Probabilistic ParaMor

Christian Monson
Kristy Hollingshead
Brian Roark

Probabilistic ParaMor

A Baseline for Unsupervised Induction

The **ParaMor** System

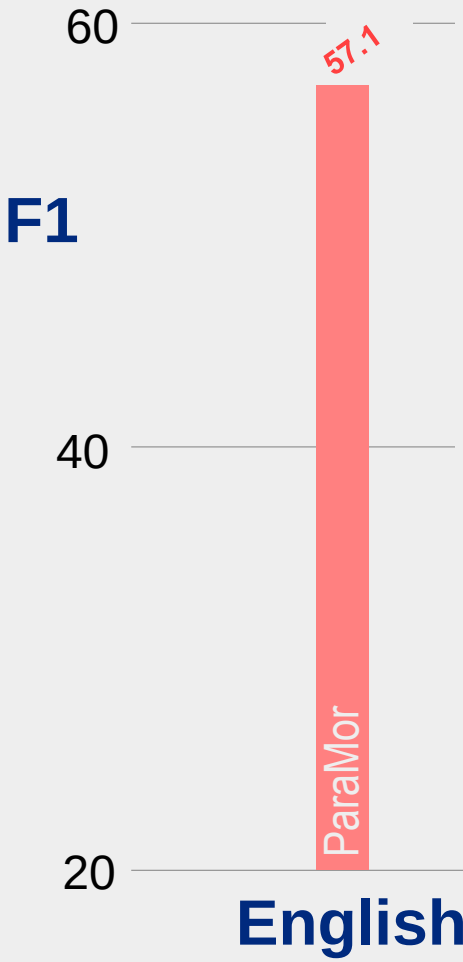
Morpho Challenge 2008

New Work

Probabilistic ParaMor

Morpho Challenge **2009**

Results of Linguistic Evaluation 2008

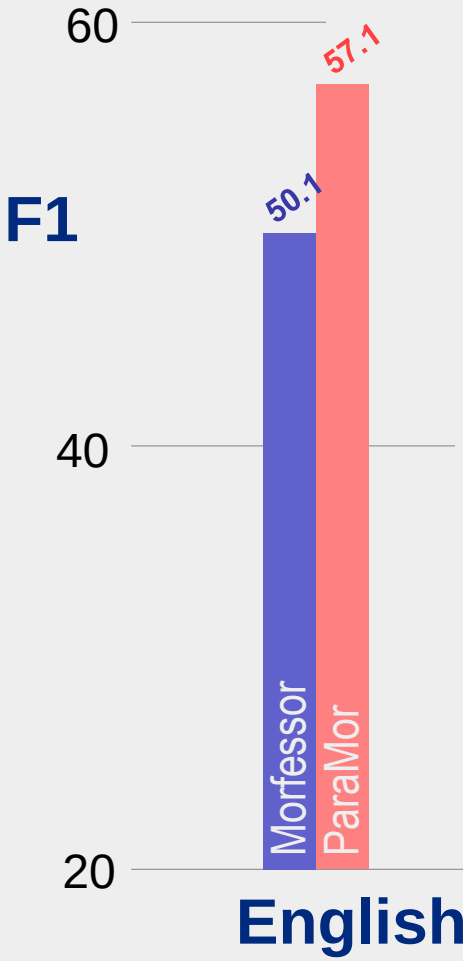


ParaMor

- Built for my Ph.D. thesis
- Rule-based
- Builds linguistically motivated paradigms

Word	ParaMor
dispute's	disput +e +'s
juried	juri +ed
losses	los +ses
…	…

Results of Linguistic Evaluation 2008



Morfessor

Baseline system for Morpho Challenge

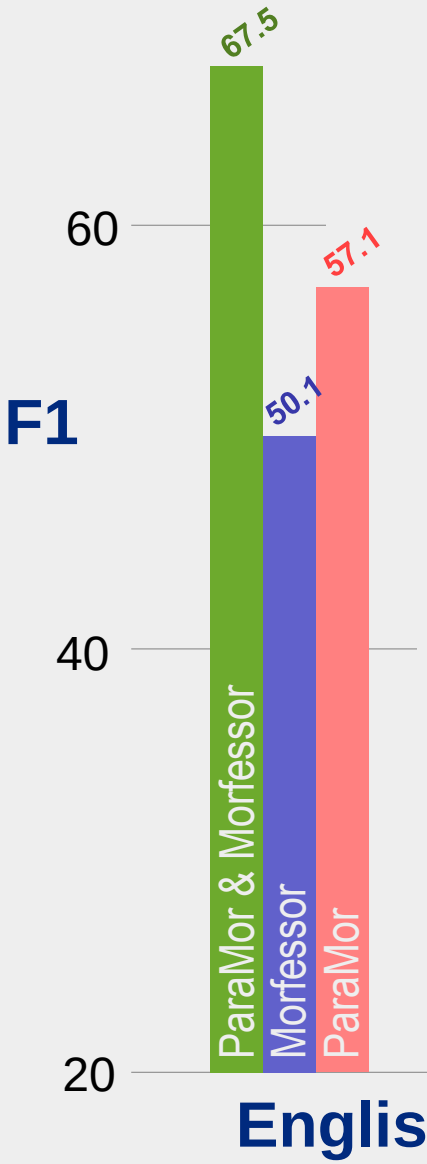
Freely available

Minimum Description Length

Black box segmentation

Word	ParaMor	Morfessor
dispute's	disput +e +'s	dispute/STM +'s/SUF
juried	juri +ed	juried/STM
losses	los +ses	losses/STM
…	…	…

Results of Linguistic Evaluation 2008



Submit **Both** analyses:

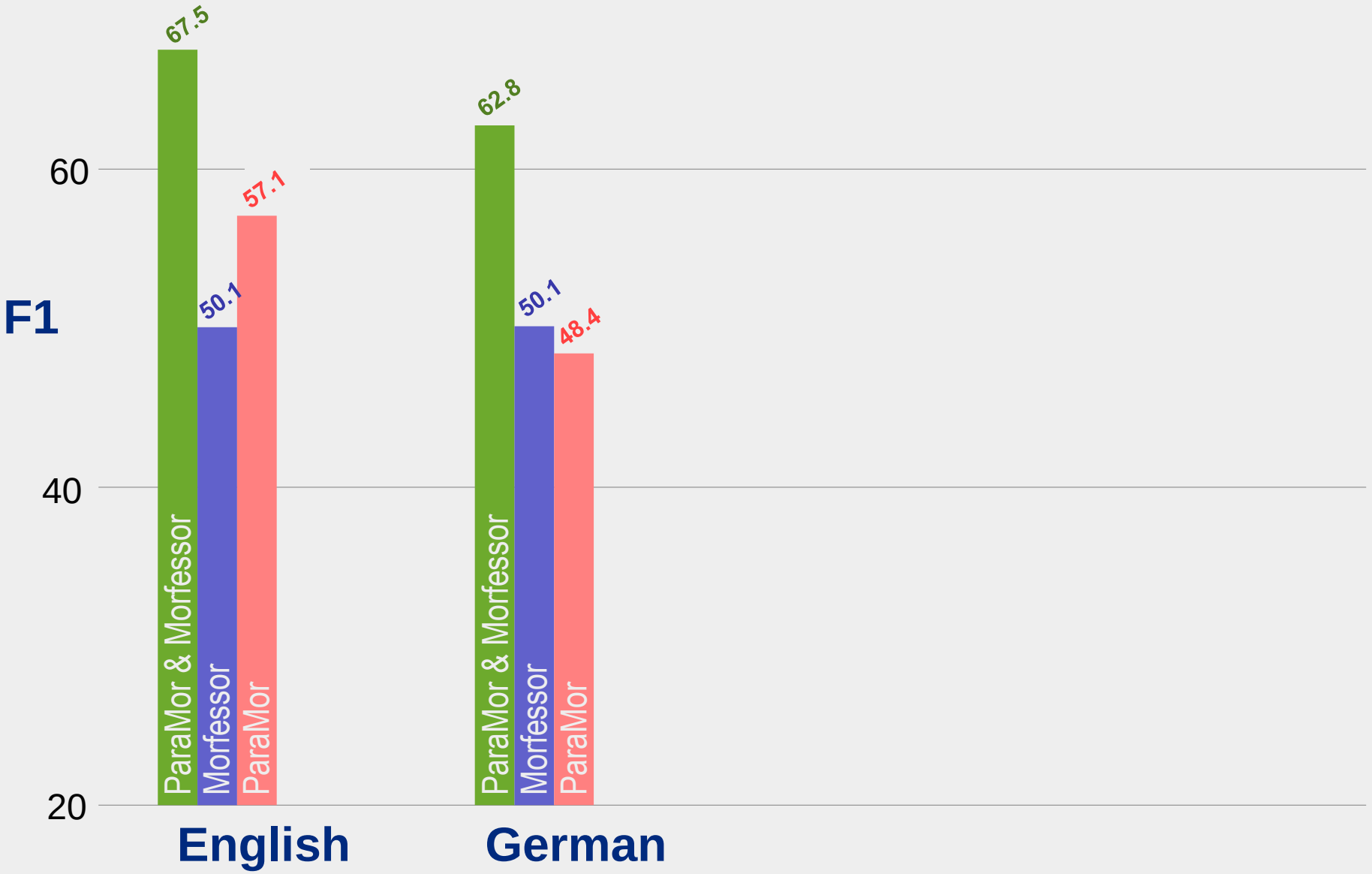
A ParaMor analysis and

A Morfessor analysis

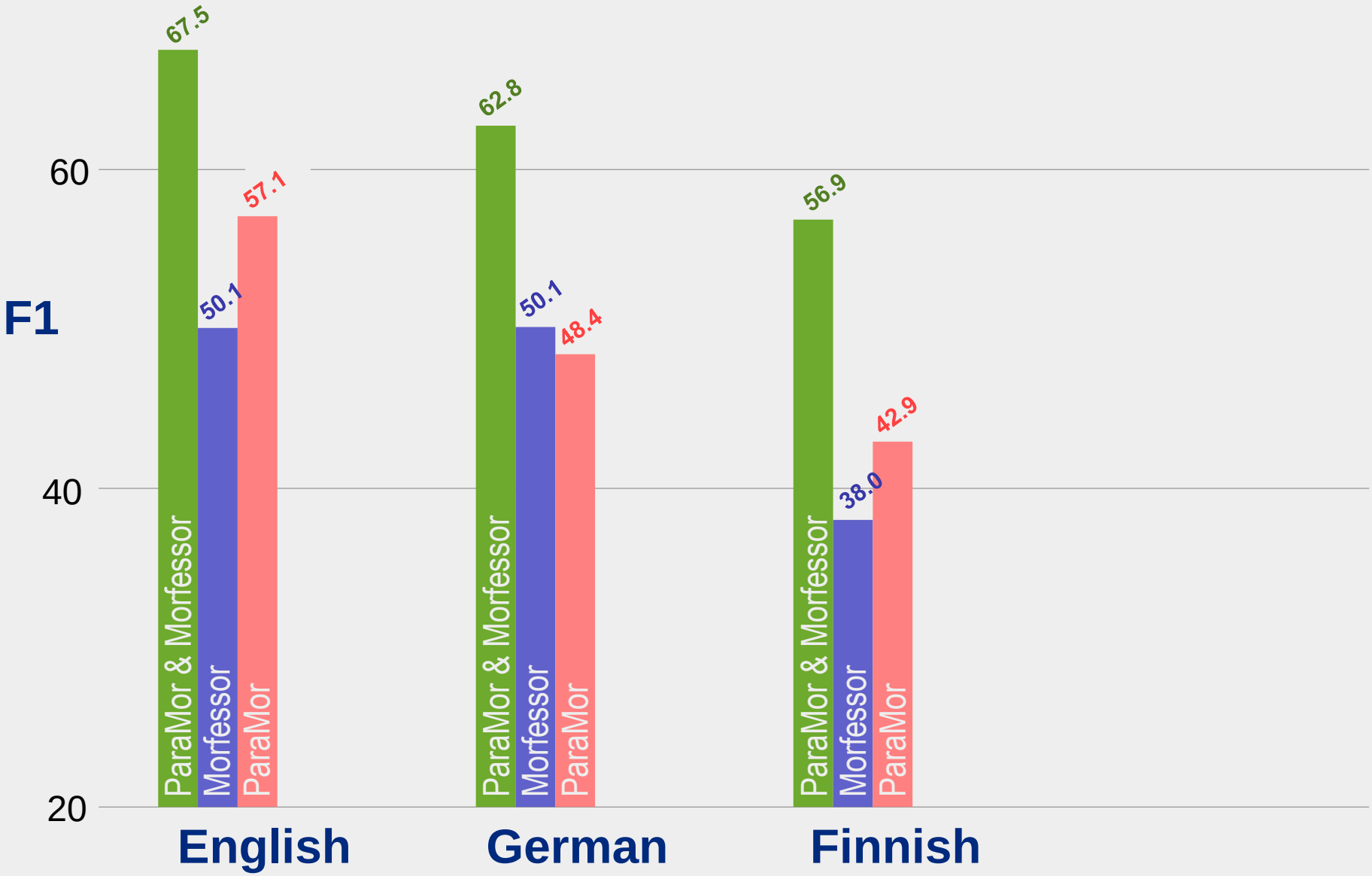
Let the evaluation script average across the submitted analyses

Word	ParaMor	Morfessor
dispute's	disput +e +'s	dispute/STM +'s/SUF
juried	juri +ed	juried/STM
losses	los +ses	losses/STM
…	…	…

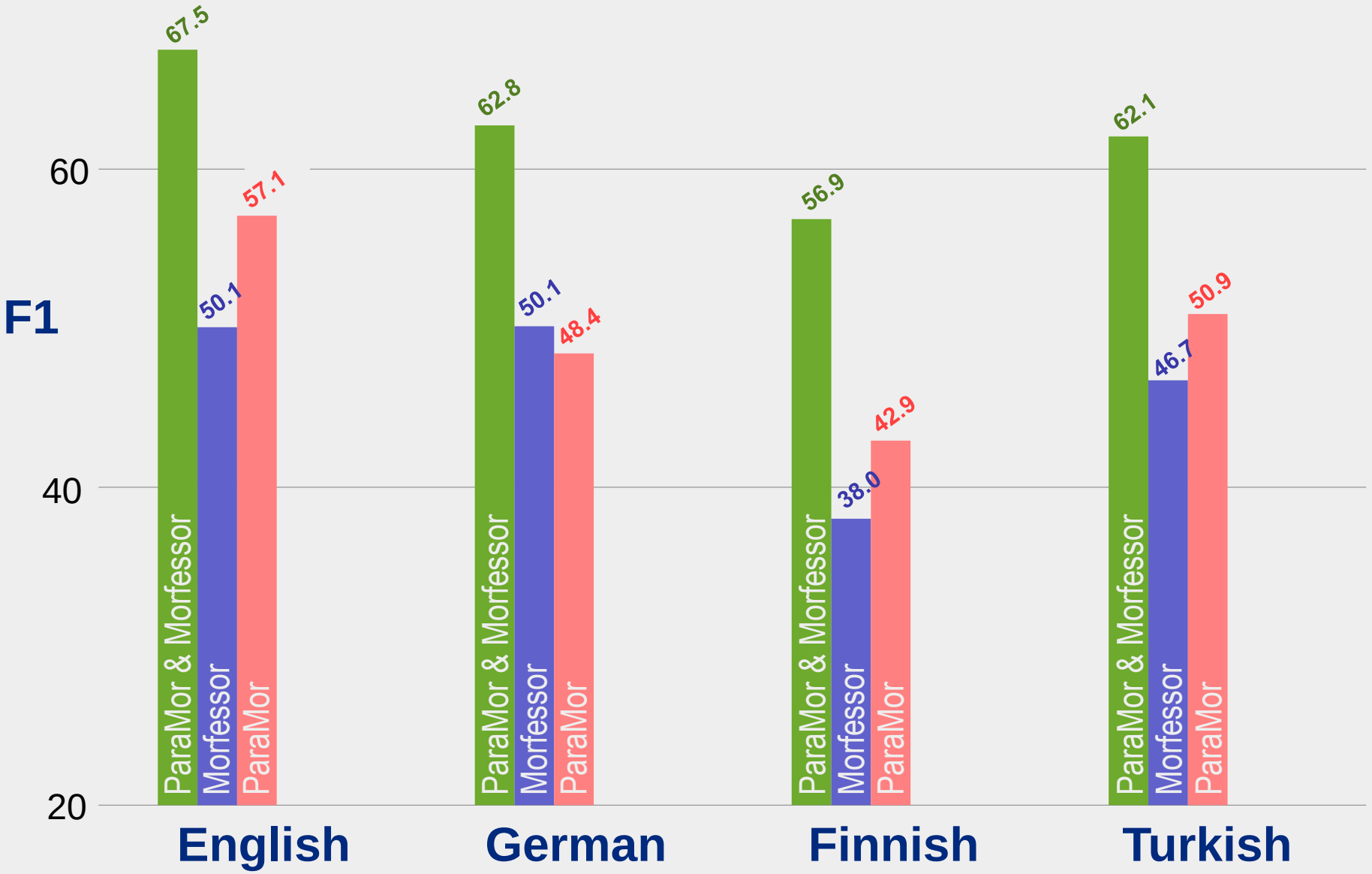
Results of Linguistic Evaluation 2008



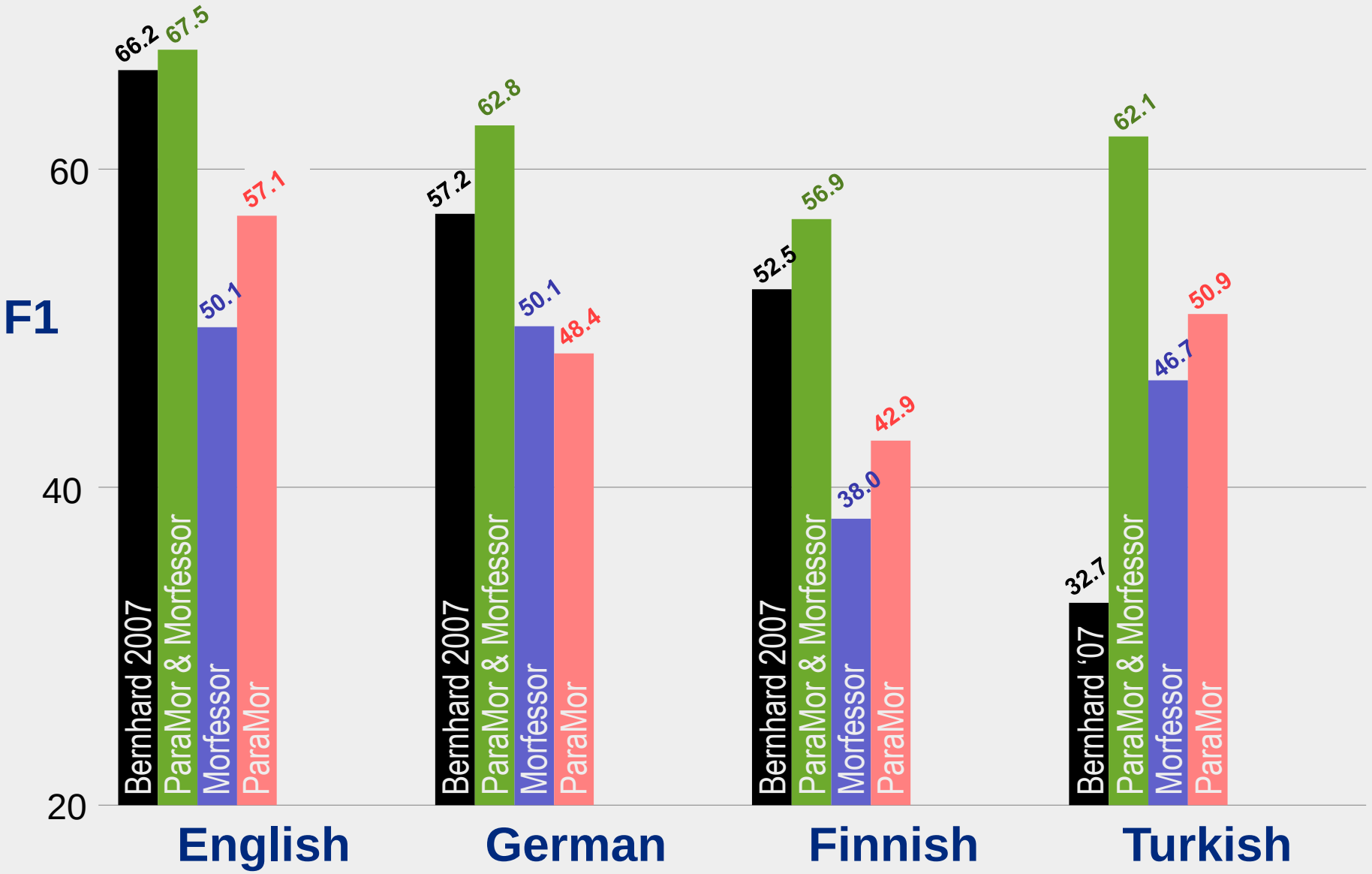
Results of Linguistic Evaluation 2008



Results of Linguistic Evaluation 2008



Results of Linguistic Evaluation 2008



Improvements to ParaMor

A Baseline for Unsupervised Induction

- The **ParaMor** System
- **Morpho Challenge 2008**

New Work

- **Probabilistic** ParaMor
- Morpho Challenge **2009**

One Joint Analysis

2 Separate Analyses is Unsatisfying

Fuse ParaMor and Morfessor

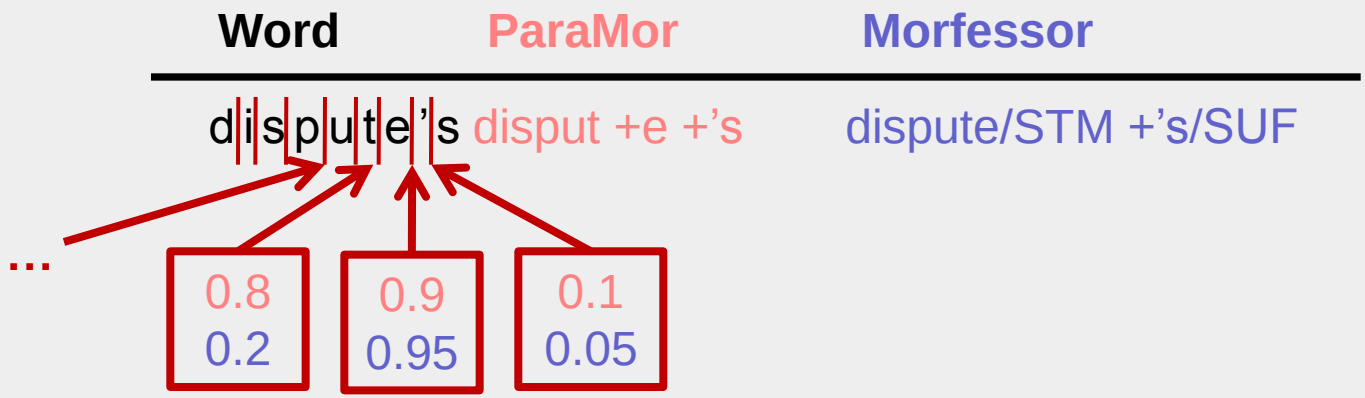
A **single** definitive segmentation

Word	ParaMor	Morfessor
dispute's	disput +e +'s	dispute/STM +'s/SUF
juried	juri +ed	juried/STM
losses	los +ses	losses/STM
…	…	…

We Want 1 Joint Analysis of Each Word

The Plan

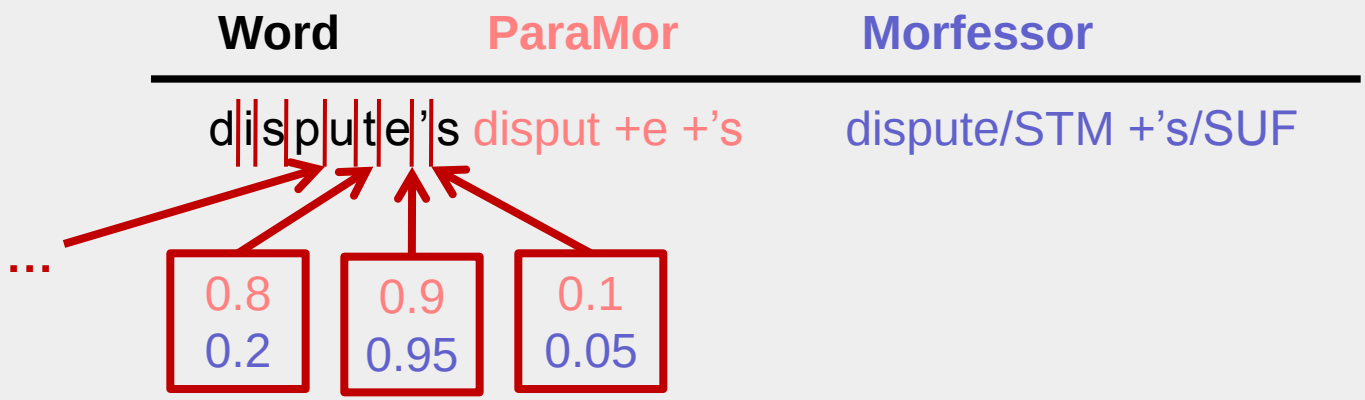
Score each character boundary, *i*



We Want 1 Joint Analysis of Each Word

The Plan

Score each character boundary, i



Compute weighted average

$Ave_i = \lambda \text{ParaMor}_i + (1 - \lambda)\text{Morfessor}_i$

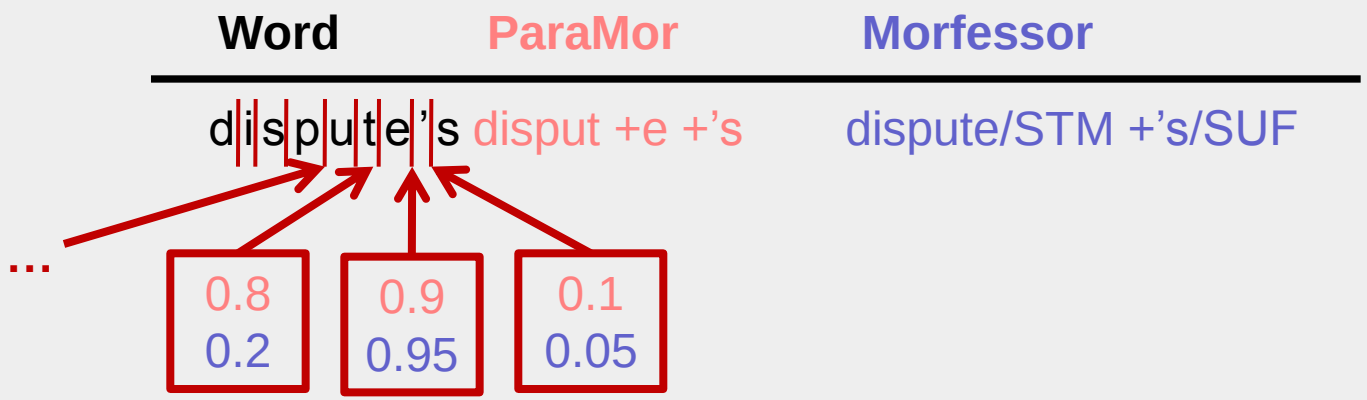
Segment when $Ave_i > \tau$,

τ : an empirical threshold

We Want 1 Joint Analysis of Each Word

The Plan

Score each character boundary, *i*



Where do these scores come from?

ParaMor is rule-based
➔ No numeric scores

Morfessor is a Blackbox
➔ No numeric scores

Tagging = Sequence Labeling

English Part-of-Speech Tagging

97.1% Accuracy

NP Chunking

F_1: 93.5%

Chinese Word-Segmentation

Above 90% Accuracy

Tagging = Sequence Labeling

English Part-of-Speech Tagging

97.1% Accuracy

NP Chunking

F_1: 93.5%

Chinese Word-Segmentation

Above 90% Accuracy

Morphological Segmentation

Tagging for Morphological Segmentation

Tag Each Character in Each Word As:

- A **Left-Edge** of a Morpheme
- A **Non-Left-Edge**

Our Tagger

- Finite State
- Log-Linear Model
  - 2nd order Markov assumption
  - Training
    - Perceptron Algorithm
    - Viterbi scoring
  - Testing
    - Forward-Backward

Now, Just Wait a Second!

Morpho Challenge

UNSUPERVISED induction

We're training a **SUPERVISED** tagger

Now, Just Wait a Second!

Morpho Challenge

UNSUPERVISED induction

We're training a **SUPERVISED** tagger

Mimic an Unsupervised Systems

Analyze data with ParaMor (or Morfessor)

Treat ParaMor's Analyses as THE TRUTH

Train a tagger to mimic ParaMor (or Morfessor)

Tagger Score = Segmentation Confidence

Hungarian Development Data

Hunglish Corpus (Varga et al., 2009)

500,000 word subset

	ParaMor		Morfessor	
Hungarian (τ=0.5)	Original	Tagger Mimic Mimic Tag Acc: **96.8%**	Original	Tagger Mimic Mimic Tag Acc: **95.5%**
Precision				
Recall				
F_1				

Hungarian Development Data

Hunglish Corpus (Varga et al., 2009)

500,000 word subset

Hunmorph (Trón et al., 2005)

Hand-Built Morphological Analyzer

	ParaMor		Morfessor	
Hungarian (τ=0.5)	Original	Tagger Mimic Mimic Tag Acc: 96.8%	Original	Tagger Mimic Mimic Tag Acc: 95.5%
Precision	64.8	63.7	72.9	72.5
Recall	30.4	32.2	31.3	28.8
F_1	41.4	42.7	43.8	41.3

We Can Do Better

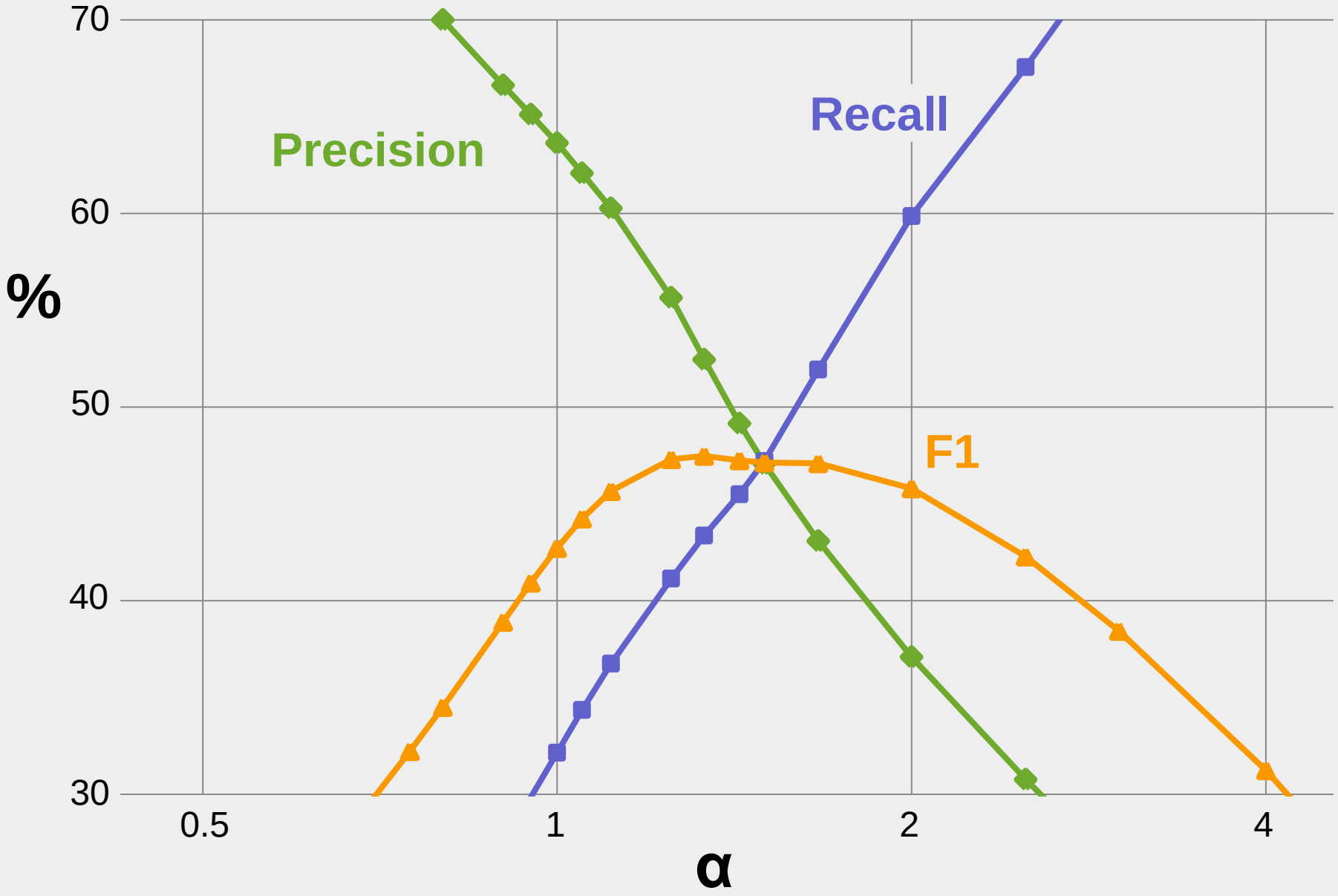
Optimize for F_1

Sort the tagger's scores of all character boundaries in the corpus

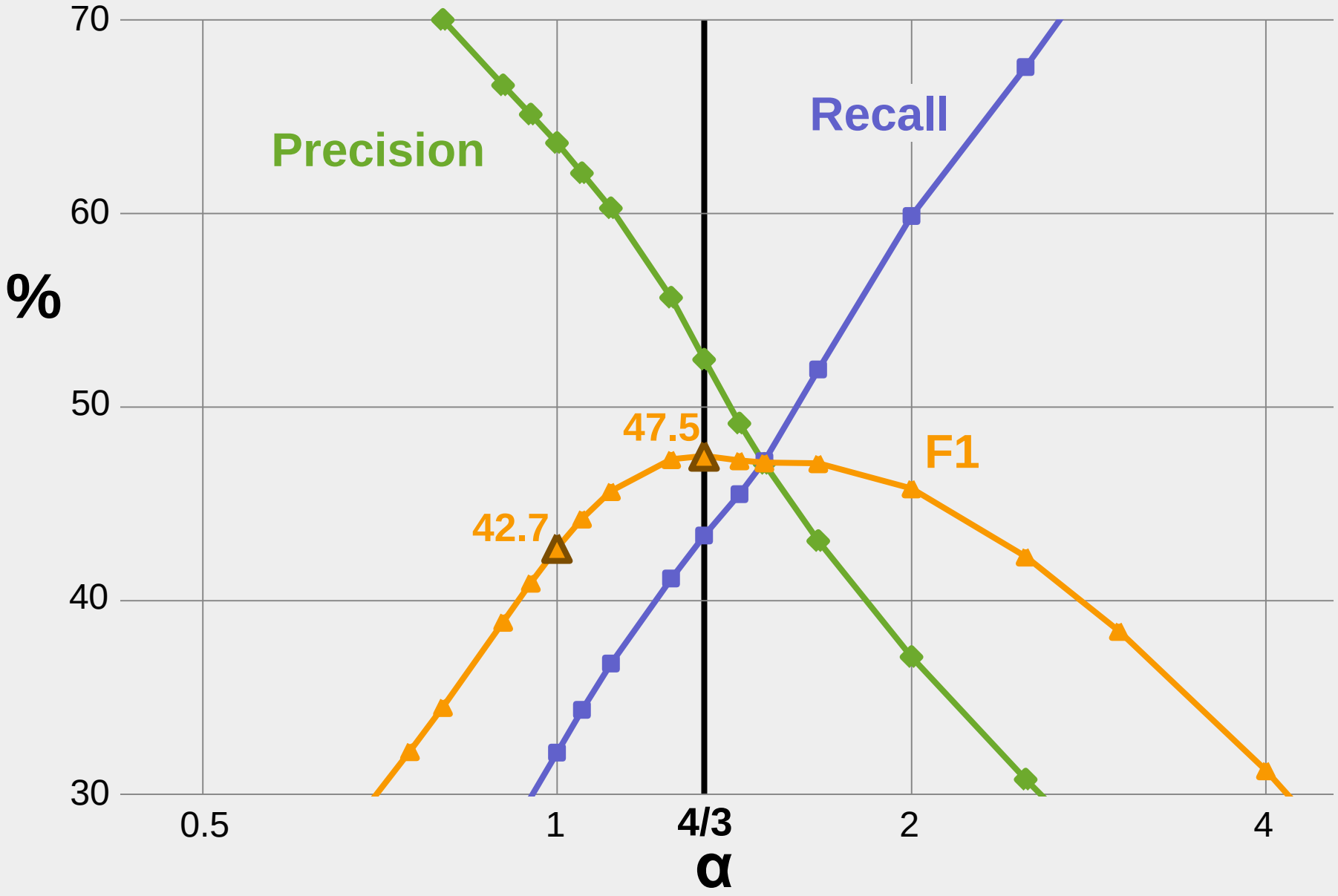
Let k be the # of segmentations at τ=0.5

Segment at $\alpha\ k$ character boundaries

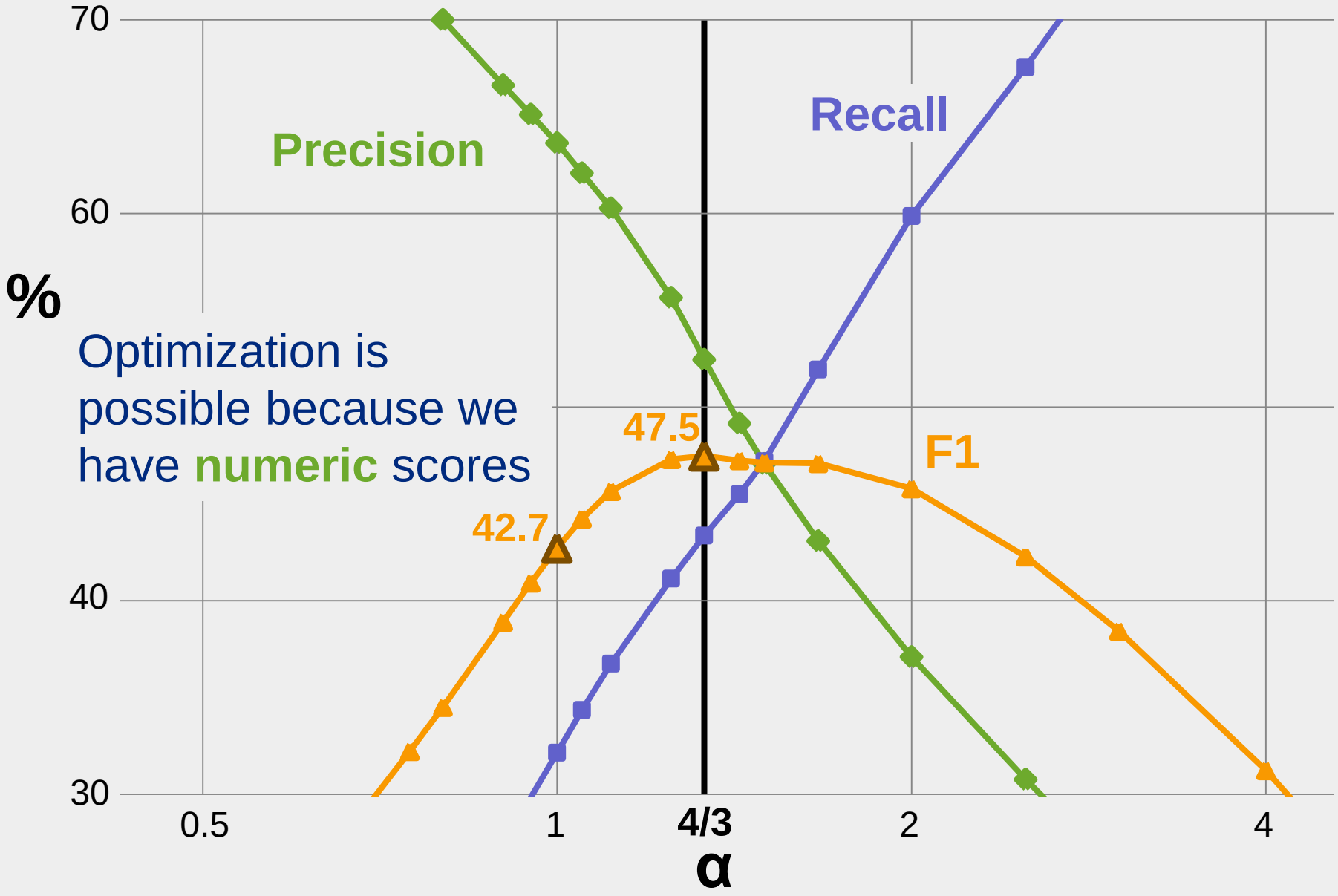
ParaMor Hungarian – Adjusting α



ParaMor Hungarian – Adjusting α



ParaMor Hungarian – Adjusting α



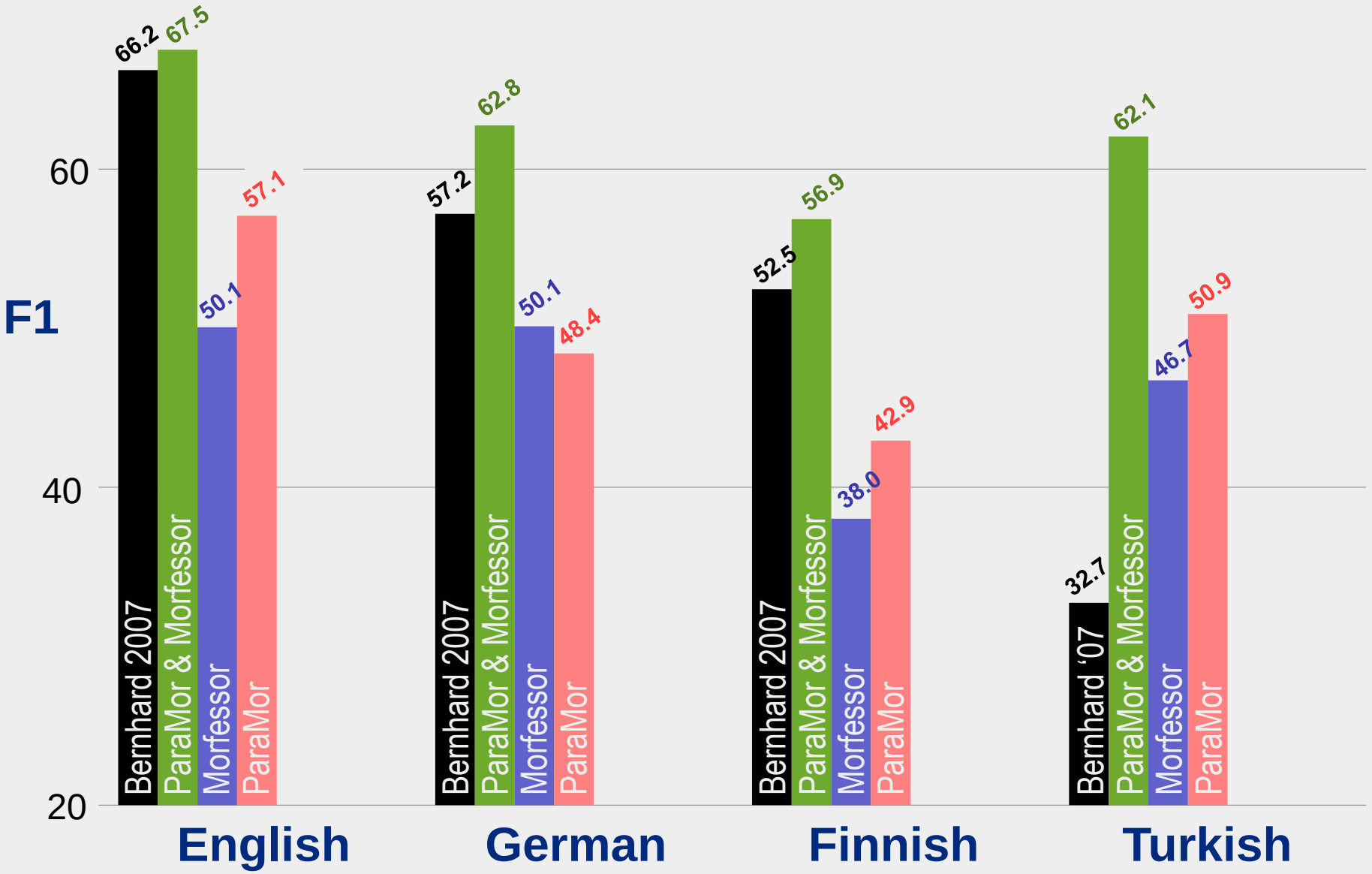
Optimize toward a Metric

A Statistical Framework Allows Optimization

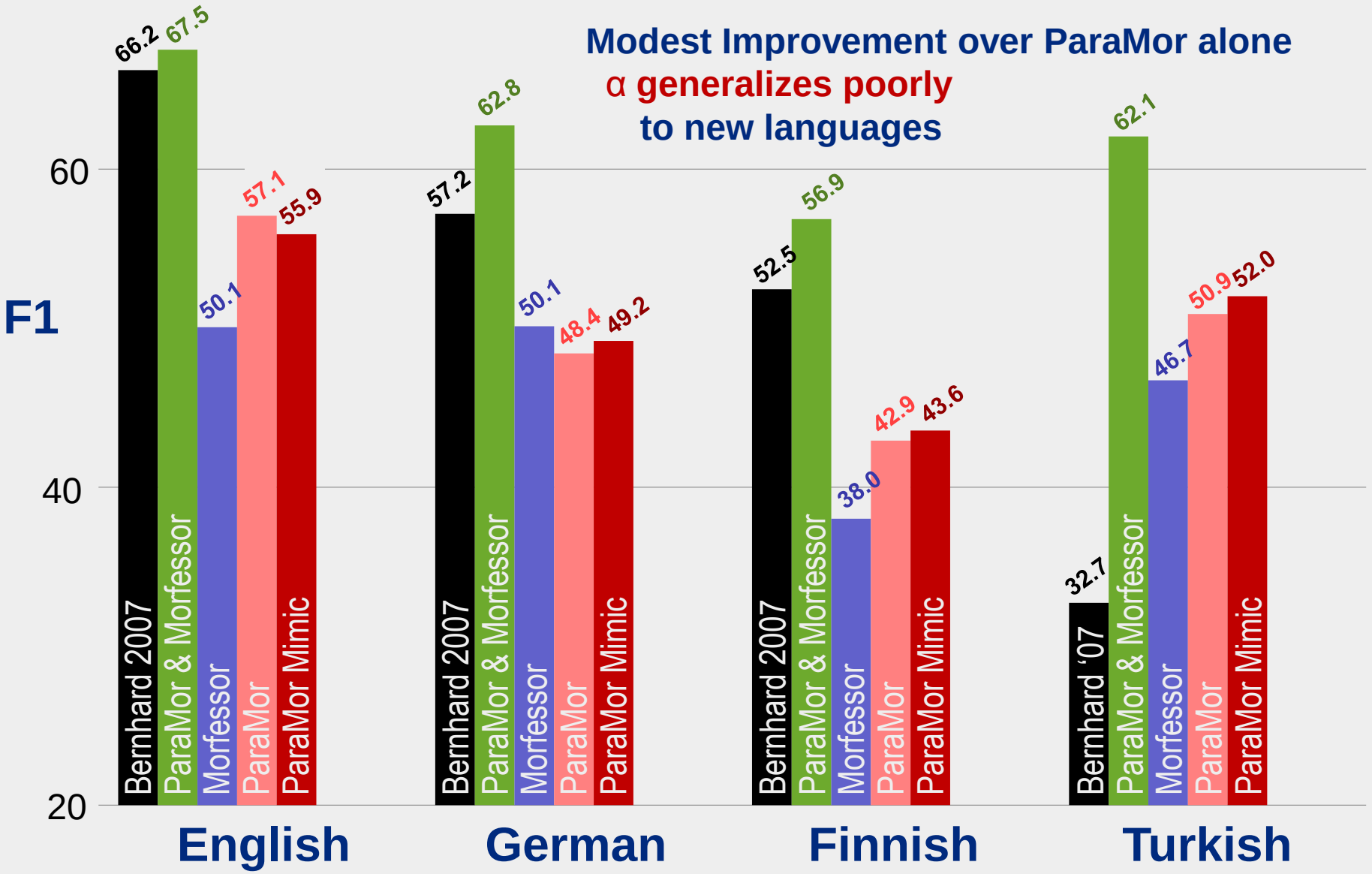
Improved over the original ParaMor and Morfessor

	ParaMor			Morfessor		
Hungarian	Original	Tagger Mimic		Original	Tagger Mimic	
		α=1	α=4/3		α=1	α=3/2
Precision	64.8	63.7	52.5	72.9	72.5	59.3
Recall	30.4	32.2	43.4	31.3	28.8	39.3
F_1	**41.4**	42.7	**47.5**	**43.8**	41.3	**47.3**

Results of Linguistic Evaluation 2009



ParaMor Mimic



ParaMor & Morfessor – Tagger Mimic

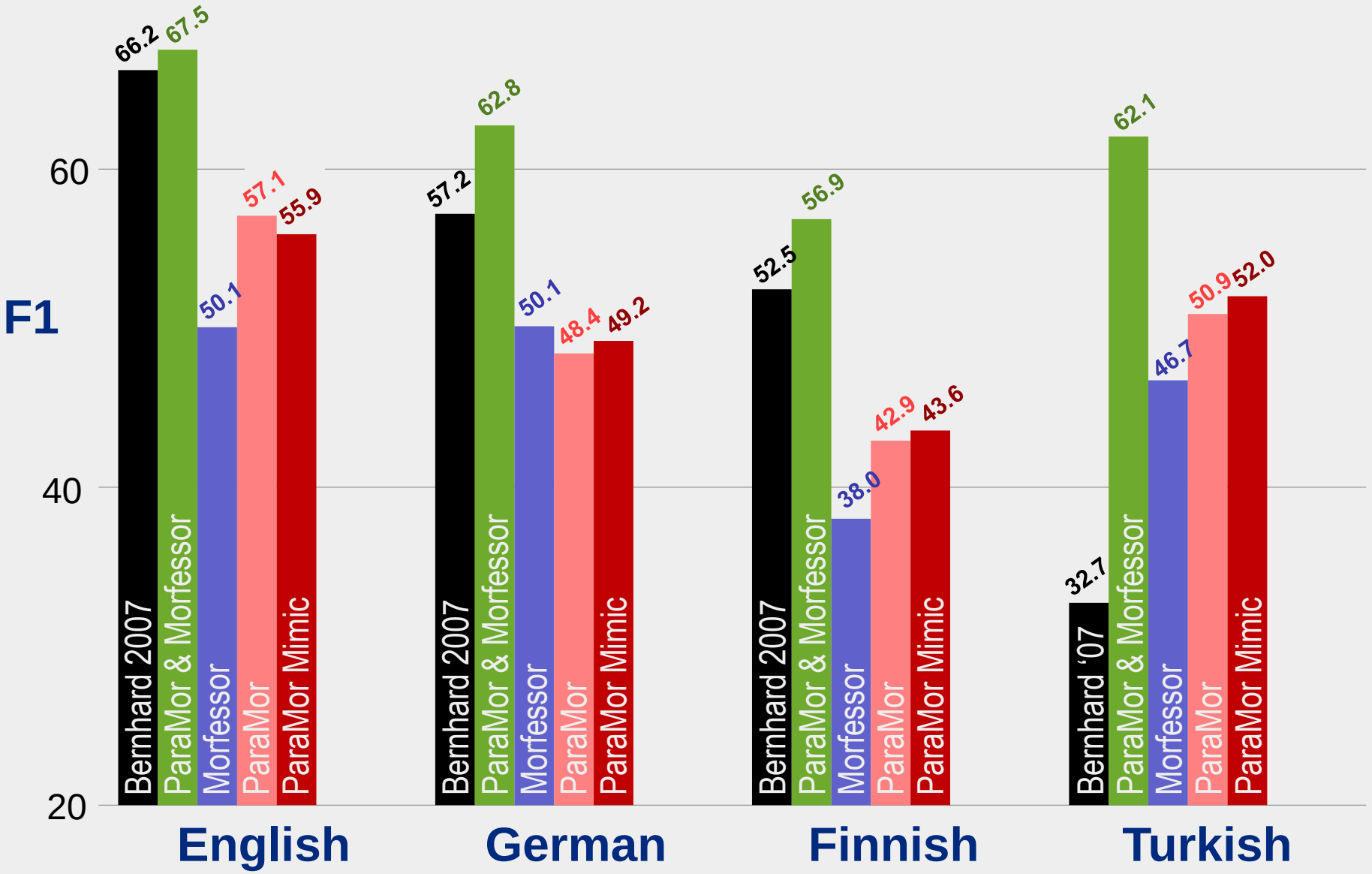
Weight ParaMor and Morfessor Equally

$\lambda = 0.5$

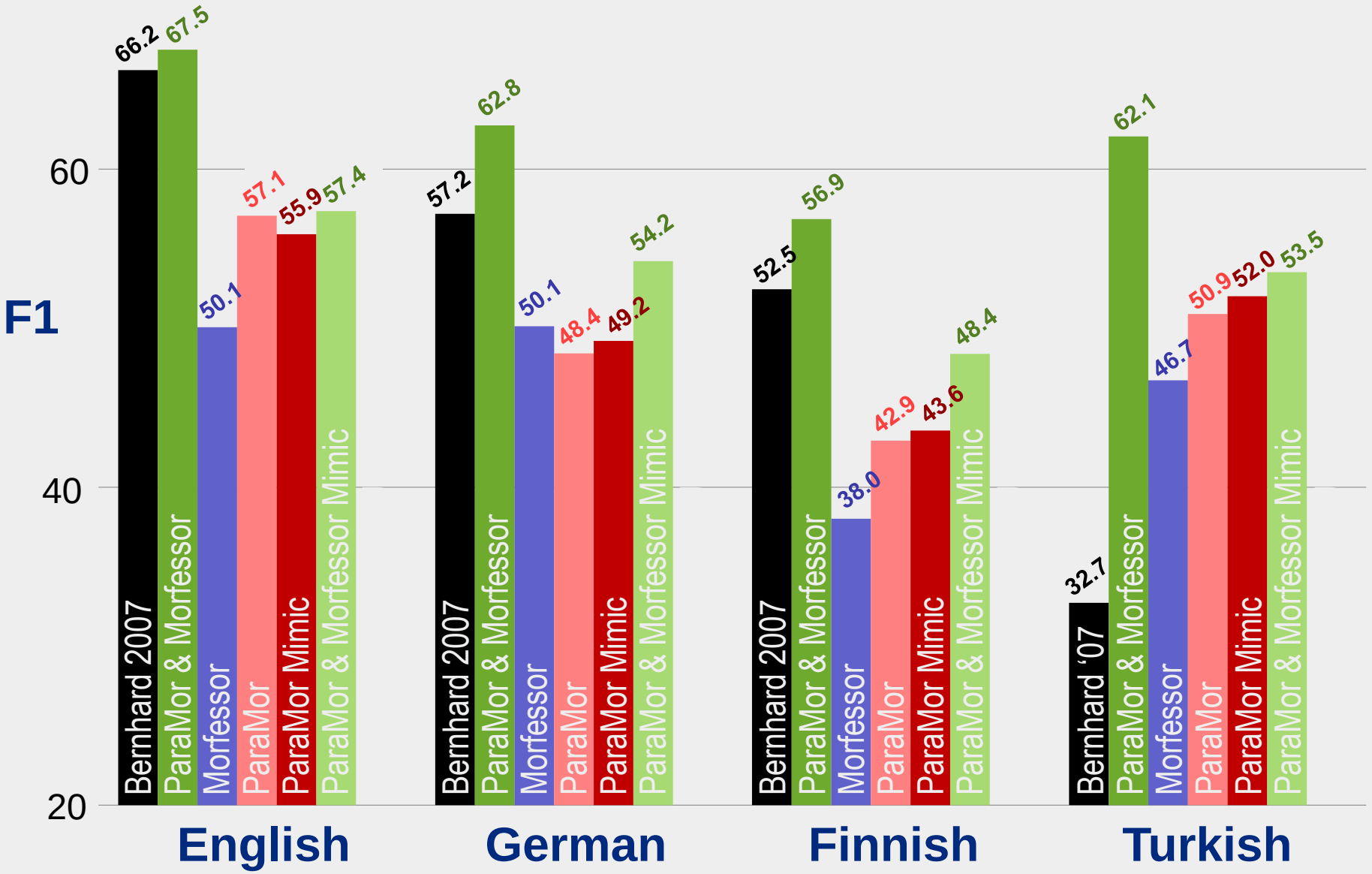
$Ave_i = \lambda \text{ParaMor}_i + (1 - \lambda)\text{Morfessor}_i$

Adjust α over Ave_i

Results of Linguistic Evaluation 2009

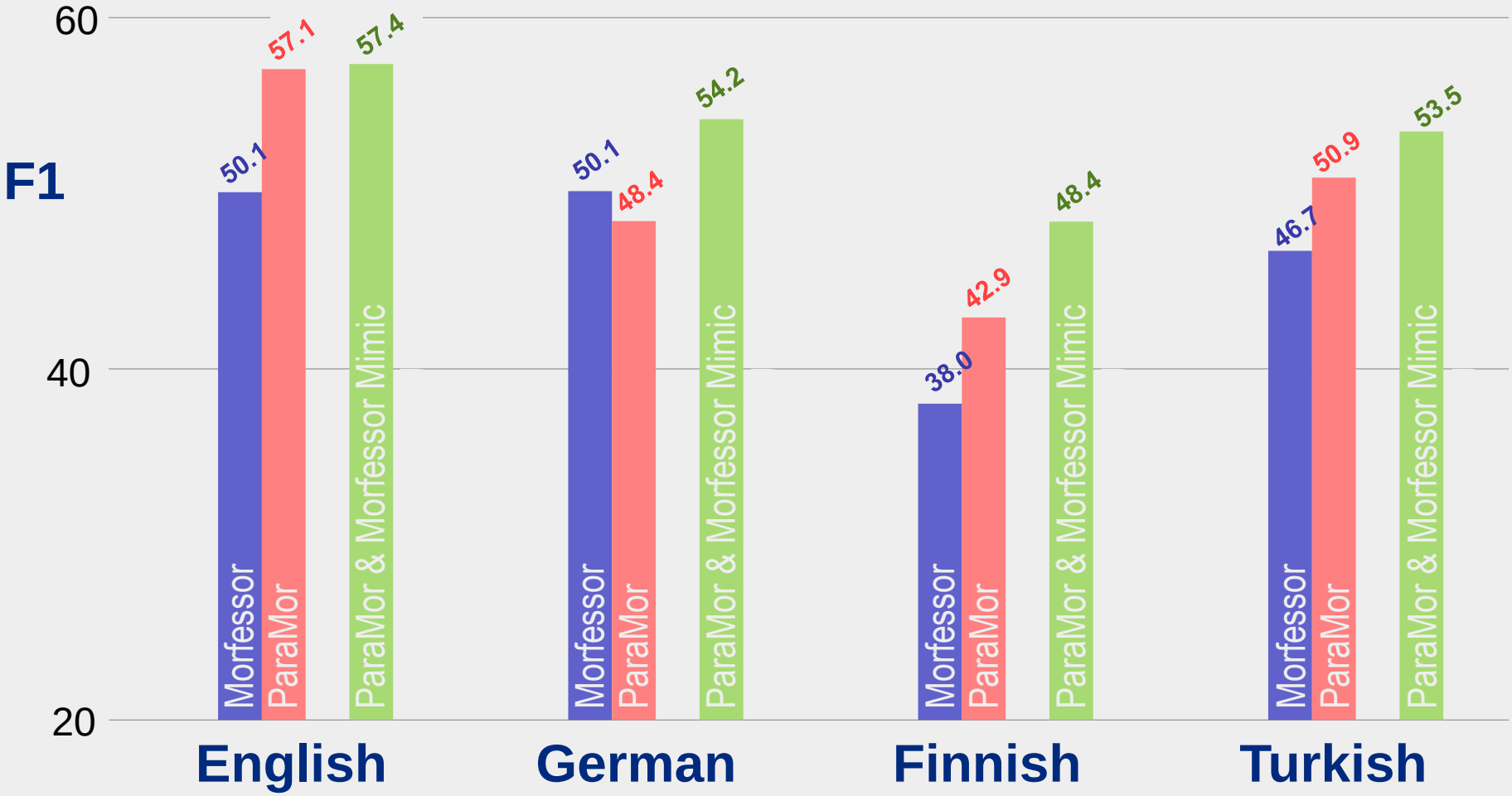


Results of Linguistic Evaluation 2009

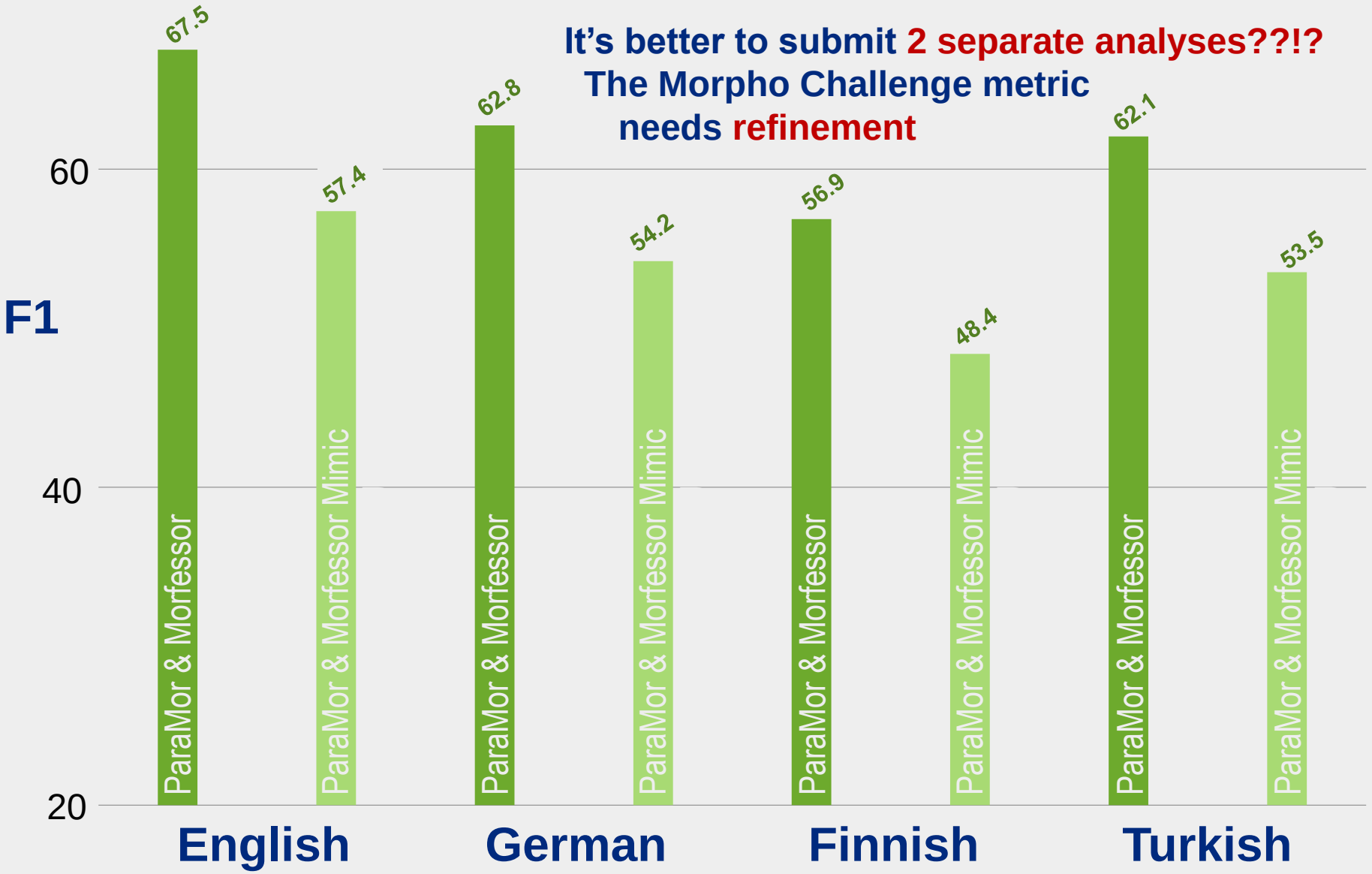


Results of Linguistic Evaluation 2009

Significant improvement over ParaMor alone and Morfessor alone

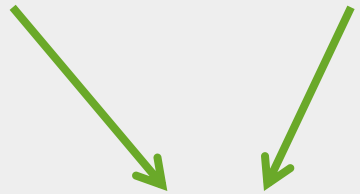


Results of Linguistic Evaluation 2009



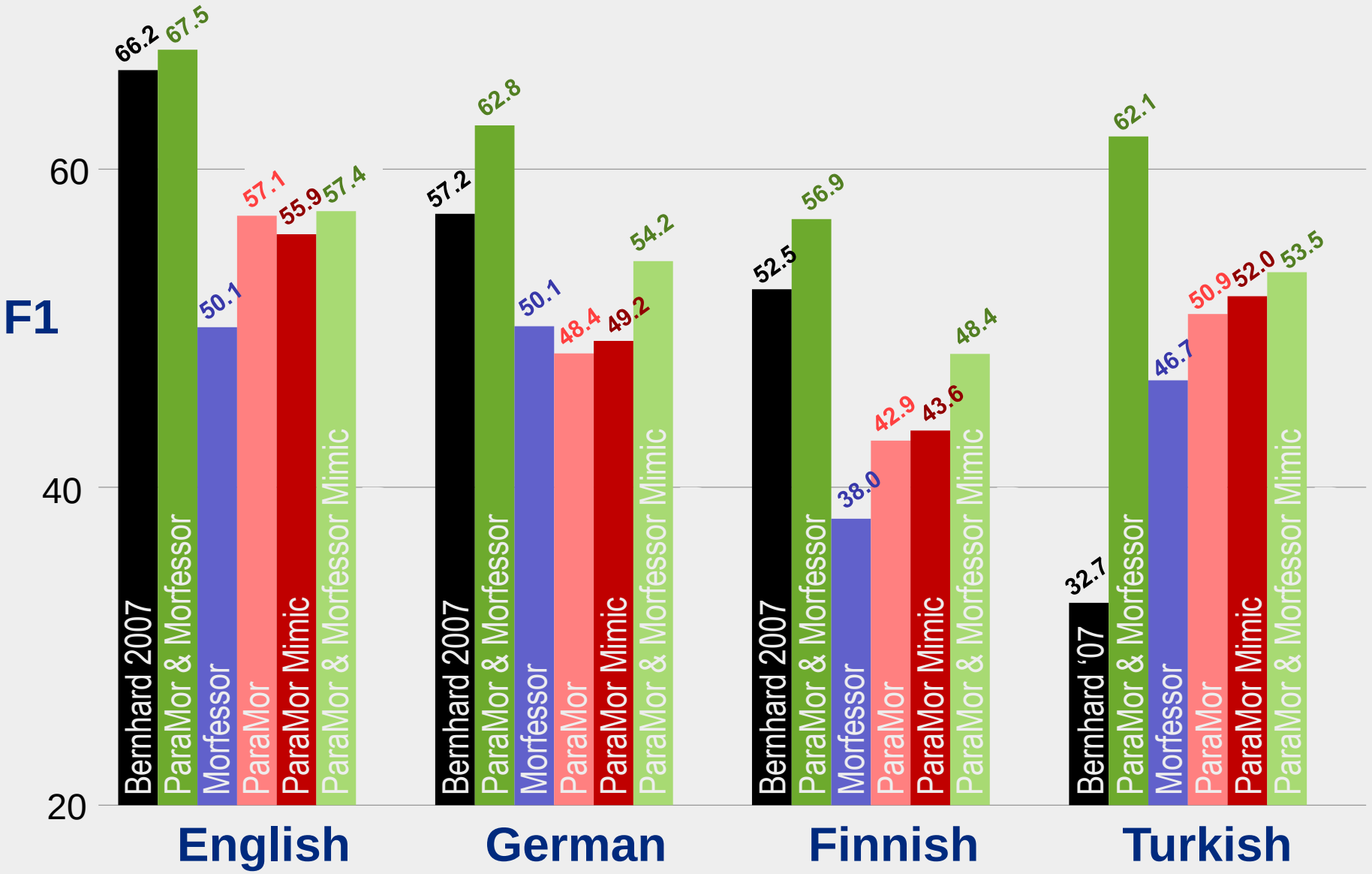
Union of ParaMor and Morfessor

Word	ParaMor	Morfessor
dispute's	disput +e +'s	dispute/STM +'s/SUF

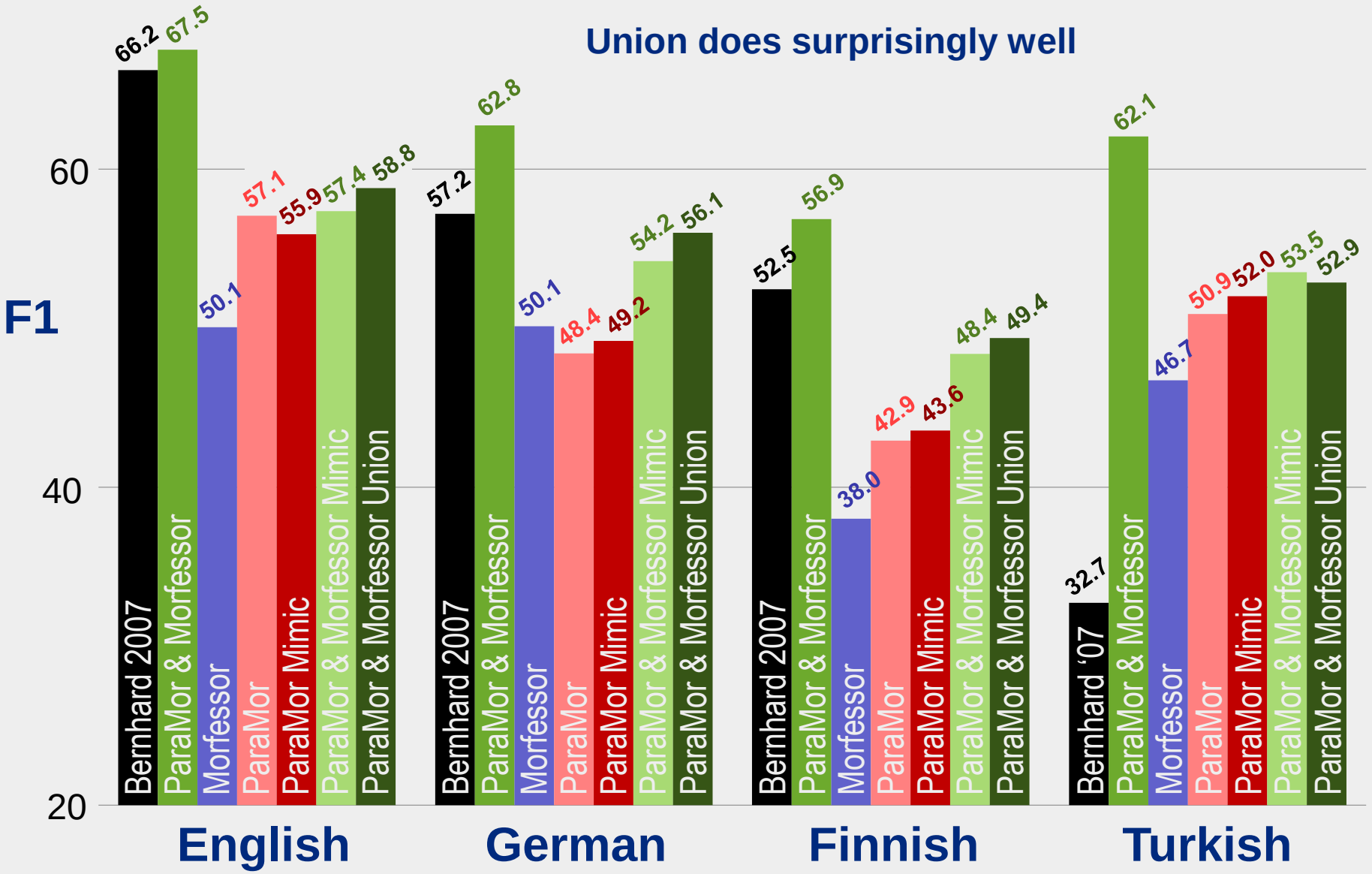


disput +e +'s

Results of Linguistic Evaluation 2009



Results of Linguistic Evaluation 2009



Tagger Mimics for Morphology Are:

Disappointing

- **Small** gains when moving to other languages
- The **Morpho Challenge metric** needs work
- Joint mimic **doesn't outperform Union**

Encouraging

- Produces **meaningful numeric scores** for both:
  - Rule-based systems (e.g. ParaMor)
  - Blackbox systems (e.g. Morfessor)
- **Enables** metric tuning

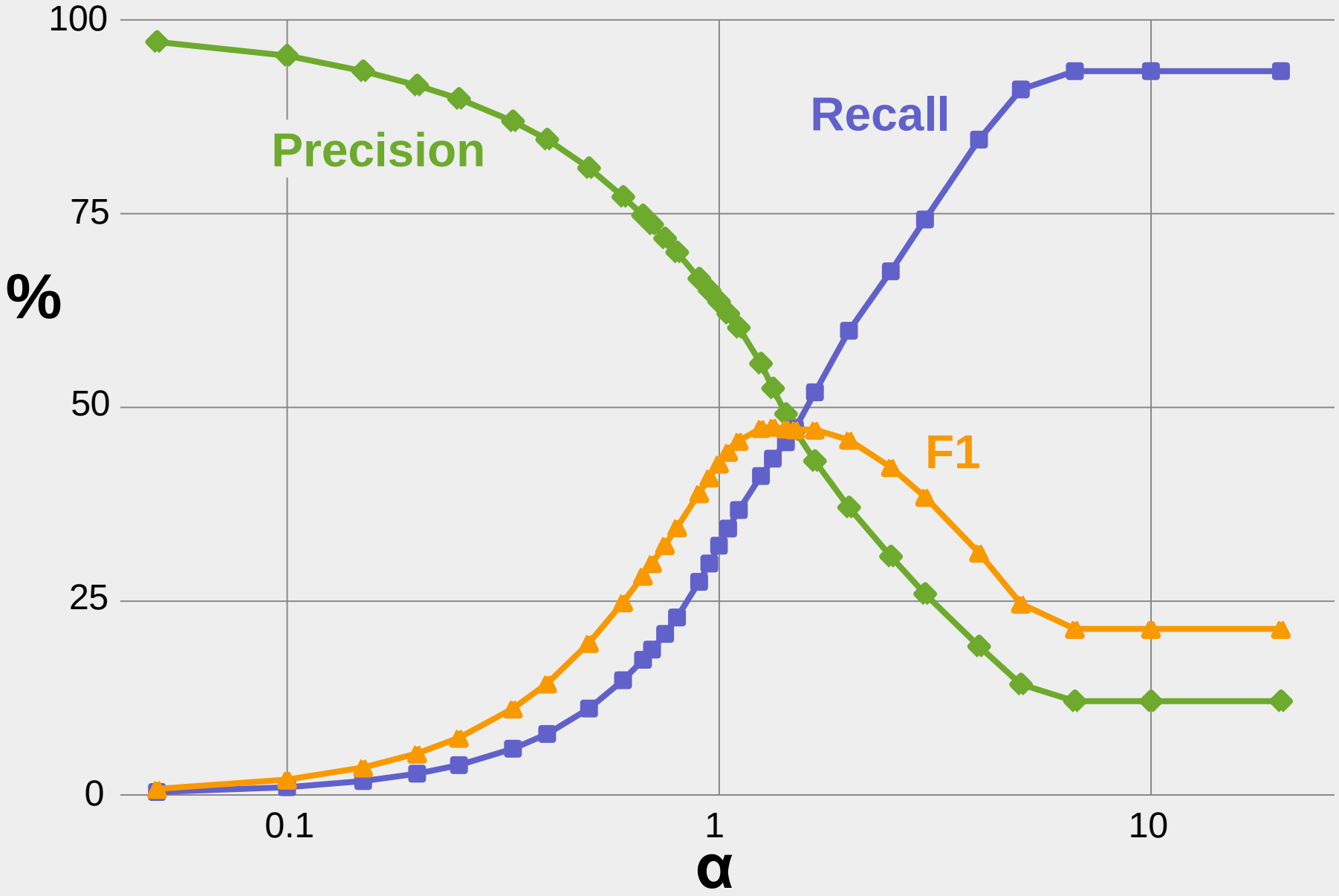
ευχαριστώ

(Thank You!)

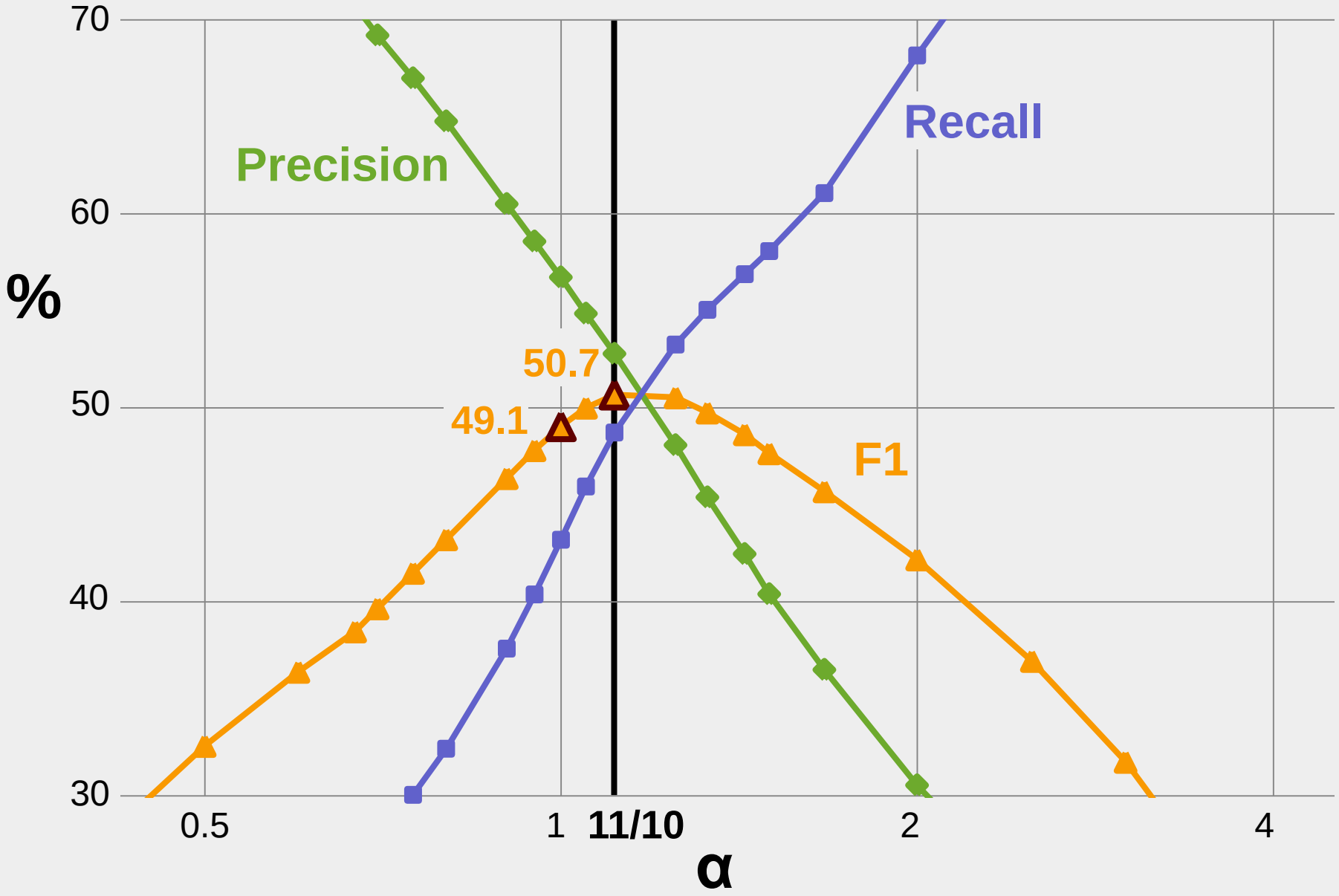
Kristy Hollingshead
Tagger-in-Chief

Brian Roark
Head Thinker

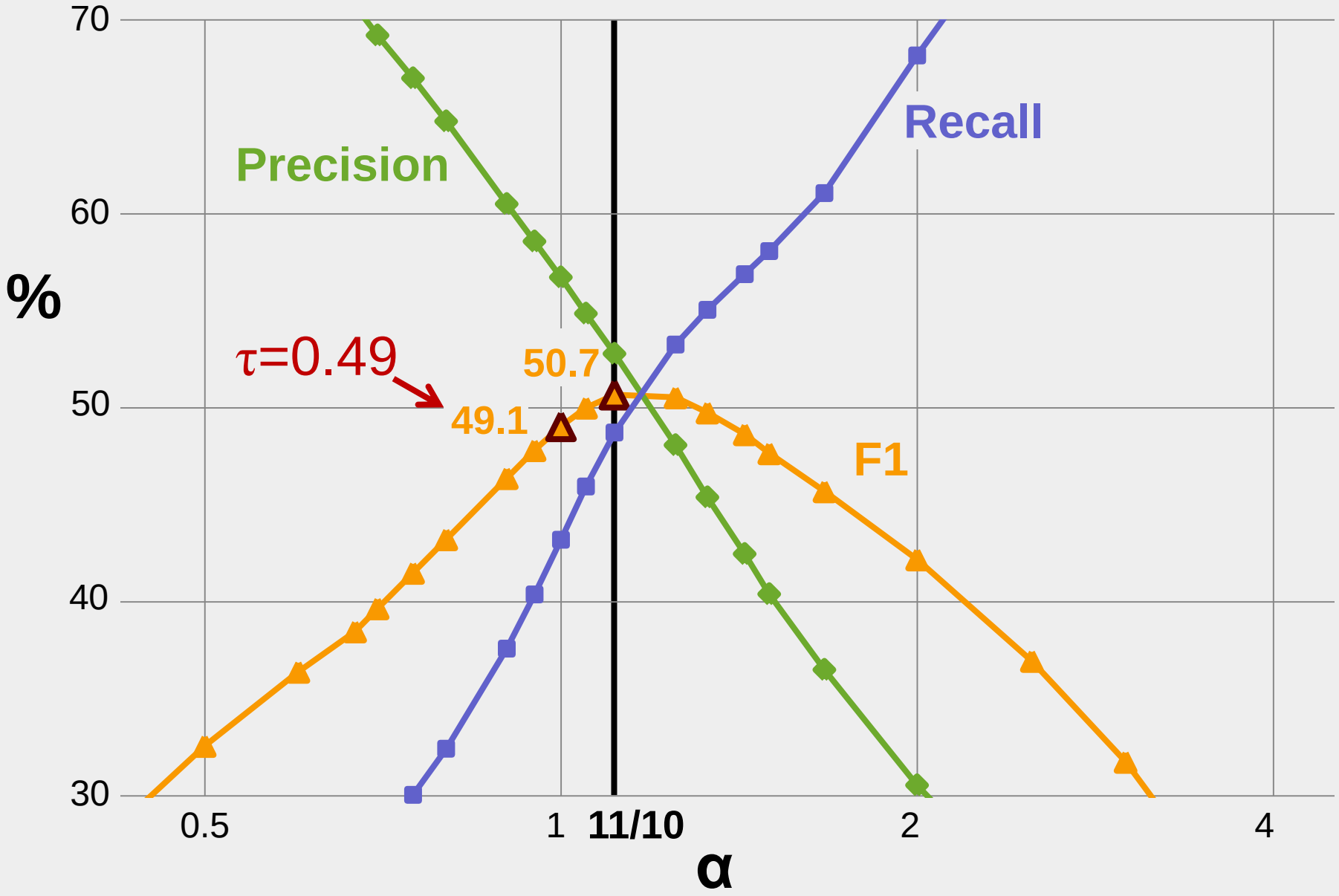
ParaMor Hungarian – Adjusting α



ParaMor & Morfessor



ParaMor & Morfessor

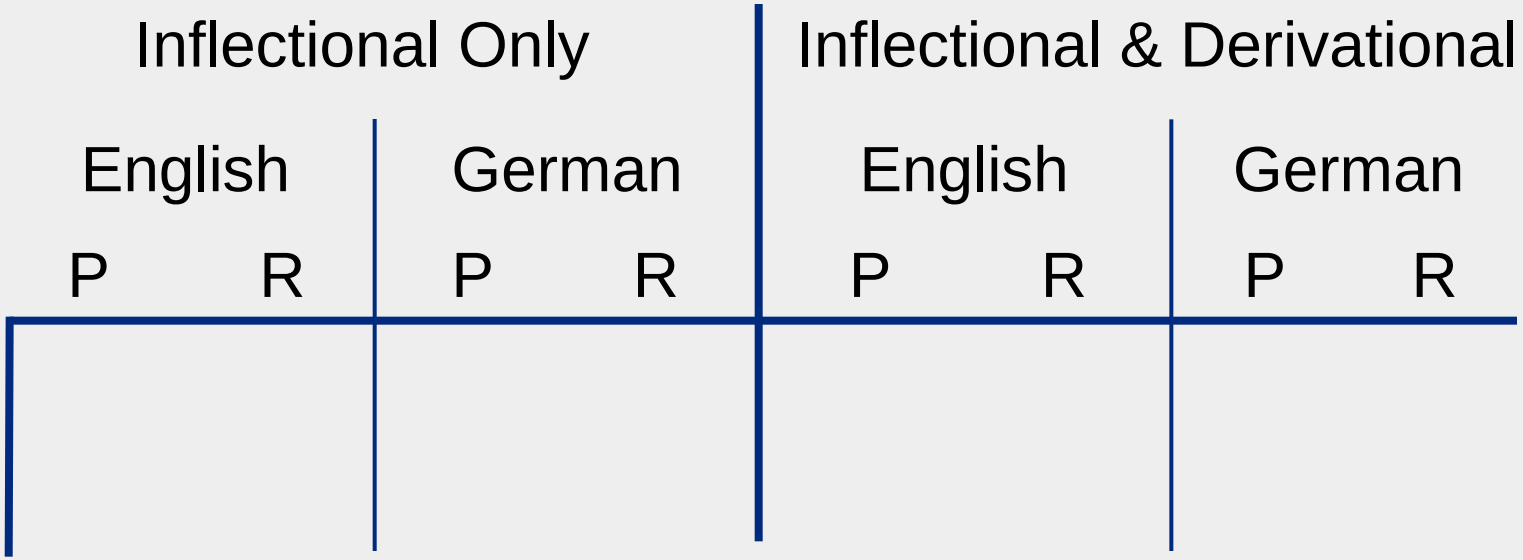


Preliminary Linguistic Evaluation

Built

2 styles of answer keys

For 2 languages



Preliminary Linguistic Evaluation

Built

2 styles of answer keys

For 2 languages

	Inflectional Only				Inflectional & Derivational			
	English		German		English		German	
	P	R	P	R	P	R	P	R
ParaMor	33.0	**81.4**	42.8	**68.6**	48.9	53.6	60.0	33.5

Preliminary Linguistic Evaluation

Built

2 styles of answer keys

For 2 languages

	Inflectional Only				Inflectional & Derivational			
	English		German		English		German	
	P	R	P	R	P	R	P	R
ParaMor	33.0	81.4	42.8	68.6	48.9	53.6	60.0	33.5
Morfessor	53.3	47.0	38.7	44.2	**73.6**	34.0	**66.9**	37.1