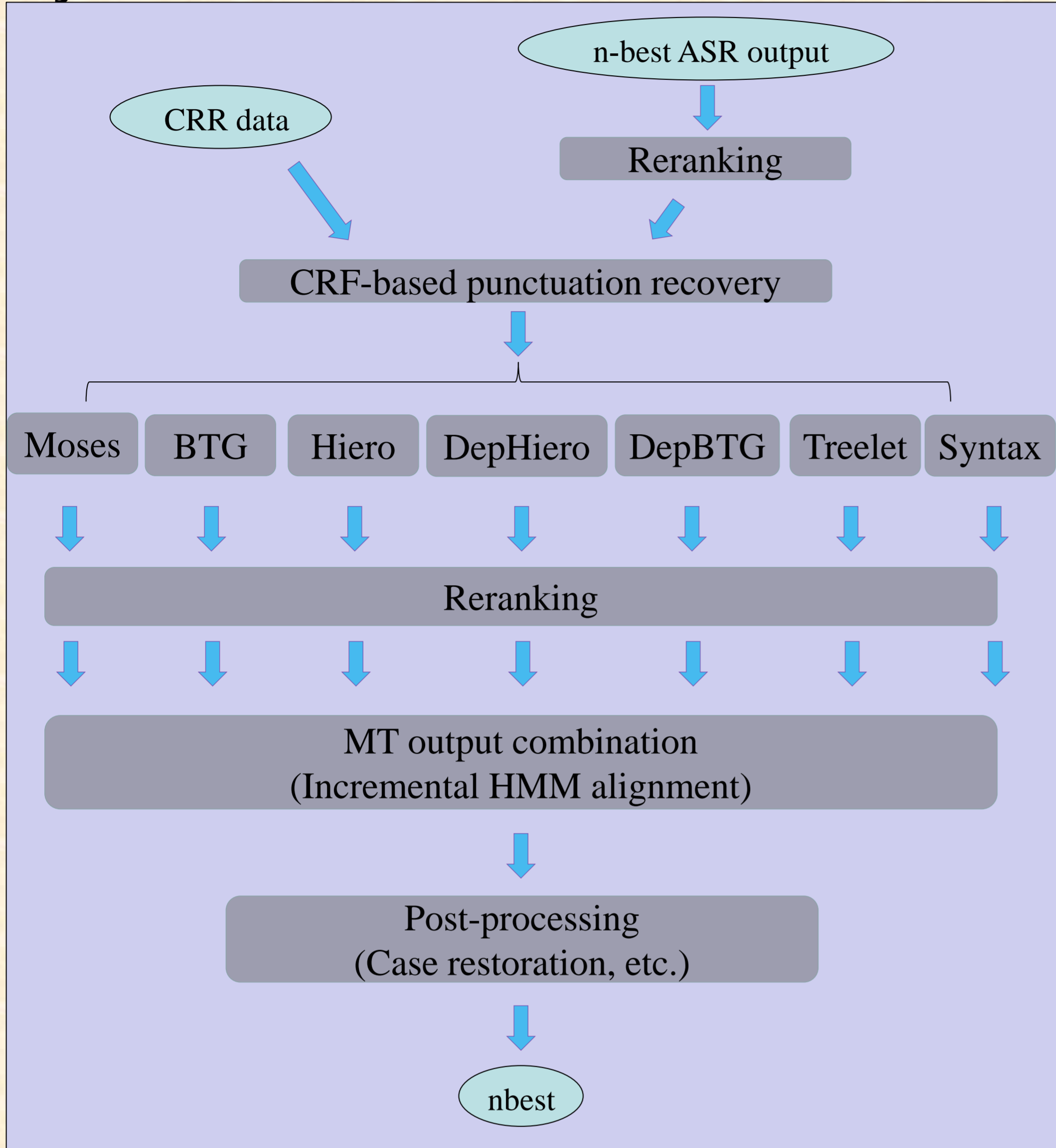


## System Architecture



## Reranking of ASR/MT Hypotheses

**Data:** a set of hypotheses with Oracles

**Objective function:** Bleu

**Training method:** Max-Bleu training

**Model:** Log-linear

**Features for ASR n-best output:**

1. ASR output scores
2. Language model probabilities
3. Count of characters/words

|                 | BTG   | Hiero | DepHiero | DepBTG | Syntax |
|-----------------|-------|-------|----------|--------|--------|
| original 1-best | 41.05 | 40.43 | 42.80    | 39.16  | 39.28  |
| reranked 1-best | 41.98 | 41.24 | 43.57    | 39.93  | 39.99  |

**Features for MT n-best output:**

1. N-gram posterior probabilities
2. Sentence length probabilities
3. Language model probabilities
4. Length ratio between source input and translation

|                 | BTG   | Hiero | DepHiero | DepBTG | Syntax |
|-----------------|-------|-------|----------|--------|--------|
| original 1-best | 45.69 | 45.45 | 48.47    | 44.15  | 44.45  |
| reranked 1-best | 46.07 | 48.42 | 47.80    | 45.74  | 45.65  |

Experiment settings:

Task = CE; Tuning set = devset8+c2e.DIALOG

Test set = devset9; Training set = everything else

## Punctuation Recovery

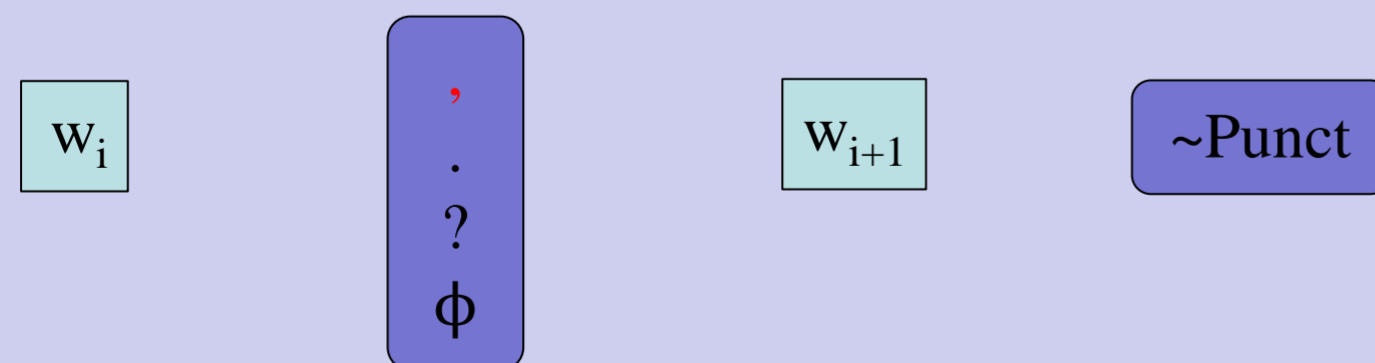
Two-stage classification:

1. decide if a word is followed by punctuation mark or not



F-score = 99.07% (on devset11 for EC/ASR)

2. decide exactly which punctuation mark it is for each punctuation position



F-score = 94.56% (on devset11 for EC/ASR)

## Experiments

With manual rules and alignment combination:

**For CE/CRR:**

|             | BTG   | Hiero | DepHiero | DepBTG | Syntax |
|-------------|-------|-------|----------|--------|--------|
| Raw output  | 46.54 | 45.45 | 48.84    | 45.43  | 46.55  |
| Reranked    | 47.76 | 48.42 | 50.02    | 47.87  | 48.42  |
| Combination | 51.43 |       |          |        |        |

**For CE/ASR:**

|             | BTG   | Hiero | DepHiero | DepBTG | Syntax |
|-------------|-------|-------|----------|--------|--------|
| Raw output  | 43.34 | 44.14 | 44.44    | 42.31  | 42.33  |
| Reranked    | 43.53 | 44.06 | 45.87    | 43.23  | 43.21  |
| Combination | 46.53 |       |          |        |        |

Experiment settings:

Task = CE; Tuning set = devset8+c2e.DIALOG

Test set = devset9; Training set = everything else