Problem 3:

SSD: correct_count = 85 out of 100, correctness = 0.85

DPD: correct_count = 97 out of 100, correctness = 0.97

Problem 4:

["bg" Bulgarian, "de" German, "el" Greek, "en" English, "es" Spanish, "fr" French, "it" Italian, "nl" Dutch, "pl" Polish, "ru" Russian]

SSD	Bulgarian	Dutch	English	French	German	Greek	Italian	Polish	Russian	Spanish
Bulgarian	8	0	0	0	0	0	0	0	2	0
Dutch	0	10	0	0	0	0	0	0	0	0
English	2	0	5	0	0	0	2	0	0	1
French	0	0	0	9	0	0	0	0	0	1
German	0	1	1	0	8	0	0	0	0	0
Greek	0	0	0	0	0	8	0	0	2	0
Italian	0	0	0	0	0	0	10	0	0	0
Polish	0	0	0	0	0	0	0	10	0	0
Russian	0	0	1	0	0	0	0	0	9	0
Spanish	0	0	0	1	0	0	1	0	0	8

DPD	Bulgarian	Dutch	English	French	German	Greek	Italian	Polish	Russian	Spanish
Bulgarian	10	0	0	0	0	0	0	0	0	0
Dutch	0	10	0	0	0	0	0	0	0	0
English	0	1	9	0	0	0	0	0	0	0
French	0	0	0	9	0	0	0	0	0	1
German	0	0	0	0	10	0	0	0	0	0
Greek	0	0	0	0	0	10	0	0	0	0
Italian	0	0	0	0	0	0	10	0	0	0
Polish	0	0	0	0	0	0	0	10	0	0
Russian	0	0	0	0	0	0	0	1	9	0
Spanish	0	0	0	0	0	0	0	0	0	10

Do you notice any patterns in your matrices?

For both matrixes there is a diagonal shows number of languages' files can be guessed correctly. And Most of the languages file can be guessed correctly.

Do the methods tend to make similar errors?

No

Are some languages more likely to be confused, or are the errors mostly evenly distributed?

SSD:

English is more likely to be confused, only 5 but of 10 English ciphertext files are guessed as English.

Russian is also confused. Even if there are 9 out of 10 Russian ciphertext files are guessed as Russian. There are also have 4 other languages' file are guessed as Russian

Italian is another one confused. If we only look at Italian files, it all can be guessed correctly, but there are 3 other language files guessed as Italian.
For DPD there is no confused language shows.

Speculate on why this might be.
The reason for the differences between SSD and DPD is because for SSD we are looking for the relative frequency of the characters in different languages. Some languages may have same or similar characters and when we look at the character frequency, it might return the similar frequency for the same character. That's why SSD's guess result has lower correctness. But for DPD we are looking for the architecture of the words in different languages. Even if two languages have the similar or same characters, the word architecture is not. The word architecture for a language has its own features, it is hard to mix two languages together by word architecture.