

# JAPANESE ANALYSIS IN PANGLOSS

DR. EDUARD HOVY, ISI  
DR. KEVIN KNIGHT, ISI  
DR. MASAYO IIDA, ISI  
MR. AKITOSHI OKUMURA, NEC  
MR. RICHARD WHITNEY, ISI  
MR. MATTHEW HAINES, ISI  
MR. ISHWAR CHANDER, USC  
MR. STEVE LUK, USC  
MR. KENJI YAMADA, USC

KBMT w/ focus on:

- SCALE
- MACHINE LEARNING
- KNOWLEDGE ACQUISITION

# CHRONOLOGY

MID-JAN

1 MONTH GOAL:

BUILD MODULES

BUILD SMALL KS's

TRANSLATE ONE TEXT,

WITH HUMAN

CONTROLLING AMBIGUITY

MID-FEB

1.5 MONTH GOAL:

INTEGRATE MODULES

BUILD KS's TO TRANSLATE

ONE TEXT CORRECTLY,

FULLY AUTOMATICALLY.

BUILD KS's TO TRANSLATE

5 TEXTS WITH

HUMAN OPERATOR

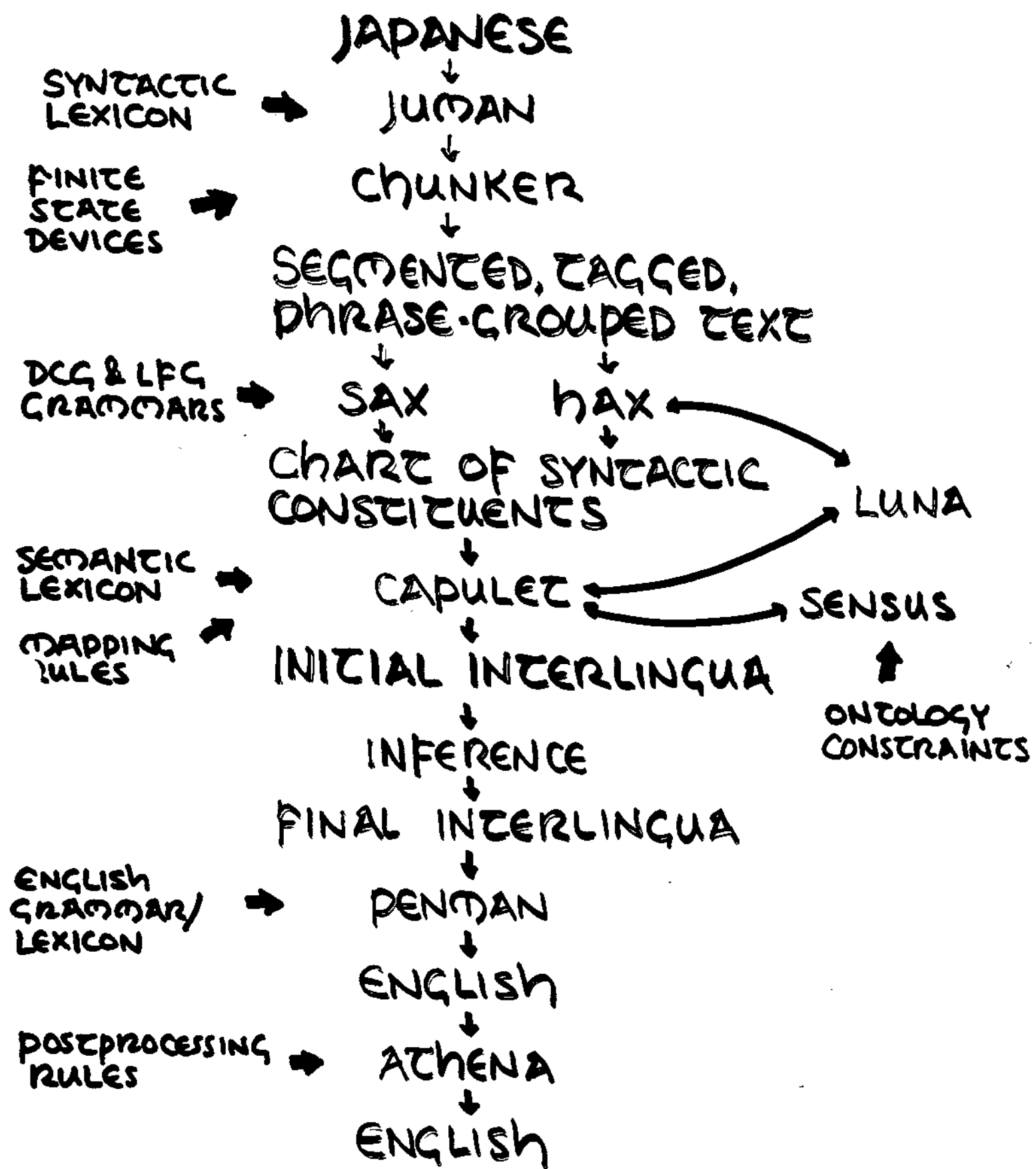
LONGER-TERM GOAL  
(being pursued concurrently)

BUILD LARGE-SCALE  
KS's AUTOMATICALLY

LDOCE → ONTOLOGY  
WORDNET

ENGLISH TEXT → ONTOLOGY CONSTRAINTS  
→ POST PROCESSING RULES

KENKYUSHA → JAPANESE LEXICON  
IPAL → JAPANESE LEXICON  
JAPANESE TEXT → JAPANESE LEXICON  
JUMAN → JAPANESE GRAMMARS



## KBMT MODULES & KNOWLEDGE SOURCES

# SAMPLE TEXT

シチズン時計は18日、米国の大手集積回路（IC）製造装置メーカー、パーキン・エルマー社と合併会社を設立、IC製造装置の生産を始めることを見通し。パーキン・エルマー社の日本にある子会社が株式の過半数を持ち、当初はICチップを製造するのに使うドライエッチング装置やステッパー（縮小投影型露光装置）を生産する計画。パーキン社は昨年かから、日本国内で、IC製造装置の販売を始めているが、IC需要が今後もかなりの成長が見込まれるため、シチズンと提携、日本国内での直接生産に踏み切る。

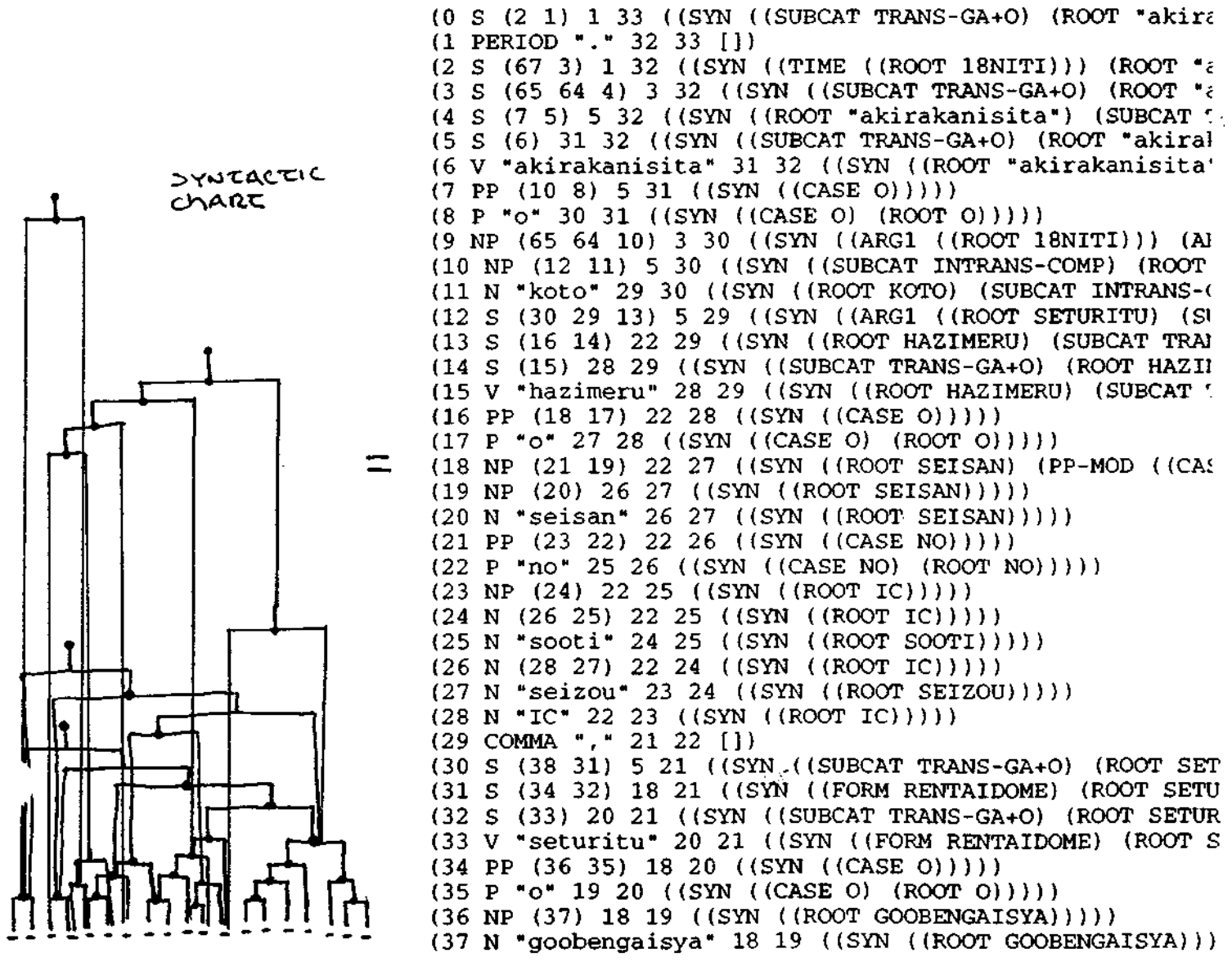
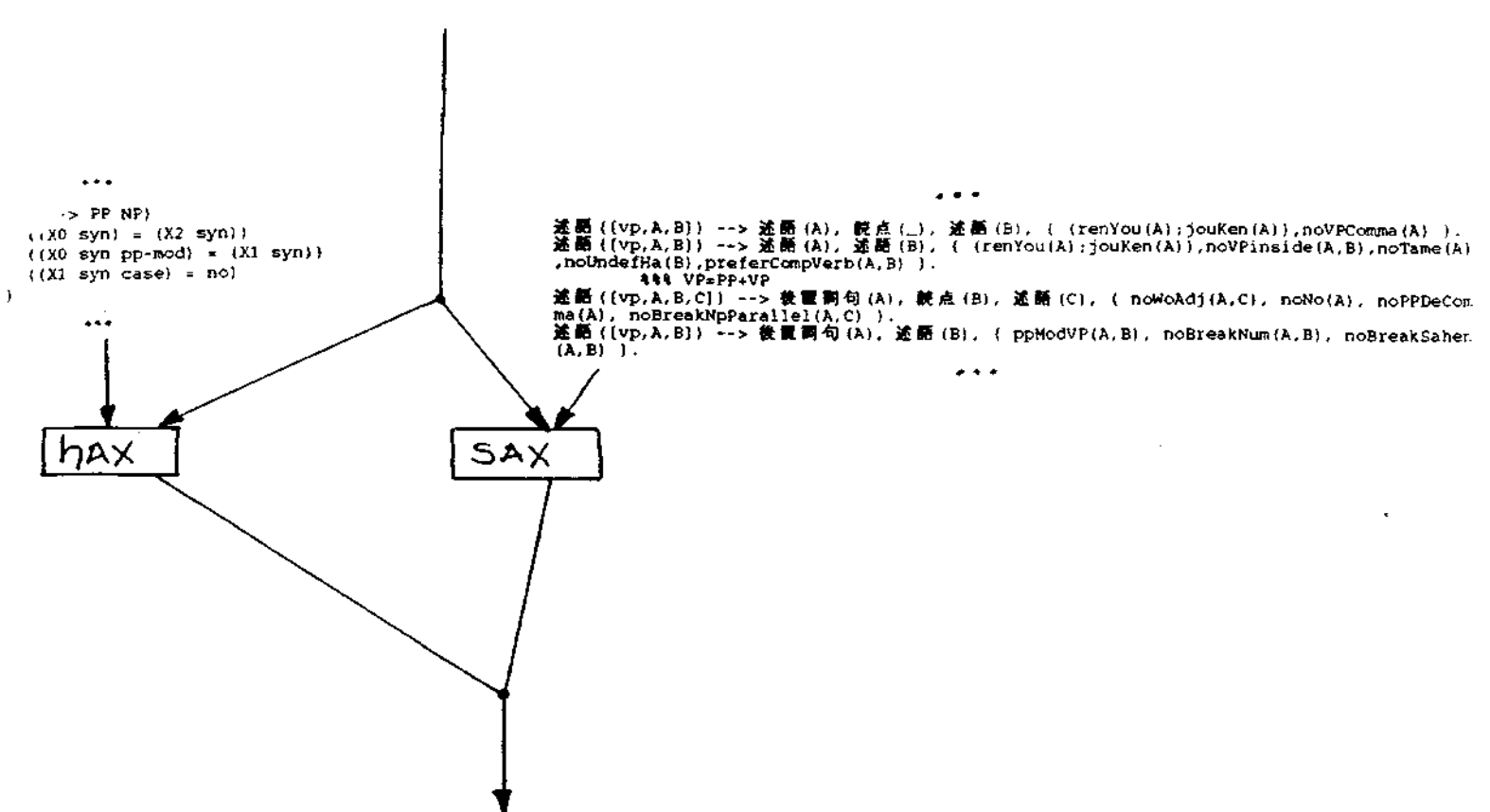
シチズン時計は 18日、米国の大手 集積 回路（IC）  
Citizen Clock-TOP 18th US-GEN major integrated circuit  
製造 装置 メーカー、パーキン・エルマー社と 合併会社を  
manufacture device maker Perkin-Elmer with joint venture-OBJ  
設立、 IC製造 装置の 生産を 始めることを  
establish IC manufacture device-GEN production-OBJ start-COMP-OBJ  
明らかにした。

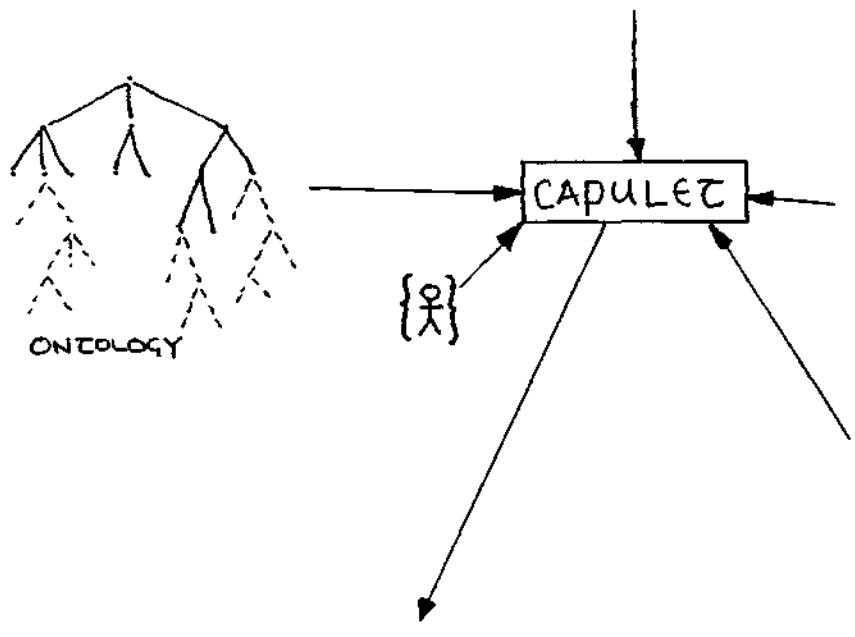
disclosed  
新会社は、 2月にも 設立される 見通し。パーキン・エルマー社の  
new company-TOP Feb-in-even be-established expect Perkin-Elmer-GEN  
日本に ある 子会社が 株式の 過半数を 持ち、当初は  
Jaapn-in exit subsidiary-SUB stock-GEN majority-OBJ hold beginning-at-TOP  
ICチップを 製造するのに 使うドライエッチング 装置 や  
IC chip-OBJ manufacture-for use dry-etching device and  
ステッパー（縮小 投影 型 露光 装置）を 生産する 計画。  
stepper (reduction projection type exposure device)-OBJ produce plan  
パーキン社は 昨年か、 日本 国内で、 IC 製造 装置の  
Perkin-TOP last\_year-from Japan country-in IC manufacture device-GEN  
販売を 始めている が、IC需要が 今後も かなりの 成長が  
sale-OBJ has-begun but IC demand-SUB in-future considerable growth-SUB  
見込まれる ため、シチズンと 提携、 日本 国内での  
be-expected because Citizen-with collaborate Japan country-inside-GEN  
直接生産に 踏み切る。  
direct production-OBJ2 decide

CITIZEN WATCH ANNOUNCED ON THE 18TH THAT IT WILL ESTABLISH A JOINT VENTURE WITH PERKIN-ELMER, A MAJOR AMERICAN MAKER OF INTEGRATED CIRCUIT MANUFACTURING DEVICES, AND WILL START PRODUCTION OF INTEGRATED CIRCUIT MANUFACTURING DEVICES.

THE NEW COMPANY IS EXPECTED TO BE ESTABLISHED IN FEBRUARY. PERKIN-ELMER'S SUBSIDIARY IN JAPAN WILL HOLD THE MAJORITY OF THE STOCK. THEY FIRST PLAN TO PRODUCE DRY-ETCHING DEVICES THAT ARE USED TO MANUFACTURE IC CHIPS, AND STEPPERS (REDUCED PROJECTION DEVICES THAT EXPOSE LIGHT). PERKIN-ELMER STARTED SELLING IC MANUFACTURING DEVICES IN JAPAN LAST YEAR. SINCE GOOD GROWTH IS EXPECTED IN THE DEMAND FOR IC'S IN THE FUTURE, THEY COLLABORATED WITH CITIZEN AND DECIDED TO PRODUCE DIRECTLY IN JAPAN.







```

...
((S -> PP S)
 (*OR*
  ((X1 syn case) = ga)
  ((X0 sem) = (X2 sem))
  ((X0 map) = (X2 map))
  ((X2 map subject-role) = X1))
  ((X1 syn case) = o)
  ((X0 sem) = (X2 sem))
  ((X0 map) = (X2 map))
  ((X2 map object-role) = X1))
  ((X1 syn case) = (*NOT* ga o))
  ((X0 sem) = (X1 sem))
  ((X0 map) = (X2 map))
  ((X0 sem) = (X2 sem))))
)
...

((V -> *securitu*)
 ((X0 sem instance) = {found, launch})
 ((X0 syn subcat) = TRANS-GA+O )
 ((X0 sem agent) = (X0 map subject-role sem))
 ((X0 sem goal) = (X0 map object-role sem))
 ((X0 syn form) = rentaidome )
)
...

```

```

(SEM
 ((SAYING
  ((OP2
    ((INSTANCE |begin|)
      (GOAL
        ((CREATED-ENTITY
          ((INSTANCE |device/instrumentality|)
            (AGENT-OF
              ((CREATED-ENTITY ((INSTANCE |microcircuit|)))
                (INSTANCE |construct, fabricate|))))))
          (INSTANCE |make, produce|))
            (AGENT [ ])))
    (OP1
      ((AGENT [ ])
        (GOAL ((INSTANCE |joint venture|))
          (INSTANCE |found, launch|)
          (INCLUSIVE
            ((RESTATEMENT-ELABORATION
              ((Q-MOD ((INSTANCE |major company|)))
                (INSTANCE |company/business|)
                (AGENT-OF
                  ((CREATED-ENTITY
                    ((AGENT-OF
                      ((CREATED-ENTITY
                        ((PAREN ((INSTANCE |microcircuit|)))
                          (INSTANCE |microcircuit|)))
                        (INSTANCE |construct, fabricate|))
                        (INSTANCE |device/instrumentality|))))))
                    (LOCATION
                      ((INSTANCE |country/administrative district|)
                        (NAME "United States")))))
                    (INSTANCE |company/business|)
                    (NAME "paakin-erumaa-sya")))))
                (INSTANCE AND)))
          (SAYER [ ])
          (INSTANCE |announce, make known|)
          TEMPORAL-LOCATING ((INDEX |eighteenth|) (INSTANCE |day/calendar day|))
          (TOPIC ((NAME "sitizuntokei") (INSTANCE |company/business|))))))

```



```

(A / |announce, make known|
:SAYER (C1 / |company/business|
        :NAME "Citizen Watch")
:TEMPORAL-LOCATING
(D1 / |day/calendar day|
  :INDEX |eighteenth|)
:SAYING
((E / |found, launch|
  :AGENT C1
  :CREATED-ENTITY (V / |joint venture|)
  :INCLUSIVE
  (R2 / RESTATEMENT
    :DOMAIN (C2 / |company/business|
             :NAME "Perkin-Elmer")
    :RANGE
    (C3 / |company/business|
      :AGENT-OF
      (M / |construct, fabricate|
        :CREATED-ENTITY
        (D / |device/instrumentality|
          :AGENT-OF
          (M2 / |construct, fabricate|
            :CREATED-ENTITY
            (I / |microcircuit|))))
      :Q-MOD (Q / |major-leading|)
      :LOCATION
      (C4 / |country/administrative district|
        :NAME "United States"))))
(ZZ / |make, produce|
  :AGENT C1
  :CREATED-ENTITY
  (YY / |device/instrumentality|
    :AGENT-OF
    (XX / |construct, fabricate|
      :CREATED-ENTITY (UU / |microcircuit|))))))

```

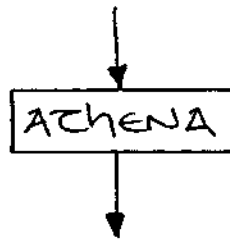
↓

DENTAN

↓

"Citizen Watch makes known on the eighteenth day that it launches a joint venture with Perkin-Elmer a major United States company that constructs a device that constructs a microcircuit and that it makes a device that constructs a microcircuit."

Among many documents a would-be public company must file with the Securities and Exchange Commission is a set of audited financial statements. Cattleguard, now expanding in four states largely as Sears, Roebuck & Co. concessionaire, was a one-restaurant company until two years ago and hadn't had auditors. Once the paper work is finished, the underwriter escorts the company's executives around the country to drum up interest in its shares. This road show often hits eight or more cities in about a week.



Among THE many documents A would-be public company must file with THE Securities and Exchange Commission is A set of audited financial statements. Cattleguard, now expanding in four states largely as XTHE Sears, Roebuck & Co. concessionaire, was XTHE one-restaurant company until two years ago and hadn't had auditors. Once THE paper work is finished, THE underwriter escorts THE company's executives around THE country to drum up interest in its shares. This road show often hits eight or more cities in about A week.

# FUTURE WORK

- OTHER DETERMINERS  
BESIDES A, AN, THE.  
ESPECIALLY NULL.
- FULL CONTEXT (88 → 96%)
- PLURALS
- OTHER AUTOMATED  
POST-EDITING TASKS