

tagcheck.txt *

TreeTaggerで解析した品詞タグチェック記録

・ NICE-NNSから10ファイルと、NICE-NSから3ファイルを対象に、TreeTaggerで付与した品詞タグをチェックし、解析ミスをしていると思われる箇所と修正案を記録しました。

(注) ここで記した品詞タグ修正方法が必ずしも正しいとは限りません。また、チェック記録はデータに反映されていません。データを利用する際の参考資料にしてください。

チェック記録の構成は以下です。

*被験者ID (ファイル名に一致): (タブ) 解析ミスを含むと考えられる箇所のセンテンス

%POSCH: (タブ) 解析ミスをしていると考えられる箇所 (タブ) 修正案

%COM: (タブ) 品詞タグチェック者 (小島) のコメント

・ コメント中にガイドラインとあるのは、Part-of-Speech Tagging Guidelines for the Penn Treebank Project (Santorini, 1991)を指します。(入手先: <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>)

*JPN002: THE JAPANESE SCHOOL EDUCATION

%POSCH: <w pos="NP" base="School">SCHOOL</w> <w pos="NN" base="School">SCHOOL</w>

%POSCH: <w pos="NP" base="Education">EDUCATION</w> <w pos="NN" base="Education">EDUCATION</w>

%COM: 文中に大文字で始まるNNは、NPと解析してしまふ。

*JPN002: Furthermore, 'total education time' is taken in all Japanese junior high school.

%POSCH: <w pos="JJ" base="<unknown>">' total</w> <punc pos="'''" base="'''">'</punc> <w pos="JJ" base="total">total</w>

%COM: 'とtotalを分けて解析していないため。

*JPN002: Firstly, many young people this time cannot read or write basic words (Japanese 'kanji.')

%POSCH: <w pos="NNS" base="<unknown>">' kanji</w> <punc pos="'''" base="'''">'</punc> <w pos="JW" base="kanji">' kanji</w>

%COM: 'とkanjiを分けて解析していないため、またはkanjiが辞書にないため。

*JPN002: So, the level of the study in technology, medicine and so is going down.

%POSCH: <w pos="RP" base="down">down</w> <w pos="RB"

tagcheck.txt *

base="down">down</w>

%COM: go straight downが可能なため。ガイドラインのRB or RP (p.21)参照。

*JPN002: So, the reduction of it means the down of possibility that the evolutional change in various field will happen.

%POSCH: <w pos="RB" base="down">down</w> <w pos="NN"

base="down">down</w>

%COM: the down of のdown (NN) をRBと解析。

*JPN006: Should English education be introduced at primary schools?

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"

base="English">English</w>

%COM: English educationやEnglish languageのEnglish (NP)はJJと解析されるてしまう。

*JPN006: The government is now considering introducing mandate English education into primary schools throughout the nation.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"

base="English">English</w>

*JPN006: There are, however, people who oppose to such a prevalence of English lessons into people of young age.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"

base="English">English</w>

*JPN006: This essay discusses the pros and cons of the introduction of English education into primary schools.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"

base="English">English</w>

*JPN006: Today, an increasing number of primary schools teach the English language as part of the school curricula.

%POSCH: <w pos="NN" base="today">Today</w> <w pos="RB"

base="oday">Today</w>

%COM: 副詞のtodayをNNに解析する傾向あり。

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"

base="English">English</w>

*JPN006: The "integrated learning period," which is a new subject integrating all the existing subjects and attempts to foster the abilities to live in real life, seems to be a good opportunity for English education.

tagcheck.txt *

%POSCH: <w pos="JJ" base="existing">existing</w> <w pos="VVG"
base="existing">existing</w>

%COM: the subjects which exist と言い換えることができる。

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

*JPN006: They use English as a tool to discover something else in
real life.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

*JPN006: For example, a school in Nagoya had their students
interview international staff when they visited the Expo in order to
make classroom newspapers afterwards about foreign people.

%POSCH: <w pos="NN" base="interview">interview</w> <w pos="VV"
base="interview">interview</w>

*JPN006: First, some people suggest that there will be less time
spent on other subjects such as Japanese and math.

%POSCH: <w pos="JJ" base="Japanese">Japanese</w> <w pos="NP"
base="Japanese">Japanese</w>

*JPN006: Moreover, many teachers are complaining that they are
not good enough to teach English to children as they are not trained to
do so.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

*JPN006: On the other hand, the proponents of English education
at primary schools maintain that the younger children start to receive
English education, the more proficient they become in the language when
they get older.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

*JPN006: One of the theories from SLA does suggests that when
they pass the critical period of age it will be difficult to master a
second language.

%POSCH: <w pos="NP" base="SLA">SLA</w> <w pos="NN" base="SLA">SLA</w>

*JPN006: Furthermore, English education brings about
international communications.

tagcheck.txt *

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

*JPN006: It is an extremely good opportunity for young children
to learn about the way people with different cultures think and behave.

%POSCH: <w pos="VV" base="behave">behave</w> <w pos="VVP"
base="behave">behave</w>

*JPN006: The both sides of the arguments sound reasonable.

%POSCH: <w pos="CC" base="both">both</w> <w pos="JJ"
base="both">both</w>

*JPN006: What we need to do now, then, is to deepen these
discussions and search for a better solution on a long term perspective.

%POSCH: <w pos="NN" base="search">search</w> <w pos="VV"
base="search">search</w>

*JPN020: For example, there could be a person who swims in the
gym after his or her work and plays and/or watches baseball on holiday
with family etc.

%POSCH: <w pos="NN" base="<unknown>">and</w> <w pos="CC"
base="or">or</w> <w pos="CC" base="and/or">and/or</w>

%COM: and/orを辞書に入れるべし。(リーダーズとジーニアスにはある)

*JPN020: First of all, we often do sports with a lot of fun.

%POSCH: <w pos="JJ" base="first">First</w> <w pos="RB"
base="first">First</w>

%COM: First of all, のfirstはJJではなくRB。

*JPN020: Playing with friends or family often make us refreshing
and talking with other person then make us close like having a chat
over a cup of coffee or having a dishes with other person.

%POSCH: <w pos="VVG" base="refresh">refreshing</w> <w pos="JJ"
base="refresh">refreshing</w>

%COM: make+0+形容詞のJJをVVGに解析。...make us very refreshingとも言え
ることからも、JJと判断。ガイドラインp.14参照。

%POSCH: <w pos="VV" base="close">close</w> <w pos="JJ"
base="close">close</w>

%COM: make+0+形容詞のJJをVVに解析。

%POSCH: <w pos="VV" base="chat">chat</w> <w pos="NN"
base="chat">chat</w>

%COM: a chat のchat (NN)をVVに解析。名詞のchatが辞書にないものと思われる。

tagcheck.txt *

*JPN020: Some of you might think it sounds wired.
%POSCH: <w pos="VVN" base="wire">wired</w> <w pos="JJ"
base="wired">wired</w>
%COM: sound+形容詞 のJJをVVNIに解析。形容詞のwiredが辞書にないものと思わ
れる。

*JPN020: Some says "Doing sports (for example, walking, swimming
etc. is called "positive rest", while taking rest (for example is
sleeping, lying on the couch etc. is called "negative rest".
%POSCH: <w pos="NP" base="Doing">Doing</w> <w pos="VVG"
base="Doing">Doing</w>
%COM: 文中で大文字で始めると、しばしばNPと解析されてしまう。

*JPN023: How many people enjoy sports by either watching them or
playing them?
%POSCH <w pos="RB" base="either">either</w> <w pos="CC"
base="either">either</w>
%COM: either A or B のeitherはCCとガイドラインp.7にあり。

*JPN023: In school days, like at elementary schools, every
student is supposed to take P.E. classes and enjoy them.
%POSCH <w pos="NP" base="P.">P.</w> <w pos="NN" base="P.">P.</w>
%POSCH <w pos="NP" base="E.">E.</w> <w pos="NN" base="E.">E.</w>
%COM: N.P.はphysical educationの略。固有名詞ではなくて普通名詞。ガイド
ラインp.19参照。

*JPN023: It continues to collage days.
%POSCH <w pos="NN" base="<unknown>">days.</w> <w pos="NNS"
base="day">days</w> <punc pos="SENT" base=".">.</punc></s>
%COM: ピリオドがdaysの一部として解析されているため。

*JPN023: Today, I would like to talk about the things that sports
give people who play it or watch it.
%POSCH <w pos="NN" base="today">Today</w> <w pos="RB"
base="today">Today</w>
%COM: 副詞のtodayを名詞に解析してしまう。
%POSCH <w pos="VV" base="watch">watch</w> <w pos="VVP"
base="watch">watch</w>
%COM: 一つのセンテンス中に複数の動詞が使用されている場合、2つ目以降の
動詞の解析に弱い。

*JPN023: First, I will talk about sports will help people to know
each other well.

tagcheck.txt *

%POSCH <w pos="NN" base="well">well</w> <w pos="RB"
base="well">well</w>

%COM: each other は、each/DT other/JJと解析し(ガイドラインのp.25参照)、
knowの後ろに名詞が無いことによるものと思われる。

*JPN023: Third, I would like to show the effect by watching
sports.

%POSCH <w pos="JJ" base="third">Third</w> <w pos="RB"
base="third">Third</w>

%COM: 副詞のthirdを形容詞に解析してしまう。

*JPN023: When I was in U.S. as an international student, I lived
in a dorm on our campus.

%POSCH <w pos="NP" base="<unknown>">U.</w> <w pos="NP"
base="S.">S.</w> <w pos="NP" base="U.S.">U.S.</w>

%COM: U.S.はU.S./NPとガイドラインのp.19にあり。

*JPN023: Though I had never played it and I could not speak
English well, I had to take part in this game as one of the dorm's
players because of lacks of members.

%POSCH <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

%COM: 名詞のEnglishをしばしば形容詞に解析してしまう。

*JPN023: I was very nervous at first.

%POSCH <w pos="JJ" base="first">first</w> <w pos="NN"
base="first">first</w>

%COM: at firstのfirstは名詞に解析されるべき。

*JPN023: So, playing sports together helps people to get along
with well.

%POSCH <w pos="RP" base="along">along</w> <w pos="RB"
base="along">along</w>

%COM: この場合のget alongは自動詞であり、間に目的語は来ない。よって
alongはRBと判断。

*JPN023: In Africa, china, Canada, Australia and so on, we can
play soccer together if we have a ball, and we can watch and talk about
soccer games.

%POSCH <w pos="NN" base="china">china</w> <w pos="NP"
base="China">china</w>

%COM: Chinaを大文字で始めなかったことによる。

*JPN023: He plays baseball in U.S. as Japanese, and broke many

tagcheck.txt *

records.

%POSCH: <w pos="NP" base="<unknown>">U.</w> <w pos="NP"
base="S.">S.</w> <w pos="NP" base="U.S.">U.S.>U.S.</w>
%COM: U.S.はU.S./NPとガイドラインのp.19にあり。

*JPN024: And today, Kotooshu, who had betrayed Asashoryu this season, prompted to the rank; Ozeki.

%POSCH: <w pos="NN" base="today">today</w> <w pos="RB"
base="today">today</w>

%COM: 副詞のtodayをNNに解析する傾向あり。

%POSCH: <w pos="NP" base="<unknown>">Kotooshu</w> <w pos="NP"
base="<Kotooshu>">Kotooshu</w>

%POSCH: <w pos="NP" base="<unknown>">Asashoryu</w> <w pos="NP"
base="<Asashoryu>">Asashoryu</w>

%POSCH: <w pos="NP" base="<unknown>">Ozeki</w> <w pos="NP"
base="<Ozeki>">Ozeki</w>

%COM: 日本語の固有名詞は大文字で始まっていればNPと解析するが、
base="<unknown>"となってしまう。

*JPN024: Either of two strong wrestlers isn't Japanese.

%POSCH: <w pos="CC" base="either">Either</w> <w pos="DT"
base="either">Either</w>

%COM: eitherやbothはCC⇔DTの解析ミス多し。

*JPN024: Asashoryu is from Mongolian and Kotooshu is from Eastern Europe.

%POSCH: <w pos="NN" base="Mongolian">Mongolian</w> <w pos="NP"
base="Mongolian">Mongolian</w>

%POSCH: <w pos="NP" base="<unknown>">Asashoryu</w> <w pos="NP"
base="<Asashoryu>">Asashoryu</w>

%POSCH: <w pos="NP" base="<unknown>">Kotooshu</w> <w pos="NP"
base="<Kotooshu>">Kotooshu</w>

%COM: MongolianはMongoliaの学習者の間違い。

*JPN024: And some of these foreign wrestlers are much strong compared to Japanese ones.

%POSCH: <w pos="RB" base="strong">strong</w> <w pos="JJ"
base="strong">strong</w>

%COM: much strongerだったら、解析ミスしなかったかも。

*JPN024: Especially Kotooshu is now the hottest wrestler in sumo.

%POSCH: <w pos="NP" base="<unknown>">Kotooshu</w> <w pos="NP"
base="<Kotooshu>">Kotooshu</w>

tagcheck.txt *

*JPN024: His good-looking face and muscular body attract Japanese middle aged women.

%POSCH: <w pos="VV" base="attract">attract</w> <w pos="VVP" base="attract">attract</w>

%COM: 主語が長くなると、動詞の解析ミス多し。

*JPN024: There are both positive aspect and negative aspect in this internationalization of sumo.

%POSCH: <w pos="DT" base="both">both</w> <w pos="CC" base="both">both</w>

%COM: both A and BのbothはCCとガイドラインにあり。

*JPN024: First, I dare to say we can be proud that one of our traditional sports is become world-wide one.

%POSCH: <w pos="RB" base="world-wide">world-wide</w> <w pos="JJ" base="world-wide">world-wide</w>

%POSCH: <w pos="CD" base="one">one</w> <w pos="NN" base="one">one</w>

%COM: 一つ目のoneはCDで良いと思われるが、二つ目のoneはNNと考えられる。ガイドラインのCD or NN参照。

*JPN024: Some sports have got the popularity from all over the world, and purveyed.

%POSCH: <w pos="VVD" base="purvey">purveyed</w> <w pos="VVN" base="purvey">purveyed</w>

%COM: haveとpurveyedが離れているため、VVDと解析。

*JPN024: We shouldn't stop the current of the internationalization of sumo.

%POSCH: <w pos="JJ" base="current">current</w> <w pos="NN" base="current">current</w>

*JPN024: Japanese are kind of the weak nation compared to foreign people.

%POSCH: <w pos="JJ" base="Japanese">Japanese</w> <w pos="NP" base="Japanese">Japanese</w>

*JPN024: If the current become over-heated, we may not able to see Japanese sumo wrestlers on TV.

%POSCH: <w pos="JJ" base="current">current</w> <w pos="NN" base="current">current</w>

%POSCH: <w pos="VVN" base="become">become</w> <w pos="VVP" base="become">become</w>

%POSCH: <w pos="VVN" base="over-heat">over-heated</w> <w pos="JJ" base="over-heat">over-heated</w>

tagcheck.txt *

base="over-heat">over-heated</w>

%COM: 学習者がbecomeに三人称単数現在のsを付けていなかったため、過去分詞と解析。。

*JPN024: So wrestles have strange hair style and play strange performance before fights)

%POSCH: <w pos="VVZ" base="wrestle">wrestles</w> <w pos="NNS" base="wrestle">wrestles</w>

%COM: wrestlesはwrestlersのスペリングミスと思われる。

*JPN025: The Death Penalty isn't Over Excuse?

%POSCH: <w pos="NP" base="Death">Death</w> <w pos="NN" base="death">Death</w>

%POSCH: <w pos="NP" base="<unknown>">Penalty</w> <w pos="NN" base="penalty">Penalty</w>

%POSCH: <w pos="NP" base="<unknown>">Excuse</w> <w pos="NN" base="excuse">Excuse</w>

*JPN025: When we consider about death penalty, we hit upon two of the mainly problems.

%POSCH: <w pos="VVD" base="hit">hit</w> <w pos="VVP" base="hit">hit</w>

*JPN025: In fact, there are many crimes which are never apologized in any persons' thought, especially in victims'.

%POSCH: <w pos="VVD" base="apologize">apologized</w> <w pos="VVN" base="apologize">apologized</w>

*JPN025: He should be arrested in the prison and try to be apologized his crime by the victims.

%POSCH: <w pos="VVD" base="apologize">apologized</w> <w pos="VVN" base="apologize">apologized</w>

*JPN025: If he practices that, I don't want to be killed by judgment.

%POSCH: <w pos="NNS" base="practice">practices</w> <w pos="VVZ" base="practice">practices</w>

%POSCH: <w pos="WDT" base="that">that</w> <w pos="DT" base="that">that</w>

*JPN025: Second;" How about the case that the judge was proved of incorrectly one."

%POSCH: <w pos="NN" base="second">Second</w> <w pos="RB" base="second">Second</w>

tagcheck.txt *

%POSCH: <w pos="RB" base="about">about</w> <w pos="IN"
base="about">about</w>

*JPN025: By these over items, I think that we don't need death
penalty and prisoners should live in prison all his life.

%POSCH: <w pos="IN" base="over">over</w> <w pos="JJ"
base="over">over</w>

%POSCH: <w pos="RP" base="in">in</w> <w pos="IN" base="in">in</w>

%COM: この学習者はBy these over items,を「これら上記の項目より」という
意味で使っていた。

*JPN030: In other words, every one of us is both good and evil.

%POSCH: <w pos="DT" base="both">both</w> <w pos="CC"
base="both">both</w>

%COM: both A and B のbothはCCとガイドラインp.7にあり。

*JPN030: When the evil aspect exceeds the good one, people commit
crimes.

%POSCH: <w pos="CD" base="one">one</w> <w pos="NN" base="one">one</w>

%COM: 形容詞+oneのoneは、NNとガイドラインp.29にあり。

*JPN030: I have not checked the result of census conducted a few
months ago; however, I strongly believe that the population of Nagoya
has also been increasing sharply.

%POSCH: <w pos="VVD" base="conduct">conducted</w> <w pos="VVN"
base="conduct">conducted</w>

%COM: 名詞+過去分詞の過去分詞を過去形に解析ミス

*JPN030: This year, Aichi prefecture in which Nagoya belongs held
the world EXPO 2005. World EXPO includes many pavilions that display
items that represent the culture of many countries all over the world.

%POSCH: <w pos="VVN" base="hold">held</w> <w pos="VVD"
base="hold">held</w>

%COM: 間に関係節をはさみ、主語と動詞が離れているための解析ミスだと考え
られる。

*JPN030: The city employed hundreds of workers from other
countries and provide them short term visa in order to make them work
legally.

%POSCH: <w pos="VVN" base="employ">employed</w> <w pos="VVD"
base="employ">employed</w>

%COM: 主語+過去形を名詞+過去分詞に解析ミス。

tagcheck.txt *

*JPN030: The EXPO ended a several months ago, those workers who have temporal visa must have returned to their home countries by now.
%POSCH: <w pos="RB" base="now">now</w> <w pos="NN" base="now">now</w>
%COM: by now のnow(名詞)を副詞に解析してしまう。

*JPN030: In conclusion, due to the nature of human beings, we cannot eliminate the crime completely.
%POSCH: <w pos="JJ" base="due">due</w> <w pos="IN" base="due">due</w>
%COM: due to のdueはdue/INとガイドライン34ページにあるが、JJやRBと解析されてしまう。

*JPN045: They don't want to be tired because it's holiday.
%POSCH: <w pos="POS" base=" ' s" '>' s</w> <w pos="VBZ" base="be">' s</w>

*JPN045: But for that, they practiced practiced practiced ...
Maybe it's over my imagine.
%POSCH: <w pos="VVN" base="practice">practiced</w> <w pos="VVD" base="practice">practiced</w>
%POSCH: <w pos="VVN" base="practice">practiced</w> <w pos="VVD" base="practice">practiced</w>
%COM: practicedが3回使われているが、最初のpracticedのみVVDと解析され、後の2つはVVNと解析。

*JPN045: I love and respect him.
%POSCH: <w pos="VV" base="respect">respect</w> <w pos="VVP" base="respect">respect</w>

*JPN045: We practiced weekdays and weekends, from morning until evening.
%POSCH: <w pos="NNS" base="weekend">weekends</w> <w pos="RB" base="weekends">weekends</w>

*JPN045: Let's play sports with every body!
%POSCH: <w pos="VBZ" base="be">' s</w> <w pos="PP" base="us">' s</w>

*JPN082: Learning English from early age have become quite popular in these days, as we can see from many private English schools with young age classes or books and televisions.
%POSCH: <w pos="NP" base="Learning">Learning</w> <w pos="VVG" base="Learning">Learning</w>

*JPN082: However I have opposite opinions with many parents, so

tagcheck.txt *

from now I'd like discuss about "learning English from early age" bases of my opinions.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP" base="English">English</w>

%COM: English educationやEnglish languageのEnglish/NPIはJJと解析される。

*JPN082: First of all, I think it is not good idea to make children start learning English before they become able to speak their own mother tongue.

%POSCH: <w pos="JJ" base="first">First</w> <w pos="RB" base="first">First</w>

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP" base="English">English</w>

*JPN082: Children start hearing their mother tongue just after they were bourn, and they started to realize those about 5 moths to 1 years old.

%POSCH: <w pos="IN" base="about">about</w> <w pos="RB" base="about">about</w>

*JPN082: English is the foreign language for us as Japanese, but not for the people in the States, for example.

%POSCH: <w pos="JJ" base="Japanese">Japanese</w> <w pos="NP" base="Japanese">Japanese</w>

%POSCH: <w pos="RB" base="as">as</w> <w pos="IN" base="as">as</w>

*JPN082: I discussed about "learning English from early age" form three points, and if I say in one word, I think we shouldn't start learning English from an early age.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP" base="English">English</w>

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP" base="English">English</w>

*JPN082: I believe there are many reasons why parents make their children to learn English even they are young, but I'd like to say that even they don't care what I think, I'd like to just say "Please listen to what your children think".

%POSCH: <w pos="NNS" base="English">English</w> <w pos="NP" base="English">English</w>

*JPN082: If children have good feeling about learning English, it might be good for them, but the worth case is that they learn because their parents said so, and they don't like to do so.

tagcheck.txt *

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

*JPN082: I believe that if people become not interested in
learning English, it is very hard to become interested in.

%POSCH: <w pos="JJ" base="English">English</w> <w pos="NP"
base="English">English</w>

*JPN201: Japan is the country that still carries out death
penalty.

%POSCH: <w pos="IN" base="out">out</w> <w pos="RP" base="out">out</w>

*JPN201: Firstly I reckon that the human being cannot judge the
human being.

%POSCH: <w pos="NN" base="human">human</w> <w pos="JJ"
base="human">human</w>

%POSCH: <w pos="VBG" base="be">being</w> <w pos="NN"
base="be">being</w>

*JPN201: Secondly the crime one has committed cannot be
compensated for the death of the one.

%POSCH: <w pos="CD" base="one">one</w> <w pos="PP" base="one">one</w>

%COM: oneが一般的な三人称単数を指す場合はPPと、ガイドラインのp. 28にあり、
一つ目のoneはPPと解析されているので問題ないと思われる。しかし、2つ目の
oneはPPかNNか？ここではPPと修正した。

*JPN201: In the current system in Japan, even if one was
sentenced to life imprisonment it is possible that the one could be
released in some years.

%POSCH: <w pos="CD" base="one">one</w> <w pos="PP" base="one">one</w>

%POSCH: <w pos="CD" base="one">one</w> <w pos="PP" base="one">one</w>

*NS001: Why I Support The Death Penalty

%POSCH <w pos="NP" base="I">I</w> <w pos="PP" base="I">I</w>

%POSCH <w pos="NP" base="<unknown>">Support</w> <w pos="NN"
base="support">Support</w>

%POSCH <w pos="NP" base="Death">Death</w> <w pos="NN"
base="death">Death</w>

%POSCH <w pos="NP" base="<unknown>">Penalty</w> <w pos="NN"
base="penalty">Penalty</w>

%COM: 文中に大文字で始まるNNは、NPと解析してしまふ。時にはIまで。

tagcheck.txt *

*NS001: I strongly believe in the clear, basic principle of "eye-for-an-eye" biblical justice.

%POSCH <w pos="NN" base="<unknown>">eye-for-an-eye</w> <w pos="JJ" base="eye-for-an-eye">eye-for-an-eye</w>

%COM: ハイフンで結ばれた修飾語はJJとガイドラインp.1にある。

*NS001: If someone deprives another person of their life through an act of premeditated murder, then they forfeit their own right to live as well.

%POSCH <w pos="VV" base="forfeit">forfeit</w> <w pos="VVP" base="forfeit">forfeit</w>

%COM: 一つのセンテンス中に複数の動詞が使用されている場合、2つ目以降の動詞の解析に弱い。

*NS001: Of course this principle would not apply to those who kill someone accidentally or commit murder because they are suffering from mental illness.

%POSCH <w pos="VV" base="commit">commit</w> <w pos="VVP" base="commit">commit</w>

%COM: 一つの節の中に複数の動詞が使用されている場合、2つ目以降の動詞の解析に弱い。

*NS001: There have been many, many cases of convicted murderers who have escaped from jail and killed again or murdered other prison inmates.

%POSCH <w pos="VVD" base="kill">killed</w> <w pos="VVN" base="kill">killed</w>

%COM: have にescaped, killed, murderedが続いていると解釈。一つの節中に複数の動詞が使用されている場合、2つ目以降の動詞の解析に弱い。しかしmurderedはVVNと解析されていた。

*NS001: If someone is executed right after being convicted that guarantees they will never kill again.

%POSCH <w pos="IN/that" base="that">that</w> <w pos="DT" base="that">that</w>

%POSCH <w pos="NNS" base="guarantee">guarantees</w> <w pos="VVZ" base="guarantee">guarantees</w>

%COM: 「そのことが」という前の内容を受けるthatを、従属節を導くthatと解析したことによる。

*NS001: I also believe that certain economic factors justify the death penalty.

%POSCH <w pos="NNS" base="<unknown>">factors</w> <w pos="NNS" base="factor">factors</w>

%COM: factorsはfactorsのスペルミスと思われる。

tagcheck.txt *

%POSCH <w pos="VV" base="justify">justify</w> <w pos="VVP"
base="justify">justify</w>
%COM: 従属節中の動詞の解析に弱い。

*NS001: Why should the public's precious tax money support murdering
monsters who threaten the very existence of our society?

%POSCH <w pos="NN" base="support">support</w> <w pos="VV"
base="support">support</w>
%COM: should からsupportが離れているための解析ミスだと考えられる。

*NS001: Not only that.

%POSCH <w pos="RB" base="that">that</w> <w pos="DT"
base="that">that</w>
%COM: 省略に弱いかも。

*NS001: I also believe that convicted murders should have some, or all,
of their property confiscated and turned over to the families of their
victims as compensation for their loss.

%POSCH <w pos="RB" base="all">all</w> <w pos="DT" base="all">all</w>

%COM: allの後にコンマがなかったら、ちゃんと解析したかも。

%POSCH <w pos="VVD" base="confiscate">confiscated</w> <w pos="VVN"
base="confiscate">confiscated</w>

%POSCH <w pos="VVD" base="turn">turned</w> <w pos="VVN"
base="turn">turned</w>

%COM: haveにconfiscated とturnedが続くと判断した。haveからこれらの動詞
が離れているための解析ミスだと考えられる。

*NS001: By the time the girl was rescued she had already been severely
crippled, both physically and mentally.

%POSCH <w pos="VVD" base="rescue">rescued</w> <w pos="VVN"
base="rescue">rescued</w>

%COM: be動詞の後なのに、解析ミス。

*NS001: On top of that her kidnapper had beaten her and, it is reported,
sexually molested her as well.

%POSCH <w pos="WDT" base="that">that</w> <w pos="DT"
base="that">that</w>

%COM: that の後ろにコンマがあったら、ちゃんと解析したかも。

*NS005: First, there is the issue of whether the person convicted has
actually committed the crime.

%POSCH <w pos="VVD" base="convict">convicted</w> <w pos="VVN"
base="convict">convicted</w>

tagcheck.txt *

%COM: 名詞＋過去分詞の過去分詞を過去形に解析ミス

*NS005: An innocent human being's life will have been taken due to a grievous error.

%POSCH <w pos="VBG" base="be">being</w> <w pos="NN" base="being">being</w>

%POSCH <w pos="RB" base="due">due</w> <w pos="RB" base="due">due</w>

%COM: due to のdueはdue/INとガイドライン34ページにあるが、RBと解析。

*NS005: In fact, this exact scenario is known to have happened multiple times in the U.S.A., and almost certainly in other countries as well.

%POSCH <w pos="NP" base="<unknown>">U.</w> <w pos="NP" base="S.">S.</w>
<w pos="NP" base="A.">A.</w> <w pos="NP" base="U.S.A.">U.S.A.</w>

%COM: U.S. はU.S./NPとガイドラインのp. 19にあり。

*NS005: Third, the argument is occasionally heard that the death penalty is necessary to prevent a string of revenge killings.

%POSCH <w pos="JJ" base="third">Third</w> <w pos="RB" base="third">Third</w>

%COM: 文頭のFirst, とSecond, は副詞に解析するのだが、文頭のThird, は形容詞に解析。

*NS005: For example, in Japan there was once a tradition of "adauchi", in which one killed the murderer of one's loved ones as a form of revenge.

%POSCH <w pos="NP" base="<unknown>">adauchi</w> <w pos="FW" base="adauchi">adauchi</w>

%COM: 学習者が日本語をローマ字にしてそのまま使うと、固有名詞(NP)と解析してしまう。ここでは、外国語のタグFWに修正した。

%POSCH <w pos="CD" base="one">one</w> <w pos="PP" base="one">one</w>

%COM: oneが一般的な三人称単数を指す場合はPPと、ガイドラインのp. 28にある。

*NS005: Although "adauchi" was formally abolished during the Meiji Era, the mentality of revenge remains a subject of discussion in Japan, and is in fact still a theme of television dramas today.

%POSCH <w pos="NP" base="<unknown>">adauchi</w> <w pos="FW" base="adauchi">adauchi</w>

%POSCH <w pos="NN" base="today">today</w> <w pos="RB" base="today">today</w>

%COM: 副詞のtodayを名詞に解析してしまう。

*NS040: One of my concerns is to do with their bilingual ability and the kind of schools they will go to from the age of eleven.

tagcheck.txt *

%POSCH: <w pos="NN" base="eleven">eleven</w> <w pos="CD"
base="eleven">eleven</w>

*NS040: As they live in Japan and are likely to be living here until at least the age of eighteen, it's important that they get a full Japanese education and I have no doubt they will be O.K.

%POSCH: <w pos="NP" base="0.">0.</w> <w pos="NP" base="K.">K.</w>
<w pos="JJ" base="O.K.">O.K.</w>

%COM: 途中でピリオドの入る単語の解析が難しい。

*NS040: I also unconsciously compare what I see and hear about here in Japan with the kind of school I went to in England, a comprehensive school in London, which was very rough indeed.

%POSCH: <w pos="RB" base="about">about</w> <w pos="IN"
base="about">about</w>

*NS040: Although I did reasonably well there, I sometimes wonder how I came out alive and not mentally scarred.

%POSCH: <w pos="VVN" base="scar">scarred</w> <w pos="JJ"
base="scared">scarred</w>

%COM: veryで強意可能なので。

*NS040: The system they have, or had, in comprehensive schools in England was perfect for someone like me, a slow learner who made his way up the different streams, from D to A (the lowest was H).

%POSCH: <w pos="NP" base="D">D</w> <w pos="NN" base="D">D</w>

%POSCH: <w pos="DT" base="a">A</w> <w pos="NN" base="A">A</w>

%POSCH: <w pos="NP" base="H">H</w> <w pos="NN" base="H">H</w>

*NS040: It allowed me to go from a non-foreign language-learning class to learning French and then German, which is what got me into university.

%POSCH: <w pos="NN" base="<unknown>">language-learning</w> <w
pos="JJ" base="language-learning">language-learning</w>

%POSCH: <w pos="JJ" base="French">French</w> <w pos="NP"
base="French">French</w>

%POSCH: <w pos="JJ" base="German">German</w> <w pos="NP"
base="German">German</w>

*NS040: As for my kids, to be honest, my main concern is with the debilitatingly heavy load that pupils suffer under here, coupled with bullying.

%POSCH: <w pos="RB" base="as">As</w> <w pos="IN" base="as">As</w>

%POSCH: <w pos="RB" base="here">here</w> <w pos="NN"
base="here">here</w>

tagcheck.txt *

*NS040: However, in the case of the excessive learning and deadening effect that you sometimes see on students nowadays, I'd like to avoid this.

%POSCH: <w pos="IN/that" base="that">that</w> <w pos="WDT" base="that">that</w>

*NS040: This might mean either having my kids go to a school such as the type I mentioned above, or abroad.

%POSCH: <w pos="VVP" base="go">go</w> <w pos="VV" base="go">go</w>

*NS040: After that, it's a dilemma, but when I think of the obvious example of my own education or of the kind that one hears about and see in American films, I wouldn't be that much keener on putting my kids through those systems.

%POSCH: <w pos="JJ" base="much">much</w> <w pos="RB" base="much">much</w>

参考資料 :

Santorini, B. (1991). Part-of-Speech Tagging Guidelines for the Penn Treebank Project.

(入手先 :

<http://www.ims.uni-stuttgart.de/projekte/complex/TreeTagger/>

)

TreeTaggerのタグセット一覧 :

<http://www.cele.nottingham.ac.uk/~ccztk/treetagger.php>

2009-3-09 KOJIMA Masumi