

關於中日機械翻譯系統中複合語的自動合成

范莉馨 任福繼 宮永喜一 板内香次
日本国立北海道大学 日本CSK

內容簡介

在機械翻譯系統中為了高效率、高精度地進行形態素解析、構文解析，往往希望尽可能地以複合語為單位進行處理。但隨著科學技術的飛速發展，許多辭彙中未登錄的新的專門用語、複合語等大量出現。使機械翻譯系統的正確率不斷下降。加之中文中含有為數眾多的多義性、多屬性單詞，且它們在形態上又無任何區別，若簡單地以複合語為單位進行登錄的話，更容易造成錯誤。

為此在本文中，我們將提出利用複合語合成規則對形態素解析結果進行中間處理，可以合並的單詞合並後作為複合語進行處理的複合語自動合成法。實驗表明：進行上述處理後，不僅可以避免形態素解析的錯誤性，還可減少或解除構文解析的曖昧性。

Automatic Composition Of Chinese Compound Words for Chinese-Japanese Machine Translation System

Fan Lixin Ren Fuji Miyayaga Yoshikazu Tochinai Kouji
Faculty of Engineering, Hokkaido University, MTDept., CSK

Abstract

In the computer processing of Chinese technical documents, there exists a problem that a number of new compound words which are not registered in a dictionary are appeared, according to the rapid development of science and technology.

In this paper, a method for automatic composition of compound words is proposed. The method extract characters or words which can form compound words and combine them in the morphological analysis of Chinese sentences, and therefore reduce or eliminate ambiguities in succeeding analysis phases. This method is especially suitable for machine translation systems, since a number of processing steps exist from the morphological analysis of input sentences to the composition of output sentences, and ambiguities on the morphological analysis will give serious influences on succeeding processing steps. Based on investigations of compound words in a large number of textbooks and technical documents in Chinese, composition rules for compound words has been extracted, and an experimental Chinese - Japanese machine translation system has been constructed. Results of experiments show that the correct translation rate is about 95%, and it is concluded that the proposed method is effective.

〒214 日本国川崎市多摩区菅馬場2-1-1-203

TEL & FAX: (081) 44-945-1719

1. 引言

目前，世界各地都在大力開展有關機械翻譯的研究工作。一些機械翻譯系統正推向實用化、商品化，但與此同時，由於原語言與目的語言在表層形式與意思對應關係上的複雜性，以及表現方法上的差異等原因，仍存在某些問題有待解決^{(1)~(3)}。此外與以英語為對象的機械翻譯系統相比，中日間機械翻譯的研究歷史不長，未開拓的部分尚有很多正待解決^{(4)~(6)}。

在研究中我們認為：中日機械翻譯系統中複合語的處理非常重要。其原因如下。

中文中，由幾個漢字結合而成的複合語較多，特別是隨着科技水平的日新月異，科技文獻中不斷出現新的複合語（或專門用語，下同）。另一方面，從形態素分析的觀點來分析，為減少構文曖昧性或多重性，往往希望尽可能的以複合語為單位進行處理^(7~9)。因此，在進行中文解析時，不僅要以固有複合語為單位進行形態素分析，還應考慮以新生複合語為單位進行抽出處理。本文中，我們把後者的處理過程定義為複合語自動合成。該方法對本中日機械翻譯系統的最大貢獻可以歸納為以下兩個方面。

1. 1 利用複合語合成減少或消除構文曖昧性

在機械翻譯系統中為能高效率、高精度地進行形態素分析、構文解析，往往希望尽可能地以複合語為單位進行處理⁽¹²⁾。若將一個複合語分別作為一個個單詞來處理的話，在進行構文解析時不僅容易產生曖昧性，還將對以後的翻譯過程以較大影響，從而降低全系統的正譯率。

例①：機器翻譯是自然語言處理的一部分。

對上文進行形態素分析，其結果如表1所示⁽¹¹⁾。

表1 形態素分析結果

Table1: The result of morphological analysis

No.	單語	屬性記号	No.	單語	屬性記号
1	機器	n 4	6	處理	v n
2	翻譯	v n	7	的	u 1
3	是	v 7	8	一	m 1
4	自然	a 2	9	部分	n 0
5	語言	n 4			

其中「翻譯」、「是」、「處理」這3個單詞均具有動詞屬性，均有可能成為句子的謂語。若直接對上述結果進行構文解析，則容易產生曖昧性、多義性。在以往的翻譯系統中，

消除這種構文曖昧性，一般是在對句子進行深層意思解析的同時，根據品詞間的連接程度等來確定相互間修飾·被修飾關係。但由於中文具有如下特徵，用上述方法進行曖昧性消除並非有效^{(12)~(16)}。

a. 構成中文的單詞與單詞之間不存在英文中用來劃分一個個單詞的空格，因此形態素分析結果與日文相同，含有多個候補，構文曖昧性較大。

b. 中文中具有多義性、多屬性的單詞很多。且兼用品詞（動名詞vn，形容動詞av，介動詞pv等）在構詞形態上無任何區別。例如，日文中的名詞「學習」與動詞「學習する」在中文中為同一表記「學習」。

c. 中文中沒有日文中由假名構成、用來提示句子成分的助詞、助動詞等機能語，且又因中文的無形態，動詞、助動詞、形容詞等均不存在語尾變化或活用現象。

為此我們根據中文語法中「多個單詞連接後可以構成同一構文要素」這一原則^{(17)~(19)}，將形態素分析結果中可以連接的相隣單詞並後作為複合語進行處理，以減少或消除可能產生的曖昧性，簡化構文解析。例如，對表1進行複合語合成，既可得表2中的「機器翻譯」、「自然語言處理」、「一部分」等名詞性複合語，從而避免了構文解析的曖昧性。

表2 複合語合成結果

Table2: The result of compound words composition

No	複合語	屬性記号	No	複合語	屬性記号
1	機器翻譯	n 9	4	的	u 1
2	是	v 7	5	一部分	n 9
3	自然語言處理	n 9			

1. 2 利用複合語合成避免形態素分析結果的錯誤

如上所述，在進行中文解析時，將複合語作為構成單詞的最小單位進行處理最為有力。但為此將所有的複合語登入辭書的話，往往會產生錯誤。其原因如下。

a. 隨着科學技術的飛速發展，新的複合語不斷產生。欲將所有的複合語登入辭書中實際上是不可能的；

b. 從構文分析的觀點來分析，固有辭書中業已登錄的複合語仍有不能作為複合語進行處理的場合。

* 該系統把中文實詞按語法分為7大類，分別記作：名詞n，動詞v，形容詞a，數詞m，量詞q，代名詞r，副詞d。且各品詞的下位區分為1,2,3...。方位詞雖然也屬於名詞範疇，但由於它在中日機械翻譯系統中具有特殊作用，我們用h來表示其屬性，以區別於普通名詞。

例②：他用機器翻譯技術文獻。

例①中“機器翻譯”為複合語，若將該複合語登入固有辭書，則在對例②進行形態素解析時就會產生如下結果：[他，用，機器翻譯，技術文獻]，顯然這是完全錯誤的。因為例②中的“翻譯”才是句子謂語，故其正確結果應為[他，用，機器，翻譯，技術文獻]。

為解決上述問題，該系統規定含有動詞性兼用品詞的複合語不直接登入辭書中，只在必要時進行自動合成。換言之：名詞“機器”和動名詞“翻譯”在例①中可以合並為名詞性複合語“機器翻譯”進行處理。而在例②中則不能進行合並，必須分別作為2個單詞進行處理。

我們從文獻[22~24]中隨機抽出49篇譯文(計1,862個句子, 76,342個漢字)，對複合語的數量和類型進行了抽樣調查。共抽出複合語1,614個，並求出了各自所屬類型²⁵。為確立複合語合成規則打下了基礎。

總上所述，本文將說明一個用於中日機譯系統的複合語自動合成法。其特徵為：①為了尽可能地避免形態素解析結果的錯誤，將含有動詞性兼用品詞的複合語不直接登入固有辭書中，而是在必要時進行自動合成處理；②為了尽可能地減少或消除構文解析結果的曖昧性，應尽可能地在初期階段進行複合語合成；③利用統計法求出的複合語合成規則進行複合語合成，以縮短中間處理所需時間。

在此基礎上，我們利用該方法構成了複合語自動合成系統，並根據實驗結果確定了該手法的有効性和可行性^{21, 25}。

下面，我們將在第2章說明複合語的自動合成，第3章說明特殊品詞的解析方法，第4章說明該手法的實驗與考察等。

2 複合語的自動合成

2.1 複合語合成規則

如上所述，中文既沒有英文中用來劃分一個個單詞的空格，也沒有日文中用假名(加、ヲ、ニ、デ...)寫成用來提示句子成分的助詞、助動詞等機能語，且不存在語尾變化。然而中文語順嚴密，主語、謂語等文法關係主要是根據單詞的排列順序來確定的。因此語順極為重要^{(13)~(20)}。

本文對從大量中文譯文中抽出的單語間連接關係進行了研究，着重考慮了以下諸問題後，進而求出了複合語合成規則(見表3)。

a. 如上所述，複合語合成是在形態素解析與構文解析之間進行的中間處理過程，參加複合語合成的單詞必須是位於形態素解析結果中的有所屬單詞，換言之其品詞屬性為已知。

b. 合成複合語的最終目的是為了簡化構文解析，減小或消除構文上的曖昧性以及錯誤等。因此考慮複合語合成規則時，只考慮在文中具有文法屬性的實詞間的連接規則。

c. 在進行機械翻譯時，應尽可能地縮短處理時間。因此

該系統中單詞間的連接順序規定為右方向性連接，從而使多重合並處理(見2.2節)成為可能。

據此，我們求出了品詞間的可能的連接關係表^{(10)~(11)}。但從中日機械翻譯系統的實際使用度以及構文解析的容易程度來考慮，以下情況將不進行複合語合成：①四字熟語及歇後語；②名詞和方位詞間的連接關係；③副詞與其他單詞間的連接關係。副詞雖然也屬於實詞範疇，但在中日機譯系統中作為謂語判定的關鍵字具有特殊作用，須單獨進行處理。④可構成獨立句子的構文構造。

表3 複合語合成規則

Table3: The rules for compound words composition

品詞名	No.	連接表記	複合語屬性	優先順序
名詞 n	1	n*/n*/n*	n 9	A
	2	n*/n*	n 9	A
	3	n*/vn	n 9	C
動詞 v	4	vn/n*	n 9	C
	5	vn/vn	n 9	C
形容詞 a	6	a1/n*	n 9	A
	7	a2/n*	n 9	A
	8	an/n*	n 9	A
	9	a*/vn	n 9	C
	10	av/n*	n 9	C
數詞 m	11	m*/n*	n 9	A
	12	m*/m*	m 9	A
	13	m*/q*	m q	B
代名詞 r	14	r*/r*	r 9	A
	15	r*/q*	r 9	A
	16	r*/h2	r 9	A

2.2 多重合並處理

如上所述，該系統中單詞間的連接順序規定為右方向性連接。且根據複合語合成規則合並成的複合語，又可根據其文法屬性再次進行複合語合成。下面用例①來說明其處理過程。

例①：機器翻譯是自然語言處理的一部分。

(1) 進行形態素解析：

其結果如表1所示，為9個單詞組成的一串碼列。

機器/翻譯/是/自然/語言/處理/的/一/部分。

(2) 根據複合語合成規則進行中間處理：
 (機器翻譯)／是／(自然語言)／處理／的／一部分。

第一層：{No.3}, {No.7}, {No.11}

機器翻譯／是／(自然語言處理)／的／一部分。

第二層：{No.3}

機器翻譯／是／自然語言處理／的／一部分。

(3) 進行構文解析與交換處理：該文為典型的「是字句」文型。很容易進行構文解析，由此得出的譯文結果如下，

中文構文：機器翻譯／是／(自然語言處理)的一部分。

↓ ↓ ↓
 (s) (v) (c)

日文譯文：機械翻譯は自然言語処理の一部分である。

2. 3 優先順序的設定

為實現多重合並處理必須對各品詞間的連接規則進行優先順序的設定。我們從中文自身特徵和實現中日翻譯系統的容易度來考慮，將連接優先順序分為3段。見表3^{11, 21, 25}。

3. 特殊品詞的解析

中文中具有如下所示的特別品詞，構成了特殊文型。本章將分析這些品詞的構詞特徵及它所構成的特殊文型，進而說明對這些文型進行中間處理時必須注意的幾個問題。

3. 1 三向動詞

3. 1. 1 定義

中文語法中對三向動詞有如下定義^{11, 19}：同時與3個名詞性構成要素相連係的動詞叫三向動詞。統計表明，常用三向動詞共有16個分別為：報告、稱、稱呼、給、告訴、選教、交、叫、借、留、賠、送、通知、問、贈¹¹。



3. 1. 2 文型及處理規則

中文中含有上述三向動詞的文型有以下兩種。

(1) SV00型 (v1為三向動詞)：

$s1 + v1 + o1 + o2 \dots \dots (1)$

例④：老師送我翻譯文獻／先生は私に翻訳の文獻を送る。

例⑤：我送給他學習用品／私は彼に學習用品を送る。

(2) 兼語式 (v2為動詞或具有動詞屬性的兼用品詞)：

$s1 + v1 + os2 + v2 + (o1) \dots \dots (2)$

例⑥：他叫我們翻譯文獻／彼は私達に文獻を翻訳させる。

例⑦：我教他包餃子／私は彼に餃子を作るのを教える。

由上述例句可知，含有三向動詞的句子有時能用複合語合成規則進行中間處理有時則不能。例如動名詞“翻譯”與名詞“文獻”在例④中可以合並為複合語，在例⑥中却不能。

為此，我們必須考慮對含有三向動詞的句子如何進行複合語合成這一問題。對含有16個常用三向動詞的文型進行分析研究後，得出了以下3個規則。

規則I：位於三向動詞前的兼用品詞與其他的品詞進行複合語合成時，表3適用。

規則II：位於三向動詞「教」「叫」「通知」「問」後的動詞性兼用品詞為句子的謂語，不能根據表3進行複合語合成。含有其他三向動詞的句子可以根據表3進行複合語合成。

規則III：含有「三向動詞+給」的句子可以根據表3進行複合語合成。

3. 1. 3 解析例

下面我們將用2種含有三向動詞的句子來說明其處理過程。

例⑧：這個翻譯程序叫我們翻譯文獻。

(1) 形態素解析：

CHO(1)=這個	CTY(1)=r 2
CHO(2)=翻譯	CTY(2)=v n
CHO(3)=程序	CTY(3)=n 4
CHO(4)=叫	CTY(4)=v 1
CHO(5)=我們	CTY(5)=r 1
CHO(6)=翻譯	CTY(6)=v n
CHO(7)=文獻	CTY(7)=n 4

其中含有3個動詞：三向動詞「叫」和2個動名詞「翻譯」。

(2) 複合語合成：

根據規則I，位於三向動詞「叫」前的動名詞「翻譯」可與相鄰的名詞「程序」按複合語合成規則No.4 {vn/n*→n9}合並成名詞性複合語{翻譯程序}：

這個{翻譯程序}叫我們翻譯文獻。

根據規則II，位於三向動詞「叫」之後的動名詞「翻譯」為句子的謂語。因此不能與其他單詞進行合並。

(3) 構文解析與交換：例⑧為典型的兼語句。其譯文如下，中文構文：這個／翻譯程序／叫／我們／翻譯／文獻。

↓ ↓ ↓ ↓ ↓
 s1 v1 os2 v2 o1

日文譯文：この翻訳プログラムは我々に文獻を翻訳させる。

例⑨：他送我學習用品。

(1) 形態素解析：

CHO(1)=他	CTY(1)=r 1
CHO(2)=送	CTY(2)=v 1
CHO(3)=我	CTY(3)=r 1
CHO(4)=學習	CTY(4)=v n
CHO(5)=用品	CTY(5)=n 4

其中共有2個動詞：三向動詞「送」和動名詞「學習」。初看該文似乎為兼語句式，若直接進行翻譯處理，將會生成錯誤譯文：“彼は私に用品を學習するのを送る”。實際上動名詞“學習”為修飾“用品”的名詞性兼用語。二者在句中構成同一構文要素，應作為複合語進行處理。

(2) 複合語合成：

根據規則II，位於三向動詞「送」之後的兼用品詞「學習」與名詞「用品」可按合成規則No.4合成複合語「學習用品」。

(3) 構文解析與交換處理：該文為SV00型。其譯文如下，

中國語構文：他 / 送 / 我 / {學習用品}。

↓ ↓ ↓ ↓
s1 v1 o1 o2

日本語譯文：彼は 私に 學習用品を 送る。

3. 2 兼用品詞

3. 2. 1 構詞特徵與複合規則

兼用品詞，特別是動詞性兼用品詞的處理是中文解析、機械翻譯中的疑點之一。因此有必要對它進行進一步的研究。

由於兼用品詞在構文形態上沒有任何變化，所以增加了構文解析的難度。

例⑩：他的生活經歷艱辛，曲折。

例⑪：他的生活經歷了兩個時代。

其中「生活」「經歷」為動詞與名詞的兼用品詞，且二者在形態上無任何區別。若直接進行複合語合成就會合成複合語「生活經歷」，在之後的構文解析中例⑩必將產生錯誤。

原形	名詞性譯語	動詞性譯語
生活	生活	暮らす
經歷	經歷	経る

據統計，常用兼用品詞共有356個，其中動名詞238，形容動詞118個。在對上述兼用品詞的構文特徵及關連情報等進行抽樣調查時，得出了其在進行複合語合成時的3個限制條件。

規則IV：當句子中只含有1個動詞性兼用品詞，且文末不是形容詞時，則該兼用品詞為句子的謂語，不能與其他單詞進行複合語合成。

規則V：當句子中只含有2個相鄰的動名詞，且文末不是形容詞時，則這2個動名詞在句中應並為複合語進行處理。

規則VI：當句中含有“用”“動”“準許”等付屬動詞時，位於其後的動詞性兼用品詞不能與其他品詞進行複合語合成。

下面為含有兼用品詞的例句。其中例⑫為不能合成複合語的例句，例⑬、⑭為可以合成複合語的例句。

例⑫：那是翻譯文獻。

譯文：それは 翻訳文獻である。

例⑬：他正在翻譯。

譯文：文獻/彼は 文獻を 翻訳している。

例⑭：他的翻譯水平很高。

譯文：彼の 翻訳レベルは かなり 高い。

3. 2. 3 解析例

本節我們將用例⑫，例⑬，⑭來說明兼用品詞的解析過程。

例⑫“他用機器翻譯技術文獻”中的名詞“機器”與動名詞“翻譯”似乎可以按照表3合成名詞性複合語“機器翻譯”。但因該句中的“用”為付屬動詞，按規則VI其後的兼用品詞“翻譯”為句子的謂語，不能與「機器」合成為複合語。

例⑬“他的生活經歷艱辛，曲折”中含有2個連續排列的兼用品詞「生活」、「經歷」，且文末為形容詞。按規則V，例⑬中的2個動名詞可按照表3中的{No.5: vn/vn→n9}合成名詞性複合語「生活經歷」。

而例⑭“他的生活經歷了兩個時代”中雖含有2個連續排列的兼用品詞「生活」、「經歷」，但文末單詞不是形容詞，所以不能按表3進行複合語合成。

上述3個句子的譯文如下，

譯文⑫：彼は 機械で 技術文獻を 翻訳する。

譯文⑬：彼の 生活經歷は 苦しく、紆余曲折している。

譯文⑭：彼の 生活は 二つの 時代を 経てきた。

4 考察

依據此法進行的實驗結果表明，該複合語合成有效地減少了構文解析中的曖昧性與多義性。此外，本文在進一步研究特殊品詞的構文特徵後利用制限條件，避免了可能產生的曖昧性，減輕了深層構文解析的難度。

且通過實驗分析表明：對具有構文與意思上多重曖昧關係的句子，不能單純判斷能否進行複合語合成，而需要進行深層意思解析和分析它與前、後共起語的關係。例如，

例⑮：他給我翻譯文獻。

該句中，不僅具有構文上的曖昧性，而且還具有意思上的曖昧性。既同時具有如下所示的2種構文樹。

構造1(兼語式)：他 給 我 翻譯 文獻。
↓ ↓ ↓ ↓ ↓
s1 v1 o1 s2 v2 o1
(日語譯文)：彼は 私に 文獻を 翻訳してくれる。
構造2(SV00型)：他 給 我 {翻譯文獻}。
↓ ↓ ↓ ↓
s1 v1 o1 o2
(日語譯文)：彼は 私に 翻訳文獻を 与える。

換言之：構造2中的動名詞「翻譯」必須與名詞「文獻」相結合，共同作為句子的謂語三向動詞「給」的選賓語o2。而在構造1中「給」用於指定授受對象，動名詞「翻譯」則為句子的謂語，與名詞「文獻」構成動賓關係，二者不能進行複合語合成。如何確定其構造屬性，必須根據其上下関連情報來決定。

5. 結言

中文中含有為數眾多的多義性、多屬性單詞，且兼用品詞在構文形態上無任何區別，加之不斷產生的複合語（專門用語），當中文與其他語言進行機械翻譯時，如果不能正確地識別和處理它們，則容易造成構文解析的曖昧性和多義性，進而影響最終譯文的正確性。

本文提出了利用複合語自動合成對形態素解析結果進行中間處理，使可以合並的單詞作為複合語進行處理，從而減少了構文解析的曖昧性與譯文的錯誤率這一方法。並在由此構建的實驗系統上進行了複合語合成試驗，確認了其有效性。

參考文獻

- 1) 田中聰積，野村浩輝：“機械翻譯”，ビット別冊，(1988)。
- 2) 長尾真：“機械翻譯サミット”，オーム社(1989)。
- 3) 牧野武則：“機械翻譯”，オーム社(1989)。
- 4) 寺下陽一，二口邦夫，鈴木悟：“並列型パーサによる中国語文解析”，情処研NL74-12(1989.09)。
- 5) 任，范，極内，宮永：“家族モデルを用いた文の分解に基づく日中機械翻訳システム”，情処論文誌Vol.32 No.10, pp.1249-1258(1991)。
- 6) 寺田栄男，孫東傑，田町常夫：“簡易型中日機械翻訳実験システムについて”，情処研NL75-8(1990.01)。
- 7) 長尾真，辻井潤一：“国語辞書の記憶と日本語文の自動分割”，論文誌Vol.19 No.6 pp514-521(1978)。
- 8) 武田浩一，藤崎哲之助：“統計的手法を用いた漢字複合語の短単位分割”，自然言語処理48-2(1985.03)。
- 9) 宮崎正弘：“係かり受け解析を用いた複合語の自動分割法”，論文誌Vol.25 No.6 pp.970-979(1984)。
- 10) 范，任，宮永，極内：“中国語文中の複合語の生成について”，情処研NL90, A190-29, pp.1-8(1990.5)。
- 11) 范：“中国語文解析システムにおける複合語の自動抽出に関する研究”，修士論文(1991.03)。
- 12) 辻井潤一：“辞書の構成と機械翻訳”，情報処理学会自然言語処理研究会Vol26, No.10(1985.10)。
- 13) M・B・ソフローノフ(橋本万太郎訳)：“中国語機械翻訳の一般原理”，中国語学(1961)。
- 14) ヴェ・エム・ソフツェフ(望月八十吉訳)：“現代中国語概論”，中国語学(1964)。
- 15) 劉永泉：“機器翻訳浅説”，中国語文12月号(1958)。
- 16) 望月八十吉：“中国語と日本語”，光生館(1974)。
- 17) 劉月華，吳文焯，他：“實用現代漢語語法”，中国外語教学与研究出版社(1986.10)。
- 18) 三野昭一：“中国語文法の基礎”，三修社(1987)。
- 19) 陳國梁：“現代漢語語法教程”西安交大出版社(1986)。
- 20) 朱美英，内田裕士：“連接關係に基づく中国語辞書データの推定”，情処研NL73-4(1989.06)。
- 21) 范，任，宮永，極内：“中国語複合語生成の実験システム”，北連大論文集286(1990.10)。
- 22) 張西祥，孫麗英：“科学与人類”機械工業出版社(1987)。
- 23) 孫久明：“科技日語第1冊”科学普及出版社(1984)。
- 24) 蘇培成：“語言病例分析”南開大学出版社(1987)。
- 25) 范，任，極内，宮永：“中日機械翻訳における中国語複合語の自動合成について”情処論文誌Vol.33, No.09 pp.1103-1113(1992.09)。

作者簡介

范莉馨：1991年日本北海道大学大学院博士前期課程修了，工学修士。現在同上博士後期課程在学。主攻方向為自然語言處理、機械翻譯等。日本電子情報通信学会、ACSSJ各會員。

任福繼：1991年日本北海道大学大学院博士後期課程修了，工学博士。現在日本CSK主任研究員、大連理工大学兼職教授。主攻方向為計算機科學、人工智能、自然語言處理、機械翻譯等。日本情報處理学会、日本人工智能学会、亞洲太平洋機械翻譯協會各會員，ACSSJ會長。

宮永喜一：1961年日本北海道大学大学院博士後期課程修了，工学博士。現在北海道大学工学部電子工学科助教授。主攻方向為並列計算機系統、數據信號處理等。日本電子情報通信学会、日本音響学会各會員。

極内番次：1964年日本北海道大学大学院博士前期課程修了。現在北海道大学工学部電子工学科教授，工学博士。主攻方向為自然語言處理、聲音情報處理及信號處理等。日本電子情報通信学会、日本情報處理学会、日本音響学会各會員，日本電子情報通信学会北海道地区支部長。