Q6.

1.

$ python3 q61.py

The program is in q61.py

When running the program, the output will be printed.

Experimental Condition	Overall Accuracy %
BigramLetterLangId	92%
BigramWordLangId-AO	60%
BigramWordLangId-GT	68%
TrigramWordLangId-KBO	67%

2.

$ python3 q61.py

The program is in q61.py

BigramLetterLangId	EN	FR	GR
EN	-	0% gold French sentences identified as English sentences	100%
FR	66.7%	-	33.2%
GR	100%	0	-

BigramWordLangId-AO	EN	FR	GR
EN	-	45.8%	54.2%
FR	0	-	0
GR	0	0	-

BigramWordLang-GT	EN	FR	GR
EN	-	0	0
FR	50%	-	50%
GR	70.3%	29.7%	-

TrigramWordLangId-KBO	EN	FR	GR
EN	-	0	100%
FR	46.2%	-	53.8%
GR	67.6%	32.4%	-

3.

Using there two equations to calculate the perplexity

$$H(X) = -\sum_{x=1}^{x=n} p(x)\log_2 p(x)$$

$$PP(W) = 2^{H(W)}$$

To achieve these goal, some code are added to related programming documents.
The result is in the following table

Experimental Condition	EN	FR	GR
BigramLetterLangId	1.00000005	1.00000006	1.00000007
BigramWordLangId-AO	1.000002	1.000001	1.000002
BigramWordLangId-GT	1.50	1.68	1.69
TrigramWordLangId-KBO	2.23	2.48	2.25

This result may not be good.
From bigram letter model to trigram word model, the perplexity is increasing, which is different from the textbook.
One possible reason is that with the limitation of training data and the model algorithm, the probability to find the aim pattern is more difficult in bigram model and trigram model.
To solve these problem, more training data and algorithm improvement is needed.