Informatické zpracování dovednosti čtení s porozuměním a úlohy odpovídání na otázky se zabývají oblastmi zpracování přirozeného jazyka a vyhledávání informací. Čtení s porozuměním je schopnost modelu číst a zpracovat text a porozumět jeho významu. Jednou z jeho aplikací je úloha odpovídání na otázky, které se zabývá vytvořením systému, který dokáže v textu automaticky najít odpověď na určitou otázku, která přímo souvisí s obsahem dokumentu. Pro angličtinu se jedná se o hojně studovanou úlohu, pro kterou existují obrovská tréninková data a spousty modelů. Pro tuto oblast však neexistují žádné modely ani data v češtině.

Tato práce se zaměřuje na vytvoření systémů pro úlohy čtení s porozuměním a odpovídání na otázky v českém jazyce, a to bez nutnosti ručně vytvářet česká data. Hlavním cílem je automatické vytvoření českých trénovacích a testovacích dat a vytvoření modelů pro úlohu odpovídání na otázky v češtině. Využívá se existujících anglických dat a modelů za pomoci překladu a mezijazykového přenosu znalostí a následného porovnání výsledků a výběru modelu s nejlepšími výsledky. Nejprve jsme přeložili volně dostupná anglická data pro úlohu odpovídání na otázky SQuAD 1.1 a SQuAD 2.0 do češtiny, abychom vytvořili trénovací a testovací data. Poté jsme přetrénovali a vyhodnotili několik základních modelů BERT a multilingválních modelů XLM-RoBERTa používaných pro tuto úlohu v angličtině. Nejlepší výsledky jsme získali s modelem XLM-RoBERTa natrénovaným na angličtině a vyhodnoceným přímo na českých datech. Tento model dosáhl velmi dobrých výsledků, podobně jako model natrénovaný na přeložených českých datech. Výsledek získaný z XLM-Roberta však považujeme za nejlepší, protože model během tréninku nevyžaduje žádná česká data. To také dokazuje, že mezijazykový přenos znalostí je v případě těchto neuronových modelů velice flexibilní a umožňuje porozumění významu a obsahu textu i v jazyce, pro který nebyl původně model trénován.