

# TIMEX2 Guideline Issues: Normalization

Temporal Expression Recognition and Normalization  
(TERN) Evaluation Workshop  
September 23, 2004

Lisa Ferro  
lferro@mitre.org

# English Guideline Issues on Normalization

- More than one way to say the same thing
  - Durations anchored on beginning or end time
  - Alternative ISO formats (XXXX-XX-XXTMO vs. TMO)
  - Set expressions: 2000-XX vs. XXXX-XX WITHIN 2000 (latter not actually allowed, but the attributes exist to do it)
- recent, past etc. – confusing!
  - “the past three weeks” is P3W anchored ENDING the reference week
  - “recent/recently” is PAST\_REF anchored on the reference date (usually the date\_time stamp on the article)
  - “over recent weeks” is PXW anchored BEFORE the reference week
  - “in recent weeks” is PAST\_REF anchored BEFORE the reference week
  - “recent weeks” can be a SET expression when the context implies a discontinuous set

## English Guideline Issues on Normalization (cont.)

- Indefinite predicate nominals: “Today is a sunny day”.
  - The guidelines say not to give “a sunny day” the same VAL as “Today”; annotators do anyways.
- Semantics not captured:
  - “earlier that day” (though we can get “early that day”)
  - p.m. (AF or EV?)

# Chinese Guideline Issues on Normalization

- Better handling and guidance needed for:
  - Idiomatic expressions
  - Chinese New Year (the day vs. the vacation period)
  - Chinese lunar years