Morphological Copying / “Reduplication”

● English plural: book -> books
● Indonesian plural: buku -> buku-buku

Significant variation in copying processes (Rubino, 2013)

(1) Total reduplication = unbounded copy (83%)
   - C: wanta-wanta-wanta
   - ‘woman’ (Indonesian)

(2) Partial reduplication = bounded copy (75%)
   - a. CV: guyon-guyon-guyon
     → ‘jest’ to ‘jest repeatedly’ (Sundanese)
   - b. CVCV: takki-tal-tal-takki
     → ‘legs’ (Agta)
   - c. CVCVC: tananu-tanana-tanana
     → ‘return’ (Dyirbal)


2-way Finite-State Transducers

2-way FSTs can reseed parts of the input string unlike 1-way FSTs (Engelfried and Hoogeboom, 2001; Shalitt, 2009, Pilot and Reymer, 2016)

● 1-way FSTs ⊂ 2-way FSTs ⊂ Turing Machines
● 2-way FSTs ⊇ MSO-definable string transductions ⊍ Streaming String Transducers

Why 2-way FSTs?

2-way FSTs can model every encountered reduplicative morpheme, including unbounded copying. 1-way FSTs cannot model unbounded copying (Engelfried and Hoogeboom, 2001).

Why 2-way FSTs?

1. 2-way FSTs can model every encountered reduplicative morpheme, including unbounded copying. 1-way FSTs cannot model unbounded copying (Engelfried and Hoogeboom, 2001).

2. 2-way FSTs directly capture linguistic generalizations of copying while 1-way FSTs do not.

3. Functions 2-way FSTs define are exactly the “regular functions,” i.e. those definable with MSO logic (Engelfried and Higgeboom, 2001).

References