6. CVCC syllables from the etymological point of view

Out of the 772 syllables of the type CVCC in the DBS, only a third, or 267 syllables (35%) come from words of Persian origin, and the remaining 505 syllables (65%) come from words of various etymological origins. Arabic is the biggest contributor with 408 syllables (53% of the DBS; 81% of the DNPS). Another significant source of CVCC syllables is French with 83 syllables (11% of DBS; 17% of DNPS). The remaining 14 syllables come from languages such as German, English and Turkish.

The Data of Basic Syllables is now divided into two etymological components: the DPS and the DNPS. We will first look at the properties of the 267 syllables in the data of Persian origin and make comparisons with the data of non-Persian origin. As before, each syllable position will be discussed separately.

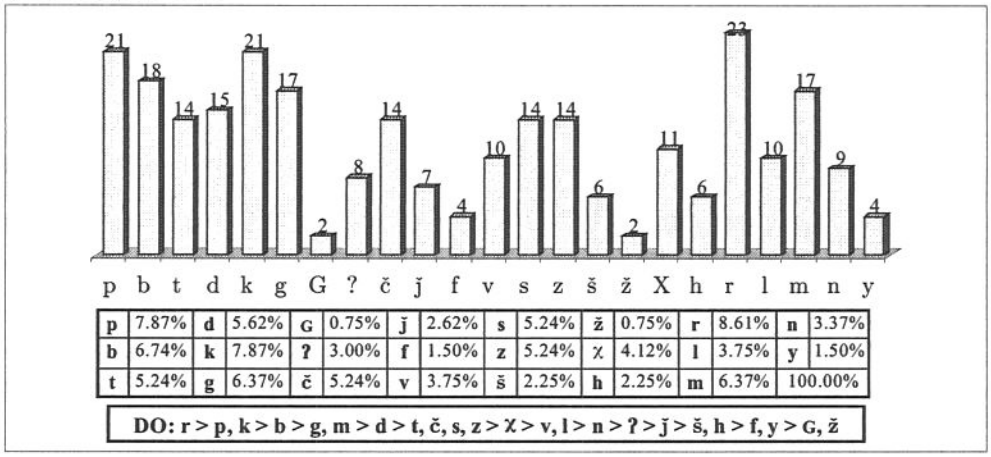
6.1 Onset

DPS

The frequencies and percentages of onset consonants in the CVCC syllables of the DPS are given in Figure 6.1. Here, as in the other syllable types, all consonants can occur in the onset. Figure 6.2 gives the frequencies and percentages of different manner subclasses. It shows that stops predominate. Together with affricates they cover about half (51%) of all onset consonants. This is a little more than their proportions in the onsets of CVC syllables of the DPS, where it is 48% (see Figure 4.13). The comparison of these tables also shows that the proportion of fricatives is somewhat lower in CVCC syllable onsets (25%) than in CVC syllable onsets (28%). As a result, the proportions of obstruents vs.

sonorants are the same in onsets of both syllable types (obstruents 76% and sonorants 24%).

Figure 6.1: *Frequencies and percentages of consonants in the onset of the CVCC syllables of the DPS*



p	7.87%	**d**	5.62%	**G**	0.75%	**ǰ**	2.62%	**s**	5.24%	**ž**	0.75%	**r**	8.61%	**n**	3.37%
b	6.74%	**k**	7.87%	**ʔ**	3.00%	**f**	1.50%	**z**	5.24%	**χ**	4.12%	**l**	3.75%	**y**	1.50%
t	5.24%	**g**	6.37%	**č**	5.24%	**v**	3.75%	**š**	2.25%	**h**	2.25%	**m**	6.37%	100.00%	

DO: r > p, k > b > g, m > d > t, č, s, z > χ > v, l > n > ʔ > ǰ > š, h > f, y > G, ž

Figure 6.2: *Frequencies and percentages of manner subclasses in the onset of the CVCC syllables of the DPS*

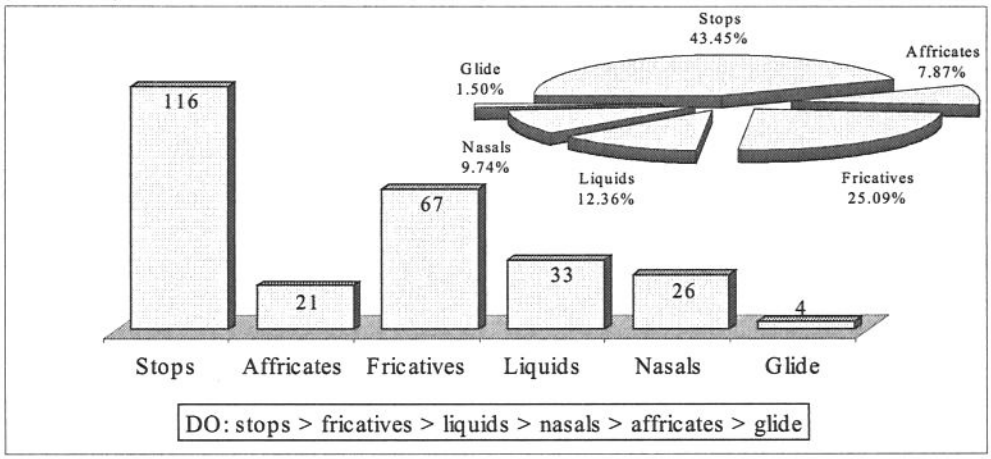
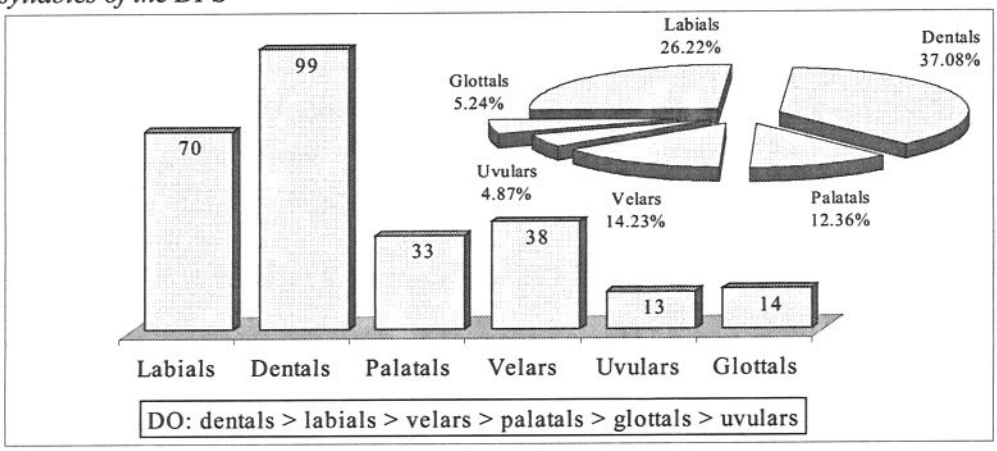


Figure 6.3 presents frequencies and percentages of different place classes. Dentals and labials are the biggest groups, with proportions of 37% and 26% respectively. The proportion of dentals is very much the same in CVCC and CVC onsets in the DPS (see Figure 4.15), but the proportion of labials is bigger in CVCC syllable onsets (26% in CVCC vs. 22% in CVC syllable onsets).

Figure 6.3: *Frequencies and percentages of place classes in the onset of the CVCC syllables of the DPS*



DNPS: Comparison of the DPS and DNPS

Let us now look at the CVCC syllables of non-Persian origin. Figure 6.4 shows onset consonant frequencies and percentages. We see that one consonant, /č/, does not occur at all, and two consonants, /g/ and /ž/, are found only in one syllable each. Two of the stops, /G/ and /ʔ/, have relatively high frequencies, and so has the fricative /s/ and the nasal /n/.

Figure 6.4: *Frequencies and percentages of consonants in the onset of the CVCC syllables of the DNPS*

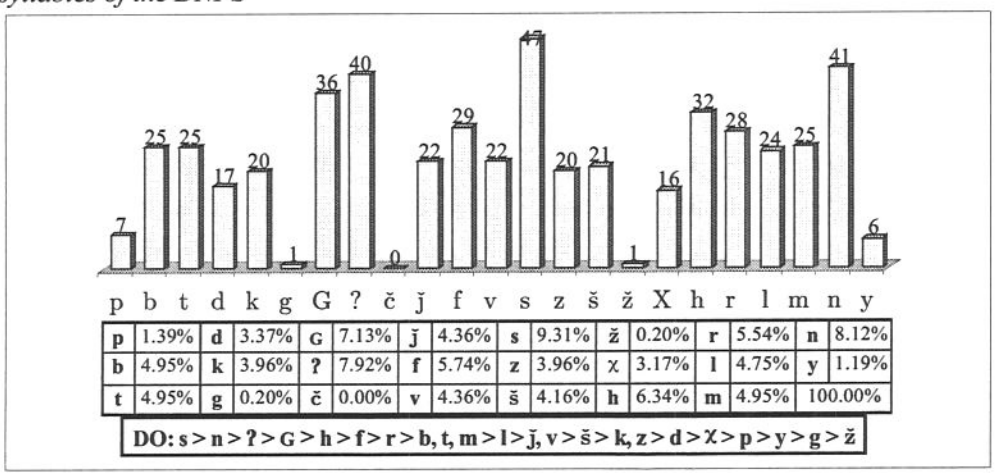
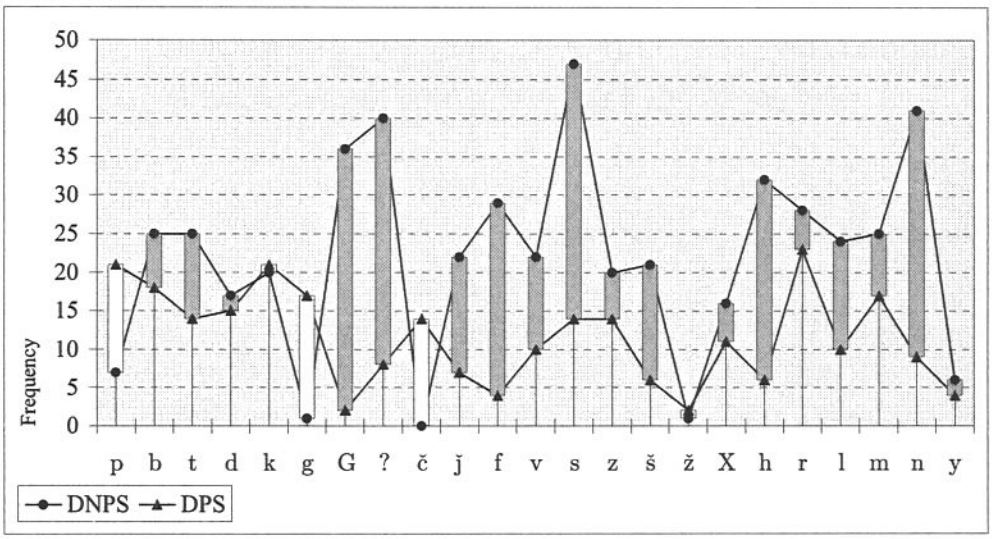


Figure 6.5 compares the frequencies of the onset consonants in the DPS and in the DNPS. The white column indicates that the consonant is more frequent in the DPS, the shaded column means that the consonant is more frequent in the DNPS. Two of the stops, /p/ and /g/, and the affricate /č/, are clearly more frequent in the DPS. On the other hand, two stops, /G/ and /ʔ/, the fricatives /f/, /s/ and /h/, and the nasal /n/ are conspicuously more frequent in the DNPS.

Figure 6.5: *Comparison of the frequencies of consonants in the onset of the CVCC syllables of the DPS and DNPS*



The differences in onset consonant frequencies can be seen better in Figure 6.6. For each onset consonant it shows the percentage of tokens coming from the DPS vs. the DNPS. Since 35% of the CVCC syllables come from words of Persian origin, the expected average ratio of onset consonant frequencies would be 35% of Persian origin and 65% of foreign origin. As Figure 6.6 shows, the results deviate from this theoretical average with the Persian proportion bigger than expected in stops, except for /G/ and /ʔ/, in the affricate /č/ – where all tokens come from the DPS – in fricatives such as /z/, /ž/ and /x/ and in the sonorants /r/ and /m/. On the other hand, the foreign proportion is over 80% in the stops /G/ and /ʔ/, in the fricatives /f/ and /h/, and in the sonorant /n/.

Figure 6.6: *Percentages of onset consonant tokens coming from Persian vs. foreign sources in the CVCC syllables of the DBS*

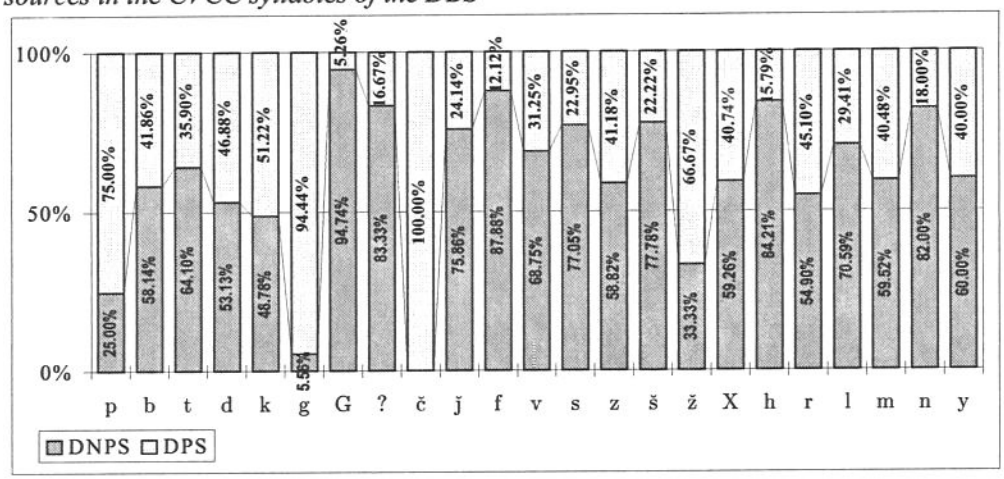


Figure 6.7 gives the percentages of manner subclasses in the CVCC syllable onsets of the DPS and the DNPS. Again, the expected average division would be 35% of the DPS and the rest in the DNPS. The figure shows that the proportion of the DPS is smaller than expected in fricatives and nasals, but in all other manner subclasses it is bigger than expected. Of affricate tokens almost half come from the DPS.

Figure 6.7: *Percentages of different manner subclasses coming from Persian vs. foreign sources in the onset of CVCC syllables of the DBS*

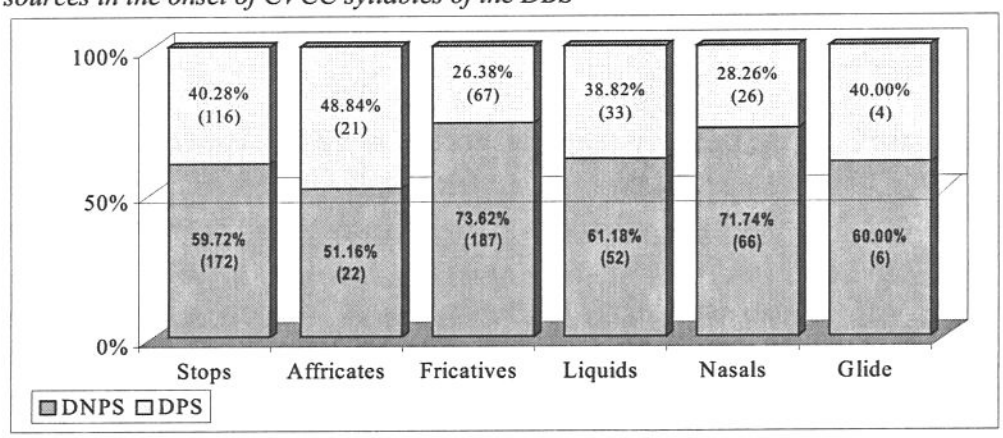
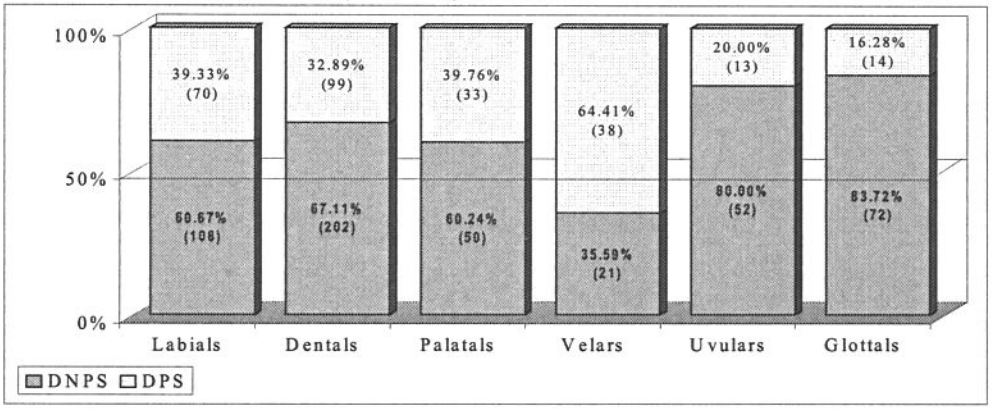


Figure 6.8 presents the percentages of different place classes of the CVCC onset consonants in the DPS and the DNPS. It shows that dentals come closer to the expected division of 35% Persian and 65% non-Persian, while the proportion of Persian is somewhat bigger than expected in labials (39%) and palatals (40%), and significantly bigger in velars (64%). On the other hand, 80% or more of uvulars and glottals in the onset of CVCC syllables come from words of foreign origin.

Figure 6.8: *Percentages of different place classes coming from Persian vs. foreign sources in the onset of CVCC syllables of the DBS*



6.2. Coda

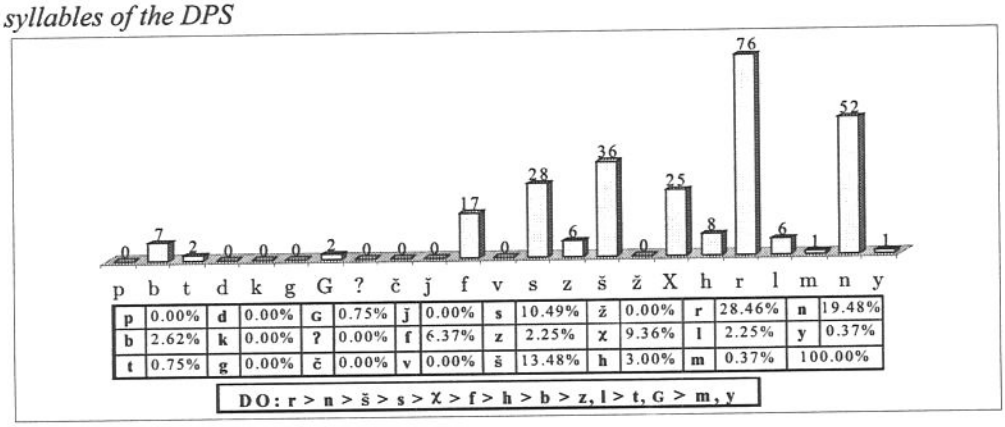
6.2.1. C_1

DPS

Figure 6.9 gives the frequencies and percentages of the first coda consonant (C_1) of CVCC syllables in the DPS. It shows that nine consonants do not occur at all in C_1: five of them are stops (/p/, /d/, /k/, /g/, /ʔ/), two are affricates (/č/, /ǰ/) and two are fricatives (/v/, /ž/). Since three of these, /p/, /č/ and /ž/, are missing in the DBS as well, the number of missing C_1 consonants specific to words of Persian origin is six. It seems that non-continuant consonants (stops and affricates) are avoided in this position in the DPS, i.e. neither of the affricates and only three (out of eight) stops can occur, namely, /b/, /t/ and /G/, but their frequencies are

very low. On the other hand, two sonorants, /r/ and /l/, cover nearly half of the C_1 occurrences; the frequencies of the other sonorants are very low.

Figure 6.9: *Frequencies and percentages of consonants in the C_1 position of the CVCC syllables of the DPS*



The frequencies and percentages of different manner subclasses of C_1 consonants of the DPS are given in Figure 6.10. It shows more clearly what Figure 6.9 suggestes that continuant consonants predominate. Fricatives have a proportion of 45%, liquids (mainly /r/) have a proportion of 31%, and nasals (in practice only /n/) have a proportion of 20%. Altogether, continuant consonants cover 96% of C_1 occurrences, while the proportion of non-continuant consonants is only 4%.

Figure 6.10: *Frequencies and percentages of manner subclasses in the C_1 position of the CVCC syllables of the DPS*

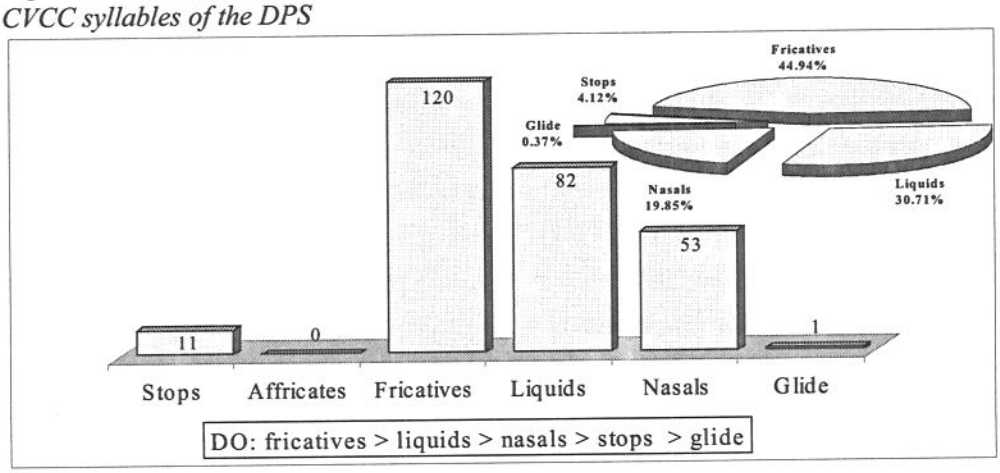
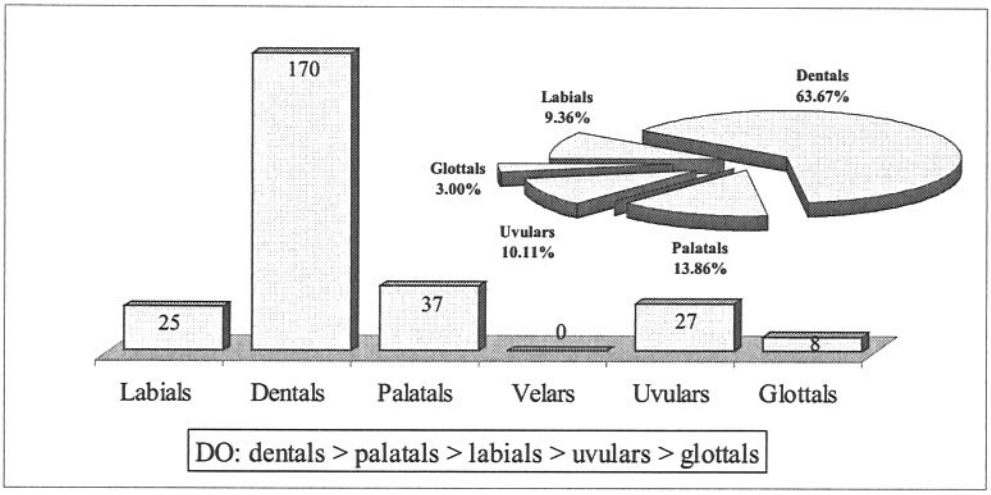


Figure 6.11 presents frequencies and percentages of place classes of C_1 consonants in the DPS. The figure shows that almost two-thirds (64%) of the occurrences are dentals, while the proportions of the remaining groups range from 14% (palatals) to 0% (velars).

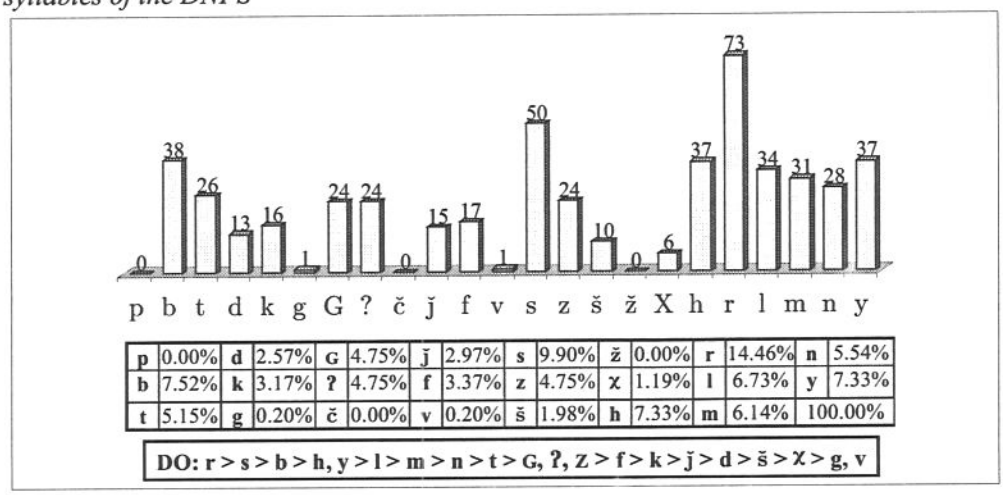
Figure 6.11: *Frequencies and percentages of place classes in the C_1 position of the CVCC syllables in the DPS*



DNPS: Comparison of DPS and DNPS

Let us now look at C_1 consonants of the DNPS. Figure 6.12 shows their frequencies and percentages. The same three consonants, namely, /p/, /č/ and /ž/, that are missing in C_1 position in the DBS, naturally do not occur in the DNPS either. Moreover, /g/ and /v/ have only one occurrence each. The most frequent consonant is /r/ with a share of 14% of C_1 occurrences; other sonorants are also fairly frequent, with proportions ranging around 6-7% each. The most frequent fricatives are /s/ (10%) and /h/ (7%). The proportions of stops and affricates are clearly bigger than in the DPS, where it is only 4%; in the DNPS, the proportion of non-continuant consonants is almost a third (31%). The most frequent of them, /b/, covers as much as 8% of all C_1 occurrences.

Figure 6.12: *Frequencies and percentages of consonants in the C_1 position of the CVCC syllables of the DNPS*



p	0.00%	**d**	2.57%	**G**	4.75%	**ǰ**	2.97%	**s**	9.90%	**ž**	0.00%	**r**	14.46%	**n**	5.54%
b	7.52%	**k**	3.17%	**ʔ**	4.75%	**f**	3.37%	**z**	4.75%	**χ**	1.19%	**l**	6.73%	**y**	7.33%
t	5.15%	**g**	0.20%	**č**	0.00%	**v**	0.20%	**š**	1.98%	**h**	7.33%	**m**	6.14%	100.00%	

DO: r > s > b > h, y > l > m > n > t > G, ʔ, z > f > k > ǰ > d > š > χ > g, v

Figure 6.13 compares the C_1 consonant frequencies in the DPS and the DNPS. The shaded column indicates that the consonant is more frequent in the DNPS and the white column indicates that it is more frequent in the DPS. The figure shows that /š/, /χ/, /n/, and to a small extent /r/, are more frequent in words of Persian origin. All the other occurring consonants are more frequent in words of foreign origin.

Figure 6.13: *Comparison of the frequencies of consonants in the C_1 position of the CVCC syllables of the DPS and DNPS*

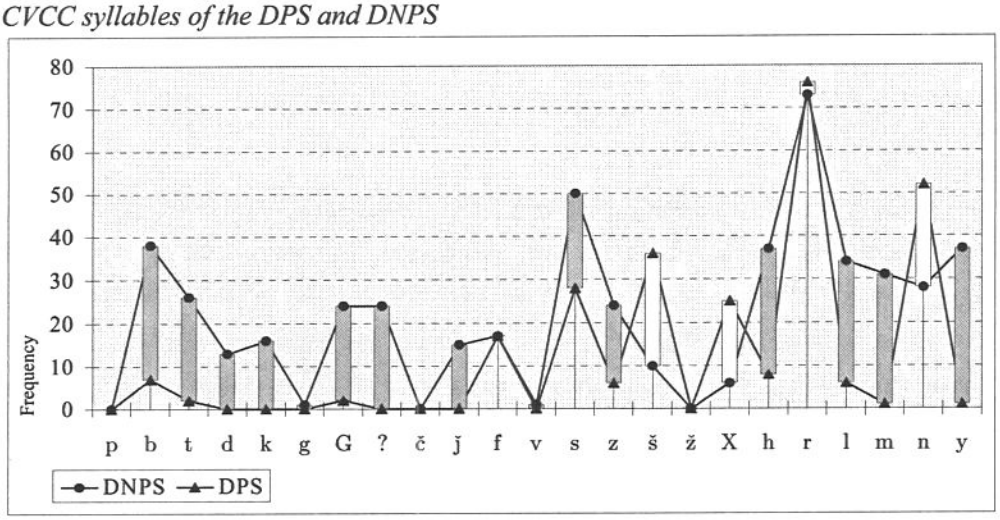


Figure 6.14 shows what percent of the C_1 tokens of each consonant comes from the DPS vs. from the DNPS. We see that 84% or more of all occurrences of non-continuant consonants come from words of foreign origin. The class of continuants is more varied. Of fricatives the proportion of tokens from the DPS is close to the expected 35% in /s/. The proportion of Persian is bigger than expected in /š/, /χ/, and /f/, and smaller than the expected in the rest. Two of the sonorants, /n/ and /r/, have more than half of their occurrences coming from words of Persian origin, while the other sonorants in C_1 have their origin in loanwords in 85% of the cases, or more.

Figure 6.14: *Percentages of C_1 consonant tokens coming from Persian vs. foreign sources in the CVCC syllables of the DBS*

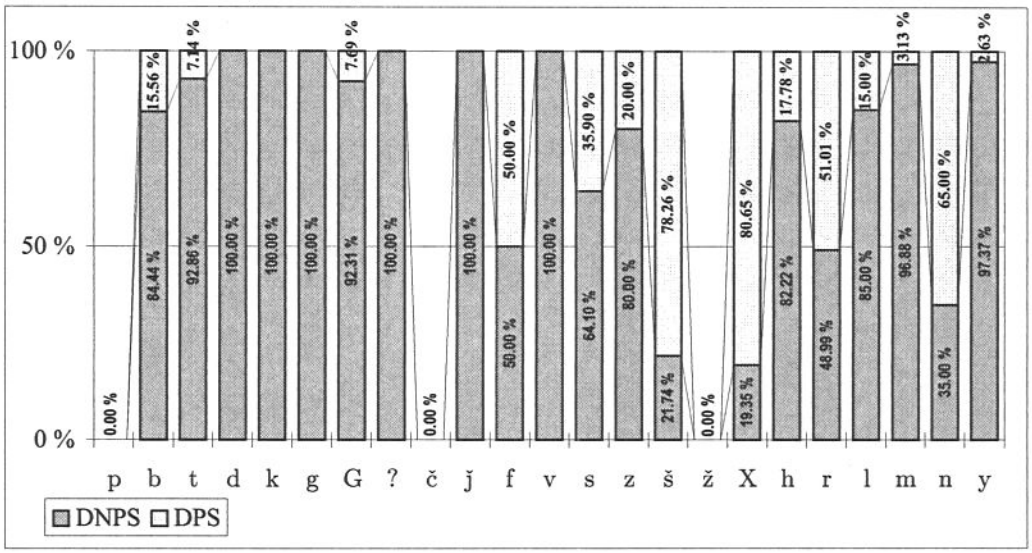


Figure 6.15 gives the percentage of C_1 consonant tokens coming from the DPS vs. the DNPS for each manner subgroup. As the previous figures have suggested, all affricates, and almost all instances of the glide (97%) and stops (93%) come from the DNPS. On the other hand, the proportions of the DPS are bigger than the expected, with 35% in fricatives, nasals and liquids. We have to remember, however, that only one of the nasals, /n/, and one of the liquids, /r/, is more frequent in the DPS.

Figure 6.15: *Percentages of different manner subclasses coming from Persian vs. foreign sources in the in C_1 position of CVCC syllables of the DBS*

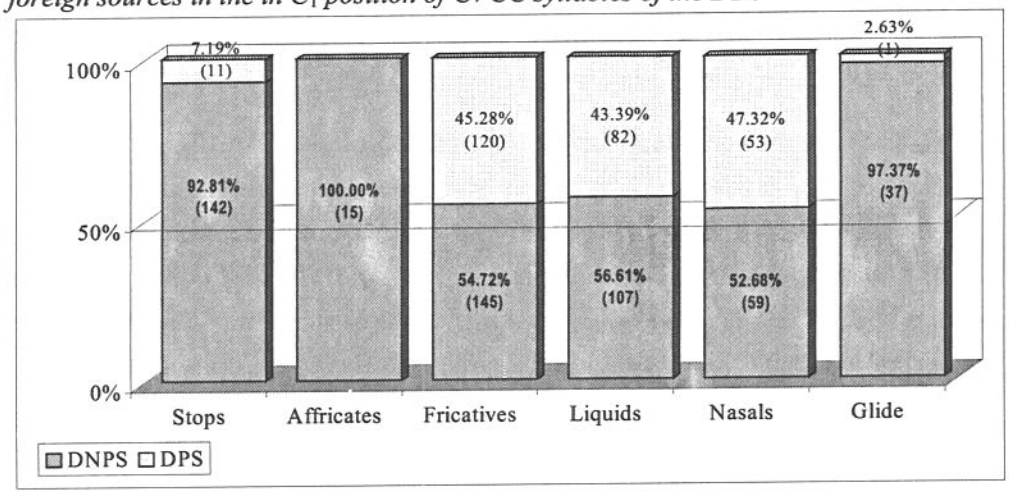
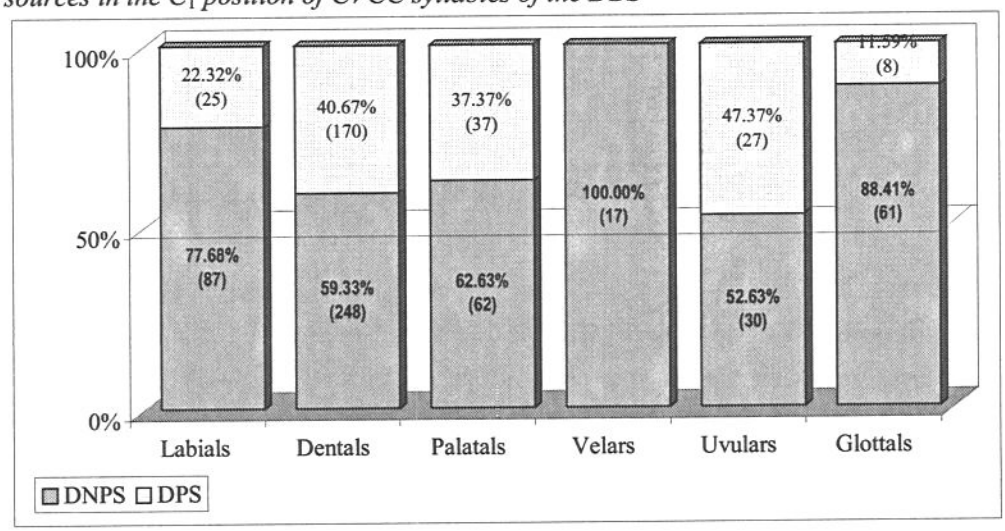


Figure 6.16 shows what percent of the C_1 consonant tokens originate in the DPS vs. the DNPS for each place class. We see that all instances of velars – which are stops – come from the DNPS, and that the proportion of the DNPS is bigger than the expected 65% in glottals and labials. On the other hand, the proportion of the DPS is bigger than the expected 35% in uvulars, dentals, and palatals.

Figure 6.16: *Percentages of different place classes coming from Persian vs. foreign sources in the C_1 position of CVCC syllables of the DBS*

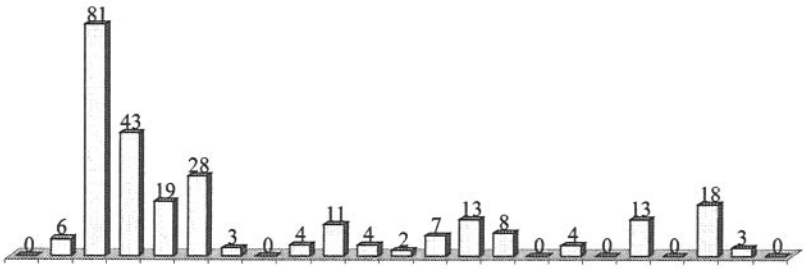


6.2.2. C_2

DPS

Let us first look at consonants in the C_2 position of the CVCC syllable in the DPS. Figure 6.17 gives their frequencies and percentages. The figure shows that six consonants do not occur in C_2 at all; four of them are continuants (/ž/, /h/, /l/, /y/) and two are non-continuants (/p/ and /ʔ/). Moreover, most consonants have low frequencies. High frequencies are found only in stops, more specifically in dentals and velars: /t/ alone comprises nearly a third (30%) of all occurrences in C_2. However, some sonorants, mainly /m/ and /r/, and some fricatives (especially /z/) have moderate frequencies in C_2 in syllables of Persian origin.

Figure 6.17: *Frequencies and percentages of consonants in the C_2 position of the CVCC syllables of the DPS*

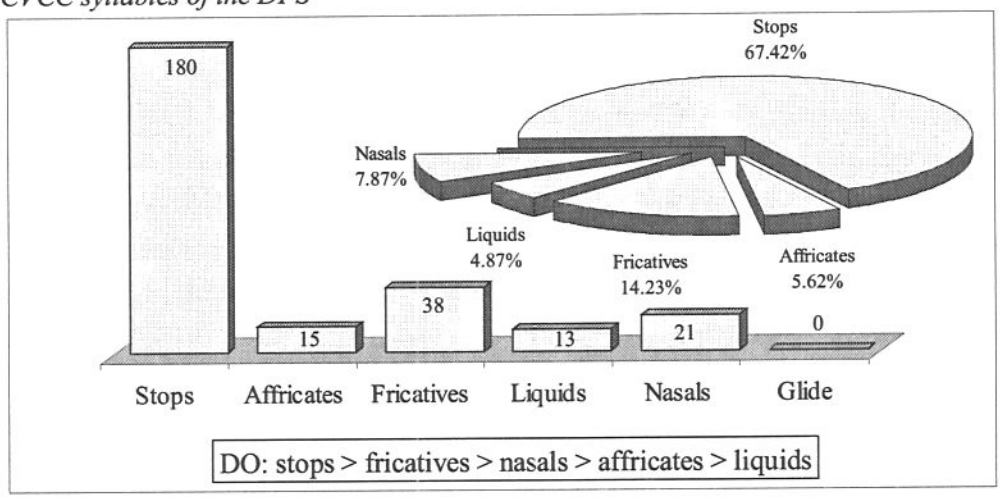


p	0.00%	d	16.10%	G	1.12%	j	4.12%	s	2.62%	ž	0.00%	r	4.87%	n	1.12%
b	2.25%	k	7.12%	ʔ	0.00%	f	1.50%	z	4.87%	χ	1.50%	l	0.00%	y	0.00%
t	30.34%	g	10.49%	č	1.50%	v	0.75%	š	3.00%	h	0.00%	m	6.74%		100.00%

DO: t > d > g > k > m > z, r > j > š > s > b > č, f, χ > G, n > v

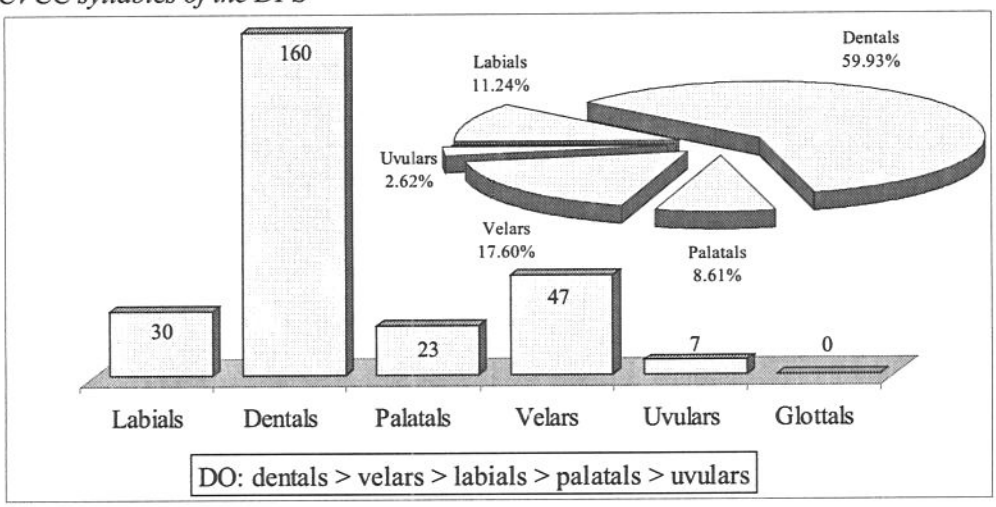
Figure 6.18 presents frequencies and percentages of different manner subclasses in C_2. It shows that stops alone cover two thirds (67%) of all C_2 tokens. Stops and affricates together, i.e. non-continuant consonants, have a proportion of nearly three-quarters (73%) of C_2 tokens in the DPS.

Figure 6.18: *Frequencies and percentages of manner subclasses in the C_2 position of the CVCC syllables of the DPS*



Frequencies and percentages of different place classes of the DPS in C_2 are given in Figure 6.19. The figure shows that 60% of the tokens are dentals; the proportions of the rest range from 18% (velars) to 0% (glottals).

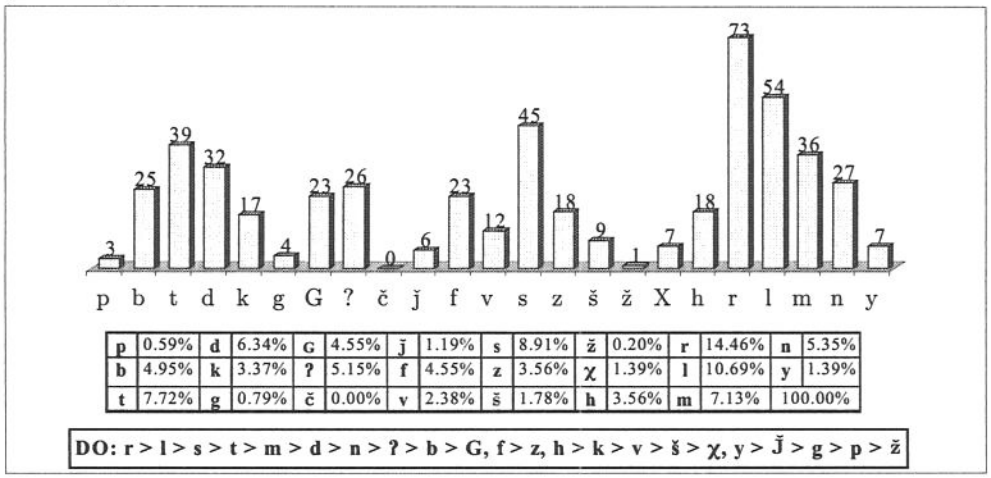
Figure 6.19: *Frequencies and percentages of place classes in the C_2 position of the CVCC syllables of the DPS*



DNPS: Comparison of the DPS and DNPS

Now let us look at consonants from the DNPS in the C_2 position and make comparisons with the DPS. Figure 6.20 shows C_2 consonant frequencies and percentages in the DNPS. We see that the profile is very different from that in Figure 6.17, which describes the data of Persian origin. The highest percentages in Figure 6.20 belong to the sonorants /r/ and /l/; the five sonorant consonants together cover 39% of C_2 tokens in the DNPS. The corresponding figure for the DPS is as low as 13% (cf. Figures 6.17-18). In the fricative group, /s/ is the most frequent one; the proportion of the whole fricative group is 26%. Non-continuants, which are the dominant group in C_2 in the DPS, with a proportion of 73%, have a proportion of only 35% in the DNPS.

Figure 6.20: *Frequencies and percentages of consonants in the C_2 position of the CVCC syllables of the DNPS*



p	0.59%	d	6.34%	G	4.55%	j	1.19%	s	8.91%	ž	0.20%	r	14.46%	n	5.35%
b	4.95%	k	3.37%	?	5.15%	f	4.55%	z	3.56%	χ	1.39%	l	10.69%	y	1.39%
t	7.72%	g	0.79%	č	0.00%	v	2.38%	š	1.78%	h	3.56%	m	7.13%	100.00%	

DO: r > l > s > t > m > d > n > ? > b > G, f > z, h > k > v > š > χ, y > J > g > p > ž

Figure 6.21 compares C_2 frequencies of the DPS and the DNPS. The white column indicates that the consonant is more frequent in the DPS, and the shaded column means that the consonant is more frequent in the DNPS. On the basis of Figures 6.17 and 6.20 we can expect that (at least some) stops are more frequent in the DPS, while sonorants are more frequent in the DNPS. This is indeed the

case since dental stops and /g/ are clearly more frequent in the DPS. On the other hand, /s/ and the liquids are drastically more frequent in the DNPS.

Figure 6.22 shows what percentage of the tokens comes from the DPS vs. the DNPS for each consonant. In stops and affricates, the Persian proportion is over 50%, except for the labials, /G/ and /ʔ/. Of the fricatives, only /z/, /š/ and /χ/ exceed the expected 35% of tokens from the DPS; in all the other fricatives, 85% or more of the tokens come from loanwords. All other sonorants except /m/ have 85% or more of their occurrences from non-Persian sources.

Figure 6.21: *Comparison of the frequencies of consonants in the C_2 position of the CVCC syllables of the DPS and DNPS*

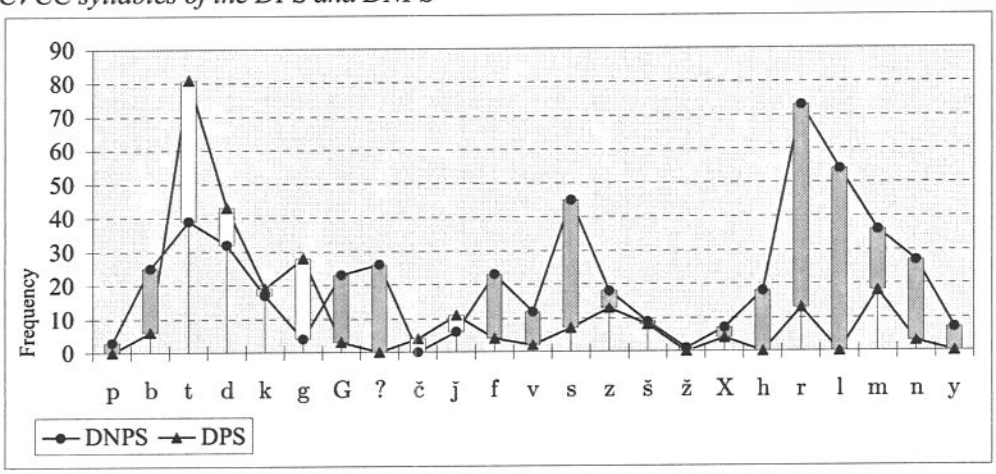


Figure 6.22: *Percentages of C_2 consonant tokens coming from Persian vs. foreign sources in the CVCC syllables of the DBS*

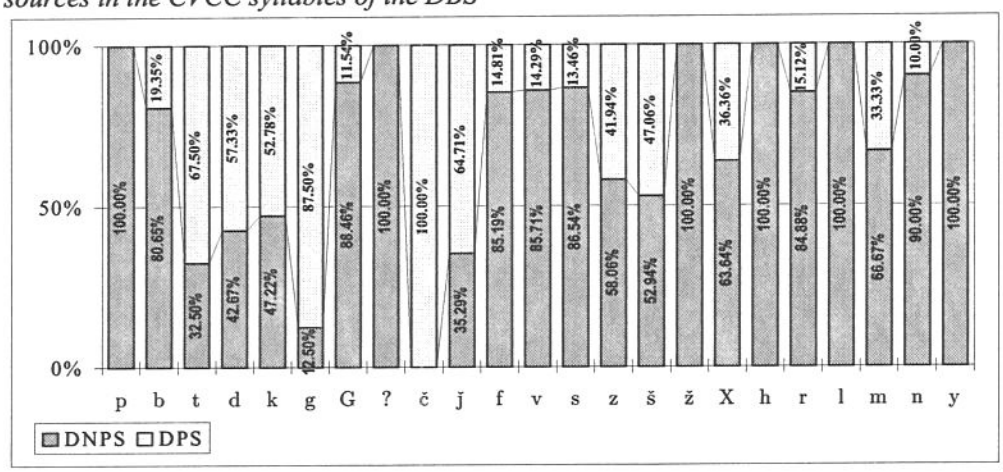


Figure 6.23 compares manner subclass frequencies and percentages in the DPS and the DNPS. It shows that more than half of stop and affricate tokens in C_2 come from words of Persian origin, while in the rest (i.e. in continuants), 75% or more of the tokens come from words of foreign origin.

Finally, Figure 6.24 shows what percentage and how many tokens of each place class come from words of Persian vs. foreign origin. It is only dentals that correspond to the expected 35%-65% division with 36% of their tokens in C_2 coming from the DPS. The proportion of the DPS is 50% or more in palatal and velar tokens and the proportion of the DPS is less than 25% in the rest.

Figure 6.23: *Percentages of different manner subclasses coming from Persian vs. foreign sources in the C_2 position of CVCC syllables of the DBS*

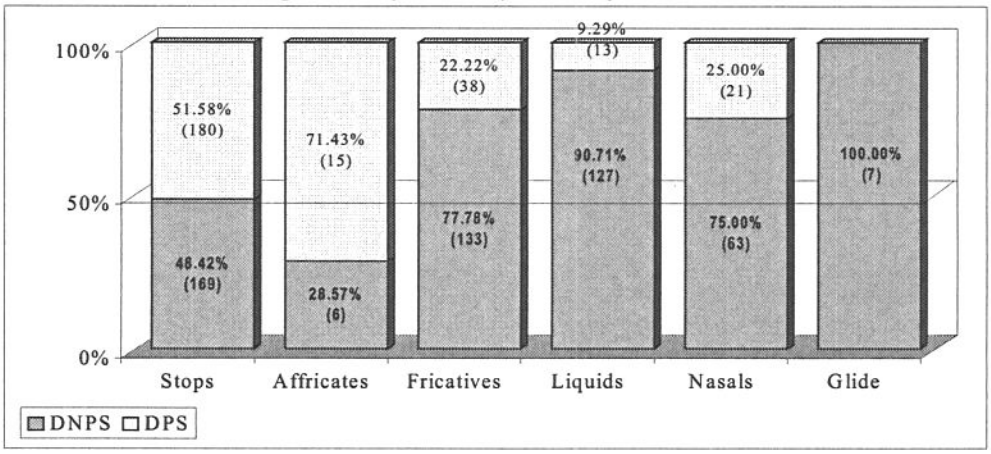
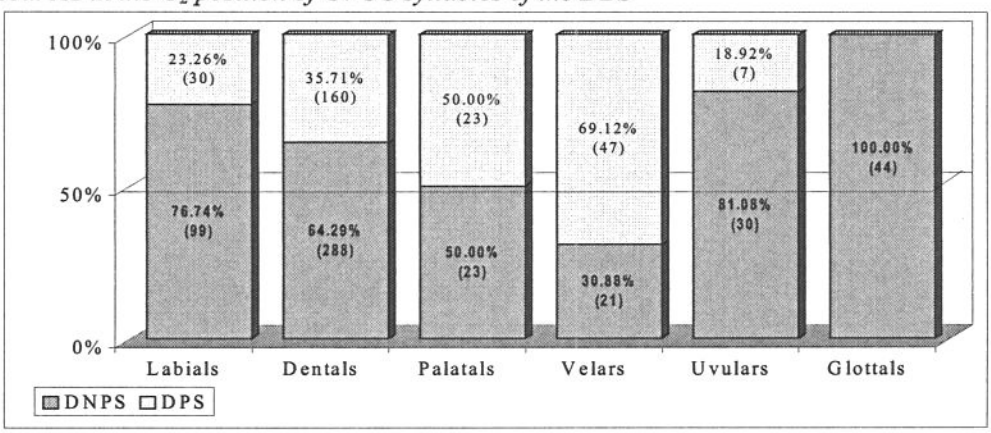


Figure 6.24: *Percentages of different place classes coming from Persian vs. foreign sources in the C_2 position of CVCC syllables of the DBS*

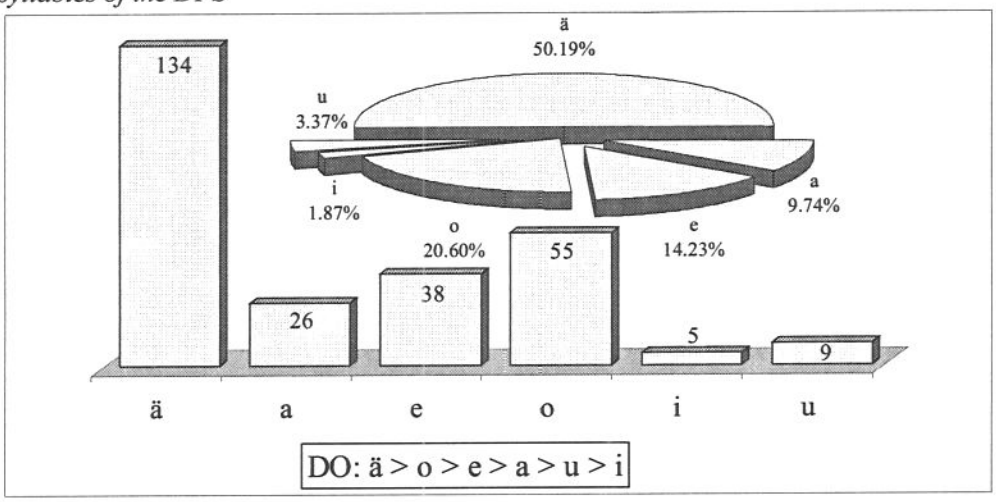


6.3. Nucleus

DPS

Figure 6.25 presents vowel frequencies and percentages in CVCC syllables in the DPS. The figure shows that half of all vowel occurrences here belong to /ä/, and one fifth to /o/. The short vowels /ä/, /e/ and /o/ dominate, together covering 86% of all vowel occurrences. The sonority of the vowel also plays a clear role, and, in both the short and the long vowel series, the more sonorous vowel is the more frequent one (i.e. in the short vowel series, /ä/ is more frequent than the mid vowels, and in the long vowel series, /a/ is more frequent than the close vowels).

Figure 6.25: *Frequencies and percentages of the vowels in nucleus of the CVCC syllables of the DPS*



The DPS is quite similar to the DBS in terms of the nuclear vowels of CVCC syllables (cf. Figure 5.10). The biggest differences are in the proportions of mid vowels: /o/ has a proportion of 21% in the DPS as against 18% in the DBS, and the proportion of /e/ is 14% in the DPS, but 19% in the DBS.

DNPS: Comparison of the DPS and DNPS

Figure 6.26 compares vowel frequencies of CVCC syllables in the DPS and the DNPS. It shows that there are about twice as many tokens of /ä/ from foreign sources as there are from Persian sources, and about three times as many tokens of /e/ from non-Persian as there are from Persian data.

Figure 6.26: *Comparison of the frequencies of nuclear vowels in CVCC syllables of the DPS and DNPS*

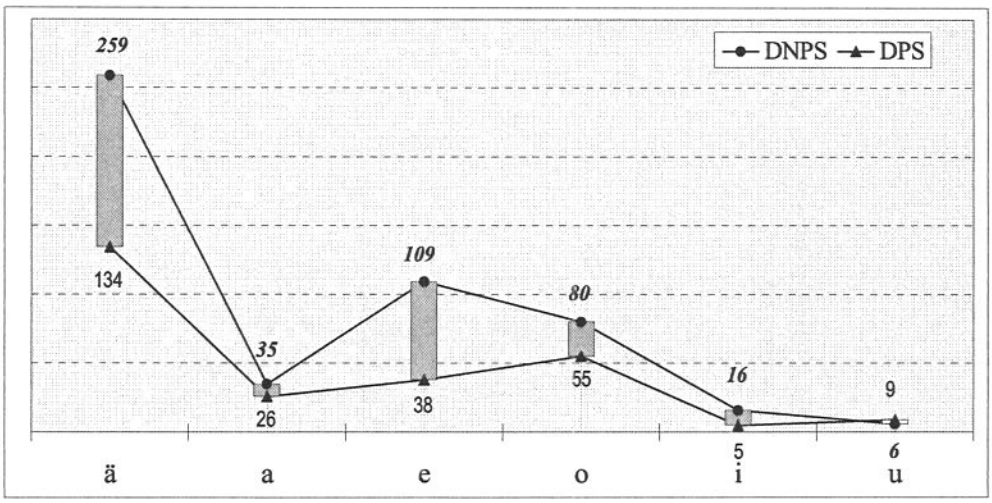
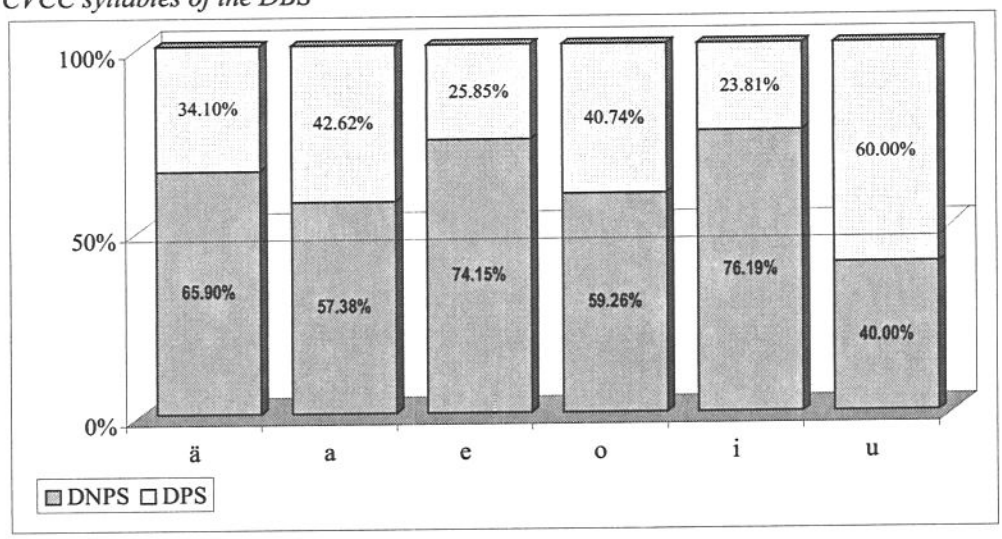


Figure 6.27 shows the percentages of tokens of Persian vs. foreign origin for each vowel in the CVCC syllables. We can recall that 35% of the CVCC syllables (and therefore also of the nuclear vowels) are of Persian origin. We see that /ä/ comes closest to the average proportion of Persian (i.e 35%). In the back vowels, /a/, /o/ and /u/ the proportion of Persian is bigger than the average, ranging between 40% and 60%. In /e/ and /i/ the proportion of Persian is smaller than the average, having about one-fourth of the tokens.

Figure 6.27: *Percentages of Persian vs. foreign sources in the frequencies of vowels in CVCC syllables of the DBS*



6.4. Comparison of the frequencies of the consonants in different positions of the syllable

Figure 6.28 presents consonants frequencies in the onset, C_1 and C_2 of the CVCC syllables in DPS. Part A of the figure compares frequencies at syllable edges, i.e. in the onset and C_2. We see that the curves are fairly similar, except for the high frequencies of dental stops in C_2. Stops seem to be somewhat more frequent in the onset and clearly more frequent in C_2 than other manner classes, although all manner classes occur in these positions. Part B of the figure shows clear differences in consonant frequencies between C_1 and C_2; stops are quite infrequent in C_1, but frequent in C_2, especially dentals. The situation is reversed for fricatives and sonorants: they have low frequencies in C_2 and higher frequencies in C_1.

Figure 6.28-A: *Comparison of the frequencies of the consonants in the onset and C_2 positions of CVCC syllables of the DPS*

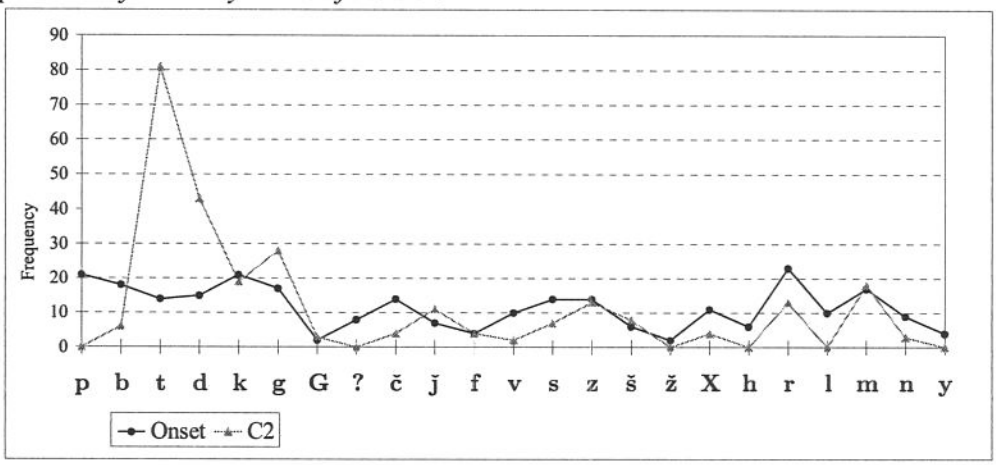


Figure 6.28-B: *Comparison of the frequencies of the consonants in the C_1 and C_2 positions of CVCC syllables of the DPS*

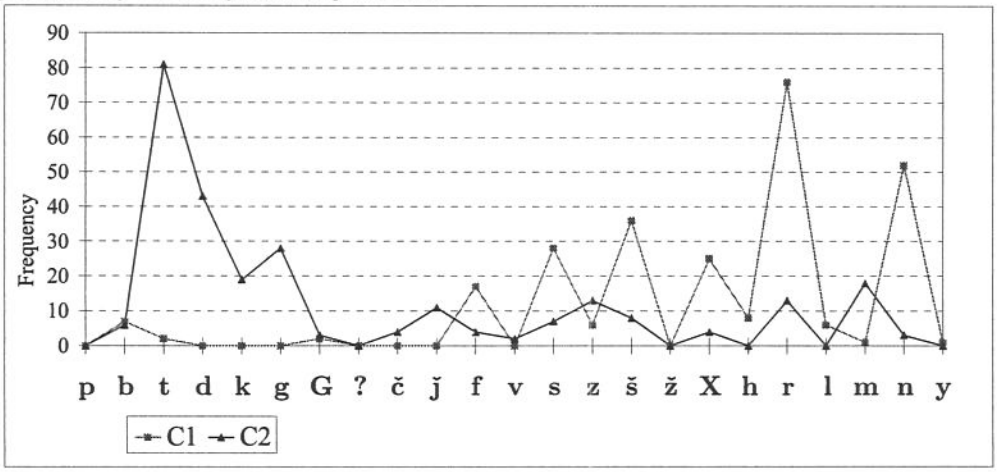


Figure 6.29 shows the frequencies of different manner subclasses in the different syllable positions. We see that stops and affricates, i.e. non-continuants, occur almost exclusively at syllable edges in the DPS. Stops favour the C_2 position in particular. Fricatives, liquids and nasals, i.e. continuants, are most frequent in the C_1 position, and least frequent in the C_2 position. The glide is quite infrequent in any position of the CVCC syllable of the DPS.

Figure 6.29: *Comparison of the frequencies of manner subclasses in the onset and coda of the CVCC syllables of the DPS*

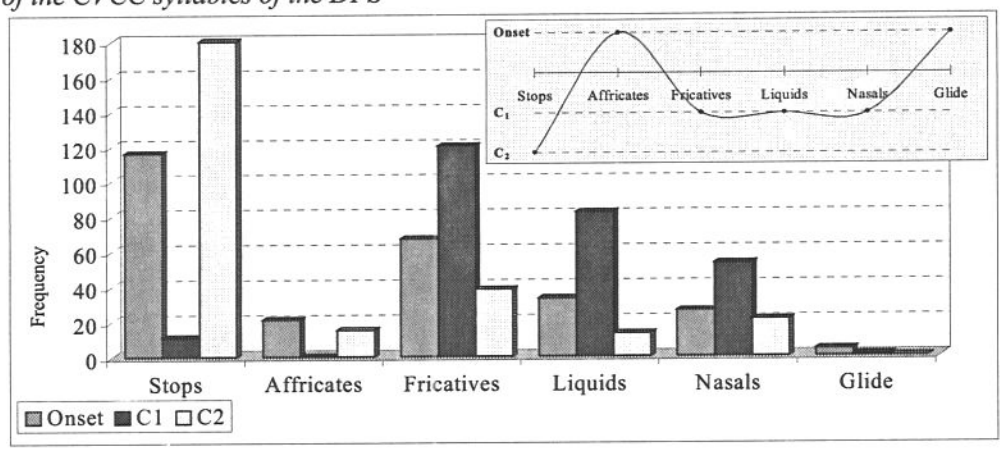


Figure 6.30 shows more clearly the difference between continuants vs. non-continuants in different syllable positions. There is practically no difference in their frequencies in the onset position, but the differences are very clear in the coda where only 4% of C_1 occurrences are non-continuants, while the proportion of non-continuants in C_2 is 73%.

Figure 6.30: *Comparison of the frequencies of non-continuants and continuants in the onset and coda of the CVCC syllables of the DPS*

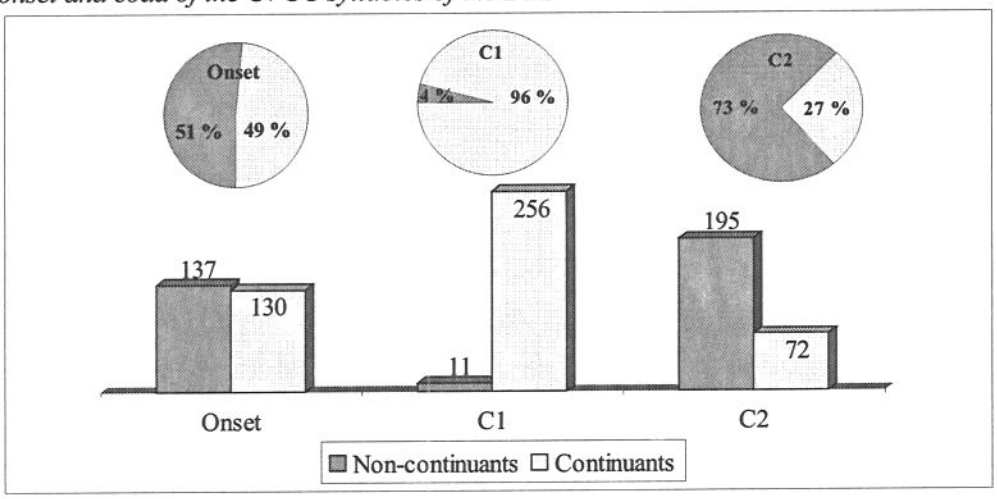


Figure 6.31 confirms the view that the sonority difference shows up in the opposition *continuant/non-continuant* rather than *voiced/voiceless*. The figure shows, first, that the different positions do not differ very much with respect to the voicing of consonants, and, second, that voiced consonants are somewhat more frequent in all positions. One reason for the bigger proportion of voiced consonants is that their number (13) is bigger than the number of voiceless consonants (10) in Modern Persian. The proportion of voiced consonants is biggest in C_1 (57%), but it is only slightly more than their proportion in the onset. Moreover, 83% of fricatives in C_1 are voiceless. Thus, it is possible to conclude that differences in voicing do not reflect differences in sonority.

Figure 6.31: *Comparison of the frequencies and percentages of voiceless vs. voiced consonants in the onset and two coda positions of the CVCC syllables of the DPS*

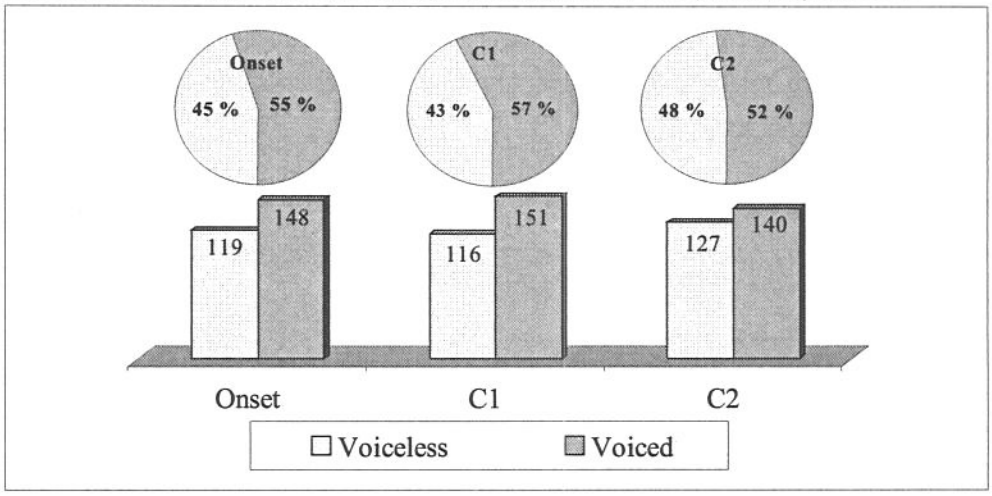
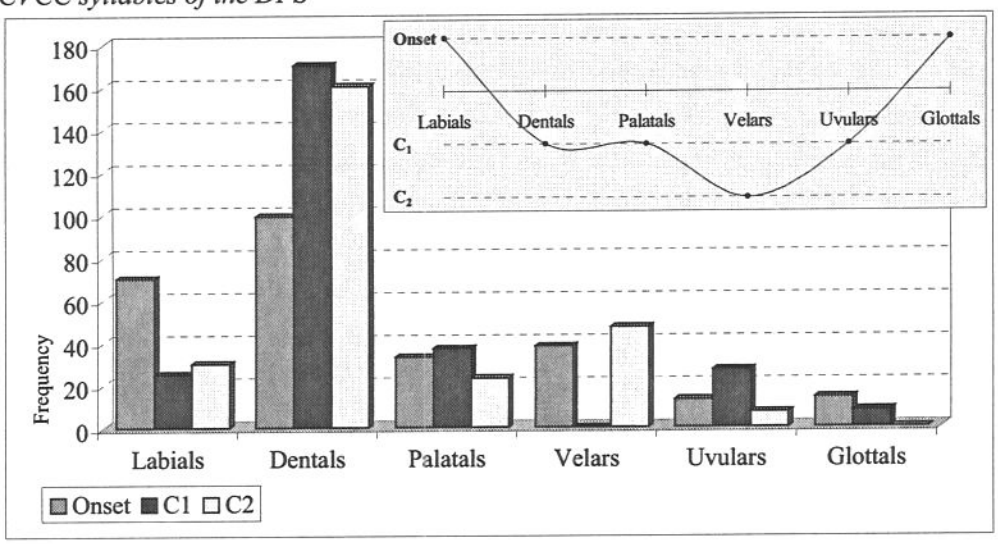


Figure 6.32 shows the frequencies of place classes in different positions of the CVCC syllable in the DPS. Dentals dominate in all positions, and in the coda especially. In other words, consonant clusters favour dentals.

Figure 6.32: *Comparison of the frequencies of place classes in the onset and coda of the CVCC syllables of the DPS*



6.5. Segment sequences in the CVCC syllable of the DPS and the DNPS

6.5.1. Onset+nucleus

As was mentioned before, the CVCC syllables of the DBS contain 106 different Co+V sequences. Out of these, 86 sequences (81%) are found in the data of Persian origin. The missing 20 sequences found only in the data of foreign origin come mainly from French and Arabic, as the following list shows.[52] The sequences are grouped according to the source languages:

(1) F (10 sequences): /bu/, /ga/, /fi/,/si/, /ša/, /ši/, /ža/, /la/, /li/, /lu/
(2) A (5 sequences): /pi/, /be/, /Go/, /ʔe/, /fe/
(3) G: /di/
(4) AFE: /te/
(5) AFT: /yo/
(6) AF: /no/
(7) FA: /fo/

[52] A = Arabic, C = Chinese, E = English, F = French, G = Greek, H = Hindi, I = Italian, L = Latin, P = Persian, R = Russian, S = Sanskrit, T = Turkish, Y = Aramo-Syriac.

Table 6.1 compares the Co+V sequences of the CVCC syllable in the DBS and in the DPS. 'P' in the slot means that the sequence occurs in the DPS, and '-' means that the sequence is missing in the DPS, but occurs in the DNPS. The shaded slot indicates that the sequence is found neither in the DPS nor in the DNPS, i.e. there is a gap in the DBS. The table shows that the sonority of the vowel has an effect on the possible Co+V sequences, i.e. all onset consonants of the DPS can occur with /ä/, while there are only 5 to 6 sequences with close nuclear vowels. The contribution per vowel of the data of foreign origin is from zero to six sequences: zero for /Cä/ sequences and six for /Ci/ sequences.

Table 6.1: *Co+V sequences, based on the DBS and DPS*

	p	b	t	d	k	g	G	ʔ	č	ǰ	f	v	s	z	š	ž	χ	h	r	l	m	n	y
ä	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
a	P	P	P	P	P	-	P	P	P		P	P	P		-	-	P		P	-	P	P	P
e	P	-	-	P	P	P	P	-	P	P	-	P	P	P	P		P	P	P	P	P	P	P
o	P	P	P	P	P	P	-	P	P	P	-	P	P	P	P		P	P	P	P	P	-	-
i	-			-				P			-	P	-	P	-				P	-		P	
u	P	-		P	P	P							P				P			-			

6.5.2. Nucleus + coda

Let us first look at $V+C_1$ sequences. We saw before that 82 sequences of $V+C_1$ are found in the DBS. Out of these, 47 sequences (57%) are found and 35 sequences (43%) are missing in the DPS. The following is a list of the 35 sequences in the DBS that were not found in the etymologically Persian syllables:

(1) A (14 sequences): /äd/, /äk/, /äʔ/, /äǰ/, /äy/, /ed/,/ek/, /eG/, /eʔ/, /eǰ/, /ez/, /od/, /oʔ/, /oǰ/
(2) F (12 sequences): /ab/, /at/, /al/, /ay/, /eg/, /ov/, /ib/, /it/, /il/, /ub/, /ud/, /uk/
(3) AF (5 sequences): /äm/, /et/, /ob/, /ok/, /om/
(4) FE (2 sequences): /am/, /in/
(5) FT: /ur/
(6) FR: /ak/

Table 6.2 compares the V+C_1 sequences in the DPS and the DNPS. As in Table 6.1, 'P' means that the sequence is found in the DPS, and '-' means that the sequence occurs only in the DNPS. Shaded slots indicate gaps in both. The table shows that no vowel occurs with all C_1 consonants; the number of different V+C_1 sequences in the DPS varies from 12 (for /äC_1/) to 2 (for /iC_1/). The short vowels /ä, e, o/ and the long vowels /a, i, u/ form two clearly separate groups here. The number of V+C_1 combinations is from 11 to 12 with a short vowel in the nucleus, but with a long vowel in the nucleus there are only from 2 to 6 different V+C_1 sequences per vowel. The number of V+C_1 sequences coming from the DNPS ranges from 8 (for /e/) to 4 (for the close vowels /i/ and /u/).

Table 6.2: *V+C_1 sequences, based on the DBS and DPS*

C_1 →	p	b	t	d	k	g	G	ʔ	č	ǰ	f	v	s	z	š	ž	χ	h	r	l	m	n	y
ä		P	P	-	-		P	-		-	P		P	P	P		P	P	P	P	-	P	-
a		-	-		-						P		P		P		P		P	-	-	P	-
e		P	-	-	-	-	-	-		-	P		P	-	P		P	P	P	P	P	P	P
o		-	P	-	-		P	-		-	P	-	P	P	P		P	P	P	P	-	P	
i		-	-										P				P			-		-	
u		-		-	-					-	P		P		P		P		-			P	

Next, let us look at the combinations of V+C_2 in the DPS and the DNPS. We saw before that there are 87 different combinations of V+C_2 in the DBS. In the DPS, the number of these combinations is 54 (62% of the DBS), i.e. 33 combinations come from the DNPS. A list of these 33 V+C_2 combinations is given below, together with the source language. We see that French is the biggest contributor here:

(1) F (17 structures): /äCp/, /aCp/, /aCš/, /aCž/, /aCr/, /aCl/, /aCm/, /eCb/, /eCv/, /oCp/, /iCk/, /iCr/, /iCl/, /iCm/, /uCs/, /uCr/, /uCl/

(2) A (12 structures): /äCʔ/, /äCh/, /äCl/, /äCy/, /eCG/, /eCʔ/, /eCh/, /eCn/, /oCʔ/, /oCv/, /oCh/, /oCn/

(3) AF (3 structures): /eCf/, /eCl/, /oCl/

(4) FE: /iCg/

Table 6.3 shows the occurring V+C_2 combinations in the DPS and the DNPS. As before, 'P' means that the V+C_2 combination occurs in syllables of Persian origin, '-' means that the combination occurs only in the data of foreign origin, and a blank indicates a combination that is missing in both. As in V+C_1, we see that short vowels and long vowels differ clearly in the DPS: in the series of short vowels, the number of V+C_2 combinations ranges from 17 (with /ä/ in the nucleus) to 11 (with /e/ in the nucleus), whereas the range of variation in the series of long vowels is from 7 combinations (with /a/ in the nucleus) to one (with /i/ in the nucleus). The number of combinations coming exclusively from foreign sources varies from 8 combinations (with /e/ in the nucleus) to 3 (with /u/ in the nucleus).

Table 6.3: *V+C+C_2 sequences, based on the DBS and DPS*

C_2 →	p	b	t	d	k	g	G	ʔ	č	ǰ	f	v	s	z	š	ž	χ	h	r	l	m	n	y
ä	-	P	P	P	P	P	P	-	P	P	P	P	P	P	P		P	-	P	-	P	P	-
a	-		P	P	P	P			P				P	P	-	-			-	-	-		
e		-	P	P	P	P	-	-		P	-	-	P	P	P		P	-	P	-	P	-	
o	-	P	P	P	P	P	P	-	P	P	P	-	P	P	P		P	-	P	-	P	-	
i			P		-	-													-	-	-		
u			P	P	P								-						-	-			

6.5.3. Coda clusters

We have seen before that the DBS contains 209 different clusters. Only one-fourth of these, i.e. 54 clusters (26%), are found in the DPS. In other words, three-quarters of the coda clusters in Modern Persian, altogether 155 clusters, come from foreign sources, mainly from Arabic. In the following, we will first look at some statistics describing the nature of the 54 different clusters, and then we will have a closer look at the clusters themselves.

Figure 6.33 compares consonant frequencies in C_1 and C_2 of the 54 different CC clusters in the DPS. The white column means that the consonant is more frequent in C_2 and the shaded column means that the consonant is more frequent in C_1. The figure shows that that C_1 frequencies range mainly from zero to 5; a notable exception is /r/, which is a C_1 member of 15 different clusters. C_2 frequencies are more evenly distributed ranging from zero to seven. Non-

continuants tend to be more frequent in C_1. This tendency is seen more clearly in Figure 6.34 which shows that stops are common in clusters as the last member, liquids as the first member. Fricatives are quite frequent in both positions, but more frequent as the first member. Nasals differ from other continuants in that they favour C_2. However, nasals are less common in clusters than either liquids or fricatives.

Figure 6.33: *Comparison of consonant frequencies in CC clusters in the CVCC syllables of the DPS*

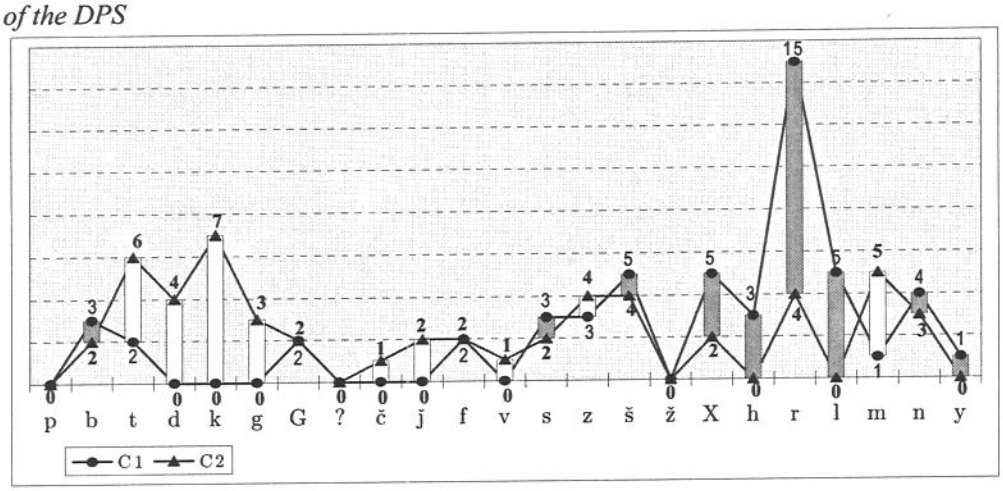
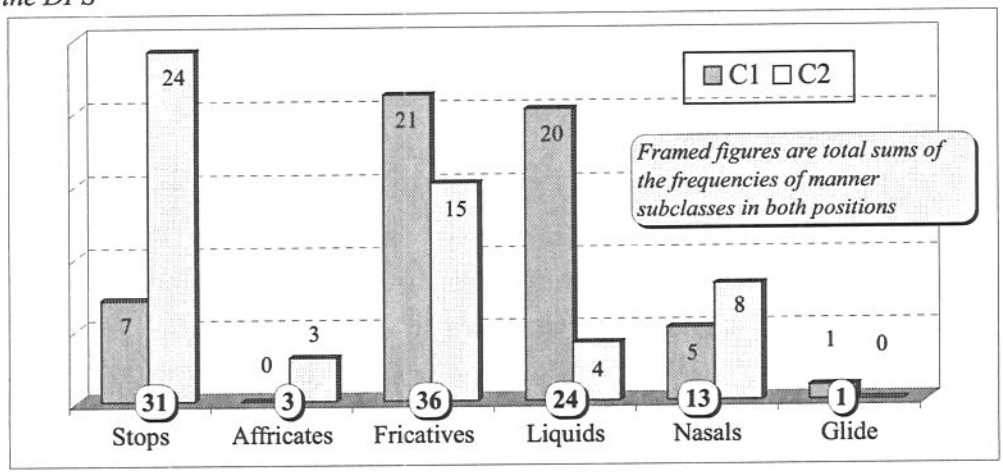


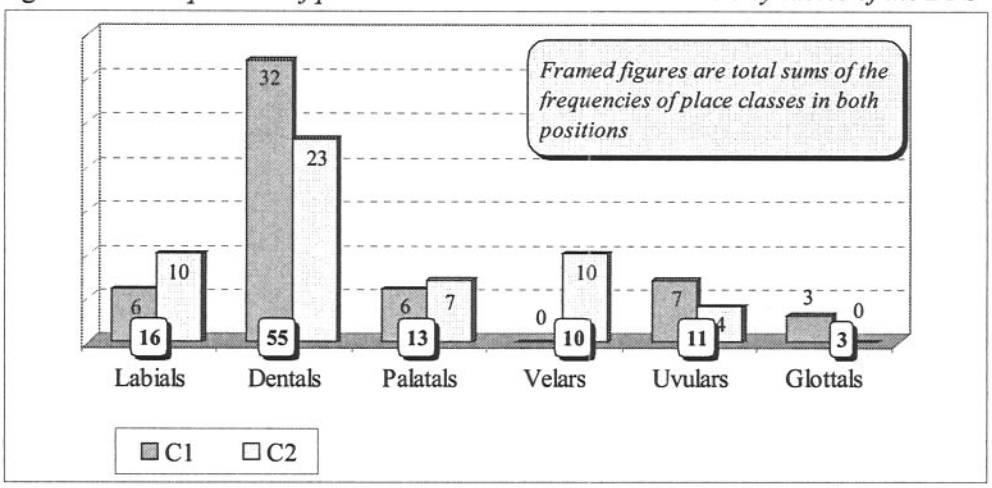
Figure 6.34: *Frequencies of manner subclasses in CC clusters of the CVCC syllables of the DPS*



There is a clear difference between the DPS and the DBS with respect to the consonants of different CC cluster types. This can be seen by comparing Figure 6.34 to Figure 5.17. Figure 5.17 shows that consonant frequencies are nearly the same in C_1 and C_2 in the DBS for fricatives, liquids and nasals and that stops are only somewhat more frequent in C_2. This means that the sonority of the consonant is not an important factor in determining whether the consonant favours C_1 or C_2 in the DBS. In the DPS, on the other hand, as Figure 6.34 shows, less sonorous (= non-continuants) favour C_2, i.e. the syllable edge.

Figure 6.35 shows frequencies of different place classes in C_1 and C_2 of the DPS. We see that dentals are by far the most frequent, and especially in C_1.

Figure 6.35: *Frequencies of place classes in CC clusters in CVCC syllables of the DPS*



As Figure 6.35 shows, here the DPS differs from the DBS (see Figure 5.20). Even though dentals are also the most common group in the DBS, the difference between the frequencies of dentals in C_1 and C_2 is quite small, and it is C_2 that has a slightly higher frequency of dentals in the DBS.

Let us now have a closer look at the clusters. Table 6.4 shows clusters that occur in the DBS. The clusters are given in two groups. Slots marked by 'P' indicate clusters that are found only in words of Persian origin. There are 54 such clusters. Slots marked by '-' indicate clusters that come from words of foreign

origin. There are 155 of these clusters. Together they form the 209 cluster types of the DBS.

Table 6.4: *Comparison of the CC clusters in the DBS and the DPS*

C_1↓ C_2→	p	b	t	d	k	g	G	ʔ	č	ǰ	f	v	s	z	š		χ	h	r	l	m	n	y
p																							
b			-	-	P			-					-	P	-		-	-	P	-		-	
t		-			P		-	-			-							-	P	-	-	-	-
d							-					-	-					-	-	-	-		
k			-										-						-	-	-	-	
g																				-			
G		-	-	P							-	-	-	P	-			-	-	-	-		-
ʔ		-	-	-							-		-	-	-				-	-	-	-	-
č																							
ǰ		-		-				-				-		-				-	-		-		
f			P				-	-				-	-	-	P		-		-	-		-	-
v																			-				
s		P	P	-	P		-	-		-	-						-	-	-	-	-	-	
z		-		P			-	-		-	-	-							-	-	P	P	
š			P	-	P	P	-				-	-							-		P	P	-
χ			P										P	-	P				P	-	P		
h			-	-								-	-	-	-				P	-	P	P	-
r		P	P	P	P	P	P	-	P	P	P	P	P	P	P	-	P	-			P	-	
l		-	P	-	P		P	-			P	-	-				P	-			-		
m	-	-	-	-			-	-					-	-	P				-	-		-	
n		-	-	P	-	P		-		P	-		-	P			-	-					
y		-	-	-	P			-			-		-	-	-		-	-	-	-	-	-	

Table 6.4 shows that, in addition to the three consonants that do not occur in C_1 in the DBS, namely, /p, ž, č/, the etymologically Persian data lacks the following six consonants in C_1: four stops, /d, k, g, ʔ/; one affricate, /ǰ/; and one fricative, /v/. This means that all clusters of the type /dC, kC, gC, ʔC, ǰC, and vC/ found in the DBS come from words of foreign origin. Moreover, five consonants are missing in the C_2 position in the DPS. Thus, clusters of the following type in the DBS come from loanwords: /Cp, Cʔ, Cž, Ch, Cl, Cy/.

Below is a list of the 155 clusters of non-Persian origin, together with the source language:

(1) A (131 clusters): /bt/, /bd/, /bʔ/, /bs/, /bš/, /bχ/, /bh/, /bn/, /tb/, /tG/, /tʔ/, /tf/, /th/, /tm/, /tn/, /tl/, /ty/, /dG/, /dv/, /ds/, /dh/, /dl/, /dm/, /km/, /kn/, /kr/, /kl/, /Gb/, /Gt/, /Gf/, /Gv/, /Gs/, /Gš/, /Gh/, /Gm/, /Gr/, /Gl/, /Gy/, /ʔb/, /ʔt/, /ʔd/, /ʔf/, /ʔs/, /ʔz/, /ʔš/, /ʔm/, /ʔn/, /ʔr/, /ʔl/, /ʔy/, /ǰb/, /ǰd/, /ǰʔ/, /ǰv/, /ǰz/, /ǰh/, /ǰm/, /ǰr/, /fG/, /fʔ/, /fv/, /fs/, /fz/, /fχ/, /fn/, /fr/, /fl/, /fy/, /sd/, /sG/, /sʔ/, /sǰ/, /sf/, /sχ/, /sh/, /sn/, /sl/, /zb/, /zG/, /zʔ/, /zǰ/, /zv/, /zf/, /zr/, /zl/, /šd/, /šG/, /šf/, /šv/, /šr/, /šy/, /χl/, /χz/, /ht/, /hd/, /hv/, /hs/, /hz/, /hš/, /hl/, /hy/, /mt/, /md/, /mG/, /mʔ/, /mz/, /mn/, /ml/, /nb/, /nʔ/, /nχ/, /nh/, /rʔ/, /lb/, /ld/, /lʔ/, /lv/, /ls/, /lh/, /yb/, /yt/, /yʔ/, /yf/, /ys/, /yz/, /yš/, /yχ/, /yh/, /ym/, /yn/, /yr/

(2) F (10 clusters): /kt/, /gl/, /vr/, /sm/, /mp/, /mb/, /nk/, /rž/, /rh/, /yd/

(3) AF (6 clusters): /dr/, /sr/, /mr/, /nf/, /rn/, /lm/

(4) AE (2 clusters): /ms/, /yl/

(5) FE: /nt/

(6) FA: /ns/

(7) AFR: /ks/

(8) FAG: /bl/

A glance at Table 6.4 shows that there are few different clusters in the DPS with a stop as the first member, and many clusters with a sonorant, especially a liquid, as the first member. Moreover, consonants from the same manner subclass tend not to cluster in the DPS. These tendencies can be seen better in Table 6.5, which shows the number of clusters with members from the same and from different manner subclasses in the DPS, DNPS, and DBS. We see in Section A of Table 6.5 that there are only six clusters with members from the same manner subclass in the DPS. Loanwords have contributed 32 additional clusters of this type to the DBS, which is over five times as many as there are such clusters in the DPS.

Table 6.5. *The number of different types of CC clusters (described as different combinations of manner classes) in the DPS, DNPS and DBS*

A. Clusters with members from the same manner subclass			
	DPS	**DNPS**	**DBS**
stop + stop	3	15	18
affricate + affricate	0	0	0
fricative + fricative	3	16	19
nasal + nasal	0	1	1
liquid + liquid	0	0	0
glide + glide	0	0	0
Total	**6**	**32**	**38**

B. Clusters with members from different manner subclasses							
1. Clusters with increasing sonority				**2. Clusters with decreasing sonority**			
	DPS	**DNPS**	**DBS**		**DPS**	**DNPS**	**DBS**
stop + affricate	0	0	0	**affricate + stop**	0	3	3
stop + fricative	2	19	21	**fricative + stop**	9	12	21
stop + nasal	0	9	9	**fricative + affricate**	0	2	2
stop + liquid	2	11	13	**nasal + stop**	2	10	12
stop + glide	0	3	3	**nasal + affricate**	1	0	1
affricate + fricative	0	3	3	**nasal + fricative**	2	6	8
affricate + nasal	0	1	1	**liquid + stop**	9	4	13
affricate + liquid	0	1	1	**liquid + affricate**	2	0	2
affricate + glide	0	0	0	**liquid + fricative**	8	5	13
fricative + nasal	7	3	10	**liquid + nasal**	1	2	3
fricative + liquid	2	10	12	**glide + stop**	1	4	5
fricative + glide	0	3	3	**glide + affricate**	0	0	0
nasal + liquid	0	2	2	**glide + fricative**	0	6	6
nasal + glide	0	0	0	**glide + nasal**	0	2	2
liquid + glide	0	0	0	**glide + liquid**	0	2	2
Total	**13**	**65**	**78**	**Total**	**35**	**58**	**93**

Section B of Table 6.5 shows the number of clusters with members from different manner subclasses. Part B1 gives clusters with increasing sonority.[53] It shows that this cluster type is not very common in the DPS. There is only one pattern which is fairly frequent, namely, *fricative+nasal,* with a frequency of 7. Altogether the DPS contains 13 clusters with increasing sonority, i.e. about twice as many as it has clusters with members from the same manner subclass. Loanwords have added 65 new clusters with increasing sonority, i.e. five times the number of such clusters in the DPS.

[53] As before, the assumption here is that all manner subclasses represent different degrees of sonority, and members of the same manner subclass have the same degree of sonority.

Part B2 of Table 6.5 presents clusters with decreasing sonority. The table shows that this is the favoured type in the DPS; there are three cluster types with a fair number of different clusters in each: *fricative+stop* (9 clusters), *liquid+stop* (9 clusters), and *liquid+fricative* (8 clusters). More generally, the type *sonorant+obstruent* is the favoured one, and about half (25 clusters) of all the 54 clusters in the DPS are of this type. The DPS has altogether 35 different clusters with decreasing sonority. This is three times as many as there are clusters with increasing sonority in the DPS. Words of foreign origin have contributed 58 additional clusters with decreasing sonority. The foreign contribution is proportionally much less here than in the previous groups with less than one and a half as many clusters of this type as there are in the DPS.

Table 6.6 summarises the number of different clusters in the three types of sonority sequences (unchanged, increasing and decreasing sonority) in the DPS the DNPS and the DBS. The table also gives the number of cluster tokens in each data group. The table shows that the number of different clusters and the number of cluster tokens give the same preference order to the different sonority sequences, with one exception. The least favoured cluster type in every data group is the one where the sonority level remains the same. The most favoured type – with one exception – is sequences where sonority decreases towards the syllable edge. The exceptional case is the number of different clusters in the DNPS, where the the highest number of different clusters is found in the type where sonority increases towards the syllable edge.

Table 6.6: *The number of different clusters and the number of cluster tokens with unchanged, increasing and decreasing sonority in the DPS, DNPS and DBS*

	Number of different clusters			**Number of cluster tokens**		
	DPS	**DNPS**	**DBS**	**DPS**	**DNPS**	**DBS**
Clusters with unchanged sonority	6	32	38	10	59	69
Clusters with increasing sonority	13	65	78	28	207	235
Clusters with decreasing sonority	35	58	93	229	239	468
Total	54	155	209	267	505	772

But even though the preference order of different sonority sequences is the same, no matter whether we count the number of different clusters or the number

of cluster tokens, the two ways of calculation give different *strengths* of preference. If we consider the number of clusters, we get roughly the following ratios of the different sonority sequences (i.e. the ratio of unchanged to increasing and decreasing sonority): for the DPS, 1:2:6 (= 6:13:35); for the DNPS, 1:2:2; and for the DBS, 1:2:2.5. On the other hand, if we consider the number of cluster tokens, we get roughly the following ratios: for the DPS, 1:3:23 (= 10:28:229); for the DNPS, 1:3:4; and for the DBS, 1:3:6.5. We see that in the DPS the portion of cluster tokens with decreasing sonority is very big (the ratio of unchanged to decreasing sonority is 1:23), while this ratio is only 1:6 if we count the number of different clusters. In other words, clusters with decreasing sonority are by far the most frequent ones in the DPS. The ratios show that it is the portion of clusters with decreasing sonority that varies most according to the data groups (DPS, DNPS, DBS) and depending on whether we count the number of different clusters or the number of cluster tokens. On the other hand, there is little variation in the ratios of clusters with unchanged sonority to clusters with increasing sonority.

Tables 6.5 and 6.6 show that there are changes in Persian CC clusters as a result of loanwords. But what exactly has changed? We saw that clusters from words of Persian origin show a strong tendency towards dissimilarity of cluster members and an even stronger tendency towards decreasing sonority in clusters. These tendencies can still be seen in Modern Persian too, but they are much weaker. The tendency towards decreasing sonority in clusters is especially weaker. Nevertheless, we can not say that loanwords have changed the basic pattern of CC clusters. The DPS allows, in addition to clusters with decreasing sonority, also clusters with members from the same manner subclass, and clusters with increasing sonority – and one such pattern (*fricative+nasal*) has a fair number of different clusters. It is not the basic pattern of CC clusters that has changed, or even the preferences (of dissimilarity and decreasing sonority of cluster members), but rather the *strength* of the preferences.

So far, we have seen what different kinds of clusters there are in the DPS and how they differ from those in the DBS. Let us finally have a look at cluster frequencies in the DPS, and compare them with those in the DBS.

Appendix 9 lists all the CC clusters in the DBS, together with their frequencies, and Appendix 15 gives the CC clusters with their frequencies in the DPS. If we take the ten most frequent clusters from both lists, we get the following top ten lists for the DBS and the DPS:

	DPS			DBS	
	Clusters	**Frequency**		**Clusters**	**Frequency**
1.	st	25	**1.**	st	33
2.	ng	22	**2.**	ng	26
3.	št	21	**3.**	rd	26
4.	nd	20	**4.**	št	22
5.	rd	19	**5.**	nd	22
6.	χt	17	**6.**	rs	17
7.	ft	14	**7.**	χt	17
8.	rz	10	**8.**	ft	15
9.	šk	9	**9.**	hr	14
10.	nǰ	9	**10.**	rz	14

We see that the top five clusters are the same on both lists, with numbers 3 to 5 in somewhat different orders. Altogether eight out of the ten most frequent clusters are the same on both lists. In all the top ten clusters in the DPS, sonority decreases towards the syllable edge. We noted before that this is also true of all but one cluster, the deviant cluster in /hr/, in the top ten DBS list. Moreover, all clusters on both lists have members from different manner classes. Thus, if we look at the most frequent clusters, we see that both the tendency towards dissimilarity of cluster members and the tendency towards decreasing sonority are strong in the DPS material as well as in the DBS data.