising is more of a gradient than a categorical phenomenon, and therefore that Musil’s representation of fronting and/or raising here by *e* may be a viable surface form of raised /a/. Since this raising rule is an established characteristic of the north-Arabian dialect group, as well as being attested in all other sources of Rwala dialect, its absence from Musil is conspicuous.

It was observed in §2.2.3 that characteristic (v), the necessity for two separate syncope rules in order to account for the different treatment of high and low vowels, is found in all sources of the Rwala dialect including Musil. A clear example of this distinction is the difference between the S-STEM A-CLASS *ḍkaraw* (/ḍakaraw/) and the I-CLASS *labsaw* (/labisaw/). We observed that in some cases /i, u/-syncope appears to affect the low vowel /a/ as well, e.g. *kbīr* (/kabiir/) but concluded that this probably represents a two step process incorporating /a/-raising, i.e. /kabiir/ > kibīr > kbīr.

The criterion in (vi) is amply evidenced in Musil’s corpus (cf. §4.2.1) where *-aw* and *-ow* occur to the exclusion of the variant suffix *-am*. The monophthongal suffix *-u* mainly occurs in the imperative paradigm (where it is transcribed as long, *-ū*) and before suffixes.

In §4.2.3 it was seen that the P-STEM suffixes are *-i*, *-īn* and *-ūn*, where *-i* is at odds with the expected suffixes according to (vii). It was also remarked there that *-īn* and *-ūn* tend to come as a pair in other dialects, and therefore the attestation of *-i* is rather anomalous. It is unclear whether this represents a type of variation not recorded in the corpus, such as speakers of a different dialect.

Feature (viii) is a negative characteristic which purely functions to distinguish the north-Arabian bedouin dialects from the sedentary dialects of Palestine and Transjordan (cf. Palva 1984: 367-9). There is some evidence of its adoption in bedouin dialects, although this mostly seems to have occurred after the time of Musil¹⁸⁶, in whose records the *b*- imperfect is not attested at all.

§3.4 addresses several functions of the adnominal linker as outlined in (ix). The usage is in line with that recorded for Najdi dialects in general. It should be noted that Cantineau, who refers to it as *tanwīn*, includes it on this list of general north-Arabian features as evidence that the group is morphologically conservative, although it is now

¹⁸⁶ The situation can be seen developing through the lens of the literature. In the 1960s Cleveland (1963: 56-8) pronounced the *b*- imperfect a characteristic of rural and urban Jordanian dialects as opposed to the southern and bedouin dialects which didn't have it. In the 1970s however at least two studies noted its attestation in bedouin dialects: Blanc (1970: 139) for the Negev bedouins, where he noted that it was a surprising feature in a bedouin dialect, and Palva (1976: 33, 48) for the dialect of the 'Aḡārma, where he remarked that it represented koineisation which had brought the dialect "closer to the sedentary dialects of the Syro-Palestinian area" (*ibid.*, 48) *vis-à-vis* the camel rearing tribes of the inner desert. In the 1980s Bani Yasin and Owens (1984: 214-216) noted that it was a new feature in the Bdūl dialect and give detailed examples showing that it was not yet well established.

generally accepted that the adnominal linker, or dialectal *tanwīn* in other descriptions, is not the descendent of something resembling CLA *tanwīn* since the usage is quite different.

6.1.3 *Attestation of central Najdi characteristics*

Of the following characteristics, (i)-(viii) are taken from Cantineau (1936-37: 234) and are specifically ‘Anazi characteristics. (ix)-(xiv) are taken from Ingham (2011; 1986: 93-94). As discussed in §1.3.1, Ingham’s incorporation of Cantineau’s ‘Anazi group has not been fully explained and as such this selection of central-Najdi features is rather idiosyncratic.

- (i) /t/ maintained in feminine suffixes
- (ii) /ǧ/ = [gʲ]
- (iii) /k, g/ > [ts, ɟ]
- (iv) no *imala* of final /aa/
- (v) low attestation of word-internal superheavy syllables
- (vi) *āna* and *ḥenna* pronouns
- (vii) /k/ and /ć/ allomorphs always after CVC (*faras-k*, *šâreb-k*)
- (viii) *aḥaḍ* and *akal*
- (ix) *sam*ʿ in the S-STEM I-CLASS
- (x) *mā* ... *b*-
- (xi) *ḍa*, *ḍi* demonstrative
- (xii) /ik/ rather than /ak/ in some dialects
- (xiii) suffixed -š meaning ‘anything’ in some dialects

The feature described in (i) is only a feature which distinguishes all Najdi dialects from the north Najdi (Šammari) group, which loses the -t of feminine suffixes. In the Rwala dialect, both in Musil and elsewhere, -t is maintained.

(ii) refers to the articulation of /ǧ/ which in the ‘Anazi dialects is recorded as [gʲ] by Cantineau. As discussed in §2.1.1.2, Musil’s evidence sheds no light on the articulation of /ǧ/, although his description does suggest that a palatalised articulation is attested. I have argued that Musil was not particularly sensitive to this distinction, and also that his transcription <ǧ> can only be interpreted as a phonemic representation rather than indicative of articulatory realisation.

The case of (iii) is similarly unanswerable from Musil’s data alone. All related evidence shows that the Rwala dialect should have [ts] and [ɟʒ] but Musil recorded [tʃ] and, uniquely, [ʒ]. The articulation [tʃ] is attested in other dialects, but alongside [ɟʒ] rather than [ʒ]; this would therefore place it in a different group. Musil’s material is therefore at variance with the group it belongs to according to the relevant isogloss, but I have argued that Musil must be wrong on this matter.

The feature outlined in (iv) is one which separates the ‘Anazi dialects from the Syro-Mesopotamian group recorded by Cantineau, which has been strongly influenced by the dialects of the neighbouring sedentary population. Musil’s Rwala material does not show a fronting or raising of word-final /aa/.

Cantineau gives examples for criterion (v) which refer to two different processes: *nāgat-i* ‘my she-camel’ as opposed to *nāgt-i*, which refers to the failure to syncopate of the feminine marker /at/, and *rās-ana* ‘our head’ as opposed to *rās-na*, which refers to epenthesis. In §2.2.3.3 it was observed that the syncope of the feminine marker /at/

described in (v) does occur in Musil, but only in a minority of cases. In most cases this type of syncope does not occur and his record of the dialect concurs with this characteristic. In §2.2.2.1 it was shown that Musil did record cases where epenthesis did not break up a word-internal VVCC or (less frequently) VCCC sequence, resulting in the attestation of word-internal superheavy syllables, as is the norm in the “innovative” dialects. These cases, however, occur alongside a majority of cases where epenthesis did occur in the way that Cantineau considered to be characteristic of ‘Anazi dialects. Although it can be said that Musil’s material generally concurs with Cantineau’s characteristic, it also contains numerous exceptions. There is therefore a discrepancy between what Cantineau considered to be a clearly observable isogloss and the much more varied picture from Musil’s data. This could be on account of the method of data collection (long term observation in Musil and elicitation in Cantineau) and sample size, as well as the known sources of variation in Musil including, perhaps especially in this case, the record of speakers from other areas. (v) presumably extends to cases where /i, u/-syncope produces forms such as *ṣāḥb-i* ‘my friend’ as well (cf. §2.2.3.1), and on this matter Musil’s data are also similar to those of Cantineau, with these forms infrequently attested.

The pronouns of (vi) are attested in Musil. *ḥenna* is given as a distinctive characteristic because it contrasts with *eḥna*, which is also attested in Musil, but as a minority variant.

Concerning the /k/ and /ć/ allomorphs of the bound objective/possessive pronouns /ak/ and /eć/ described in (vii), the situation is mixed. As described in §5.2.1.3, the pronouns /k/ and /ć/ are attested as allomorphs, although they freely vary with their counterparts /ak/ and /eć/ in the contexts where the former could occur.

The forms *akal* and *aḥaḍ* outlined in (viii) are the common variants in Musil, although forms such as *kala* and *ḥaḍa*, which Cantineau considered to be markers of the northern Najdi (Šammari) and Syro-Mesopotamian dialects, are also attested in a minority of cases.

The remaining features are from Ingham, and are more problematic because Ingham's central Najdi group, in contrast to Cantineau's more precise 'Anazi group, is somewhat generalised and the features may apply only to the dialects of central Najd rather than the bedouin emigré of the Syrian Desert. Despite positing the existence of the central Najdi group, and claiming that the 'Anazi dialects belong to it, Ingham has not (to the best of my knowledge) clearly outlined the linguistic criteria which differentiate this group.

There is reason to reject Ingham's claim represented in (ix) that central Najdi dialects exclusively have the form *sam^c* in the S-STEM I-CLASS. The discussion in §4.2.4 clearly demonstrates that the I-CLASS shows considerable variation throughout the region and, as mentioned in that section, this is a topic for further study. It was recorded in §4.2.2.2 that the form of the I-CLASS in Musil is usually *šireb*, although the first vowel is not always raised, giving forms such as *labes*.

Concerning (x), it was shown in §5.1.9 that the negative expression *mā* (+ pronoun) *b-* is attested in Musil, although the form without *b-* also occurs.

With regard to (xi), it is not entirely clear why Ingham claims that *ḍa* and *ḍi* are specifically central Najdi terms, and he appears to claim that dialects of another group (possibly northern Najdi although this is unconfirmed) only have the composite forms *hāḍa* and *hāḍi*, which he describes as widespread (Ingham 1986: 93). As established in §5.2.2, the *hā* element of the demonstrative may be dropped in Musil.

The case of (xii) demonstrates that Ingham's conception of the central Najdi group shows variation in this regard and he acknowledges that, while /ik/ is a feature which may only occur in central Najdi dialects, many of these dialects, including the 'Anazi dialects of the Syrian desert, have /ak/ (Ingham 1979: 54). The pronoun /ik/ is therefore a feature of Najdi dialects located further south than the territory of the Rwala and the form in Musil is always /ak/ or /k/.

Feature (xiii) is not attested in Musil, although Ingham does say that it only occurs in some dialects.

It can be seen that the Rwala dialect according to Musil fits the categories in this section less completely than the bedouin-type and north-Arabian type categories in the previous two sections. While it is possible that this is because the categories themselves are more specific—and some, such as (ii) and (iii), find little support either way in Musil—it also seems that it would be incorrect to assume that the Rwala dialect does not fit into this

type. I would suggest that the absorption of Cantineau's 'Anazi type into the central Najdi type requires further clarification: one has the impression that in Ingham's classification the central Najdi group has become a default category which includes all dialects which do not attest the idiosyncrasies associated with either the northern Najdi (Šammari) type or the southern Najdi type. This has produced two suboptimal results for the central Najdi group: firstly, the group has few positive characteristics by which to classify its members (excluding Cantineau's 'Anazi criteria given above, and it is uncertain how fully these are representative of the central Najdi group in general); secondly, the group attests a significant level of variation, as the criteria from Ingham demonstrate. I therefore suggest that a positive system of classification for the central Najdi group is a desirable avenue for further study.

6.2 Attestation of non-Najdi characteristics

Excluding the criteria discussed in the previous section, there are remarkably few other areas in which Musil's Rwala material attests characteristics which do not conform to the expected patterns of the Najdi group. The characteristics in question are mostly from the verbal system and were discussed in Chapter 4, but are summarised again here:

- (i) “vowel harmony” of prothetic vowel
- (ii) *a-e* melody in the P-STEM of forms V-VI
- (iii) P-STEM of doubled verbs
- (iv) lack of I-CLASS for final-weak verbs
- (v) *yi-* prefix in FORM-I
- (vi) *u-i* passive melody
- (vii) lack of analytic genitive

In §2.2.2.2 it was observed that Musil’s Rwala material shows a type of vowel harmony in the prothetic vowel, as expressed in (i). This is a two way system in which the prothetic vowel is a high vowel which is round (i.e. /u/) if the following vowel is round (i.e. also /u/) and not round (i.e. /i/) if the following vowel is not round (i.e. /a/ or /i/); this situation is exactly the same as that of the dialect of the Bani Şaḡar recorded by Palva (1980: 128). This is a bit of an outlier in the central Najdi group, where Prochazka (1988: 38-39) records these dialects as having only /i/ as the prothetic vowel, and also differs from the Hijazi dialects where a full system of vowel harmony is exhibited (i.e. /a/ may also occur as a prothetic vowel before a following /a/, cf. Prochazka *ibid.*).

(ii) refers to the situation described in §4.8.1.4-5 where Musil has recorded *y(e)takasser*, which seems to represent an underlying form /yetakassir/ on account of the contrast in other forms, e.g. in FORM-II between the imperative *kasser* (< /kassir/) with the S-STEM 3MS *kassar* (< /kassar/). Other descriptions show that in central Najdi the form is *y(e)takassar*. This situation is hard to account for, and it is possible that Musil generalised the *a-e* P-STEM melody from the other derived forms to forms V-VI; since Musil’s forms show a similarity to CLA, it is also possible that they were generalised on this basis. The situation is

complicated, however, because it is the generally recorded melody *a-a* which has produced non-contrastive imperatives and S-STEM 3MS forms, while Musil's system maintains a contrast, and the implications of this change have not been discussed in any of the sources I have consulted.

Similarly to (ii), (iii) describes a case of Musil's verbal forms differing from what is generally attested in the central Najdi group. It was shown in §4.10.4 that Ingham (2011) and Prochazka (1988: 57), among others, record the P-STEM forms of FORM-I doubled verbs as having initial stress and concomitant degemination of the final consonant, e.g. *yírid*, while Musil's forms show no degemination, e.g. *yerudd*, which may also suggest word-final stress. As before, it is difficult to discern whether these forms are genuine or whether they have been contrived to fit Musil's expectations of the verbal system, especially because his forms more closely resemble those of CLA.

(iv) refers to the apparent lack of an I-CLASS in the S-STEM of the FORM-I final-weak verbs, i.e. **mišī*, alongside the A-CLASS *laga*. As discussed in §4.10.3 it is difficult to prove the existence of an I-CLASS, although this does not mean that it did not exist, since there is not much data to prove either way and some forms do tentatively suggest that one existed. It would be unprecedented for there to be no I-CLASS in this category, although it was noted that in some related dialects there is evidence that the I-CLASS has somewhat merged into the A-CLASS, and its lack of attestation may be a reflection that the merger of the two categories was already underway at the point at which Musil recorded the dialect.

Feature (v) is one of the most puzzling and I have found no parallels in related dialects to shed light on the situation. We saw in §4.3.1 that the P-STEM prefix /ya/ is recorded as *ya-*, *ye-* and *yi-*, where *ya-* and *ye-* probably show the usual *a~e* variation found in Musil but the *yi-* variant is altogether different. Given that *yi-* occurs but not *ti-* or *ni-* it was considered that a form of assimilation of /y/ features could explain this. The case is problematic because it has led to a breakdown of contrast in the verbal system, since other records of the Rwala dialect show that the FORM-I P-STEM form *yaCCiC* and the FORM-IV P-STEM form *yiCCiC* are differentiated primarily on the basis of the prefix vowel; this situation is even more complex in Musil since the FORM-I *yaCCaC* and *yaCCiC* forms are both (often) transcribed with *e* in the verbal stem, meaning that a majority of FORM-I P-STEM forms are indistinguishable from those of FORM-IV. Since this situation is not attested in other dialects, it is very difficult to know whether to accept Musil's record here.

Concerning (vi), it was noted in §4.7.1 (and sporadically in §4.8 for some of the derived forms) that Musil's records attest an internal passive with the vocalic melody *u-i* alongside other passive melodies including the expected Najdi melody *i-i*. It is possible that this is another case of Musil conforming his data to match CLA, but we have also seen that *u-i* is attested as a passive melody in other related dialects and, furthermore, there is really too little data for the internal passive to make a conclusive point here.

It was mentioned in §3.3 that feature (vii), the analytic genitive, is absent from Musil but recorded for other Najdi dialects (e.g. Ingham 1994a: 54 and 57-9). This is something of

an argument *ex silentio* in that the absence of the feature from Musil's records cannot prove that it did not exist. Nonetheless, the lack of a single attestation of the feature in a corpus of this size is fairly convincing evidence that there was not one. The possibility that the analytic genitive could be a more recent innovation in the dialects concerned is a matter for further study, however.

6.3 Evidence of shared lexicon

This thesis has addressed the categories of phonology, morphology and function words in order to both describe and assess the Rwala dialect as it is presented in Musil. Many of the sources for north-Arabian dialects already cited also use the category of shared vocabulary as a means of dialect classification, e.g. the discussion of shared "bedouin" vocabulary in §6.1.1. Since the lexicon is the least structured part of the grammar, I have chosen not to use it as a source for the classification of the dialect. There are two reasons why the lexicon is less useful for classification, outlined here.

Firstly, according to Hetzron, it is only the category of shared morpholexical items which constitutes the basis for establishing genetic affiliation in historical linguistics, rather than shared lexical items in general (Hetzron 1976: 98). Individual lexemes may be borrowed, replaced, or simply disappear, in a way which may have nothing to do with the relationship of the dialect in question to other members of its group. Nonetheless, an assessment of shared vocabulary may be appropriate given the linguistic situation described

in §1.3.1, where it was observed that while the dialects of “large bedouin tribes” are analysable in a genetic framework (cf. Ingham 1982a: 30) they are nonetheless part of a *sprachbund* in which areal effects are easily observable. Having established that the Rwala dialect is certainly a member of the central Najdi group enables us to make more confident assertions about the attestation of any particular lexical item.

Secondly, it is an obvious restriction of this approach that, since the data available comes from a closed corpus, the non-attestation of any lexeme is an argument *ex silentio* and does not constitute evidence that the lexeme does not exist. The categories of phonology and morphology affect the whole lexicon, and the function words presented in Chapter 5 are very unlikely not to occur in any excerpt of linguistic production, but this is not true of content words. A full analysis of the shared lexical inventories of Najdi dialects is a matter for many more studies, but here we will briefly consider the topic from some example lists in published sources. This section addresses examples of shared vocabulary given by other researchers for bedouin dialects in general, for Najdi dialects, and for the *nabaṭī* poetic tradition.

6.3.1 *Shared bedouin lexicon*

It was noted above (cf. §6.1.1) that all four of Palva’s (2011) examples of shared bedouin vocabulary are attested in Musil. A larger example list of typically bedouin

vocabulary is provided by Rosenhouse (2011), where it can be seen that most nouns are attested in Musil but for other parts of speech the situation is mixed:

Nouns:

Found in Musil

yed(d) ‘hand, handle’
manaḥīr, ḥašm ‘nose’
barāṭīm, ‘mouth¹⁸⁹, lips’
maṭar ‘rain’
zimūl, ḍūn ‘riding animals’
álbil ‘camels’
dabbe ‘hill’
bēt ša‘r ‘tent’
ḥōša ‘knife’
maḡrāfe, miḡrafa ‘spoon’
‘aḡawīd ‘good, fine men’
gōm ‘group, tribe, enemy’
nišāma ‘tribe’s warriors’
ganaṣ ‘hunting’
m‘azzeb ‘host’
‘aḡiyy ‘child’
‘iyāl ‘children’
zāda ‘food for the journey’

Not found in Musil

ḡnima ‘ewe’¹⁸⁷
šifāif ‘mouth¹⁸⁸, lips’
šīta ‘rain’
rwēsa ‘hill’¹⁹⁰
ḍūf ‘children’
zuwwāda ‘food for the journey’

Adjectives

zēn ‘good’
šēn ‘bad’

zḡayyir ‘small’¹⁹¹
glayyil ‘a little’

¹⁸⁷ There is no word for ‘ewe’ attested in Musil and it is likely that *ḡnima* could be derived from *ḡanam* ‘sheep [coll.]’, although the same form could be the singular (not feminine) form from the same collective.

¹⁸⁸ Rosenhouse translates this as ‘mouse’ which I assume is a mistake.

¹⁸⁹ See previous note.

¹⁹⁰ *rās* is attested meaning ‘summit’ and it seems logical that the diminutive (feminine) form *rwēsa* could also mean ‘hill’.

¹⁹¹ This adjective and the following one, *glayyil* ‘a little’, are attested in the adjectival CaCiC template but not in diminutive forms. Voicing of /š/ > z is also attested.

grayyib ‘near’

rḥamān ‘merciful, kind’¹⁹²

Verbs

gōṭar ‘to go’

ṭabb ‘to reach, enter’

nišad (‘*ala*) ‘to ask, look for’

šarrag ‘to go eastward’

widd- (+ pron.), *rād* ‘to want, like’

dahḥag, dahrağ ‘to look’

kital ‘to kill’¹⁹³

hawwad, awṭa ‘to go down’

gannaş ‘to hunt’¹⁹⁴

gawwaş ‘to shoot at, snipe’

‘*agad* ‘*gādo* ‘*alēha* ‘to wed, marry’¹⁹⁵

Adverbs

(*h*)*alḥin* ‘now’

bāčir ‘tomorrow’

hnāk, gād ‘there’

hina, hniy ‘here’

balḥayl ‘much, strongly, very’

dīma, hēdīkt alḥin ‘then’

kiḍiy ‘thus’

ḥāf-aḷḷah ‘maybe’

hassā ‘now’

ğodwa ‘tomorrow’

hnūh ‘there’

Particles

wēš ‘what?’

*min, men, min-hu, min-h*¹⁹⁶ ‘who?’

mata ~ mita ‘when?’

‘*ayy* ‘which?’

‘*amma ~ ’umma, mār* ‘but’

‘*ād* ‘so, then’

mār ~ umār ~ mēr ‘so, but’

mint ‘(you are) not’

‘*alāma, alāš ~ alēš, alawēš* ‘why?’

‘*andinma, lamman ~ lammin*

‘when (temporal adverb)’

wlinnih ‘and there (it was)’

wlinn, wilin, win, winno ‘and lo (a sudden turn in the narrative)’

‘*aminno* ‘because’

šnū, šnī ‘what?’

¹⁹² This adjective is attested but not as a *ghawa*-form.

¹⁹³ This verb is attested but not with the devoicing of /g/ > k.

¹⁹⁴ This verb is attested but as FORM-I, not FORM-II.

¹⁹⁵ The verb ‘*agad* is used meaning ‘to marry’ in Musil but not in this particular idiom.

¹⁹⁶ This form is given by Rosenhouse as *min-hu-minh*; I have taken this as a misprint and assumed that two separate forms were implied here, namely *min-hu* and *min-h*, rather than a compound particle.

<i>ar</i> ^c ~ <i>har</i> ^c , <i>ar</i> ^c <i>ih</i> ‘here (he is)’	<i>wagtēš</i> ‘when?’
<i>in</i> , <i>čān</i> ~ <i>kān</i> , <i>inkān</i> ¹⁹⁷ ‘if’	<i>yāt(u)</i> , <i>wein</i> ‘which?’
	<i>iza</i> ‘if’

6.3.2 Shared Najdi lexicon

Other lists are specific to the Najdi group rather than to bedouin dialects in general. In Chapter 5 we saw examples of individual lexical items (function words) which are often held to be particularly characteristic of the north- and northwest-Arabian dialect groups, such as *mār* ‘but’ and *yowm* ‘when’. Ingham (2011) provides twenty-four examples of characteristically Najdi items (content words), of which exactly half are found in Musil:

Found in Musil	Not found in Musil
<i>baġa</i> , <i>yabi</i> ‘to want’	<i>amda</i> , <i>yimdi</i> ‘to be time for’
<i>bāg</i> , <i>ybūg</i> ‘to steal, betray’	<i>bēg</i> ‘muffler for holding a coffeepot’
<i>ḍall</i> , <i>yḍill</i> ‘to act dishonourably’	<i>bḍa</i> , <i>fīḍa</i> ‘here’ (northern <i>bhāḍa</i>)
<i>ḥaya</i> ‘rain, grass’	<i>ḍarrab</i> , <i>yḍarrib</i> ‘to do a thing properly’
<i>‘ilm</i> pl. <i>‘lūm</i> ‘news’	<i>tigahwa</i> , <i>ytigahwa</i> ‘to take coffee’
<i>mizin</i> ‘rain cloud’	<i>haga</i> , <i>yhaga</i> ‘to think, hold an opinion’
<i>nāḍar</i> , <i>ynāḍir</i> ‘to look at’ ¹⁹⁸	<i>igt</i> ‘weys’
<i>sām</i> , <i>ysūm</i> ‘to make an offer’	<i>ḥaḍaf</i> , <i>yḥaḍif</i> ‘to throw’
<i>sawwa</i> , <i>ysawwi</i> ‘to do, make’	<i>nida</i> ‘grass’
<i>tār</i> , <i>yṭūr</i> ‘to stand up’	<i>sanna</i> ^c , <i>ysanni</i> ^c ‘to do a thing properly’
<i>ṭawwar</i> , <i>yṭawwir</i> ‘to shoot a gun’	<i>šanag</i> ‘side’
<i>zēn</i> ‘good’	<i>zaham</i> , <i>yizham</i> ‘to call’

¹⁹⁷ The form in Musil is *en čān*.

¹⁹⁸ In the article in question Ingham uses <ḍ> as the symbol for the pharyngealised voiced interdental fricative so these are transcribed by him as *nāḍar/ynāḍir*.

It is perhaps more useful to consider function words than content words as a point of comparison. Ingham (*ibid.*) also provides the following list of Najdi modals¹⁹⁹:

<i>baġa, yabi</i>	future intent/imminent future	‘will’
<i>lēt-, ya lēt</i>	desiderative/precativ	‘would that, please’
<i>‘asa-</i>	optative	‘I hope’
<i>ġi‘il-</i>	optative	‘may’
<i>ħal-, ħall-</i>	jussive	‘let’
<i>da‘-</i>	jussive	‘let’
<i>illa</i>	necessitive	‘must’
<i>kūd-</i>	optative resultant	‘mayhap, in case’
<i>yallah</i>	qualificatory	‘just, barely’
<i>ćān</i>	obligative unfulfilled	‘should have’
<i>ćān b-</i>	unfulfilled past intent	‘was going to’
<i>ćinn-</i>	speculative, polite inquiry	‘it seems’
<i>tigil</i>	speculative	‘it seems, perhaps’
<i>l-</i>	jurative	‘I swear that’

With the sole exception of *ćān b-*, all of these forms are attested in Musil with only the following slight departures in usage: *da‘* only occurs once and its context does not mark it out as definitively jussive, and the use of *ćān* meaning ‘should have’ is also only attested in doubtful contexts.

6.3.3 *Shared poetic lexicon*

According to Holes, “the bedouin natural, social and emotional world has its own specialised poetic lexicon” (2011a) and the fact that the *nabaṭī* poetry comprises a core of shared vocabulary is well established. What makes this particular set of shared lexical items

¹⁹⁹ Several forms marked as distinctively northern (i.e. Šammari) have been removed from this list. They are: *kūd* necessitive ‘must’; *aġadi-* optative resultant ‘mayhap, in case, perchance’; *widd-ak* obligative remote ‘ought to’.

so fascinating is that it contains many lexemes found in the Classical poetry but which are no longer attested in either MSA or other (non-bedouin) dialects. It has been observed that “the lexical relationship of the modern bedouin dialects to that of early Classical poetry is a vast, yet almost completely untouched area of research” (Holes, *ibid.*), and the inclusion of some comparative vocabulary here can do no more than demonstrate that from any given list of characteristic vocabulary the material in Musil generally attests the same items.

Holes (2011a) provides an example list of characteristic vocabulary of the *nabaṭī* poetry. This vocabulary is compiled for the three themes of animals, weather, and warfare and weaponry, all of which are major *topoi* of the oral poetic tradition and hence sources of specific vocabulary. Since many of the terms in this list refer to “different types of (e.g. camel)” the precise definition of each term has been omitted:

Found in Musil

Animals

rīl, *rabda* pl. *ribd*, *hēg* ‘ostrich’
(of different ages, types or colours)
maḥlūl ‘camel calf that carries pegs in its nostrils’
ḍalūl, *rabāʿ*, *zamil*, *ʿēra*, *fāṭir*, *ṭarš*, *hiġin* ‘camel’ (of different age, type and size)
ḥurr, *ḍāmīr*, *hiḥāf*, *mahzūl*, and *halīma* ‘camel’ (of different pedigree or quality)
bāz, *šihān*, *nadāwi*, *ṣagar* ‘falcon’ (of different types)

Not found in Musil

ḥūr ‘she-camels producing plenty of milk’
ḥaškal a type of ostrich
mišġir a type of camel
šāfi, *ḥirsās*, *ḥalḥīl* ‘camel’ (of different pedigree or quality)

Weather

marāhīš, muzna, himlūl, dēm, sēl, wasm
different types of cloud or rain
rašš describes a way the rain falls
mafla, fēḏa, manādi, marbaʿ, and mirbāʿ
‘pasture’ (of different types)
ḥamāda, mafāza, maḥāwil ‘desert’ (of
different types)
tall, ḥēd, ḥazim, riḡim, ‘hill, high place’
margab ‘a high place used as a lookout’

tʿūl lit: ‘pouring down in sheets’
ginif pl. *gunūf* ‘cumulus’
maḥāyil ‘cloud showing signs of
rain’
ḡēn, ṭašš ‘types of clouds or rains’
hataf describes a way the rain
falls
tūr, gardūd ‘hill, high place’

Warfare and weaponry

al-kōn ‘battle’
al-waḡā, al-ḥarāyib ‘battle’
naḥa, fazaʿ verbs for seeking or
offering help in war
naxwa ‘war cry’
bārūd ‘rifle, gun’
tār ‘to shoot’

kāwan, iktawan verbs for fighting
māda ‘battle-field’
fāziʿ, fazzāʿ ‘helper in time of
war’
mismār, mwazzan ‘bullet’
nawāšīn ‘rifle sights’
mrabbaʿ ‘cross-hairs’
ḥašab ‘rifle stock’
mišṭ ‘cartridge clip’
baṭīn ‘magazine’
(yōm) al-warīd, al-mlāga, al-laḡḡa
‘battle’
mʿaddila, tumēdi ‘rifle, gun’
ṭaḥḥ, ‘*addal* ‘to shoot’

In this case it can be observed that the majority of animal and weather related vocabulary is attested in Musil, while the warfare and weaponry vocabulary is mixed. Holes also provides the following list of function words characteristic of the poetic tradition, all of which are attested in Musil in at least one of their variant forms:

kūd (in) ~ kūdin ‘perchance’, ‘except for’, ‘by force’
la badd ma, sbidd (< lēs bidd) ma, lās badd ma
‘it is inevitable that...’
la- ‘I swear that...’
wa la hagwati... ‘I don’t think...’ and similar phrases

<i>ǧa'al ~ ǧi'l ~ yi'al (allah)</i>	'may God ...' (+ imperfect)
<i>(ya) kam... + noun</i>	'many a...', 'how...'
<i>ma... and yāma...</i>	'many a...', 'how...' (+ adj. or verb)
<i>ya šēn, ya šūm and ya zēn, ya zīn</i>	'how bad...!' and 'how good...!'
<i>tugul, tugūl, tigūl²⁰⁰ etc.</i>	'you'd say'
<i>činn- ~ kinn-, činn- and kama</i>	'as if'
<i>zē, zayy, zi</i>	'like'

The only item from this list not attested in Musil is *wila, ila, la-, lē* 'lo and behold...'.²⁰⁰

The list in Holes also includes the ethic dative *-k*, which is attested in Musil in forms such as *ǧā-k* 'he came [to you]'; these forms are held to be used rhetorically to "involve" the listener and are more a stylistic feature of the poetic register than the other items listed here which, in Musil at least, occur frequently in conversation.

Also writing about the tradition of oral poetry, Sawayan (1992) provides some examples of words from his corpus which he considers to be so specific to bedouin life as to render them nearly untranslatable. These are listed here as further examples of core vocabulary²⁰¹:

²⁰⁰ The form in Musil is *tegel*.

²⁰¹ The author includes these terms in order to describe the resonances they carry when used in poetry; these cannot be included here but can be consulted in his work. As one example, *halūǧ* ostensibly refers to a wailing camel (she wails because her calf has been taken away from her; this is the usual custom when a camel is kept for milk, but in this case the camel is aware of the existence of the calf and becomes depressed), but according to Sawayan constitutes "a symbol of the ultimate in grief" (Sawayan 1992: 75).

Found in Musil

<i>midda</i>	‘blood payment’
<i>ḥāwih</i> ²⁰²	‘a bedouin institution whereby a weak tribe seeks the protection of a stronger tribe through the payment of a regular tax in the form of camels, sheep, butter, etc.’
<i>nḥawih</i>	‘appeal; tribal solidarity’
<i>man</i> ^c	‘a pledge by the pursuer to spare the life of the pursued and protect him from others, the pursued in turn surrenders his mount and weapons to the pursuer’
<i>waḡh</i>	‘patronage, protection’
<i>ḥalūḡ</i>	‘wailing camel’
<i>ḥāyil</i>	‘barren’
<i>ḥāyif</i>	‘a camel rustler on foot who sneaks around the edge of the camp when it is dark and people are asleep to steal one or two camels and scurry away’
<i>‘aḡid</i>	‘raid leader’
<i>imānih</i>	‘responsibility, trust’
<i>rif</i>	‘verdure, opulence, life of plenty’

Not found in Musil

<i>ḍimmih</i>	‘covenant, trust, safe-keeping, guardianship, conscience, scrupulousness’
<i>ḥarḡ</i>	‘unlawful, forbidden, in critical position’
<i>ziriḡ</i>	‘camels that are being led quickly and stealthily’
<i>māḥūdih</i>	‘plundered, spoiled, usually used as a deprecation meaning useless, wretched, worthless, miserable’ ²⁰³

In all three categories described in this section function words are almost all attested, while content words give much more varied results. As predicted in §6.3, this is likely to be on account of the corpus rather than evidence that the unattested content words did not

²⁰² In Sawayan’s Šammari material, the suffix *-ih* is the feminine ending rendered *-a* in Musil.

²⁰³ While the passive participle ‘captured’ is amply attested in Musil, it is included here with other unattested forms because the general use of the word as a deprecation is not attested. The exception to this may be the use of *māḥūda* ‘a woman who has lost her virginity’, since it is unclear whether this is a specific meaning within the semantic range of the word or whether this shows the type of deprecation described by Sawayan.

exist. On the basis of the shared function words, we may consider that Musil's records show very close alignment with the Najdi group, and therefore that shared lexicon (though only briefly touched on here) confirms the classification from the previous sections.

Bibliography

Abbreviations

<i>BASOR</i>	<i>Bulletin of the American Schools of Oriental Research</i>
<i>BEO</i>	<i>Bulletin d'études orientales</i>
<i>BSL</i>	<i>Bulletin de la Société de Linguistique de Paris</i>
<i>BSOAS</i>	<i>Bulletin of the School of Oriental and African Studies</i>
<i>EALL</i>	<i>Encyclopedia of Arabic Language and Linguistics</i> . Print edition: Versteegh, K., M. Eid, A. Elgibali, M. Woidich and A. Zaborski (eds) (2005-9). Vols. I-V. Online edition: managing editors L. Edzard and R. de Jong (first published 2011). Leiden : Brill.
<i>JIPA</i>	<i>Journal of the International Phonetic Association</i>
<i>JSS</i>	<i>Journal of Semitic Studies</i>
<i>ZDPV</i>	<i>Zeitschrift des Deutschen Palästina-Vereins</i>
<i>ZDMG</i>	<i>Zeitschrift der Deutschen Morgenländischen Gesellschaft</i>
<i>ZAL</i>	<i>Zeitschrift für Arabische Linguistik</i>
<i>ZSVG</i>	<i>Zeitschrift für Semitistik und verwandte Gebiete</i>

List of works

AAAS (American Association for the Advancement of Science) (1928). 'Presentation of the Charles P. Daly gold medal to Professor Alois Musil', in: *Science* 67, 289-290.

Abboud, P. F. (1979). 'The verb in northern Najdi Arabic', in: *BSOAS* 42 (2), 467-499.

——— (1978). 'The vowel of the imperfect prefix in Najdi Arabic', in: Jazayeri, Polome and Winter (1978), 129-138.

——— (1964). *The syntax of Najdi Arabic*. Diss. Texas, Austin.

Abu-Haidar, F. (1991). *Christian Arabic of Baghdad*. Wiesbaden : Harrassowitz.

AGS (American Geographical Society) (1946). 'Obituary : Alois Musil', in: *The Geographical Review* 36 (4), 686-687.

- Aguadé, J. and M. Woidich (eds) (2004). *Estudios de dialectología norteafricana y andalusi : Homenaje a Peter Behnstedt en su 60. aniversario (= EDNA 8)*. Zaragoza : Instituto de estudios islámicos y del oriente próximo.
- Aguadé, J., M. Marugán and F. Corriente (eds) (1994). *Actas del congreso internacional sobre interferencias lingüísticas arabo-romances y paralelos extra-íberos : Celebradas en Madrid del 10 al 14 de diciembre de 1990*. Zaragoza.
- Alhawary, M. T. (ed.) (2005). *Perspectives on Arabic linguistics XVII-XVIII : Papers from the seventeenth and eighteenth annual symposia on Arabic linguistics*. Amsterdam : J. Benjamins.
- Al-Hazmi, A. (1975). *A critical and comparative study of the spoken Arabic of the Ḥarb tribe in Saudi Arabia*. Unpublished Ph.D. thesis, Leeds.
- Al-Mozainy, H. Q. (1981). *Vowel alternations in a bedouin Hijazi dialect : Abstractness and stress*. Ph.D. Austin, Texas.
- , R. Bley-Vroman and J. J. McCarthy (1985). 'Stress shift and metrical structure', in: *Linguistic Inquiry*, 16 (1), 135-144.
- Al-Rasheed, M. (2010). *A history of Saudi Arabia*. 2nd ed. Cambridge : Cambridge University Press.
- Al-Sweel, A. I. (1990). 'Some aspects of Najdi Arabic phonology', in: *ZAL* 21, 71-82.
- (1987). 'Verbal and nominal forms of Najdi Arabic', in: *Anthropological Linguistics* 29 (1), 71-90.
- Al-Wer, E. (2011). 'Jordanian Arabic (Amman)', in: *EALL* Online edition. Consulted online on 18 March 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol2_0065.
- Ángeles, V. (2011). 'Relative pronoun (Arabic dialects)', in: *EALL* Online edition. Consulted online on 21 March 2018. http://dx.doi.org/10.1163/1570-6699_eall_EALL_SIM_0117.
- Ashby, P. (2011). *Understanding phonetics*. London : Hodder Education.
- Badawi, E. M., M. G. Carter, A. Gully and M. Awad (2016). *Modern written Arabic : A comprehensive grammar*. Second edition. London : Routledge.

- Bailey, C. (1991). *Bedouin poetry from Sinai and the Negev : Mirror of a culture*. Oxford : Clarendon Press.
- Bani Yasin, R. and J. Owens. (1984). 'The Bduul dialect of Jordan', in: *Anthropological Linguistics* 26, 202–232.
- Behnstedt, P. (1985). *Die nordjemenitischen Dialekte, i. Atlas*. Wiesbaden : Reichert.
- and M. Woidich (2005) *Arabische Dialektgeographie : Eine Einführung*. Leiden : Brill.
- Bellem, A. and J. C. E. Watson (2014) 'Backing and glottalization in three SWAP language varieties', in: Giolfo (2014), 169-207 [accessed online at <http://dro.dur.ac.uk/14890/1/14890.pdf>, page numbers from this version].
- Benmamoun, E. and R. Bassiouney (eds) (2018). *The Routledge handbook of Arabic linguistics*. Routledge.
- Bergsträsser, G. (1915). 'Sprachatlas von Syrien und Palästina', in: *ZDPV* 38, 169-222.
- Bettega, S. (2014). 'The use of dialectal *tanwīn* in Qatari Arabic' in: *Ricognizioni : rivista di lingue, letteratura e culture moderne* 2 (1), 15-25.
- Betts, A. (1989). 'The Solubba : nonpastoral nomads in Arabia', in: *BASOR* 274, 61-69.
- Bezold, C. (ed.) (1906). *Orientalische Studien Theodor Nöldeke zum Siebzigsten Geburtstag, 2. März 1906, gewidmet von Freunden und Schülern*. Vol. 1. Gieszen : Töpelmann.
- Blanc, H. (1970). 'The Arabic dialect of the Negev bedouins', in: *Proceedings of the Israel Academy of Sciences and Humanities*, vol. IV no. 7, 112-150.
- (1964). *Communal dialects in Baghdad*. Cambridge, Mass. : Harvard University Press.
- Blau, J. (1999). *The emergence and linguistic background of Judaeo-Arabic : A study of the origins of Neo-Arabic and Middle Arabic*. 3rd rev. ed. Jerusalem : Ben-Zvi Institute for the Study of Jewish Communities in the East.
- Brockelmann, C. (1908–1913). *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. 2 vols. Berlin : Reuther und Reichard.

- Broselow, E. (2018). 'Syllable structure in Arabic dialects', in: Benmamoun and Bassiouney (2018), 32-47.
- (1992a). 'Parametric variation in Arabic dialect phonology', in: Broselow, Eid and McCarthy (1992), 7-45.
- (1992b). 'Transfer and universals in second language epenthesis', in: Gass and Selinker (1992), 71-86.
- (1982). 'On predicting the interaction of stress and epenthesis', in: *Glossa*, 16 (2), 115-132.
- , M. Eid and J. J. McCarthy (eds) (1992). *Perspectives on Arabic linguistics IV : Papers from the fourth annual symposium on Arabic linguistics*. Amsterdam, Philadelphia : John Benjamins.
- , S. Chen and M. Huffman (1997). 'Syllable weight : convergence of phonology and phonetics', in: *Phonology*, 14 (1), 47-82. Cambridge University Press.
- , S. Chen, M. Huffman and R. Hsieh (1995). 'The timing structure of CVVC syllables', in: Eid (1995), 119-138.
- Cadora, F. J. (1992). *Bedouin, village and urban Arabic : An ecolinguistic study*. Leiden : E. J. Brill.
- Cantineau, J. (1960). *Etudes de linguistique arabe : Mémorial Jean Cantineau (Cours de phonétique arabe)*. Paris : C. Klincksieck.
- (1939). 'Remarques sur les parlers de sédentaires syro-libano-palestiniens', in: *BSL* 40, 80-88.
- (1936-1937). *Études sur quelques parlers de nomades arabes d'Orient*. Paris : Larose.
- Chomsky, N. and M. Halle (1968). *The sound pattern of English*. New York : Harper and Row.
- Clements, G. N. (1990). 'The role of the sonority cycle in core syllabification', in: Kingston and Beckman (1990), 283-333.
- Cleveland, R. (1963). 'A classification for the Arabic dialects of Jordan', in: *BASOR* 171, 56-63.

- Cohen, D. (1962). 'Koinè, langues communes et dialectes arabes', in: *Arabica* 9, 119–144.
- Comrie, B. (1989). *Language universals and linguistic typology : Syntax and morphology*. 2nd ed. Oxford : Basil Blackwell.
- Cortés, J. (1996). *Diccionario de árabe culto moderno : Árabe-español*. Madrid : Gredos.
- Dankovičová, J. (1999). 'Czech', in: IPA (1999), 70-73.
- Donzel, E. van (1997). 'Sirḥān', in: *Encyclopaedia of Islam*, 2nd ed. Leiden : Brill. Volume IX, 673.
- Doughty, C. M. (1979). *Travels in Arabia Deserta* (reprint of the 1936 ed.). New York : Dover Publications.
- Eid, M. (ed.) (1995). *Perspectives on Arabic linguistics VIII : Papers from the eighth annual symposium on Arabic linguistics*. Amsterdam and Philadelphia : John Benjamins.
- and J. J. McCarthy (eds) (1990). *Perspectives on Arabic linguistics II : Papers from the second annual symposium on Arabic linguistics*. Amsterdam, Philadelphia : John Benjamins.
- Eisenstein, H. (1990). 'Die Semantik des ix. Verbalstammes im Klassisch-Arabischen', in: Forstner (1990), 67-77.
- Elihay, J. (2004). *The olive tree dictionary : A transliterated dictionary of conversational eastern Arabic (Palestinian)*. Jerusalem : Minerva.
- Euting, J. (1906). 'Der Kamels-Sattel bei den Beduinen', in Bezold (1906), 393-398.
- al-Faraǧ, Ḥālīd ibn Muḥammad ad-Dawsarī (1952). *Dīwān al-nabaṭ*. 2 volumes. Damascus.
- Ferrando, I. (2018). 'The adnominal linker *-an* in Andalusī Arabic, with special reference to the poetry of Ibn Quzmān (twelfth century)', in: Holes (2018), 96-111.
- Fery, C. and R. van de Vijver (eds) (2003). *The syllable in optimality theory*. Cambridge : Cambridge University Press.

- Finnish Oriental Society (Suomen Itämainen Seura) (ed.) (1995). *Dialectologia Arabica : A collection of articles in honour of the sixtieth birthday of Professor Heikki Palva*. Helsinki : Finnish Oriental Society.
- Fischer, W. (2002). *A grammar of classical Arabic*. J. Rodgers (ed.) 3rd rev. ed. New Haven, Conn. ; London : Yale University Press.
- and O. Jastrow (eds) (1980). *Handbuch der arabischen Dialekte*. Wiesbaden : Harrassowitz.
- Fleisch H. (1944). *Les verbes à allongement vocalique interne en sémitique*. Paris : Institut d'ethnologie.
- Foley, J. (1972). 'Rule precursors and phonological change by meta-rule', in: Stockwell and Macaulay (1972), 96-100.
- Forstner, M. (ed.) (1990). *Festgabe für Hans-Rudolf Singer : Zum 65. Geburtstag am 6. April 1990 überreicht von seinen Freunden und Kollegen, Teil 1*. Frankfurt am Main : Peter Lang.
- Gass, S. M. and L. Selinker (eds) (1992). *Language transfer in language learning*. Rev. ed. Amsterdam : John Benjamins.
- Giolfo, M. E. B. (ed.) (2014). *Arab and Arabic linguistics : Traditional and new theoretical approaches*. Oxford : Oxford University Press.
- Goldenberg, G. and A. Shisha-Halevy (eds) (2009). *Egyptian, Semitic and general grammar : Studies in memory of H. J. Polotsky*. Jerusalem : Israel Academy of Sciences and Humanities.
- Goldsmith, J. A. (ed.) (1995). *The handbook of phonological theory*. Oxford : Blackwell.
- Gordon, M. K. (2016). *Phonological typology*. Oxford : Oxford University Press.
- Grandlauney, R.-V. du (2011). 'Nisba' in: *EALL* Online edition. Consulted online on 09 February 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol3_0230.
- Grigore, G. and G. Bițună (eds) (2016). *Arabic varieties far and wide : Proceedings of the 11th international conference of AIDA, Bucharest 2015*. Bucharest : Editura Universității din București.

- al-Ġuhaymān, ‘Abd al-Karīm (1982-1983). *al-Amṭāl al-ša‘biyya fī qalb al-Ġazīra al-‘Arabiyya*. 10 volumes. Riyadh : Dār Ašbāl al-‘Arab.
- Ġunaydil, Sa‘d ibn ‘Abdallah ibn (1987). *Ḥawāṭir wa-nawādir turāṭīyah*. Riyadh.
- (1981). *Min a‘lām al-adab al-ša‘bī, šu‘arā’ al-‘āliyah*. Riyadh.
- Haak, M., R. de Jong and K. Versteegh (eds) (2004). *Approaches to Arabic dialects : A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*. Leiden : Brill.
- Ḥamīs, ‘Abdallah ibn Muḥammad ibn (1982). *al-Adab al-ša‘bī fī Ġazīrat al-‘Arab*. 2nd ed., Riyadh.
- Ḥanḍal, Fāliḥ (1978). *Mu‘jam al-’alfāḍ al-‘ammiyya fī al-‘Imārāt al-‘Arabiyya al-Muttaḥida*. UAE.
- Harrell, R. S. (ed.) (1966). *A dictionary of Moroccan Arabic : Moroccan–English*. Washington, DC : Georgetown University Press.
- Harrigan, P. (2009). ‘From Moravia to Arabia’, in: *Saudi Aramco World* (magazine), November/December 2009 edition, 8-16. Accessed online at: <https://web.archive.org/web/20131202223849/http://www.saudiaramcoworld.com/issue/200906/from.moravia.to.arabia.htm>.
- Hava, J. G. (1951). *al-Farā’id al-durriyya*. Beirut.
- Hayes, B. (2009). *Introductory phonology*. Malden, MA ; Oxford, Eng. : Wiley-Blackwell.
- (1995). *Metrical stress theory : Principles and case studies*. Chicago ; London : University of Chicago Press.
- (1989). ‘Compensatory lengthening in moraic phonology’, in: *Linguistic Inquiry* 20, 253-306.
- (1980). *A metrical theory of stress rules*. Unpublished MIT dissertation. Cambridge, Massachusetts.

- Henkin, R. (2010). *Negev Arabic : Dialectal, sociolinguistic, and stylistic variation*. Wiesbaden : Harrassowitz.
- (2009). ‘How interdialectal is peripheral oral bedouin poetry?’, in: Goldenberg and Shisha-Halevy (2009), 239-269.
- Hess, J. J. (1938). *Von der Beduinen des innern Arabiens*. Zurich and Leipzig : Max Niehans Verlag.
- Hetzron, R. (ed.) (1997). *The Semitic languages*. London : Routledge.
- (1976). ‘Two principles of genetic reconstruction’, in: *Lingua* 38, 89-108.
- Hinds, M. and El.-S. Badawi (1986). *A dictionary of Egyptian Arabic*. Beirut : Librairie du Liban.
- Holes, C. (ed.) (2018), *Arabic historical dialectology : Linguistic and sociolinguistic approaches*. Oxford : Oxford University Press.
- (2016). *Dialect, culture, and society in eastern Arabia. Volume 3 : Phonology, morphology, syntax, style*. Leiden : Brill.
- (2011a). ‘Nabaṭī poetry, Language of’, in: *EALL* Online edition. Consulted online on 11 February 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_SIM_000031.
- (2011b). ‘Gulf States’, in: *EALL* Online Edition. Consulted online on 04 August 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol2_0032.
- (2011c). ‘Omani Arabic’, in: *EALL* Online Edition. Consulted online on 19 August 2020. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol3_0242.
- (2005). *Dialect, culture, and society in eastern Arabia. Volume 2 : Ethnographic texts*. Leiden : Brill.
- (2004a). *Modern Arabic : Structures, functions, and varieties*. Rev. ed. Washington, D.C. : Georgetown University Press.
- (2004b). ‘*Tanwīn* in the Arabic dialects of eastern Arabia’ in: Aguadé and Woidich (2004), 89-97.

- (2001). *Dialect, culture, and society in eastern Arabia. Volume 1 : Glossary*. Leiden : Brill.
- (1990). *Gulf Arabic*. London : Routledge.
- and S. S. Abu Athera (2009). *Poetry and politics in contemporary bedouin society*. Reading NY : Ithaca Press.
- Ingham, B. (2011). 'Najdi Arabic', in: *EALL Online Edition*. Consulted online on 05 April 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol3_0224.
- (1997). *Arabian diversions : Studies on the dialects of Arabia*. Reading : Ithaca Press.
- (1995). 'Texts in the dialect of the Rwalah of northern Arabia', in: *Finnish Oriental Society (1995)*, 121-140.
- (1994a). *Najdi Arabic : Central Arabian*. Amsterdam : John Benjamins.
- (1994b). 'The effect of language contact on the Arabic dialect of Afghanistan' in: Aguadé, Marugán and Corriente (1994), 105-118.
- (1991). 'Subordinate clauses of time and condition in bedouin dialects', in: *BSOAS* 54 (1), 42-62.
- (1990). 'Camel terminology among the Āl Murrah bedouins', in: *ZAL* 22, 67-78.
- (1986). 'Notes on the dialect of the Āl Murra of eastern and southern Arabia', in: *BSOAS* 49 (2), 271-289; reproduced in: Ingham (1997), 87-116.
- (1982a). *North east Arabian dialects*. London ; Boston : K. Paul International.
- (1982b). 'Notes on the dialect of the Ḍhafir of north-eastern Arabia', in: *BSOAS* 45 (2), 249-297; reproduced in: Ingham (1997), 66-86 [page numbers from this version].
- (1979). 'Notes on the dialect of the Muṭair of eastern Arabia', in: *ZAL* 2, 23-35; reproduced in: Ingham (1997), 52-65 [page numbers from this version].
- IPA (1999). *Handbook of the International Phonetic Association : A guide to the use of the International Phonetic Alphabet*. International Phonetic Association, issuing body. Cambridge : Cambridge University Press.

IPA Chart (2018). <http://www.internationalphoneticassociation.org/content/ipa-chart>, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2018 International Phonetic Association.

Jazayeri, M. A., E. C. Polome and W. Winter (eds) (1978). *Linguistic and literary studies in honour of Archibald A. Hill : II : Descriptive linguistics*. The Hague ; Paris ; New York : Mouton.

Johnstone, T. M. (1981). *Jibbāli lexicon*. Oxford ; New York : Oxford University Press.

——— (1967a). *Eastern Arabian dialect studies*. London Oriental Series 17. London.

——— (1967b). 'Aspects of syllabication in the spoken Arabic of 'Anaiza', in: *BSOAS* 30 (1), Fiftieth Anniversary Volume, 1-16.

——— (1965). 'The sound change $j > y$ in the Arabic dialects of peninsular Arabia', in: *BSOAS* 28, 233-441.

——— (1964). 'Further studies on the Dōsiri dialect of Arabic as spoken in Kuwait', in: *BSOAS* 27, 77-113.

——— (1963). 'The affrication of "kāf" and "gāf" in the Arabic dialects of the Arabian Peninsula', in: *JSS* 8 (2), 210-226.

——— (1961). 'Some characteristics of the Dōsiri dialect of Arabic as spoken in Kuwait', in: *BSOAS* 24 (2), 249-297.

Jones, D. (1917). *An English pronouncing dictionary (on strictly phonetic principles)*. Lond. &c.

Jong, R. de (2011a). *A grammar of the bedouin dialects of central and southern Sinai*. Leiden, Boston : Brill.

——— (2011b) 'Gahawa-syndrome', in: *EALL* Online edition. Consulted online on 09 March 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_SIM_vol2_0008.

——— (2000). *A grammar of the bedouin dialects of the northern Sinai littoral : Bridging the linguistic gap between the eastern and western Arab world*. Leiden : Brill.

- Jurafsky, D. (1996). 'Universal tendencies in the semantics of the diminutive', in: *Language* 72 (3), 533-578.
- Kager, R. (1995). 'The metrical theory of word stress', in: Goldsmith (1995), 367-402.
- Kaye, A. S. and J. Rosenhouse (1997) 'Arabic dialects and Maltese', in: Hetzron (1997), 263-311.
- Keenan, E. L. and B. Comrie (1977). 'Noun phrase accessibility and universal grammar', in: *Linguistic Inquiry* 8 (1), 63-99.
- Kenstowicz, M. (1983). 'Parametric variation in accent assignment', in: Richardson, Marks and Chukerman (1983) (page numbers unknown).
- Kingston, J. and M. Beckman (eds) (1990). *Papers in laboratory phonology 1 : Between the grammar and physics of speech*. Cambridge : Cambridge University Press.
- Kiparsky, P. (2003). 'Syllables and moras in Arabic', in: Fery and van de Vijver (2003), 147-182.
- Kořutová, K. (2018). *Prof. Alois Musil, alias Sheikh Mūsā*. First English edition. Vyřkov : Karel Dvořáček Library and Academic Society of Alois Musil.
- Kurpershoek, P. M. (2015). 'Praying mantis in the desert : The Najdi poet Ibn Subayyil consumed with love for the bedouin', in: *Arabian Humanities* 5. Consulted at: <https://journals.openedition.org/cy/2962>.
- (2005). *Voices from the desert : Glossary*. Leiden ; New York : Brill.
- (2002). *A Saudi tribal history : Honour and faith in the traditions of the Dawāsir*. Leiden : Brill.
- (1999). *Bedouin poets of the Dawāsir tribe : Between nomadism and settlement in southern Najd*. Leiden : Brill.
- (1995). *The story of a desert knight : The legend of Šlēwīḥ al-‘Aṭāwī and other ‘Uṭaybah heroes : an edition with translation and introduction*. Leiden : Brill.
- (1994). *The poetry of ad-Dindān : A bedouin bard in southern Najd : An edition with translation and introduction*. Leiden : Brill.

- La Rosa, C. (2016). 'Le relateur -Vn en arabe de Sicile : exemples et remarques linguistiques' in: Grigore and Bițună (2016), 359-367.
- Ladefoged, P. and I. Maddieson (1996). *The sounds of the world's languages*. Oxford : Blackwell.
- Lancaster, W. (1997). *The Rwala bedouin today*. 2nd ed. Prospect Heights : Waveland.
- and F. Lancaster (1997). 'Šhararāt', in: *Encyclopaedia of Islam*, 2nd ed. Leiden : Brill. Volume IX, 316-317.
- (1995). 'Ruwala', in: *Encyclopaedia of Islam*, 2nd ed. Leiden : Brill. Volume VIII, 643-645.
- (1988). 'Thoughts on the bedouinisation of Arabia', in: *Proceedings of the Seminar for Arabian Studies* 18, 51-62.
- Landberg, C. de (1940). *Glossaire de la langue des bédouins 'Anazeh*. Ed. K. V. Zetterstéen. Uppsala : A. B. Lundequistska Bokhandeln.
- (1920-42). *Glossaire Daḡnais*. 3 volumes. Leiden : E. J. Brill.
- (1919). *Langue des bédouins 'Anazeh : Texte arabe avec traduction, commentaire et glossaire*. Part 1: *Texte arabe et traduction*. (Following parts were not published, but the glossary was edited by Zetterstéen and published as Landberg 1940). Leiden.
- (1905-13). *Études sur les dialectes de l'Arabie méridionale. Deuxième volume : Daḡnah*. 3 volumes. Leiden : E. J. Brill.
- (1901). *Études sur les dialectes de l'Arabie méridionale. Premier volume : Ḥaḡramoût*. Leiden : E. J. Brill.
- (1895). *Critica Arabica* 3. Leiden : E. J. Brill.
- Lane, E. W. (1984). *Arabic-English lexicon*. Cambridge : Islamic Texts Society.
- Leben, W. (1973). *Suprasegmental phonology*. Ph.D. dissertation, MIT.
- Lehn, W. (1967). 'Vowel contrasts in Najdi Arabic', in: Stuart (1967), 123-131.

- Littman, E. (1926). 'Zwei seltenere arabische Nominalbildungen (*qaitul* und *quttail*)', in: *ZSVG* 4, 24–41.
- Maamouri, M. (ed.) (2013). *The Georgetown dictionary of Iraqi Arabic, Arabic-English, English-Arabic*. Washington D.C. : Georgetown University Press.
- Macdonald, M. C. A. (2015). 'Was there a "bedouinization of Arabia"?', in: *Der Islam* 92, 42-84.
- Mandaville, J. P. (2011). *Bedouin ethnobotany : Plant concepts and uses in a desert pastoral world*. Tucson : University of Arizona Press.
- McCarthy, J. J. (2005). 'The length of stem-final vowels in colloquial Arabic', in: Alhawary (2005), 1-26.
- (1986). 'OCP effects: Gemination and antigemination', in: *Linguistic Inquiry* 17, 207-263.
- (1979a). *Formal problems in Semitic phonology and morphology*. Ph.D. dissertation, MIT; published as McCarthy (1985), New York ; London : Garland Publishing.
- (1979b). 'On stress and syllabification', in: *Linguistic Inquiry* 10, 443-465.
- and A. Prince (1990). 'Prosodic morphology and templatic morphology', in: Eid and McCarthy (1990), 1-54.
- Meeker, M. E. (1979). *Literature and violence in North Arabia*. Cambridge, Eng. ; New York : Cambridge University Press.
- Mitchell, T. F. (1960). 'Prominence and syllabication in Arabic', in: *BSOAS* 23, 369–389.
- Monahan, P. J. (2011) 'Vowel harmony', in: *EALL* Online edition. Consulted online on 07 March 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0373.
- Montagne, R. (1935). 'Contes poétiques bédouins (recueillis chez les Šammar de Ğeziré)' in: *BEO* 5, 33-119.
- Moore, J. (1990). 'Doubled verbs in Modern Standard Arabic', in: Eid and McCarthy (1990), 55-93.

Musil, A. (1928a). *The manners and customs of the Rwala bedouins*. New York : American Geographical Society.

——— (1928b). *Northern Neġd, a topographical itinerary*. New York : American Geographical Society.

——— (1928c) *Palmyrena, a topographical itinerary* (Oriental Explorations and Studies 4). New York : American Geographical Society.

——— (1927a). *The middle Euphrates, a topographical itinerary* (Oriental Explorations and Studies 3). New York : American Geographical Society.

——— (1927b). *Arabia Deserta, a topographical itinerary* (Oriental Explorations and Studies 2). New York : American Geographical Society.

——— (1926). *The northern Heġâz, a topographical itinerary* (Oriental Explorations and Studies 1). New York : American Geographical Society.

Owens, J. (2006). *A linguistic history of Arabic*. Oxford : Oxford University Press.

——— (1993). 'Nigerian Arabic in comparative perspective' in: *Sprache und Geschichte in Afrika* 14, 85-176.

Palva, H. (2011). 'Northwest Arabian Arabic', in: *EALL* Online edition. Consulted online on 30 March 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol3_0233.

——— (1993). 'Metrical problems of the contemporary bedouin *qaṣīda* : A linguistic approach', in: *Asian Folklore Studies* 52 (1), 75-92.

——— (1992). *Artistic colloquial Arabic : Traditional narratives and poems from al-Balqā' (Jordan)*. Helsinki : Finnish Oriental Society.

——— (1984-86). 'Characteristics of the Arabic dialect of the Ḥwēṭāt tribe', in: *Orientalia Suecana*, 33-5, 295-312.

——— (1984). 'A general classification for the Arabic dialects spoken in Palestine and Transjordan', in: *Studia Orientalia* 55, 359-376.

- (1980). 'Characteristics of the Arabic dialect of the Bani Şaxar tribe', in: *Orientalia Suecana* 29, 112-139.
- (1978). *Narratives and poems from Ḥesbān : Arabic texts recorded among the semi-nomadic al-ʿAğārma tribe (al-Balqāʾ district, Jordan)*. Göteborg, Sweden : Acta Universitatis Gothoburgensis.
- (1976). *Studies in the Arabic dialect of the semi-nomadic al-ʿAğārma tribe (al-Balqāʾ district, Jordan)*. Gothenburg : Acta Universitatis Gothoburgensis.
- Paradis, C. and J.-F. Prunet (1989). 'On coronal transparency', in: *Phonology* 6 (2), 317-348.
- Parkinson, F. B. (1992). 'The feature pharyngeal in Rwaiili Arabic : A case for long distance multiple linking', in: *Working papers in linguistics*, no. 41, 101-134. Ohio State University : Department of Linguistics.
- Piamenta, M. (1990-1991). *Dictionary of post-classical Yemeni Arabic*. Leiden : Brill.
- Procházka, S. (2011). 'Prepositions', in: *EALL* Online edition. Consulted online on 21 March 2018. http://dx.doi.org/10.1163/1570-6699_eall_EALL_SIM_vol3_0108.
- (2004). 'Unmarked feminine nouns in modern Arabic dialects', in: Haak, de Jong and Versteegh (2004), 237-262.
- Prochazka, T. (1988). *Saudi Arabian dialects*. London : Kegan Paul International.
- Qafisheh, H. A. (1997). *NTC's Gulf Arabic-English dictionary*. Lincolnwood, Ill. : NTC Publishing Group.
- Redhouse, J. W. (1856). *An English and Turkish dictionary*. London : Bernard Quaritch.
- Reinhardt, C. (1894). *Ein arabischer Dialekt gesprochen in ʿOmān und Zanzibar*. Stuttgart &c.
- RGS (The Royal Geographical Society with the Institute of British Geographers) (1910). 'Musil's explorations in northern Arabia, 1908-9', in: *The Geographical Journal* 35 (5), 579.
- Richardson, J., M. Marks, and A. Chukerman (eds) (1983). *Proceedings of the nineteenth regional meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, University of Chicago, Chicago, Illinois.

- Rosenhouse, J. (2011). 'Bedouin Arabic', in: *EALL* Online edition. Consulted online on 08 April 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0037.
- (1984). *The bedouin Arabic dialects : General problems and a close analysis of North Israel bedouin dialects*. Wiesbaden : O. Harrassowitz.
- (1982). 'An analysis of major tendencies in the development of the bedouin dialects of the north of Israel', in: *BSOAS* 45 (1), 14-38.
- Ryding, K. C. (2011). 'Iḏāfa', in: *EALL* Online edition. Consulted online on 12 August 2020. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol2_0043.
- (2005). *A reference grammar of Modern Standard Arabic*. Cambridge : Cambridge University Press.
- Šimáčková, Š., V. J. Podlipský and K. Chládková (2012). 'Czech spoken in Bohemia and Moravia', in: *JIPA* 42 (2), 225-232.
- Simpson, St. J. (1994). 'Gazelle-hunters and salt-collectors : a further note on the Solubba', in: *BASOR* 293, 79-81.
- Socin, A. (1900-1901). *Diwan aus Centralarabien*. 3 volumes. 1. Texte nebst Glossen und Excurse (1900); 2. Übersetzung (1900); 3. Einleitung, Glossar und Indices (1901). Leipzig.
- Sowayan, S. A. (ed.) (2000). *al-Ṭaqāfah al-taqlīdiyyah fī al-Mamlakah al-ʿArabiyyah al-Suʿūdiyyah*. 12 volumes. Riyadh.
- (1992). *The Arabian oral historical narrative : An ethnographic and linguistic analysis*. Wiesbaden : O. Harrassowitz.
- (1985). *Nabaṭi poetry : The oral poetry of Arabia*. Berkeley ; London : University of California Press.
- (1982). 'A poem and its narrative by Riḏa ibn Ṭārif aš-Šammari', in: *ZAL* 7, 48-73.
- Steingass, F (2007). *A comprehensive Persian-English dictionary*. 1892, reprinted 2007 New Delhi : Manohar.

- Stewart, F. H. (1990). *Texts in Sinai bedouin law. Part 2 : The texts in Arabic ; Glossary*. Wiesbaden : Harrassowitz.
- (1986). 'Review of Lancaster (1997) *The Rwala bedouin today*', in: *International Journal of Middle East Studies* 18 (2), 243-246.
- Stockwell, R. P. and R. K. S. Macaulay (eds) (1972). *Linguistic change and generative theory*. Bloomington, Ind : Indiana University Press.
- Stuart, D. G. (ed.) (1967). *Linguistic studies in memory of Richard Slade Harrell*. Washington D.C. : Georgetown University Press.
- Subayyil, 'Abd Allāh ibn (2018). *Arabian romantic : Poems on bedouin life and love* (edited and translated by M. Kurpershoek). New York : New York University Press.
- al-Suwaydā', 'Abd ar-Raḥmān ibn Zayd (1988). *Min šu'arā' al-ğabal al-‘āmmiyyīn*. 3 volumes. Riyadh.
- (1987). *Faṣīḥ al-‘āmmi fī šamāl Nağd*. 2 volumes. Riyadh : Dār al-Suwaydā'.
- Sweet, L. (1983). 'Review of Lancaster (1997) *The Rwala bedouin today*', in: *American Ethnologist*, 10 (1) 212-213.
- Taine-Cheikh, C. (2011). 'Numerals', in: *EALL* Online edition. Consulted online on 10 February 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_vol3_0238.
- Tapper, R. (1983). 'Review of Lancaster (1997) *The Rwala bedouin today*', in: *BSOAS* 46 (1), 152-153.
- Teeple, D. (2011). 'Syncope', in: *EALL* Online edition. Consulted online on 12 August 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0328.
- Trask, L. and R. M. Millar (2015). *Trask's historical linguistics*. 3rd edition. Abingdon ; New York : Routledge.
- Versteegh, K. (2011). 'Interrogative pronouns', in *EALL* Online edition. Consulted online on 24 May 2018. http://dx.doi.org/10.1163/1570-6699_eall_EALL_SIM_vol2_0028.
- (1984). *Pidginization and creolization : The case of Arabic*. Amsterdam ; Philadelphia : J. Benjamins.

- Wallin, G. A. (1855). 'Über die Laute des Arabischen und ihre Bezeichnung', in: *ZDMG* 9, 1-69.
- (1851-2). 'Probe aus einer Anthologie neuarabischer Gesänger in der Wüste gesammelt', in: *ZDMG* 5 (1851), 1-23, and 6 (1852), 190-218.
- Waltisberg, M. (2011). 'Conjunctions', in: *EALL* Online edition. Consulted online on 31 May 2018. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0065.
- Watson, J. C. E. (2011a). 'Dialects of the Arabian Peninsula' in: Weninger, Khan, Streck and Watson (2011), 897-908.
- (2011b). 'Adverbs', in: *EALL* Online edition. Consulted online on 27 March 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0005.
- (2011c). 'Şanʿānī Arabic', in: *EALL* Online edition. Consulted online on 14 August 2020. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0293.
- (2007). 'Syllabification patterns in Arabic dialects : long segments and mora sharing', in: *Phonology*, 24, 335–356.
- (2006). 'Arabic morphology : diminutive verbs and diminutive nouns in Sanʿani Arabic', in: *Morphology* 16 (2), 189-204.
- (2002). *The phonology and morphology of Arabic*. Oxford : Oxford University Press (reprint 2007).
- (1993). *A syntax of Şanʿānī Arabic*. Wiesbaden : Harrassowitz.
- Wehr, H. (1994). *A dictionary of modern written Arabic : (Arabic-English)*. ed. J. M. Cowan, 4th edition, considerably enlarged and amended by the author. Urbana, IL : Spoken Languages Services, Inc.
- Weninger, S., G. Khan, M. P. Streck and J. C. E. Watson (eds) (2011). *The Semitic languages : An international handbook*. Berlin ; Boston : De Gruyter Mouton.
- Werbeck, W. (2001). *Laut- und Formenlehre des nordjemenitisch-arabischen Dialekts von Manāḥa*. Münster : Rhema.

- Wetzstein, I. G. (1868). 'Sprachliches aus den Zeltlagern der syrischen Wüste', in *ZDMG* 12, 69-194.
- Wright, J. (1927). 'Northern Arabia : the explorations of Alois Musil', in: *Geographical Review* 17 (2), 177.
- Wright, W. (1967). *A grammar of the Arabic language*. Two volumes. Translation of original by C. P. Caspari. 3rd ed. rev. by W. Robertson Smith and M. J. de Goeje. Cambridge : Cambridge University Press.
- Yassin, M. A. F. (1977). 'Bi-polar terms of address in Kuwaiti Arabic', in: *BSOAS* 40 (2), 297-301.
- Zemánek, P. (2011a). 'Assimilation', in: *EALL* Online edition. Consulted online on 11 March 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0030.
- (2011b). 'Root', in: *EALL* Online edition. Consulted online on 05 August 2019. http://dx.doi.org/10.1163/1570-6699_eall_EALL_COM_0291.