Representing the Continuum between Arguments and Adjuncts within Predicate-Frames
Pierre Marchal, Thierry Poibeau, Yves Lepage

To cite this version:
Pierre Marchal, Thierry Poibeau, Yves Lepage. Representing the Continuum between Arguments and Adjuncts within Predicate-Frames. NINJAL International Symposium on “Valency Classes and Alternations in Japanese”, Aug 2012, Tokyo, Japan. hal-00783722

HAL Id: hal-00783722
https://hal.archives-ouvertes.fr/hal-00783722
Submitted on 1 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Representing the Continuum between Arguments and Adjuncts within Predicate-Frames

Pierre Marchal\(^1\),\(^2\), Thierry Poibeau\(^2\), Yves Lepage\(^2\)
\(^1\) ER-TIM, INALCO, Paris, France  
\(^2\) IPS, Waseda University, Kitakyûshû, Japan

\texttt{pierre.marchal@inalco.fr, thierry.poibeau@ens.fr, yves.lepage@waseda.jp}

\section*{Introduction}

In the context of an automated task of acquisition of predicate-frames in Japanese, we introduce a method to observe and evaluate distinction criteria between arguments and adjuncts on a large scale. Rather than a binary classification, we are interested in a representation of the continuum between arguments and adjuncts.

\begin{itemize}
  \item \textit{complements} = \textit{arguments} ∪ \textit{adjuncts}
\end{itemize}

\section*{Method and resources}

We aim at modeling the two extremes of the continuum between arguments and adjuncts.

\begin{itemize}
  \item \textbf{Prototypical argument}: a type of complement which appears in every predicate-frame (i.e. usage) of a given verb.
  \item \textbf{Prototypical adjunct}: a type of complement which never appears in any predicate-frame of a given verb.
\end{itemize}

Here, the goal is to build two lists of verbs, one list to retrieve prototypical arguments, one another to retrieve prototypical adjuncts. We rely on an existing resource: 动詞項構造シソーラス. We consider two case markers: wo (prototypical argument), and de (prototypical adjunct).

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
  & Verbs & Predicate-frames \\
\hline
wo-list of verbs & 2,560 & 4,761 \\
  & 49 \% & 46 \% \\
\hline
de-list of verbs & 4,954 & 9,416 \\
  & 95 \% & 91 \% \\
\hline
動詞項構造シソーラス & 5,190 & 10,364 \\
  & 100 \% & 100 \% \\
\hline
\end{tabular}
\caption{Comparison of the two lists of verbs with 动詞項構造シソーラス}
\end{table}

For each list of verbs we retrieve instances of predicate-frames. We process about 1.9M sentences from the BCCWJ with a dependency parser to build examples as in figure 1.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Instance of a predicate-frame}
\end{figure}

We build two lists of examples. In the wo-list, every complement marked with the particle wo is considered as an argument. In the de-list, every complement marked with the particle de is considered as an adjunct.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
  & Arguments / adjuncts \\
\hline
wo-list of examples & 1,041,818 & 504,391 (wo) \\
  & 504,391 (wo) \\
de-list of examples & 1,890,151 & 144,481 (de) \\
  & 144,481 (de) \\
\hline
\end{tabular}
\caption{The two lists of examples}
\end{table}

\section*{Experiment}

We apply two existing argument/adjunct distinction criteria to our lists of examples to compute a measure of autonomy expressing the placement of complements along the continuum. The autonomy measure ranges from 0 to 1. 0 corresponds to an argument, 1 corresponds to an adjunct.

\begin{itemize}
  \item \textbf{Ordering}: arguments tend to be closer to the verb than adjuncts.
    \begin{equation}
    \text{autonomy}_{\text{ordering}}(c) = \frac{\text{distance}(c)}{\text{number of complements}}
    \end{equation}
  \item \textbf{Head dependence}: arguments tend to cooccur with a narrower range of verbs than adjuncts.
    \begin{equation}
    \text{autonomy}_{\text{head dependence}}(c) = \frac{\text{productivity}(c)}{\text{count}(c)}
    \end{equation}
\end{itemize}

Complements are sorted according to their degree of autonomy toward the verb.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Ordering}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Head dependence}
\end{figure}

With both criteria it appears that prototypical arguments tend to have a lower autonomy than prototypical adjuncts.

\section*{Evaluation}

We evaluate distinction criteria on 31,531 examples, all containing at least one complement marked with the wo particle and one complement marked with the de particle.

We add an extra criterion which is the combination of ordering and head dependence (i.e. the product of their degree of autonomy).

Our evaluation method rely on the idea of a continuum between arguments and adjuncts. For a criterion to be relevant to this task, it needs to order properly the complements along the continuum. That is, even if a criterion assigns a low autonomy to an adjunct it can still be accurate if it assigns – in the same context – an even lower autonomy to an argument (as shown in table 3).

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
  & デ格 (wo) & デ格 (de) \\
\hline
Ordering (C1) & 0 & 0.5 \\
Head dependence (C2) & 0.29 & 0.32 \\
C1 + C2 & 0 & 0.16 \\
\hline
\end{tabular}
\caption{Evaluation on the example of figure 1}
\end{table}

It appears that accuracy can be slightly improved by combining different criteria.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|}
\hline
  & Accuracy \\
\hline
Ordering (C1) & 81.40 \\
Head dependence (C2) & 75.07 \\
C1 + C2 & 81.72 \\
\hline
\end{tabular}
\caption{Results of the evaluation}
\end{table}