I. Style Guidelines

Abbreviations. Latin abbreviations (except for *et al.*) are used only in material enclosed within parentheses; in running text, English equivalents such as *that is, for example,* and *compare* are used.

Abstracts. The article should begin with an informative abstract of 150–250 words. It should state the objectives of the work, summarize the results, and give the principal conclusions and recommendations. It is preferable that the abstract not be in the first person, and it should not contain any mathematical notation or cite references. Work planned but not completed should not appear.

**Boldface.** Boldface is used for the first occurrence of a term: *<The agreement predicates are defined solely over unordered sets of features.>*

Double quotation marks. Double quotes (“*x*”) are used for
1. Quotations (citations) within the text: *<He asserted that “no man is an island.”>*
2. A coining or a special use of a word or phrase: *<The word “fractal” suggests something that is “fractured.”>*

Footnotes. Whenever it does not impede the logic or readability of the article, footnote material should be integrated into text.

In-text lists. In-text lists are introduced with (1), (2), (3), and so on.

**Italics.** Italics are used for
1. Emphasis: *<We want to determine just why this happens.>*
2. Words or sentences used within the text: *<For example, persuade controls the subject of its complement, as in We persuaded John to leave.>*
3. Foreign words or phrases not in common use in English: *<One would italicize *pieta* but not per se.>*
4. Book titles: *<... as described in Chomsky’s *Aspects of the Theory of Syntax.*>*

Punctuation
1. If three or more items are conjoined, a comma appears before the *and* that precedes the last item: *<a, b, and c.>*
2. There is a comma after *i.e.* and *e.g.*
3. There is no terminal punctuation following displayed equations.
4. There is a comma in numerals 1,000 and above.
5. Commas and periods appear inside double quotation marks; commas and periods appear outside single quotation marks (except in the colloquial English translation that follows a numbered, glossed non-English example).
Semicolons and colons appear outside both single and double quotation marks.

6. Decade names are written without an apostrophe: <the 1990s>.

**Percentages.** Percentage is expressed with the percentage symbol (%), always with a numeral, even for percentages less than 10: 95%, 8%.

**Relative pronouns.** *That* is used to introduce restrictive relative clauses; *which* is used to introduce nonrestrictive relative clauses.

**Single quotation marks.** Single quotes (‘x’) are used for the definition of a phrase or a foreign word/sentence: <One usually defines *etre* as ‘to be’.*>

**Spelling and capitalization**
1. American spelling conventions (e.g., *behavior* rather than *behaviour*, *criticize* rather than *criticise*) are observed throughout the journal.
2. Full sentences following a colon begin with a capital letter.

**Word choice.** *Article* rather than *paper* refers to works within *Computational Linguistics* (<The research reported in this article> rather than <The research reported in this paper>). *Paper* is acceptable in reference to works other than the current one, if it can be appropriately applied (particularly in respect to papers presented at conferences and the like).

II. References

**Text references**
1. If the author’s name occurs in the text, the date is enclosed in parentheses:
   <Hobbs (1978) first proposed that . . .>
   <. . . first proposed in Hobbs (1978)>

2. When the reference itself is within parentheses, and the parentheses enclose nothing other than references, the phrase *e.g.* or *cf.*, or the words *see* or *see also*, the date is not enclosed in parentheses (note that no comma separates the author’s name from the date):
   <(Hobbs 1978)>
   <(e.g., Hobbs 1978)>   <(cf. Hobbs 1978)>
   <(see Hobbs 1978)>   <(see also Hobbs 1978)>

3. If the parentheses enclose other material, the date is enclosed in square brackets rather than parentheses:
   <(e.g., Cassell et al. [1994] and much research since then)>

4. The word *page* is spelled out in citations:
   <(Stuckard 2000, page 240)>

5. For works with one, two, or three authors, all authors’ surnames are given in the in-text citation. For works with four or more authors, *et al.* (in roman type) replaces the surnames of all authors except the first. (The names of all
authors are provided in the corresponding reference entry, regardless of the number of authors.)

<(Smith 2000)>
<(Smith and Jones 2000)>
<(Smith, Jones, and Wexler 2000)>
<(Smith et al. 2000)>

Reference list. References should be listed alphabetically by author at the end of the article according to the following style. All authors must (where possible) have first names specified.

1. Article in journal:

2. Book:

3. Article in edited collection/Chapter in book:
Jurafsky, Daniel, and James H. Martin. 2000. _Speech and Language Processing_, chapter 1. Prentice Hall.

4. Technical report:

5. Thesis or dissertation:
6. Unpublished item:

7. Conference proceedings:

8. Paper published in conference proceedings:

III. Word List

A-chain (A-bar-chain)
A-position
Altavista
ambiguity-preserving generation
analog
anaphora generation (n., adj.)
anaphora resolution methods
Appendix, Appendices
**AP Treebank**
artificial intelligence (n., adj.)
attribute-value grammars
automatic speech recognition (n., adj.)

bound variable (n.)
bound-variable (adj.)
breadth-first search
Brown corpus
byproduct
canceled, canceling
case 1, case 2 (etc.)
case frame patterns
Center Continuation
Center Establishment
Center Retain
Center Shift
centering
centering model
centering theory
chapter 7 (etc.)
chart parser
chi-square test
Chomsky adjunction
Chomsky normal form
chunk parser
class-based interpolated
closed class (n.)
closed-class (adj.)
code set
coexist
cognitive science (n., adj.)
collative semantics
combinatory categorial grammar
Common Lisp
common sense (n.)
commonsense (adj.)
compile time (n.)
compile-time (adj.)
complete-link clustering
complex NP assumption
computational linguistics (n., adj.)
computer-assisted language learning
non-normalized

noun phrase antecedent
(an) NP
NPs (plural of NP)
NP’s (possessive of NP)
NP-complete

off-line
one-sense-per-discourse heuristic
on-line
on-the-fly (adj.)
on the fly (adv.)
ontology-engineering architecture
ontology-learning architecture
open class (n.)
open-class (adj.)
optimality theory
outperform

parameter estimation algorithm
Pareto-optimal
Pareto ranking (n., adj.)
parse forest
parser-output trees
part of speech (n.)
part-of-speech (adj.)
part-of-speech-tagged corpus
pattern-matching method
Penn Treebank (Penn-II Treebank, Penn-III Treebank)
Ph.D.
Ph.D. thesis
phonological-rule induction algorithm
phrase structure grammar
phrase structural position
POS-language models
POS tagging
POS-tagging errors
Prague Dependency Bank
predicate-argument structure
pre-fixed (meaning “fixed in advance”)
prepositional phrase (n.)
prepositional-phrase (adj.)
present-tense (adj.)
probabilistic context-free grammars
probabilistic feature grammars
probability mass function
proper noun (n.)
proper-noun (adj.)
proto-allophones
proto-phonemes
pseudo-disambiguation task
public-domain programs
push-back operation
pushdown (n., adj.)
push down (v.)
Q-structure
qualia structure (n., adj.)
quantifier-raising approach
quasi synonyms
query processing (n.)
query-processing (adj.)
question answering (n.)
question-answering (adj.)
random number generator
range concatenation grammar
real time (n.)
real-time (adj.)
real-world programming
re-create
red-herring debate
reentrancy
resolution rate
Retain [in centering]
rhetorical structure theory
right-branching tree
right-sibling
Rough Shift
route follower
route giver
rule set
rule-induction technique
run time (n.)
run-time (adj.)
S-dominated C-structure tree
search space (n.)
search-space (adj.)
Section 1.1 (etc.)
semantic cohesion value
semantic distance metric
semantic similarity measure
semantic space (n.)
semantic-space (adj.)
semantic type coercion
the Semantic Web
SemCor
semiautomatically
sense-clustering algorithm
sense-tagged corpus
sense-tagging corpora
sentence-aligned parallel bilingual corpus
sentence alignment techniques
sentence boundary disambiguation
sentence-initial position
sentence-level grammatical function
set-theoretic (adj.)
signaled, signaling
shift-reduce parser
shortest-path problem
single-link clustering
single-string automaton
small-worlds property
Smooth Shift [in centering]
source language (n., adj.)
sparse-data problem
spell-checker
spell-checking (n., adj., v.)
spreading activation mechanism
standard letter-tree recognizer
statistical alignment method
statistical language models
statistical parsing approach
step 1, step 2, etc.
stop list
stopword
stress acquisition model
structure-building module
sub-sequence
support-verb constructions
surface-scope-preserving representations
syntactic-function-based
Table 3 (etc.)
tail-recursive parses
target language (n., adj.)
term expansion (n., adj.)
term extraction (n., adj.)
test data (n., adj.)
test set (n., adj.)
text analysis task
text data mining
text generation process
text-planning process
text-processing program
thematic-relation hypothesis
Theorem 1, Theorem 2, etc.
theorem proving (n.)
thesauruses
third-person pronoun
time series (n., adj.)
top level (n.)
top-level (adj.)
topic prominence
topic-linked concentrated word usage
training data (n.)
training-data (adj.)
training set (n.)
training-set (adj.)
tree-adjoining grammar
tree-adjoining parsing
treebank
treebanking
tree-configurational relationship
tree cut model
tree search algorithm
tree-sentence pair
tree set
tree substitution grammar
trigrams
t-test
two-level transducer
type-checking system
unigram language model
UNIX
unknown-word (adj.)
user model info
vector space (n., adj.)
verb-forming processes
very-high-dimensional spaces
voice mail
vowed

Wall Street Journal (italicized when the
publication itself, specifically, is referred
to)
Wall Street Journal corpus (no italics)
Wall Street Journal Treebank (no italics)
Web
Webmaster
Web-mining (adj.)
Web mining (n.)
Web site
weighted deduction system
weighted deductive parsing
weighted majority algorithm
wh-movement
white space
wide-coverage pure unification grammars
wide-scope brackets
Wizard-of-Oz dialogue (models, experiments,
etc.)
word alignment (n., adj.)
word-based n-gram models
word boundary (n., adj.)
word class (n., adj.)
word-for-word translation
word formation (n., adj.)
word-frequency distribution
word list (n., adj.)
WordNet
word object (n., adj.)
word reordering (n.)
word-reordering (adj.)
word segmentation (n., adj.)
word segmenter (n., adj.)
word sense (n., adj.)
word stream (n., adj.)
word string cover relation
word token (n., adj.)
word type (n., adj.)
World Wide Web
workhorse

X-bar schema

zeroth