This dissertation investigates the challenges faced by the Irish language, biases within Irish open source datasets, the efficacy of gloss analysis in determining datasets to improve Automatic Speech Recognition (ASR) systems performance, and the impact of transfer learning on ASR performance. Research on this topic reveals bias against low-resource languages specifically related to dialects and a lack of technological support due to limited research and funding, as well as underlying bias that exists across ASR systems for any language and the damaging effects of them. Analysis of the Mozilla Common Voice datasets was performed to uncover bias and to link the performance of models trained on these datasets to the data. Under-representation of certain groups was found, particularly women and those over the age of 60. Models were fine-tuned on a set of datasets before being fine-tuned again on Irish to prove that transfer learning on selected languages is a promising approach to enhance ASR performance, with findings indicating a 9.5% improvement in Word Error Rate (WER) through pre-training on English data before fine-tuning on Irish. I also found that bias in the models could not be solely linked back to the metadata from the datasets, meaning that a more in-depth investigation must be done into where the discrimination in the model performance is coming from. Surprisingly, the study finds a lower direct correlation between dataset size and model performance than expected, highlighting the importance of dataset selection. Gloss analysis offers some insights into suitable datasets for pre-training but can't distinguish a dataset that will boost the performance of a model from one that will negatively impact it.