2 Phonology

The synchronic description of the grammar of Maltese Π follows the traditional patterns in the sense that we proceed from phonology via morphology and syntax to semantics. For all of these sections, the division of the members of BLOMP 2.0 in (1.188) into two major classes based on the language of origin of the Πs – Semitic vs Italo-Romance – serves as a common reference point. Saari (2003: 15–16 and 44–45) distinguishes 29 Semitic from four Italo-Romance Πs, i.e. in his account, 88% of the Maltese Πs have a Semitic background as opposed to only 12% which are borrowings from Italo-Romance. Our own investigations yield strikingly different percentages.

In Figure 2.1, a third and a fourth class labelled MIXED and UNCERTAIN are featured, respectively. The sole representative of MIXED is (Π35) *minflok* 'instead of' whose hybrid character historically results from the combination of Semitic *minn* 'from', Semitic *fi* 'in' and the Italo-Romance noun *lok* 'place'.[1] UNCERTAIN too contains only a single element, namely (Π10) *daqs* 'equal to' which is assumed to have either a dialectal Arabic or a Greek origin (Aquilina 1987: 249).

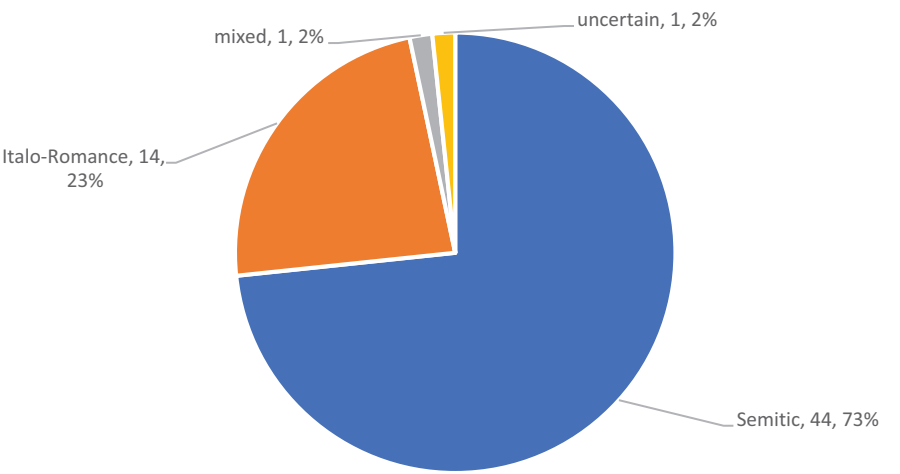


Figure 2.1: Share of Maltese Πs according to language of origin I.

The picture which emerges from Figure 2.1 closely resembles that presented in Schmidt/Vorholt/Witt (2020: 249) and Vorholt (2023a). For the purpose of the

1 As argued in Section 1.6.4.1 (v).

https://doi.org/10.1515/9783111331591-002

ensuing discussion, it is, however, unnecessary to keep the minor classes MIXED and UNCERTAIN. Since in the latter case, the possibility of a dialectal Arabic source cannot be disproved, we classify (Π10) *daqs* 'equal to' as a Semitic Π. Serracino-Inglott (1975b: 19) claims that the proposed Arabic and Greek origins of *daqs* are too far off the mark semantically. His own proposal assumes the univerbation of the PP *ta' qies* 'of (a) measure' to *daqs*. This diachronic development requires several intermediate steps, namely first the proper univerbation of Π and complement which is then followed by the shift of the stress site from *qies* to *ta'*. This shift is necessary to motivate the reduction and deletion of the long high front vowel in the second syllable. Independent of the disappearance of this vowel, the initial voiceless denti-alveolar plosive /t/ must have undergone sonorization to yield /d/. This means that Serracino-Inglott's hypothesis rests on the chronology of several subprocesses none of which can be proved by philological evidence – and thus remains hypothetical. Even if the Greek hypothesis is correct, it is still possible to argue that this Π belongs to the non-Italo-Romance component of the inventory. This gives us the opportunity to subsume (Π35) *minflok* 'instead of' under the rubric of the Italo-Romance Πs of Maltese on the understanding that a Maltese Π is considered Italo-Romance in origin if at least an identifiable part of its form can be traced back etymologically to an Italo-Romance variety. In this way, UNCERTAIN is added to Semitic whereas MIXED and Italo-Romance are united to yield the distribution presented in Figure 2.2.

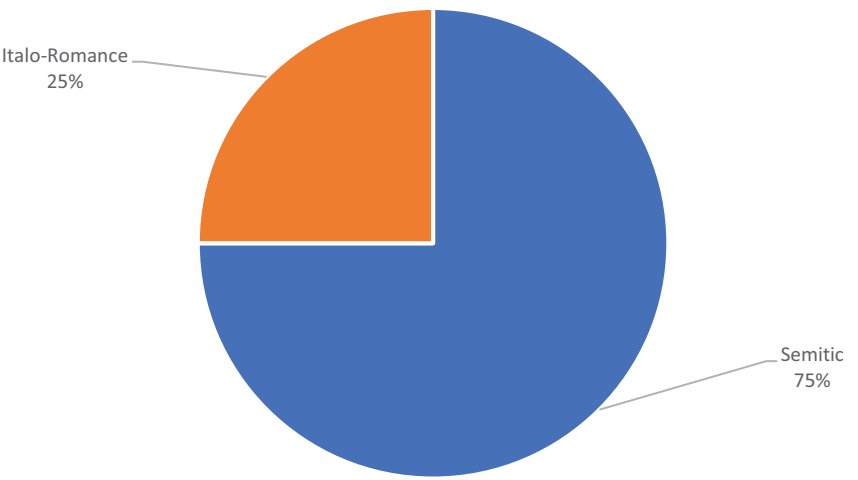


Figure 2.2: Share of Maltese Πs according to language of origin II.

There is a now a majority of 45 Semitic ΠΙs alongside a minority of 15 Italo-Romance ΠΙs – these absolute numbers correspond to a three-to-one ratio. For the interpretation of the phenomena presented and discussed subsequently, this quantitative difference provides a kind of yardstick insofar as we check for each phenomenon whether the above shares are replicated. The Italo-Romance share of the Maltese Π-inventory is for instance smaller than the Italo-Romance share in the Maltese lexicon in general but much bigger than the Italo-Romance share in the domain of function words (Schmidt/Vorholt/Witt 2020: 250–251). According to Comrie/Spagnol (2016: 325 and 328), some 30% of the Maltese lexicon can be attributed to Italo-omance (as opposed to 56% classified as Arabic = Semitic) whereas with 85% the Semitic component predominates in the domain of function words where Italo-Romance covers only 6% of the inventory. Thus, ΠΙs occupy a middle position between the sizable Italo-Romance share in the general lexicon and the almost negligible Italo-Romance impact on function words. This means that ΠΙs differ from other subclasses of function words in terms of the processes of contact-induced transfer. In this section, among other things, we test whether Semitic and Italo-Romance ΠΙs behave differently as to their phonological properties.

To address this and other issues, we start from the suprasegmental level in Section 2.1 which is dedicated to syllables. Section 2.2 looks into the segmental side of the phonological form of Maltese ΠΙs. The facts are analysed both qualitatively and quantitatively. In the latter case, we try to correlate the token frequency of the Maltese ΠΙs to their complexity in terms of the number of syllables and segments which constitute their lexical form.[2] To facilitate understanding of what we are about to argue in the subsequent sections, we provide an overview over the ΠΙs and some of their relevant properties in Table 2.1. The ΠΙs are ordered according to alphabetical principles. For each Π, the phonetic realisation is disclosed in broad IPA-based transcription. In those cases where final devoicing neutralises the phonation contrast to the detriment of [+voice], the cell is divided in two with the right sub-cell hosting the word-final phoneme. For polysyllables, we add the syllable boundaries separately using standard orthography. The number of syllables is indicated as well as that of segments. Note that long vowels count as two segments in Table 2.1. In the rightmost column, the language of origin of the Π is identified.

2 Morphonological processes (including external sandhi) which result in allophony and segmental reduction of the ΠΙs are discussed in Section 3.

Table 2.1: Overview over Пs and their phonologically relevant properties.

#	П	IPA		syllabified	syllables	segments	origin
П01	*apparti*	[ɐppɐrtɪ]		ap.par.ti	3	7	Romance
П02	*a skapitu*	[ɐskɐpɪtʊ]		as.ka.pi.tu	4	8	Romance
П03	*barra*	[bɐrrɐ]		bar.ra	2	5	Semitic
П04	*bejn*	[bɛɪn]			1	4	Semitic
П05	*bħal*	[phɐːl]			1	5	Semitic
П06	*bħala*	[phɐːlɐ]		bħa.la	2	6	Semitic
П07	*bi*	[bɪ]			1	2	Semitic
П08	*biswit*	[bɪswiːt]		bis.wit	2	7	Semitic
П09	*bla*	[blɐ]			1	3	Semitic
П10	*daqs*	[dɐʔs]			1	4	Semitic
П11	*dwar*	[dwɐr]			1	4	Semitic
П12	*faċċata*	[fɐʧʧɐtɐ]		faċ.ċa.ta	3	7	Romance
П13	*favur*	[fɐvʊr]		fa.vur	2	5	Romance
П14	*fejn*	[fɛɪn]			1	4	Semitic
П15	*fi*	[fɪ]			1	2	Semitic
П16	*fost*	[fɔst]			1	4	Semitic
П17	*fuq*	[fʊʔ]			1	3	Semitic
П18	*ġewwa*	[ʤɛʊwɐ]		ġew.wa	2	5	Semitic
П19	*ġo*	[ʤɔː]			1	3	Semitic
П20	*għajr*	[ɐːɪr]			1	5	Semitic
П21	*għal*	[ɐːl]			1	4	Semitic
П22	*għand*	[ɐːnt]	/d/		1	5	Semitic
П23	*ħdejn*	[hdɛɪn]			1	5	Semitic
П24	*ħlief*	[hlɪːf]			1	5	Semitic
П25	*inkluż*	[ɪnklʊːs]	/z/	in.kluż	2	7	Romance
П26	*kif*	[kɪf]			1	3	Semitic
П27	*kontra*	[kɔntrɐ]		kon.tra	2	6	Romance
П28	*lejn*	[lɛɪn]			1	4	Semitic
П29	*lil*	[lɪl]			1	3	Semitic
П30	*ma'*	[mɐ]			1	2	Semitic
П31	*madwar*	[mɐdwɐr]		ma.dwar	2	6	Semitic
П32	*maġenb*	[mɐʤɛmp]	/b/	ma.ġenb	2	6	Semitic
П33	*matul*	[mɐtʊːl]		ma.tul	2	6	Semitic
П34	*minbarra*	[mɪnbɐrrɐ]		min.bar.ra	3	8	Semitic
П35	*minflok*	[mɪnflɔk]		min.flok	2	7	Romance
П36	*mingħajr*	[mɪnɐːɪr]		min.għajr	2	8	Semitic
П37	*mingħala*	[mɪnɐːlɐ]		min.għa.la	3	8	Semitic
П38	*mingħand*	[mɪnɐːnt]	/d/	min.għand	2	8	Semitic
П39	*minħabba*	[mɪnhɐbbɐ]		min.ħab.ba	3	8	Semitic
П40	*minkejja*	[mɪnkɛjjɐ]		min.kej.ja	3	8	Semitic
П41	*minn*	[mɪnn]			1	4	Semitic
П42	*mintul*	[mɪntʊːl]		min.tul	2	7	Semitic
П43	*mnejn*	[mnɛɪn]			1	5	Semitic

Table 2.1 (continued)

#	Π	IPA		syllabified	syllables	segments	origin
Π44	*oltre*	[oltrɛ]		ol.tre	2	5	Romance
Π45	*permezz*	[pɛrmɛʦʦ]		per.mezz	2	7	Romance
Π46	*qabel*	[ʔebɛl]		qa.bel	2	5	Semitic
Π47	*qalb*	[ʔelp]	/b/		1	4	Semitic
Π48	*qrib*	[ʔrɪp]	/b/		1	4	Semitic
Π49	*quddiem*	[ʔʊddɪːm]		qud.diem	2	7	Semitic
Π50	*rigward*	[rɪgwert]	/d/	ri.gward	2	7	Romance
Π51	*sa*	[sɐ]			1	2	Semitic
Π52	*sforz*	[sfɔrʦ]			1	5	Romance
Π53	*skont*	[skɔnt]			1	5	Romance
Π54	*sotta*	[sɔttɐ]		sot.ta	2	5	Romance
Π55	*ta'*	[tɐ]			1	2	Semitic
Π56	*taħt*	[tɐht]			1	4	Semitic
Π57	*versu*	[vɛrsu]		ver.su	2	5	Romance
Π58	*vičin*	[vɪʧɪn]		vi.čin	2	5	Romance
Π59	*waqt*	[weʔt]			1	4	Semitic
Π60	*wara*	[werɐ]		wa.ra	2	4	Semitic

The grey-shaded rows host Πs whose segmental chain is phonetically by one unit shorter than its phonological equivalent because of the hidden presence of the abstract phoneme *għajn*[3] How we calculate the number of segments is explained below.[4]

2.1 Syllables

We look at the syllable in connection to the Maltese Π from two different angles. For a start, we determine the patterns which arise in the domain of syllabicity, i.e. the number of syllables which are contained in the lexicon form of a given Π. This is the task of Section 2.1.1. In Section 2.1.2 the focus is on the syllable structure, especially of those Πs which only count one syllable. The latter section forms the bridge to the discussion of the segmental structure of the Πs to be addressed in Section 2.2.

3 This abstract phoneme is defined in Section 2.2.1.

4 Cf. in Section 2.2.2.

2.1.1 Syllabicity

Saari (2003) does not treat syllabicity specifically. Schmidt/Vorholt/Witt (2020: 248–249) provide a ternary taxonomy of Maltese ΠΙs which comprises the categories monosyllabic, disyllabic (= "bisyllabic"), and polysyllabic (for ΠΙs with more than two syllables). According to their count, monosyllables form the largest group with thirty members followed by disyllabic ΠΙs with 21 cases. The smallest turnout goes to polysyllables which account for only six cases. In contrast to Schmidt/Vorholt/Witt (2020), we revert to a binary taxonomy with monosyllables and polysyllables. Figure 2.3 suggests that the shares are absolutely identical. Both monosyllables and polysyllables cover half of BLOMP 2.0. The average syllabicity ratio is 1.63 per ΠΙ.

Figure 2.3: Shares of monosyllabic and polysyllabic ΠΙs.

Superficially, one could mistake these identical shares to mean that nothing much can be gained from inquiring further into syllabicity. Figure 2.4 however, proves this assumption wrong.

It transpires from Figure 2.4 that, in terms of syllabicity, it makes a difference whether a ΠΙ stems from Semitic or from Italo-Romance. About 62% of all Semitic ΠΙs are monosyllables. As to those ΠΙs which have been borrowed from Italo-Romance, the share of monosyllables is down to slightly more than 13%. The importance of this imbalance is further strengthened by the shares the two languages of origin have of the different categories of syllabicity as shown in Figure 2.5.

The Italo-Romance contribution to the subclass of monosyllabic ΠΙs is small since it hardly reaches the 7%-mark. In contrast, Italo-Romance is responsible for 43% of all polysyllabic ΠΙs. This share is remarkably bigger than that of Italo-Romance

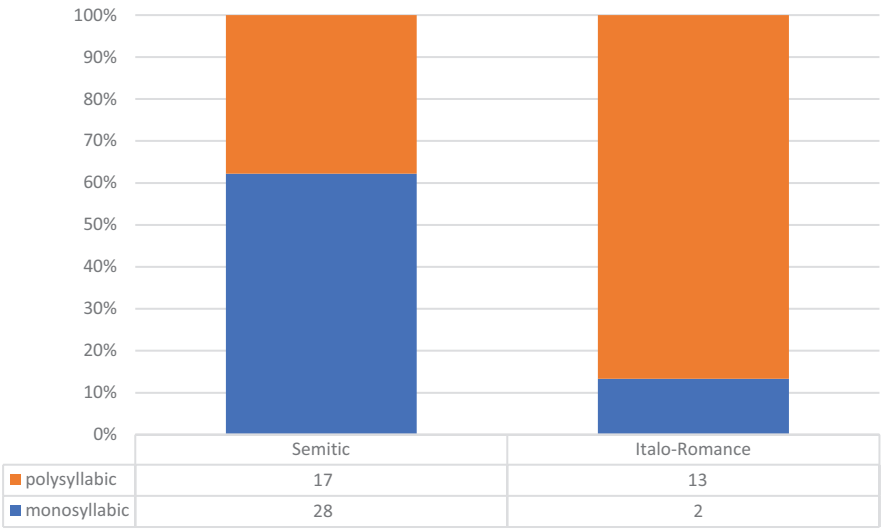


Figure 2.4: Shares of monosyllabic and polysyllabic Πs per language of origin.

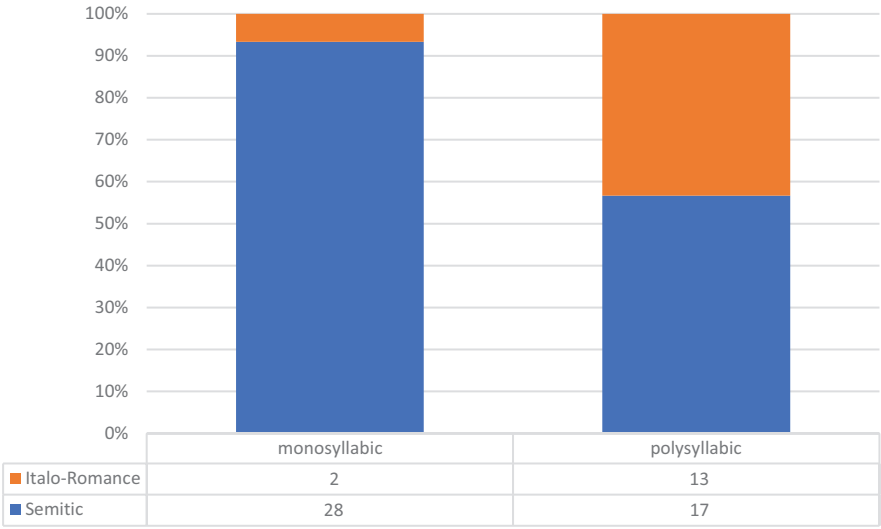


Figure 2.5: Shares of language of origin per syllabicity class.

in the general inventory of Πs as shown in Figure 2.1. This means that Italo-Romance origin is strongly associated with polysyllabicity whereas monosyllabicity and Semitic origin go together very well. Note that we are talking about preferences and not about categorical differences.

The 30 polysyllables divide into two types, namely 23 disyllabic (~ 77%), six trisyllabic Πs (~ 20%) and a single Π with four syllables (~ 3%). With ten disyllabic Πs, Italo-Romance accounts for 43% of all disyllables. The Italo-Romance share is 33% in the category of trisyllables. Moreover, the only tetrasyllabic Π is also of Italo-Romance origin. Each of these shares exceeds the share of Italo-Romance as computed for the entire Π-inventory and thus, they are in line with the above observation that the Italo-Romance component is especially strong in the domain of polysyllables.

With polysyllables, it remains to decide which syllable bears primary stress. According to Borg/Azzopardi-Alexander (1997: 320), penultimate stress is the default in Maltese. The stress site is located in the ultimate however, provided the vowel in this syllable is long or a stress-shifting suffix is added or there is a heavy coda with heavy meaning the presence of a geminate consonant or a consonant cluster.[5] The distribution of the polysyllabic Maltese Πs over the possible two stress sites can be gathered from Figure 2.6.

All six of the trisyllabic Πs and the tetrasyllabic Π give evidence of penultimate stress. The remaining ten cases of Πs with penultimate stress divide into six Semitic and four Italo-Romance Πs. There are altogether thirteen Πs with stress on the final syllable. In this subclass of the disyllables, the Italo-Romance share is even higher (though only by a very narrow margin) as shown in Figure 2.7.

Eleven of the 30 polysyllables are Πs with initial *ma'-* or *minn-*.[6] Six of this group of eleven display ultimate stress. What can be concluded with regard to polysyllabic Maltese Πs is that penultimate stress is far less dominant than in the general lexicon of the language. The increased importance of ultimate stress results primarily from the integration of Italo-Romance Πs and the lexicalisation of complex Πs. It is worth noting that ultimate stress in Italo-Romance Πs is secondary in the sense that the

5 Borg/Azzopardi-Alexander (1997: 320) assume tautosyllabic ternary consonant clusters in CODA position. This assumption contradicts the same authors' model of the canonical syllable structure which allows only for binary clusters in this position (Borg/Azzopardi-Alexander 1997: 313) to which we refer in Section 2.1.2. Whether syllable-final CCC-sequences result from processes of resyllabification at the syllable contact in polysyllables cannot be ascertained since our source does not provide examples of ternary clusters in polysyllables. A possible pattern from which final CCC-sequences arise is discussed however in connection with the negative clitic *-x* /ʃ/ which can be attached to words ending in CC. Since this situation arises typically with perfective verbs in the 1st/2nd singular marked by the suffix *-t*, Borg/Azzopardi-Alexander (1997: 307) argue that the phonological sequence /t/ + /ʃ/ yields the affricate [ʧ] on the phonetic level so that it is possible to reanalyse the ternary cluster as a binary one with the affricate constituting the offset. In any case, ternary clusters in the coda are possible only if a combination of inflection and cliticization occurs. From the lexicon however, extra-heavy CODAS seem to be excluded.

6 As discussed in Section 1.6.4.2.

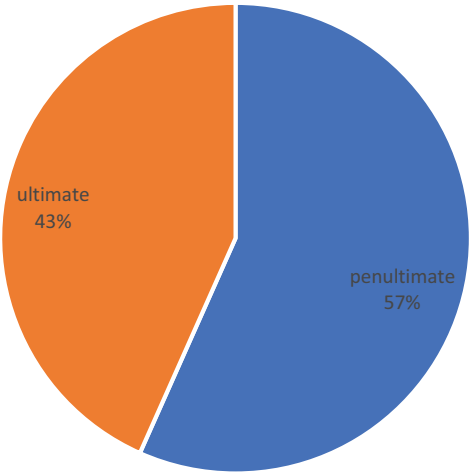
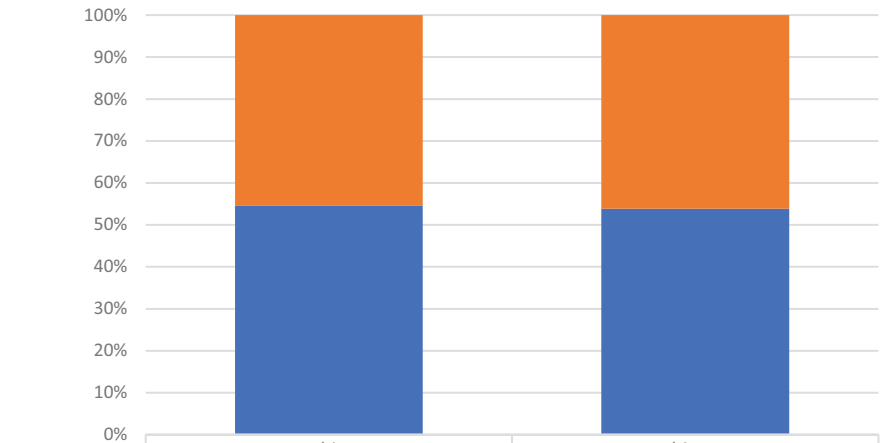


Figure 2.6: Distribution of polysyllabic Пs over stress sites.



	penultimate	ultimate
Italo-Romance	5	6
Semitic	6	7

Figure 2.7: Shares of language of origin per stress site in disyllables.

original penultimate stress remained in place also after the unstressed final vowel underwent apocope.[7]

7 Cf. Section 1.6.6 on this and related processes of adaptation.

What else can be said about polysyllables? There are only four Πs whose initial segment is a vowel, namely (Π1) *apparti* 'apart from', (Π2) *a skapitu* 'at the expense of', (Π25) *inkluż* 'including', and (Π44) *oltre* 'beyond' all four of which are Italo-Romance. All other Πs, be they polysyllabic or monosyllabic, are equipped with an initial consonant. This means that the following implication applies (cf. Figure 2.8).

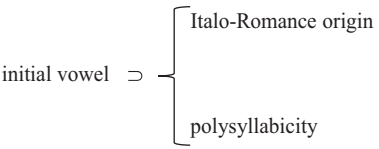


Figure 2.8: Implicational pattern of vowel-initial Πs.

Furthermore, it is possible to complement this with another implication which we disclose in Figure 2.9.

Semitic origin ⊃ initial consonant

Figure 2.9: Implicational pattern of Semitic Πs.

Why there are no genuine cases of vowel-initial Semitic Πs is discussed below.[8]

When we extend the discussion to the right margin of the Πs, we notice that no tri- or tetrasyllabic Π ends in a consonant. Thus, Figure 2.10 features a third implication.

σσσ(σ) ⊃ final vowel

Figure 2.10: Implicational pattern of tri- and tetrasyllabic Πs.

BLOMP 2.0 contains seven monosyllables which end in a vowel. All of these Πs are Semitic, viz. (Π7) *bi* 'with', (Π9) *bla* 'without', (Π15) *fi* 'in', (Π19) *ġo* 'in', (Π30) *ma'* 'with', (Π51) *sa* 'till', and (Π55) *ta'* 'of'. Accordingly, monosyllabicity, final vowel, and Semitic origin are correlated to each other implicationally as shown in Figure 2.11.

monosyllabic & final vowel ⊃ Semitic origin

Figure 2.11: Implicational pattern of monosyllabic Πs with final vowel.

This implication guides us directly to Section 2.1.2 where we scrutinise among other things the structure of monosyllables.

8 Cf. Section 2.2.

2.1.2 Syllable structure

Saari (2003: 101), as well as Schmidt/Vorholt/Witt (2020: 248), do not differentiate between the different syllabicity classes. The structures of the IIs are represented as sequences of C-symbols and V-symbols resulting in a plethora of different patterns glossing over syllable boundaries. For the purpose of this study, we prefer to separate monosyllables from polysyllables in order to make meaningful statements about the syllable structure patterns of the former. In what follows in this section, we only make statements about IIs whose segmental extension amounts to exactly a syllable.

We take Vennemann (1988: 5–6) as our frame of reference for the analysis of syllable structures. This means that we assume the hierarchical model of the syllable presented in Figure 2.12. The model is meant to hold universally. Thus, it should also be applicable to Maltese. The immediate constituents of the syllable are HEAD and RHYME. The latter can be complex provided there is a CODA. The red circles and the green triangle single out pairs of constituents of the syllable which are believed to interact with each other systematically but which do not form proper constituents according to the logic of the architecture of Vennemann's model. The syllable BODY comprises HEAD and NUCLEUS. HEAD and CODA form the syllable shell. The only obligatory constituent is the NUCLEUS whose slot is normally filled by a vowel or a syllabic consonant whereas the positions to the right and the left of the NUCLEUS are optionally occupied by consonants which may give rise to clusters. The leftmost consonant is labelled ONSET, the consonant at the opposing extreme is called OFFSET. Consonants positioned between ONSET/OFFSET and the NUCLEUS constitute the SLOPE. In polysyllables, the syllable CONTACT involves those segments between which the boundary separating two consecutive syllables in the same segmental chain has to be drawn.

Borg/Azzopardi-Alexander (1997: 313) postulate the schema in Figure 2.13 as the canonical syllable type of Maltese. We have added the square brackets and indexes.

This canon is largely in line with the model proposed by Galea/Ussishkin (2018: 56) whose additions to the list of possible syllable structures do not affect our subsequent interpretation of the facts attested within the domain of IIs. A Maltese syllable thus consists minimally of a vocalically filled syllable NUCLEUS whereas both HEAD and CODA can optionally be filled by three or two consonants, respectively. According to Maddieson's (2005: 54) ternary typology, Maltese – though being absent from Maddieson's sample – belongs to the (second most frequently attested) class of languages with complex syllable structure. This further means that there are uncovered-open, covered-open, open-closed, and uncovered-closed syllables in the general lexicon of Maltese. Word-finally, only binary consonant clusters are

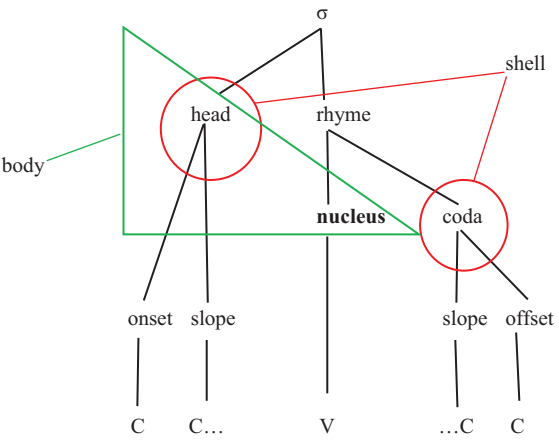


Figure 2.12: Internal structure of syllables according to Vennemann (1988).

$[[(C_{ONSET})\ (C_{SLOPE})\ (C_{SLOPE})]_{HEAD}\ [V_{NUCLEUS}\ [(C_{SLOPE})\ (C_{OFFSET})]_{CODA}]_{RHYME}]_{\sigma}$

Figure 2.13: Maltese canonical syllable structure.

legit whereas the maximum size of initial clusters is three with severe restrictions as to the onset consonant in ternary clusters (Borg/Azzopardi-Alexander 1997: 308–310). Purely vocalic monosyllables like the conjunction *u* 'and' are rare. Some of the other patterns occur more frequently. Examples of monosyllables realising the above patterns in the general lexicon are: VC *id* 'hand', VCC *art* 'ground', CV *la* NEGATION, CCV *fra* 'lay brother', CCCV *stma* 'esteem', CVC *dar* 'house', CVCC *borġ* 'heap', CCVC *flus* 'money', CCCVC *xprun* 'spur', CCVCC *spint* 'too uninhibited', and CCCVCC *żgwinċ* 'oblique' (cf. also Borg/Azzopardi-Alexander 1997: 306). Not all of these patterns are reflected in the Π-inventory.

From Figure 2.5, we know already that there are thirty monosyllables amongst the Maltese Πs. Figure 2.8 tells us that no monosyllabic Π starts with a vowel. This means that uncovered monosyllables are counted out generally so that V, VC, and VCC are impossible patterns for Maltese Πs. The tetrachoric table in Table 2.2 reveals the facts on the highest possible level of syllable structure. We will zoom in on the details below.

The overwhelmingly preferred pattern is that of the covered-closed syllable which accounts for 77% of all monosyllabic Πs. These monosyllables represent the set of Πs which are equipped with a syllable SHELL. According to Figure 2.11, the seven instances of monosyllables with empty CODA (= Πs which consist only of a syllable BODY) must refer to Semitic items. As to the complexity of the syllable margins, Figure 2.14 shows that HEAD and CODA have the same leanings towards the presence

Table 2.2: Head-coda combinations in monosyllables.

		coda		sum
		empty	**filled**	
head	empty	0	0	0
	filled	7	23	30
total		7	23	30

of a single consonant. Clusters have a slightly stronger position in the domain of CODAS than they have in the domain of the syllable HEAD.

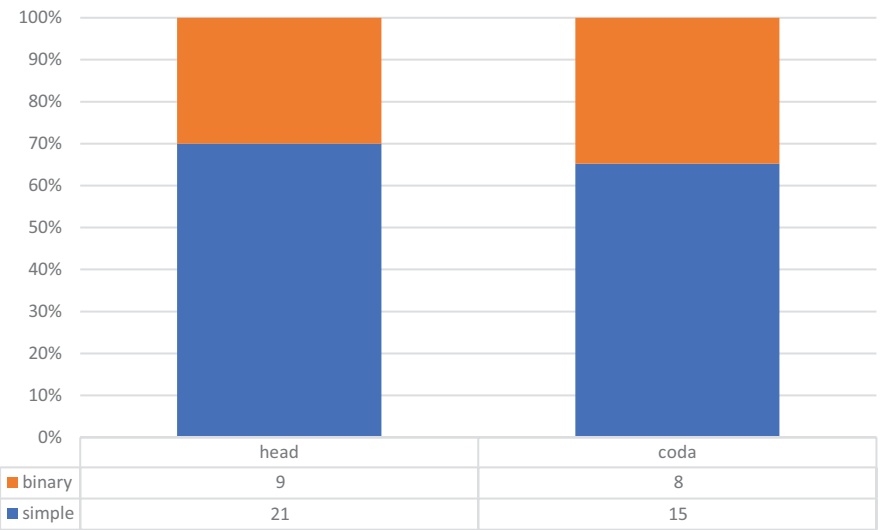


Figure 2.14: Simple vs binary heads and codas of monosyllables.

Of the thirty monosyllables which attest to a filled HEAD nine (= 30%) give evidence of an initial binary cluster, i.e. there is a SLOPE. As to the coda, we find eight (= 35%) instances of binary clusters of the SLOPE-OFFSET kind. Only two ΠS allow for clusters at both margins, namely the Italo-Romance borrowings (Π52) *sforz* 'thanks to' and (Π53) *skont* 'according to'. No Semitic Π tolerates two clusters in the syllable SHELL. In Figure 2.15 SHELLS of this kind are classified as heavy.

$[CC. . .CC]_{SHELL}$ ⊃ Italo-Romance origin

Figure 2.15: Implicational pattern of monosyllabic Πs heavy SHELLS.

We emphasise again the fact that the final cluster in the two Italo-Romance Πs is secondary because the original final vowel /o/ (raised to /u/) was lost in the borrowing process. The most common SHELL contains one consonant per margin. The combination C_{HEAD}-C_{CODA} covers nine different Πs. The SHELLS CC_{HEAD}-C_{CODA} and C_{HEAD}-CC_{CODA} with asymmetrically filled margins are attested six times each. No Π with a ternary cluster on either side of the NUCLEUS has been identified. Monosyllables with a syllable SHELL outnumber monosyllables which consist only of the syllable BODY by a ratio of three-to-one.

The NUCLEUS does not pose many problems. Two dozen monosyllabic Πs involve a simple nucleus consisting of a vowel (either short or long). The segmental chain of only six Πs contains a diphthong as nucleus, namely (Π4) *bejn* 'between', (Π14) *fejn* 'near', (Π20) *għajr* 'except', (Π23) *ħdejn* 'beside', (Π28) *lejn* 'towards', and (Π43) *mnejn* 'from near' all of which are Semitic. Note that only two of seven Maltese diphthongs are attested with monosyllabic Πs (Borg/Azzopardi-Alexander 1997: 299). The same diphthongs /ɛɪ/ and /ɐɪ/ are also attested with polysyllabic Semitic Πs in Maltese. Figure 2.16 reveals that there is another interesting pattern connected to the language of origin.

diphthong ⊃ Semitic origin

Figure 2.16: Implicational pattern of Πs with diphthong.

Turning our attention back to polysyllables we notice that there are sixteen Πs with heterosyllabic medial clusters (excluding ambisyllabic geminates) which emerge at the syllable contact. This means that the majority of 55% of all polysyllabic Πs displays word-internal consonant clusters. Seven (= 44%) of these sixteen cases have an Italo-Romance origin. The share of Italo-Romance in this class of polysyllables is higher than the share Italo-Romance Πs have of all polysyllables (cf. Figure 2.5). The Italo-Romance impact on the shape of Maltese Πs is even stronger if we differentiate the medial clusters according to the number of segments they involve. Only CC and CCC are attested. There are twelve binary clusters which cover 75% of all medial clusters. Of these dozen Πs, nine are Semitic and three Italo-Romance. As to the four ternary clusters, the Semitic share is nil since all four Πs are classified as Italo-Romance according to our criteria. These Πs are (Π25) *inkluż* 'including', (Π27) *kontra* 'against', (Π35) *minflok* 'instead of', and (Π44) *oltre* 'beyond' (cf. Figure 2.17).

CCC_{medial} ⊃ Italo-Romance origin

Figure 2.17: Implicational pattern of polysyllabic Πs with ternary medial clusters.

In all four cases, the syllable boundary runs between the single CODA consonant of the initial syllable and the binary HEAD of the following (= final) syllable (lower case $ marks the internal syllable boundary): *in$kluż, kon$tra, min$flok, ol$tre.* The CODA position of the initial syllable is filled by sonorants (nasal, lateral approximant) whereas the HEAD of the final syllables consists of an obstruent (plosive, fricative) as ONSET and a liquid (lateral approximant, rhotic) as SLOPE. These patterns raise the question of whether the Maltese Πs display any recurrent properties on the segmental level too.

2.2 Segments

In comparison to the suprasegmental level, the segments which are involved in the phonology of Maltese require a number of additional explanations before the analysis can be executed. To this end, we divide this section into two parts. We start with the individual phonemes in Section 2.2.1 and round off the description with Section 2.2.2 which takes account of the length of the segmental chains.

2.2.1 Individual phonemes

For a start, we look at the Maltese consonants in Section 2.2.1.1. Maltese vowels are discussed further below in Section 2.2.1.2. Since the approaches to Maltese phonology reflect different theories and models they tend to disagree upon the exact spell-out of the phonemics of Maltese. To circumvent the theory-related pitfalls of the debate, we adopt again the neutral position of phonological typology which is indebted largely to the ideas expressed by Gordon (2016: 43–122).

2.2.1.1 Consonants

Borg/Azzopardi-Alexander (1997: 299) postulate an inventory of phonemic consonants which comprises either 22 or 23 units. This oscillation between two options has to do with the uncertain status of the Italian loan phoneme /ʤ/. In Table 2.3 we present our version of the inventory of Maltese consonantal phonemes.

Table 2.3: Maltese consonantal phonemes.

	bilabial	labio-dental	denti-alveolar	post-alveolar	palatal	velar	glottal
plosive	**p** **b**		**t** **d**			**k** **g**	**ʔ**
nasal	**m**		**n**				
trill			**r**				
fricative		**f** **v**	**s** **z**	**ʃ**		*	**h**
affricate			**ʦ**	**ʧ** **ʤ**			
approximant	**w**				**j**		
lateral approximant			**l**				

This inventory counts 23 units although we have excluded all recent loan phonemes of which there are more than the above descriptive grammar assumes (Stolz/Levkovych 2021: 583).[9] Apart from the fact that no loan phoneme is relevant for the phonology of Maltese ΠS, Table 2.3 raises another and more important issue. There is the grey shaded cell for the velar fricative (of unspecified phonation). In lieu of the expected IPA-symbols /x/ or /ɣ/ the asterisk * is used. Why is that?

Alexander Borg (1997: 249) claims that the Maltese "sound system has an 'abstract phoneme' represented by the orthographical symbol *għ* (called [a:yn]) in the M[altese] alphabet) and representing the reflex of the [Old Arabic]fricatives */ʕ/ and */ɣ/". This abstract phoneme has the effect of lengthening adjacent low (= /ɐ/ → /ɐ:/) or mid-high vowels (= /ɛ/ > /ɛ:/, /ɔ/ →/ɔ:/) whereas it diphthongises the high vowels /ɪ/ and /ʊ/ to /ɛɪ/ ~ /ɐɪ/ and /ɐʊ/ ~ /ɔʊ/, respectively. Hume et al. (2009: 43) conclude that the lengthening effect of the *għajn* is restricted to certain environments under certain prosodic conditions. Cassola (2014: 16) assumes that at least the voiced velar fricatives was still "alive and kicking" in the 16th century. A full account of the diachrony and diatopic variation of the *għajn* and sundry lost phonemes is given in Puech (2018). According to Alexander Borg (1997: 261), the voiced velar fricative realisation [ɣ] has survived in "[a] few conservative rural communities" on Gozo. What is more, the phoneme under scrutiny is realised as voiceless velar fricative [x] or voiceless pharyngeal fricative [ħ] under certain morphonological conditions (Alexander Borg 1997: 276) some of which are also relevant for Maltese ΠS.[10] For the time being, we follow Cohen (1970: 129–131) and Vanhove

9 Alexander Borg (1997: 249) admits not only /ʤ/ but also the equally marginal loan phoneme /ʒ/ (borrowed from English) on the list of phonemes of Maltese.

10 These cases will be discussed in Section 3.

(1993: 9) who treat the *għajn* as fully-blown consonant even in those cases where it has no segmental status of its own. We consider this case to be different from the supposed zero-Π rebutted above[11] because the *għajn* produces identifiable effects on neighbouring segments whereas the absence of the Π has no such morpho-syntactic repercussions.

Not only does our treatment of the *għajn* explain why there are no vowel-initial Πs of Semitic origin in BLOMP 2.0 but this abstract consonant is also responsible for the relative homogeneity of the syllable structures. Moreover, it has to be counted on a par with other segments when we try to determine to what extent (if at all) the differences in segmental length of the Maltese Πs reflect systematic patterns. The role played by *għajn* in this context will come to the fore subsequently.[12]

Every consonant featured in Table 2.3 is attested in word-initial position in the general lexicon of Maltese (Borg/Azzopardi-Alexander 1997: 308). The situation is different in the domain of Πs. In Table 2.4, we mark those consonants which occupy the leftmost slot of the segmental chains of Maltese Πs. The colour code is as follows: pink identifies consonants which are admissible in this position exclusively with Semitic Πs whereas yellow highlights the opposite case, i.e. those consonants which function as word-initial segments only in Italo-Romance Πs. Green is reserved for consonants which are attested word-initially with Semitic as well as Italo-Romance Πs. Unmarked cells host consonants which are never attested as onset or head of any Π.

Table 2.4: Consonants in word-initial position.

	bilabial	labio-dental	denti-alveolar	post-alveolar	palatal	velar	glottal
plosive	p b		t d			k g	ʔ
nasal	m		n				
trill			r				
fricative		f v	s z	ʃ		*	h
affricate			ʦ	ʧ ʤ			
approximant	w				j		
lateral approximant			l				

With sixteen consonants being attested on the left margin, Maltese Πs make use of only 70% of the options offered by the inventory of phonemic consonants. If

11 Cf. Section 1.6.3.

12 In Sections 2.2.2 and 2.3.

we subtract the three consonants which can occupy the leftmost slot only in Italo-Romance ΠΙs, the share goes down to 57%. We interpret this as a clear sign of the special behaviour of ΠΙs in comparison to the general lexicon of the object language.

The individual consonants which are eligible for the word-initial position do not distribute evenly over the 56 consonant-initial Maltese ΠΙs. Figure 2.18 reveals that the three most frequently attested initial consonants share the place feature [labial]. Word-initial consonants with the same frequency share a bar.

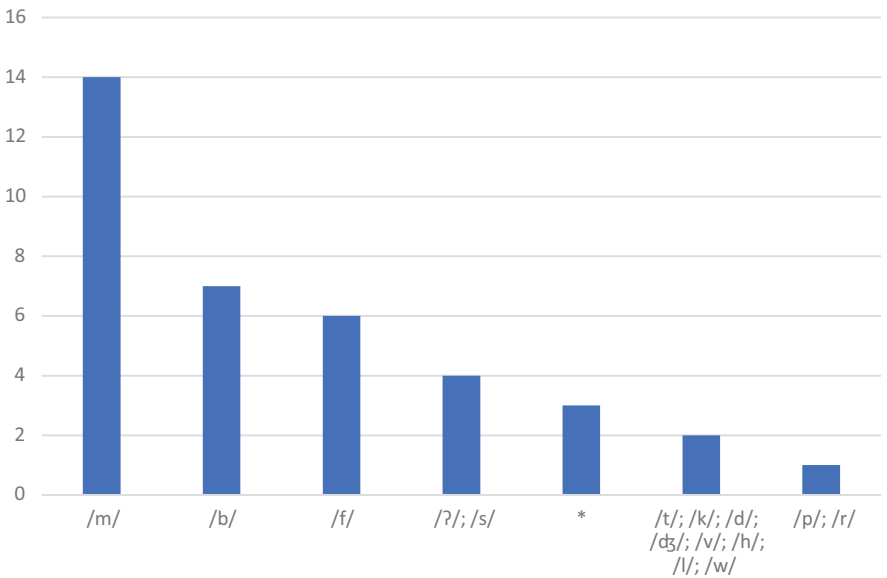


Figure 2.18: Frequency of word-initial consonants.

The bilabial nasal /m/, the voiced bilabial plosive /b/, and the voiceless labiodental fricative /f/ account for 27 of the 56 cases under scrutiny. If we add the results for the other consonants which display the same place feature – the voiced labio-dental fricative /v/, the bilabial approximant /w/,[13] and the voiceless bilabial plosive /p/ – the

13 The exact classification of Maltese /w/ is difficult. The phoneme chart provided by Alexander Borg (1997: 249) does not identify places and manners of articulation terminologically. Its organisation is however suggestive of the symbol /w/ being used to represent the bilabial approximant. The author characterises each of the sound classes except that of the approximants. With reference to /w/, Borg/Azzopardi-Alexander (1997: 303) speak of a voiced labio-velar approximant (for which the IPA has the symbol /ʋ/, however). It is likely that /w/ represents the voiced labial-velar approximant. Whatever the correct phonological analysis might be our argument remains unaffected because the feature [labial] is present either way.

share of labial consonants in word-initial position rises to 57% (with 32 out of 56 Πs). No other place feature can compete with [labial] because [denti-alveolar] with eleven cases, [glottal] with six cases, [velar] with five cases, and [postalveolar] with only two cases can lay claim only to considerably smaller shares. These differences are visible in Figure 2.19. The absence of the palatal place of articulation is worth noting.

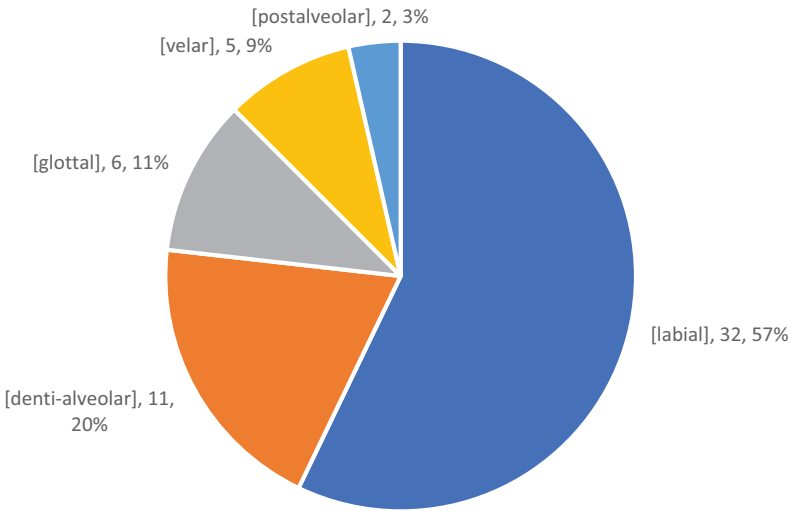


Figure 2.19: Shares of place features with initial consonants.

It is true that in the general lexicon of Maltese too, the distribution of initial consonants over the places of articulation is not egalitarian. However, if we impressionistically judge the relations by way of counting the pages in Aquilina's (1991) dictionary which feature /m/-initial entries (pp. 765–879 = 114 pages) and those containing /s/-initial entries (pp. 1244–1380 = 136 pages) it becomes doubtful that the phonology of the Maltese Πs simply replicates the phonology of the general lexicon of the same language.

Twelve of the fourteen /m/-initial Πs are lexicalised complex Πs with *ma-* or *minn-* as left constituent. This means that the synchronically biggest group of Πs is secondary in the sense that most of its members have arisen in the course of the univerbation of erstwhile sequences of different Πs or other syntagms. As a by-product of this development, /m/-initial Πs are preferably polysyllables and, except (Π35) *minflok* 'instead of', also of Semitic origin. Similarly, many /b/-initial Πs (with four out of seven cases) and to a much lesser extent also those with initial /f/ (one isolated case) result from univerbation processes. The preference for labial onsets/heads therefore is not so much phonologically motivated but the effect of

diachronic processes in the morphosyntactic domain. Italo-Romance contributes five ΠS with an initial labial to the inventory. Borrowing is also responsible for the participation of the voiceless bilabial plosive /p/ and the voiced labiodental fricative /v/ in the dominant pattern.

In the general lexicon of Maltese, the voiced obstruents /b/, /d/, /g/, /z/, /v/, and /ʤ/ are barred from occurring in word-final position – this is the opinion of Borg/Azzopardi-Alexander (1997: 307). One might take issue with this interpretation because the above ban on the voiced obstruents can also be described in terms of obligatory final devoicing (mentioned by the same authors on the same page). If vowel-initial affixes are added to words whose segmental chain ends phonetically in a voiceless obstruent, the neutralised voice distinction is activated again as e.g. in minimal pairs like (noun) SG *salt* 'jump → PL *salti* vs (adjective) SG *sald* 'sound' → PL *saldi* with both singulars being phonetically indistinguishable, namely [sɛlt]. We assume a classical neutralisation analysis according to which the voiced obstruents are represented in word-final position by their voiceless positional allophones (Hall 2011: 97–99). Accordingly, pace Borg/Azzopardi-Alexander (1997: 307), we claim that, except *għajn*, all consonantal phonemes of Maltese are admissible to the word-final position with the proviso that /j/ and /w/ occur word-finally only as closing part of diphthongs.[14]

For the representation of the word-final consonants in Table 2.5 we employ the same colour code as before. With thirteen out of the 22 word-final consonants of the general lexicon, Maltese ΠS select only 59% for this position in their inventory. Half a dozen of these phonemes are attested only in Semitic ΠS. Four consonants occur both in Semitic and Italo-Romance ΠS whereas three consonants are restricted to Italo-Romance ΠS. There is no evidence of a Π hosting an approximant in word-final position. Likewise, the postalveolar and palatal places of articulation are unattested in this position. Were it not for the isolated case of (Π35) *minflok* 'instead of', the entire region between the denti-alveolar place of articulation and the glottal place of articulation would not be involved in word-final consonants of Maltese ΠS. Italo-Romance influence is responsible for disallowing this generalisation.

There are 37 ΠS which end in a consonant. These ΠS cover 63% of the Π-inventory. Like in the case of the word-initial consonants, those which qualify as offset or coda do not distribute equally over the ΠS as shown in Figure 2.20. Consonants with identical frequencies share a bar.

14 Cf. Section 2.2.1.2.

Table 2.5: Consonants in word-final position.

	bilabial	labio-dental	denti-alveolar	post-alveolar	palatal	velar	glottal
plosive	p b		t d			k g	ʔ
nasal	m		n				
trill			r				
fricative		f v	s z	ʃ			h
affricate			ʦ	ʧ ʤ			
approximant	w				j		
lateral approximant			l				

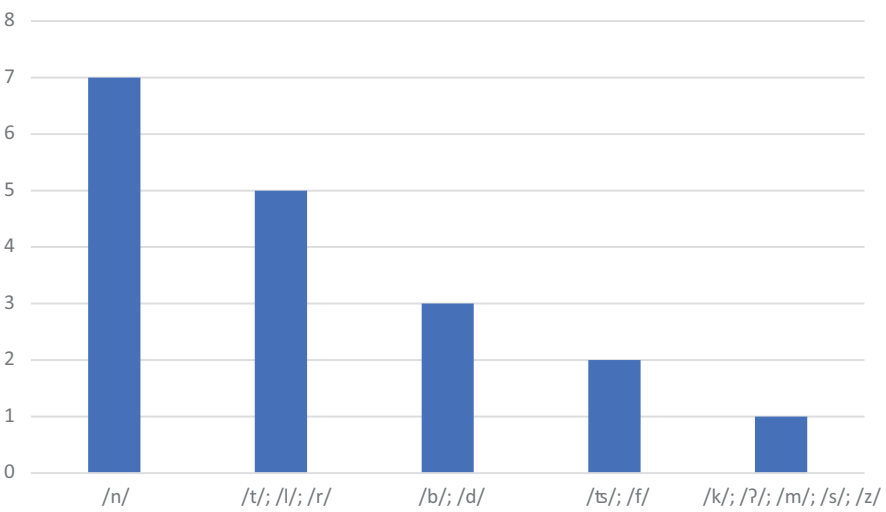


Figure 2.20: Frequency of word-final consonants.

The highest rank goes to the denti-alveolar nasal /n/ which is attested on the right margin of 19% of all IIs which end in a consonant. This is remarkable insofar as this phoneme is not among those consonants which are accepted in the word-initial position. The distribution of the denti-alveolar nasal is thus restricted. It is almost needless to say that, in the general lexicon of Maltese, /n/ is relatively common as onset or head of lexemes. In Aquilina's (1991: 879–929) dictionary, the representation of /n/-initial words requires fifty printed pages. Moreover, the four most frequently encountered word-final phonemes are all produced at the

denti-alveolar place of articulation, namely the nasal /n/, the voiceless plosive /t/, the lateral approximant /l/, and the trill /r/. In word-initial position, none of these four consonants is attested with more than two ΠS. In contrast, their type frequency is 5–7 word-finally. Together they account for 22 ΠS which is equivalent of 59% of all ΠS with a final consonant. If we add up the frequencies of all word-final consonants which share the place feature [denti-alveolar] the number of affected ΠS increases by seven. With 29 cases, the place feature [denti-alveolar] is present in 78% of all ΠS ending in a consonant. Figure 2.21 visualises the outstanding position of this privileged place feature.

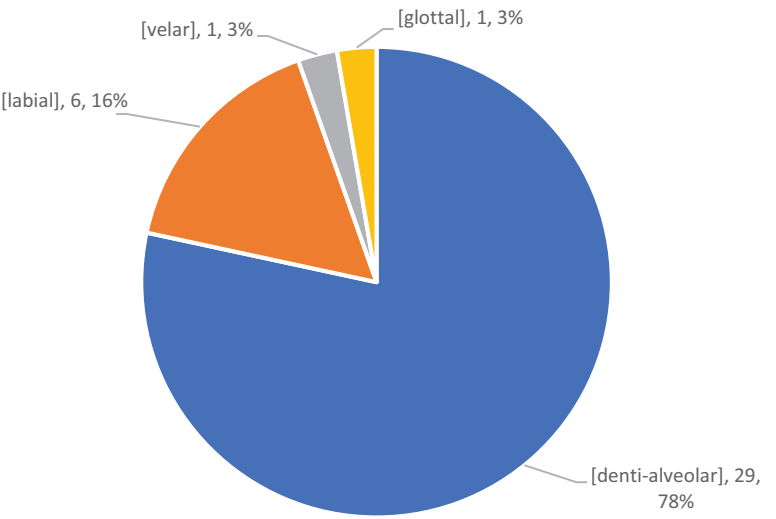


Figure 2.21: Shares of place features with final consonants.

The comparison of Figures 2.19 and 2.21 clearly shows that the opposing edges of the segmental chains of words differ as to the preferences they have in the domain of the places of articulation. It strikes the eye that the nasal manner of articulation ousts its competitors both on the left margin and on the right margin. ΠS with a final labial are Semitic. Both Semitic and Italo-Romance contribute to the strong position of the denti-alveolar place of articulation.

Besides the word-initial and word-final consonants discussed in the previous paragraphs, there are also those consonants which occur in the SLOPE either to the right or the left of the NUCLEUS. If there is a SLOPE, there is also a consonant cluster. Borg/Azzopardi-Alexander (1997: 308–310) report on the restrictions to which clusters in the syllable HEAD and syllable CODA are subject. If we understand the authors correctly they assume that the voice harmony rules do not permit that phonemes

with different phonation can be direct neighbours of each other in the segmental chain. There cannot be any doubt as to the existence of the voice harmony rule in Maltese. However, in contrast to our source, we conceive of voice harmony in terms of regressive assimilation which affects phonemes in such a way that they are represented by different allophones according to the phonological context. As in the above case of final devoicing, the otherwise phonemic phonation values are neutralised under assimilation. Assimilation is often triggered morphonologically as e.g. *kiteb* [kɪtɛp] 'he wrote' → *jiktbu* [jɪgdbʊ] 'they write'. For the purposes of this study, we distinguish between the abstract phonological and the concrete phonetical level. Only the former is of interest to us in the present context.

In the word-initial HEAD of Maltese ΠS, the following nine types of binary clusters are attested (only the first is attested twice): /b/ + /h/, /b/ + /l/, /h/ + /d/, /h/ + /l/, /**s**/ + /**f**/, /**s**/ + /**k**/, /ʔ/ + /r/, /d/ + /w/, and /m/ + /n/ all of which are also attested more or less frequently in the general lexicon. The sibilant-initial clusters are Italo-Romance (marked out in boldface). The cluster /s/ + /k/ is also the only example which involves a voiceless plosive in the SLOPE. In the other cases, the SLOPE contains segments with a higher degree of sonority (Vennemann 1988: 9), namely the voiced plosive /d/, the voiceless fricatives /f/ and /h/, the liquids /l/ and /r/, or the approximant /w/. These SLOPE consonants constitute only a small part of the consonants which are allowed in this position in the general lexicon where voiceless plosives for instance are frequent in SLOPES (Borg/Azzopardi-Alexander 1997: 309). In the word-final CODA, Maltese ΠS attest to ten different binary clusters (only the first is attested twice): /n/ + /d/, /n/ + /b/, /**n**/ + /**t**/, /**r**/ + /**d**/, /**r**/ + /**ʦ**/, /ʔ/ + /t/, /ʔ/ + /s/, /l/ + /b/, /s/ + /t/, and /h/ + /t/. Boldface identifies those clusters which are exclusive to the Italo-Romance component of the inventory. The remaining seven patterns are Semitic. These clusters fit the description of word-final consonant clusters given in Borg/Azzopardi-Alexander (1997: 310) except that the Maltese grammarians consider final voiced obstruents to be representatives of the corresponding voiceless phonemes. What we learn from the descriptive grammar is that the number of different phonotactic combinations on the right margin of Maltese words is much higher than the small number of clusters attested in the Π-inventory.

Borg/Azzopardi-Alexander (1997: 311) state that "[a]ll the consonant clusters permissible as word-final clusters are permissible word-medially." We assume that this quote refers to binary clusters. The word-medial clusters mentioned in our source are heterosyllabic sequences. The ΠS give evidence of eleven different types of word-medial clusters (only the first is attested three times), namely /n/ + *, /n/ + /t/, /n/ + /k/, /n/ + /h/, /n/ + /w/, /d/ + /w/, /s/ + /w/, /**s**/ + /**k**/, /**g**/ + /**w**/, /**r**/ + /**m**/, and /**r**/ + /**s**/. Boldface is used to identify the Italo-Romance cases. The denti-alveolar nasal /n/ occurs in combination with five different consonants, the approximant /w/ is

attested in four different clusters, the trill /r/ is the first consonant in two different clusters. Except /n/ + /t/ none of these clusters is also attested in the word-final position of Πs. Figure 2.17 has revealed already that in polysyllables ternary clusters occur only in Italo-Romance borrowings. There are four different CCC-clusters, namely /n/ + /f/ + /l/, /n/ + /k/ + /l/, /n/ + /t/ + /r/, and /l/ + /t/ + /r/. The syllable boundary separates the first consonant from the following two consonants so that the former functions as CODA of the first syllable and the other two consonants form the HEAD of the second syllable. None of these HEADs is attested as word-initial cluster in the Π-inventory.

These findings are hardly surprising since the limited size of BLOMP 2.0 does not provide many opportunities for clusters to emerge. Nevertheless, it is interesting to see that the denti-alveolar nasal /n/ is generally strong in post-nuclear positions whereas it is only marginally attested as SLOPE consonant in pre-nuclear position. The bilabial nasal /m/ is almost the mirror image of its fellow nasal in the sense that the stronghold of /m/ is on the pre-nuclear side as opposed to its negligible status on the right of the nucleus. At the extremes of the segmental chain, only the place features [labial] and [denti-alveolar] display a certain prominence. The denti-alveolar nasal also stands out in word-medial clusters. If the trill forms part of a cluster word-medially or word-finally, the Π is Italo-Romance (cf. Figure 2.22).

$(C)/r/(C)_{\text{medial/final}}$ ⊃ Italo-Romance origin

Figure 2.22: Implicational pattern of Πs with clusters involving /r/.

Both Maltese and Italian have geminate consonants. It is therefore unsurprising to find geminates also in the Maltese Π-inventory. The Maltese Πs of this kind are polysyllabic so that another unilateral implication can be put forward (cf. Figure 2.23).

geminate ⊃ polysyllabic

Figure 2.23: Implicational pattern of Πs with geminates.

Except for (Π45) *permezz* 'by means of' the geminate is always ambisyllabic across syllable boundaries. Of the ten Πs whose segmental chain involves geminates four are of Italo-Romance, namely (Π1) *apparti* 'apart from', (Π12) *faċċata* 'opposite', (Π45) *permezz* 'by means of', and (Π54) *sotta* 'under'. The six Semitic Πs are (Π3) *barra* 'outside', (Π18) *ġewwa* 'inside', (Π34) *minbarra* 'except', (Π39) *minħabba* 'on account of', (Π40) *minkejja* 'in spite of', and (Π49) *quddiem* 'in front of'. Πs with geminates account for 34% of all polysyllables in BLOMP 2.0.

2.2.1.2 Vowels

In Table 2.6, we present the vowel phonemes (monophthongs) of Maltese according to Borg/Azzopardi-Alexander (1997: 299).

Table 2.6: Maltese vowel phonemes (monophthongs).

	front		near-front		central		near-back	back	
close	i:							u:	
near-close			ɪ	ɪ:			ʊ		
open-mid	ɛ	ɛ:						ɔ	ɔ:
near-open					ɐ	ɐ:			

Except for the close high vowels and the near-close near-back vowel, all vowel qualities come in short and long versions. Quantity distinctions are phonemic. In addition to the above monophthongs, Maltese has a system of seven diphthongs which Borg/Azzopardi-Alexander (1997: 299) classify as phonetic: [ɐʊ], [ɐɪ], [ɛʊ], [ɛɪ], [ɪʊ], [ɔɪ], and [ɔʊ]. On the phonological level, they assume that the closing component of the diphthong allophonically represents the palatal or labial-velar approximant. Thus, all diphthongs with final [ɪ] are understood as sequences of V + /j/ while those ending in [ʊ] count as sequences of V + /w/. This decomposition of the phonetic diphthongs into a vocalic and a consonantal component produces a higher number of complex CODAs because in the Π-inventory a diphthong is always followed by a tautosyllabic sonorant. In contrast to Borg/Azzopardi-Alexander (1997: 299) but in accordance with Alexander Borg (1997: 270–271), we assume complex nuclei for these cases, i.e. the diphthongs are not only a phonetic phenomenon but also constitute a phonological unit.

There are seven Πs whose segmental chain contains a diphthong. This is equivalent to only 12% of BLOMP 2.0. Moreover, only the diphthongs /ɐɪ/ and /ɛɪ/ are attested. Diphthongs which involve (near-)back vowels are alien to the Π-inventory. The seven Πs which give evidence of diphthongs are Semitic without exception. On this basis, it is possible to formulate the complex implication in Figure 2.24.

$$\text{Diphthong} \supset \left\{ \begin{array}{c} /\text{ɐɪ}/ \\ \\ /\text{ɛɪ}/ \end{array} \right\} \supset \_\_\ \text{C}_{[\text{SONORANT}]} \supset \text{Semitic}$$

Figure 2.24: Implicational pattern of Πs with diphthongs.

As to the monophthongs, it turns out that nine of the eleven phonemes featured in Table 2.6 also occur in the Πs-inventory. As Table 2.7 reveals only the long quantities of the open-mid vowels (marked by grey shading) do not turn up in BLOMP 2.0.

Table 2.7: Attested vs unattested vowels (monophthongs).

	front		near-front		central		near-back	back	
close	iː							uː	
near-close			ɪ	ɪː			ʊ		
open-mid	ɛ	ɛː						ɔ	ɔː
near-open					ɐ	ɐː			

The segmental chains of Maltese Πs yield a total of ninety monophthongs 55 of which bear stress whereas 33 occur in atonic positions. We simplify matters by way of lumping together several categories of Table 2.7 to determine the impact of the language of origin on the frequency of the attested vocalic monophthongs in Figure 2.25.

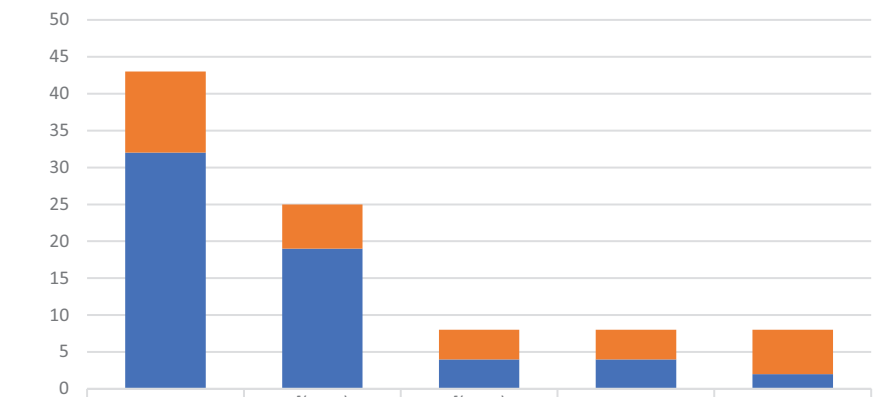


Figure 2.25: Vowel qualities (monophthongs) and language of origin.

The Semitic Πs overwhelmingly dictate the distribution of /ɐ(ː)/, /ɪ(ː)/, and /iː/. On the understanding that polysyllables account for several vowels at one go, there are 51

cases which involve Semitic Πs whereas only seventeen involve Italo-Romance Πs. Semitic outnumbers Italo-Romance by a ratio of 3-to-1. In contrast, the distribution of the less frequently attested vowels /uː/, /ʊ/, /ɔ/, and /ɛ/ is marked by an increase of Italo-Romance Πs. The latter account for thirteen cases as opposed to only ten cases which are found in Semitic Πs. Thus, the ratio of 1.3-to-1 is to the advantage of the Italo-Romance Πs.

In the foregoing sections, we have learned that the segmental phonemes of Maltese do not behave as one when it comes to forming Πs. Some phonemes are more often involved than others so that preferences emerge which are not necessarily the same as in the general Maltese lexicon. The differences which characterise the individual phonemes qualitatively tell only part of the story of the phonology of Maltese Πs. An important episode in the story remains to be told, namely the size of the segmental chains.

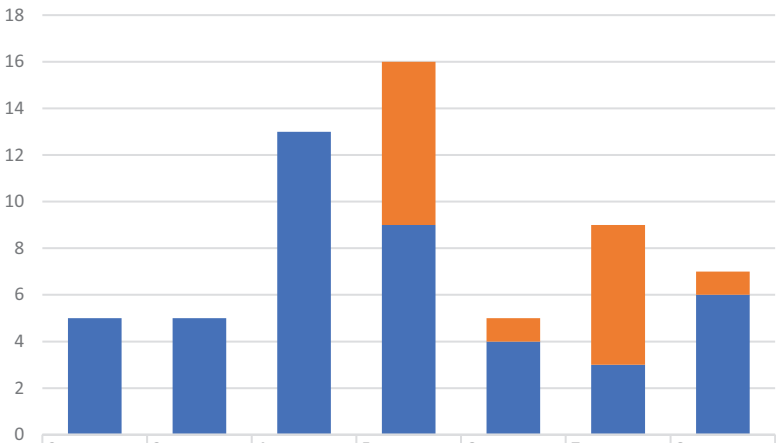
2.2.2 Segmental chains

With a view to determining the size of the segmental chains, we have to stipulate certain basics whose in-depth discussion would lead us too far away from the main topic of this study. We measure the size of a segmental chain by way of adding up the slots which are – virtually or phonetically – filled by a phonological unit. To keep the procedure as simple as possible, we adopt a very crude arithmetic according to which each slot-filler pair has the value '1'. Geminates, diphthongs, and long vowels spread over two consecutive slots so that they are counted as 2 x 1 = 2. The abstract phoneme *għajn* is assumed to fill a slot too.[15] Thus, the Italo-Romance Π (Π45) *permezz* 'by mean of' for instance, is registered as comprising seven segments.

The size of the segmental chains ranges from minimally two segments to a maximum of eight. If we correlate the different sizes and their type frequency with the languages of origin an interesting picture emerges as shown in Figure 2.26.

The 23 Πs with less than five segments are the monopoly of Semitic. Beyond this limit, the shares of Italo-Romance (15 Πs) and Semitic (22 Πs) are relatively close to each other with 48% and 52%, respectively. Accordingly, the implicational pattern in Figure 2.27 can be postulated.

15 Cf. Section 2.2.1.1.



	2 segments	3 segments	4 segments	5 segments	6 segments	7 segments	8 segments
Italo-Romance				7	1	6	1
Semitic	5	5	13	9	4	3	6

Figure 2.26: Length of segmental chains and language of origin.

Italo-Romance ⊃ n > 4 segments

Figure 2.27: Implicational pattern of Italo-Romance Πs with n = 5–8 segments.

The longest segmental chains – i.e. those which comprise more than five segments – yield polysyllables. Since some polysyllables count less than six segments, it is possible again to put forward a unilateral implication shown in Figure 2.28.

n > 5 segments ⊃ polysyllabic

Figure 2.28: Implicational pattern of Πs with n = 6–8 segments and syllabicity.

Furthermore, Πs whose phonological form involves geminates are never shorter than five segments (cf. Figure 2.29).

geminate ⊃ n > 4 segments

Figure 2.29: Implicational pattern of Πs with geminates and length.

The average number of segments per Π is 5.1. With 21 Πs about 35% of the members of BLOMP 2.0 reach or exceed this mark whereas the majority of 39 Πs (= 65%) remains below the average. In terms of types, it is therefore more normal for a

Maltese Π to involve less than six segments. In the subsequent section, we investigate whether this preference for shorter segmental chains is also reflected in the domain of token frequency.

2.3 Frequency and length

Table 1.11 in Part A informs about the token frequency of the members of BLOMP 2.0. in the Korpus Malti 3.0. We repeat that it was necessary to apply double book keeping to the items under scrutiny because many of them are commonly used also in functions other than those of proper Πs. In this section, however, it is possible to return to the undifferentiated general token frequency. This is possible because we want to determine to what extent (if at all) token frequency and length are correlated to each other. Given that many expressions are members of several word classes, we have to abstract away from the focus on Π-uses. The correlatability of token frequency and length is independent of the word-class membership of the elements. Thus, in what follows, the frequencies given refer to the overall occurrences of the expressions and not to their occurrences as Πs alone.

Schmidt/Vorholt/Witt (2020: 245–247) distinguish ten frequency classes in their study of the token frequency of the members of the original version of BLOMP in the Korpus Malti 3.0. Vorholt (2022) looks into the relationship between token frequency and length of Πs for a sample of languages to which Maltese belongs.[16] We take up the quantitative issues raised in these previous publications by way of comparing the token frequency of the members of BLOMP 2.0 with the syllabicity and number of segments of the same Πs. In Table 2.8, the Πs are ranked top-down according to their decreasing frequency. The two columns on the right disclose how many syllables and how many segments form part of the word form which represents a given Π in the lexicon. Like in the case of Table 1.11, only full syntactic words are taken account of. Reduced and cliticised allomorphs are excluded from the count. Grey shading highlights the first dozen ranks of the hierarchy which host those Πs which exceed the average of 262,900 tokens as calculated for column B in Table 1.11. The rows marked out in yellow feature Italo-Romance Πs.

16 Cf. Section 1.5.9.

Table 2.8: Token frequencies of Πs in the Korpus Malti 3.0, syllabicity, and number of segments.

rank	Π	tokens	syllables	segments
1	(Π55) *ta'* 'of'	5,096,341	1	2
2	(Π41) *minn* 'from'	1,256,789	1	4
3	(Π17) *fuq* 'on'	1,151,992	1	3
4	(Π21) *għal* 'for'	1,036,925	1	4
5	(Π26) *kif* 'as'	787,519	1	3
6	(Π11) *dwar* 'about'	651,499	1	4
7	(Π14) *fejn* 'near'	489,321	1	4
8	(Π60) *wara* 'after'	444,289	2	4
9	(Π29) *lil* 'to'	404,833	1	3
10	(Π30) *ma'* 'with'	355,717	1	2
11	(Π6) *bħala* 'as'	328,932	2	6
12	(Π15) *fi* 'in'	319,840	1	2
13	(Π4) *bejn* 'between'	236,899	1	4
14	(Π46) *qabel* 'before'	229,852	2	5
15	(Π56) *taħt* 'under'	203,834	1	4
16	(Π28) *lejn* 'towards'	171,669	1	4
17	(Π59) *waqt* 'at the time of'	170,317	1	4
18	(Π27) *kontra* 'against'	167,406	2	6
19	(Π39) *minħabba* 'on account of'	159,460	3	8
20	(Π5) *bħal* 'like'	157,848	1	5
21	(Π49) *quddiem* 'in front of'	154,607	2	6
22	(Π33) *matul* 'during'	149,416	2	6
23	(Π7) *bi* 'with'	141,102	1	2
24	(Π3) *barra* 'outside'	136,814	2	5
25	(Π51) *sa* 'till'	130,402	1	2
26	(Π45) *permezz* 'by means of'	115,212	2	7
27	(Π36) *mingħajr* 'without'	107,066	2	8
28	(Π50) *rigward* 'concerning'	98,661	2	7
29	(Π31) *madwar* 'around'	95,815	2	6
30	(Π16) *fost* 'amongst'	92,813	1	4
31	(Π40) *minkejja* 'in spite of'	82,667	3	8
32	(Π9) *bla* 'without'	80,061	1	3
33	(Π53) *skont* 'according to'	70,443	1	5
34	(Π13) *favur* 'in favour of'	60,542	2	5
35	(Π35) *minflok* 'instead of'	45,642	2	8
36	(Π18) *ġewwa* 'inside'	45,315	2	5
37	(Π22) *għand* 'at s.o.'s place'	35,070	1	5
38	(Π38) *mingħand* 'from s.o.'	32,167	2	8
39	(Π25) *inkluż* 'including'	32,143	2	7
40	(Π34) *minbarra* 'except'	30,050	3	8
41	(Π48) *qrib* 'near'	29,384	1	4
42	(Π1) *apparti* 'apart from'	28,630	3	7
43	(Π10) *daqs* 'equal to'	28,592	1	4

Table 2.8 (continued)

rank	Π	tokens	syllables	segments
44	(Π24) *ħlief* 'except'	25,346	1	5
45	(Π47) *qalb* 'amidst'	18,935	1	4
46	(Π19) *ġo* 'in'	17,759	1	3
47	(Π58) *viċin* 'near'	16,751	2	5
48	(Π23) *ħdejn* 'beside'	16,489	1	5
49	(Π20) *għajr* 'except'	8,489	1	5
50	(Π52) *sforz* 'thanks to'	7,730	1	5
51	(Π12) *faċċata* 'opposite'	7,415	3	7
52	(Π44) *oltre* 'beyond'	2,984	2	5
53	(Π8) *biswit* 'facing'	2,958	2	7
54	(Π43) *mnejn* 'from near'	1,949	1	5
55	(Π32) *maġenb* 'close to'	1,720	2	6
56	(Π2) *a skapitu* 'at the expense of'	1,367	4	8
57	(Π57) *versu* 'towards'	228	2	5
58	(Π37) *mingħala* 'in s.o.'s opinion'	1	3	8
59	(Π42) *mintul* 'all along'	0	2	7
60	(Π54) *sotta* 'under	0	2	5

Superficially, the quantitative data in Table 2.8 fail to yield a clear picture. Monosyllabic Πs for instance, can be found in the upper half as well as in the lower half of the table. Phonological chains which comprise five segments share the same fate, in a manner of speaking. At the same time, we notice that of the twelve Πs which are attested more frequently than the average only one – (Π6) *bħala* 'as' on rank #11 – exceeds the average of 5.1 computed for segments. The same Π and (Π60) *wara* 'after' on rank #8 are the only examples of disyllables in the top section of the hierarchy where monosyllables are the rule.

The bird's eyes view shows that by and large the Maltese Πs confirm Zipf's Law in the sense that high token frequency goes hand in hand with relative shortness of an expression (Zipf 1935: 38–39) whereas particularly long segmental chains are found more often with low-frequency Πs. Figure 2.30 reveals that we are dealing with (strong) tendencies rather than with strict rules. The x-axis measures the token frequency whereas the y-axis determines the length of the expressions in terms of syllables (blue dots) and segments (orange dots). As to syllabicity, the blue dots remain close to the value 1 from the highest frequency down to token frequencies of half a million or less. Below this threshold, syllabicity starts to oscillate first between 1 and 2 and at the bottom of the hierarchy between 1 and 3. This means that polysyllables are by far more common with low-frequency Πs than with high-frequency Πs. Monosyllables are typical of the high-frequency section. Similarly, the orange dots

first rise from the value 2 to that of 4 with the top-ranking ΠΙs. With less frequent ΠΙs, the number of segments is bigger than 2 in most of the cases. At the bottom of the hierarchy, particularly long segmental chains dominate the mean values.

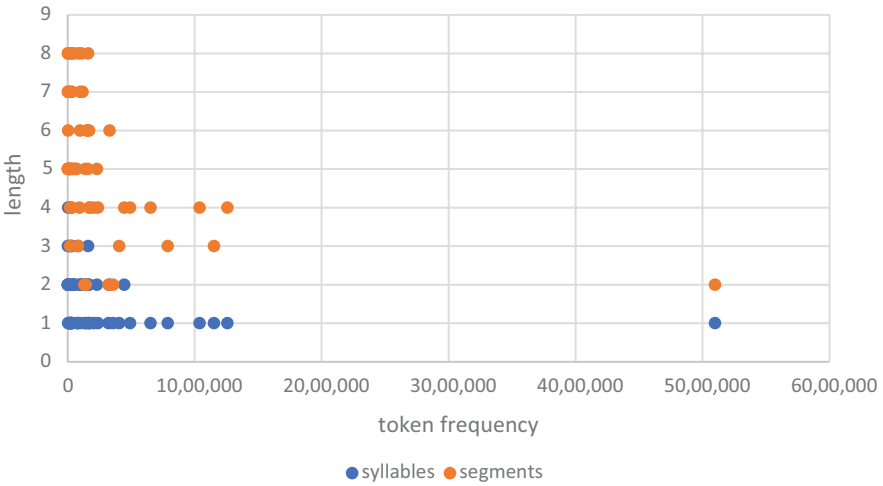


Figure 2.30: Frequency correlated to length (syllabicity and segments).

The Maltese ΠΙs as a class are thus well-behaved because they meet the expectations of Zipf's Law.

Schmidt/Vorholt/Witt (2020: 251) state that "the four highest frequency classes [. . .] are only represented by Semitic prepositions". Our study replicates this finding since the highest-ranking Italo-Romance Π is (ΠΙ27) *kontra* 'against' on rank #18. Except for this Π and (ΠΙ45) *permezz* 'by means of' on rank #26, none of the Italo-Romance ΠΙs has a token frequency of more than 100,000. The vast majority of the Italo-Romance ΠΙs (twelve out of fifteen) have token frequencies below 80,000. There are only fifteen Semitic ΠΙs in this low-frequency section of the hierarchy, i.e. it is typical for Italo-Romance ΠΙs to be used relatively infrequently. This observation also holds for the two Italo-Romance monosyllables (ΠΙ53) *skont* 'according to' on rank #33 and (ΠΙ52) *sforz* 'thanks to' on rank #50. This means that monosyllabicity alone is not enough to promote a Π to the status of high-frequency Π. It is the other way around: high frequency increases the probability that a given Π is monosyllabic.

2.4 The canonical Π – phonology

In this section, we use the insights we have gained in the previous sections on phonological issues to complement the characterization we have given of the canonical Maltese Π in Section 1.6.1. The suprasegmental and segmental properties of Maltese Πs vary. Nevertheless, it is possible to formulate a number of generalisations which capture the most typical aspects of the members of BLOMP 2.0. The following preferences circumscribe some of the properties which are associated with the canonical Π:

1. Semitic > Italo-Romance
2. monosyllabic > polysyllabic
3. covered > uncovered
4. closed > open
5. [n < 5 segments] > [n > 5 segments]
6. [n < 262,900 tokens] > [n > 262,900 tokens]

According to preferences 2–4, the canonical Π should be a monosyllable equipped with both a HEAD and a CODA. Figure 2.14 suggests that SLOPES are generally not preferred so that the margins should host only single consonants. According to Figure 2.25, the best candidate for the NUCLEUS is a monophthong whose backness is zero. The most frequent vowel employed in segmental chains of Πs has the feature [low]. As to the quantity of the nuclear segment, short and long vowels obey a "complimentary quantity restriction" (Galea/Ussishkin 2018: 55) according to which a doubly filled CODA requires the vowel to be short whereas long vowels fit in with single CODA consonants. Since the canon has the structure CVC, the nucleus should bear the feature [long]. Note that this automatic rule adds a slot to the segmental chain so that CVC has to be reinterpreted as CV:C with the length symbol occupying a slot of its own. For the head, the feature [labial] is crucial whereas in the CODA [denti-alveolar] is the preferred choice. Since at both margins the most frequently attested manner of articulation is [nasal] the canonical Π looks like the hypothetical form in Figure 2.31.

C	V	C
m	*a:*	*n*

Figure 2.31: Segmental chain of the canonical Π.

The fact that there is no Maltese Π **ma:n* is perfectly in line with the philosophy of the canonical approach propagated by Corbett (2005). The canon is only an abstract yardstick which need not be attested at all. What we can do with this

yardstick is measure the distances which separate the realised word forms from the canon.

To this end, we separate monosyllables from polysyllables since only the former display the same syllabicity as the canon. We stipulate eleven criteria which can be either fulfilled or not. Fulfilment is signalled by '1' in Table 2.9. If a given criterion is not met the corresponding cell hosts a '0'. Since all monosyllables are consonant-initial, the presence of a HEAD is not distinctive and can thus be ignored as such. What makes a difference for HEADS in relation to the canon however is whether there is a singleton consonant or a cluster. If there is a consonant, the question arises whether it realises the place feature [labial] and the manner feature [nasal]. For the NUCLEUS it is important to know whether it is a monophthong, lacks backness, can be characterised as [low] and [long]. Not all monosyllables are closed. Thus, it must be determined whether the CODA is filled or not. If there is a CODA, we need to know whether it is simple or complex. The consonant under inspection must have the place feature [denti-alveolar] and the manner feature [nasal]. If there is a SLOPE, the criteria are tested only for the onset or offset. The rightmost column in Table 2.9 reveals how many features a given Π shares with the canon. The bottom row gives the total of how many Πs share a given feature with the canon. Grey shading identifies those Πs which count among the most frequently attested ones in Table 2.8. The colour yellow marks the two Italo-Romance Πs. The Πs are ordered top-down according to the decreasing number of features shared with the canon.

Table 2.9: Canonicity of monosyllables.

	head			nucleus				coda				sum
	simple	labial	nasal	monophthong	-backness	low	long	filled	simple	denti-alveolar	nasal	
(Π41) *minn* 'from'	1	1	1	1	1	0	0	1	0	1	1	8
(Π21) *għal* 'for'	1	0	0	1	1	1	1	1	1	1	0	8
(Π5) *bħal* 'like'	0	1	0	1	1	1	1	1	1	1	0	8
(Π59) *waqt* 'at the time of'	1	1	0	1	1	1	0	1	1	1	0	8
(Π22) *għand* 'at s.o.'s place'	1	0	0	1	1	1	1	1	0	1	0	7
(Π43) *mnejn* 'from near'	0	1	1	0	1	0	0	1	1	1	1	7
(Π4) *bejn* 'between'	1	1	0	0	1	0	0	1	1	1	1	7
(Π14) *fejn* 'near'	1	1	0	0	1	0	0	1	1	1	1	7

Table 2.9 (continued)

	head			nucleus				coda				sum
	simple	labial	nasal	monophthong	-backness	low	long	filled	simple	denti-alveolar	nasal	
(Π29) *lil* 'to'	1	0	0	1	1	0	1	1	1	1	0	7
(Π11) *dwar* 'about'	0	0	0	1	1	1	1	1	1	1	0	7
(Π17) *fuq* 'on'	1	1	0	1	0	0	1	1	1	0	0	6
(Π26) *kif* 'as'	1	0	0	1	1	0	1	1	1	0	0	6
(Π30) *ma'* 'with'	1	1	1	1	1	1	0	0	0	0	0	6
(Π28) *lejn* 'towards'	1	0	0	0	1	0	0	1	1	1	1	6
(Π20) *għajr* 'except'	1	0	0	0	1	1	0	1	1	1	0	6
(Π10) *daqs* 'equal to'	1	0	0	1	1	1	0	1	0	0	1	6
(Π23) *ħdejn* 'beside'	0	0	0	0	1	0	0	1	1	1	1	5
(Π24) *ħlief* 'except'	0	0	0	1	1	0	1	1	1	0	0	5
(Π48) *qrib* 'near'	0	0	0	1	1	0	1	1	1	0	0	5
(Π56) *taħt* 'under'	1	0	0	1	1	1	0	1	0	0	0	5
(Π16) *fost* 'amongst'	1	1	0	1	0	0	0	1	0	1	0	5
(Π47) *qalb* 'amidst'	1	0	0	1	1	1	0	1	0	0	0	5
(Π15) *fi* 'in'	1	1	0	1	1	0	0	0	0	0	0	4
(Π55) *ta'* 'of'	1	0	0	1	1	1	0	0	0	0	0	4
(Π7) *bi* 'with'	1	1	0	1	1	0	0	0	0	0	0	4
(Π51) *sa* 'till'	1	0	0	1	1	1	0	0	0	0	0	4
(Π9) *bla* 'without'	0	1	0	1	1	1	0	0	0	0	0	4
(Π19) *ġo* 'in'	1	0	0	1	0	0	1	0	0	0	0	3
(Π53) *skont* 'according to'	0	0	0	1	0	0	0	1	0	1	0	3
(Π52) *sforz* 'thanks to'	0	0	0	1	0	0	0	1	0	0	1	3
total	21	12	3	24	25	13	10	23	15	15	8	

No Π reaches the potential maximum of eleven shared features. Four Πs realise eight of the eleven features of the canon. The Italo-Romance Πs yield low turnouts in terms of canonical features so that both candidates occupy ranks near the bottom of the hierarchy. At the top end of the hierarchy, there are two representatives of the highly frequent Πs. Six further members of this class share six to seven features with the canon. None of the features is shared by all thirty Πs. The absence of backness in the NUCLEUS is typical of 25 Πs. The nasal manner of articulation of the CODA consonant is the least common shared feature since it is realised only with eight Πs.

A different choice of criteria might have brought about slightly different results. We insist however on the fact that our own choice is not completely arbitrary. The outcome is interesting insofar as it shows that no Π comes very close to the canon

because the highest number of shared features leaves a gap of three between shared and canonical features. Low frequency seems to correspond relatively often with a reduced number of shared features whereas high frequency results more often in degrees of canonicity which oscillate between 36% and 73%. It is clear that the grammar of Maltese ΠΙs is by no means exhausted because apart from phonological properties, there are also morphological, syntactic, and semantic aspects to be taken account of before the canonical Maltese Π can ultimately be determined. This means that Figure 2.31 is but a small ingredient which needs to be combined with others to yield the desired result. The next ingredient will be presented in Section 3 on morphological issues.