Finding Language-Specific Features

Grammatical Feature

Morphosyntactic Feature

Spanish Corpus

Yo soy estudiante

SAE-LAPE → FFN ⤏ Decompose

Morphological Feature

Syntatic Feature

Chinese Corpus

我 是 学生

SAE-LAPE → FFN ⤏ Decompose
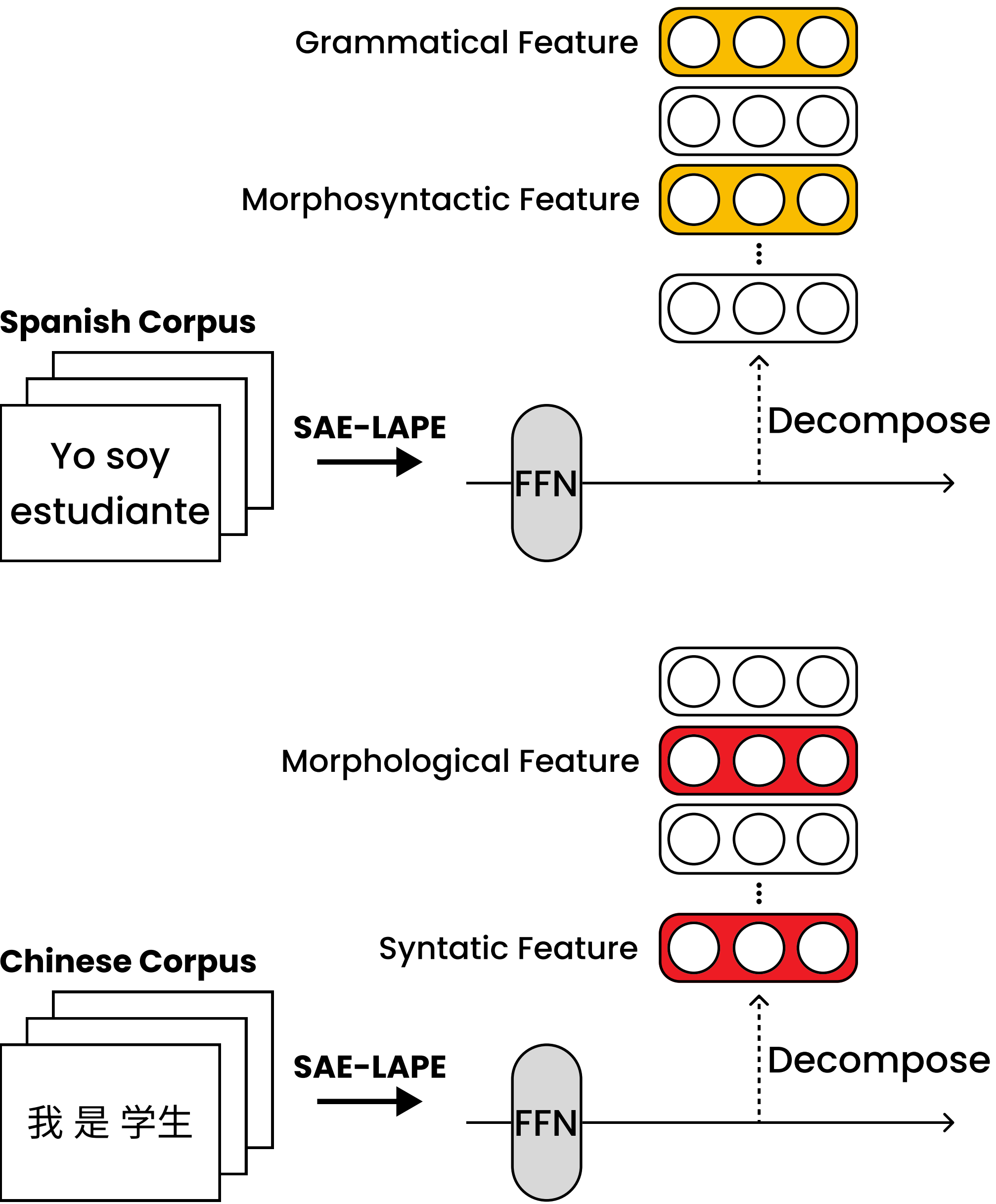
Steering with Language-Specific Features

[BOS] — FFN → print("Hello, World!")

Multiple Spanish Features

× alpha

[BOS] — FFN ⊕ → print("¡Hola, Mundo!")

Multiple Chinese Features

× alpha

[BOS] — FFN ⊕ → print("你好，世界!")
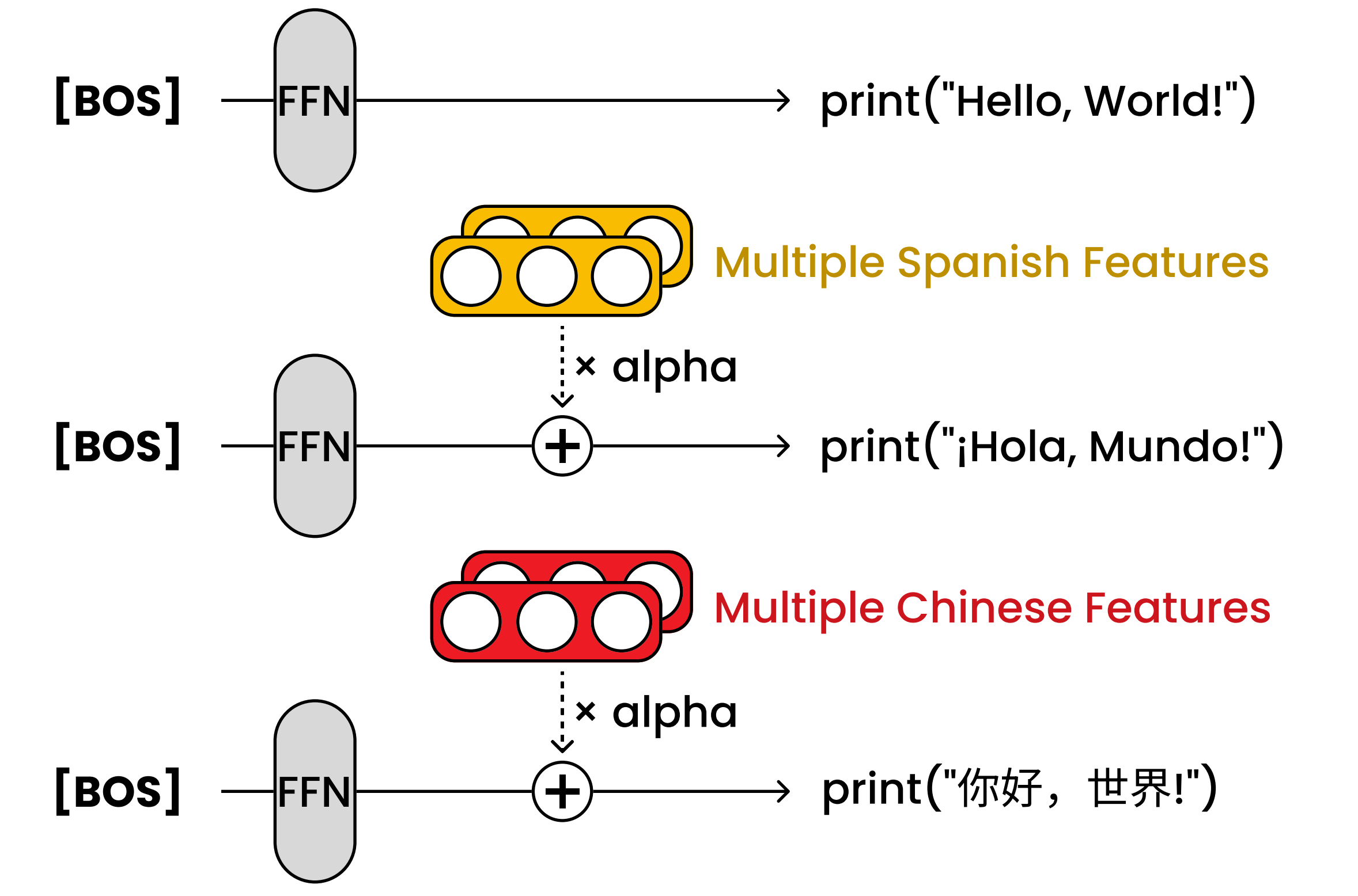
Classification with Language-Specific Features

2 occurrences

0 occurrence

Classify

Muchas gracias — FFN ⤏ Count

Spanish