特　集「機械翻訳」

座談会「機械翻訳における中間言語方式をめぐって」

1989. 5. 26.　10：00～12：00　於：オーム社

●出　席　者●

司　会　田中穂積（東京工業大学）
出席者　石崎　俊（電子技術総合研究所）　天野真家（東芝）
内田裕士（富士通）　梶　博行（日立）
村木一至（日本電気）　中村順一（九州工業大学）

司会（田中）　本日はお忙しいところお集まりいただきましてありがとうございます．我が国は，機械翻訳に関する研究・開発の経験につきましては世界のトップレベルにあるといってもよいと思います．機械翻訳と一口に言いましても，実現の仕方がいろいろあります．特に中間言語方式とトランスファ方式のうち，どちらが良いかということが最近議論の対象になっていると思います．そこで本日は中間言語方式とは何か，トランスファ方式とは何か，両者の特徴，歴史的位置づけ，中間言語方式ですとそれをどのように設計するか，中間言語方式とトランスファ方式の合流はあるか，などという問題をお話いただきたいと思います．それでは初めに，中間言語方式とは何かということについて，厳密な定義は難しいと思いますが御説明いただきたいと思います．

中間言語方式とは

内　田　機械翻訳といいますのは，翻訳すべき言語（原言語）の文章を，翻訳したいと思う言語（目標言語）の文章に機械的に変換することですが，この言語間の変換を行う方式として，中間言語方式とトランスファ方式とがあります．機械翻訳を行うとき，原言語の文章を解析し計算機にわかるような中間表現に直してから目標言語に変換する（目標言語を生成する）わけですが，中間表現として，さまざまな言語に対して共通のものが得られたら，この一つの中間表現から，さまざまな目標言語に変換することができる．つまり解析は原言語と中間表現の対，一方，生成は目標言語と中間表現の対のみを考えればよく，例えばn個の言語間の翻訳を行うことを考えますと，作成すべき翻訳システムの数はnのオーダであるということになります．このとき，中間表現は意味構造であると考えてもよいので，この意味構造を記述する言語が中間言語であると我々は考えております．中間言語で表現された意味構造には，不必要な曖昧性が含まれてはいけません．以上をまとめますと，中間言語方式とは，このような中間言語による中間表現を介して原言語を目標言語に変換する方式のことで，多言語間翻訳に都合のよい方式であるということになります．

村　木　外国人との身振り手振りによるコミュニケーションにより，何とか通じ，よく通じたなと思うことがあります．わずかな語彙で事物をさし，身振り手振りですから不十分ではありますが，たとえ互いに異なる言語を使っても，似た環境で生活している二人の間には共通にわかり合える言語があるのではないかと考えられます．そうしますとその共通にわかり合える言語を介して，さまざまな言語に変換し翻訳する方式が考えられます．これを中間言語方式

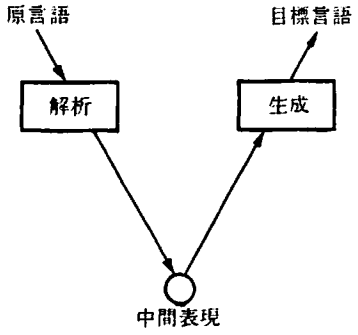

図 1　中間言語方式の概念図

と呼びます．中間言語としてさまざまなものが考えられますが，私どもの翻訳システムで使っている中間言語のことをPIVOT言語と呼んでいます．PIVOT言語を用いて辞書の意味を記述し，原言語の文章を解析してPIVOT言語で記述した中間表現を抽出し，その中間表現から目標言語の文章を生成する方式が，私どもの考えている中間言語方式です．

石　崎　トランスファ方式と対比させる形で中間言語方式とは何かを述べてみたいと思います．後で梶さん，天野さん，中村さんから詳しい説明があると思いますが，端的に言ってトランスファ方式というのは，言語対ごとに必ず変換部分がありますから，n言語間の翻訳を行う場合には，${}_nC_2$通りの翻訳システムを作らなければならない．これはn^2のオーダになりますから，nの数が多い多言語間翻訳システムを作成する場合に問題になります．ある言語の文の意味を表す中間表現があるとき，中間表現を書く道具が中間言語ですから，中間言語は，日本語，英語，フランス語に近いというより，コンピュータ言語に近いと考えたほうがよいと思います．そして中間表現は曖昧性のない概念を組み合わせて構成します．普通，自然言語の文には曖昧性がありますよね．エスペラント語でさえ文脈に依存した曖昧性が含まれます．中間言語で表現された中間表現には，曖昧性が含まれないものが土台になっている．したがって，中間言語は自然言語とは異なる側面を持っていることを知っていただきたいと思います．

司　会　石崎さんも言われましたが，中間言語方式も実はトランスファ方式の定義を聞くと一層際立ちます．そこで今度はトランスファ方式とは何かについてお話しいただきたいと思います．

トランスファ方式とは

梶　理想的な中間言語方式を実現するためには，日本語や英語などといった言語に独立な表現が必要ですが，実際にはそれはほとんど不可能といってよいのではないでしょうか．そこで，機械翻訳を行うとき，原言語に依存した内部表現と，目標言語に依存した内部表現の二つを設け，二つの表現の間で，語彙項目や構造の変換を行う過程を設ける方式が考えられます．これがトランスファ方式です．つまり原言語の文を原言語に依存した内部表現に変換する「解析」，原言語に依存した内部表現から目標言語に依存した内部表現に変換する「トランスファ」，目標言語に依存した内部表現を目標言語の文に変換する「生成」の三つのフェーズが，トランスファ方式には含まれます．トランスファ方式は，内部表現の違いによって，構文トランスファ，意味トランスファとに分類されることもあります．

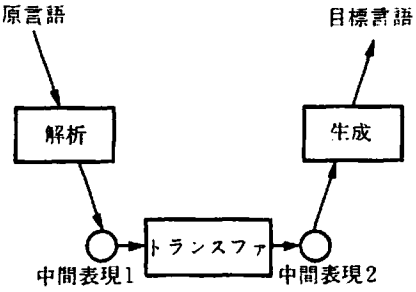

図 2　トランスファ方式の概念図

天　野　歴史的に言いますと，中間言語方式は，1960年代にフランスのグルノーブル大学のCETA研究所で行われていました．そこでの中間言語は，とにかく現存する自然言語に独立な概念の集合からなる人工言語であったと記憶しています．人間なら人間，椅子なら椅子という概念が英語，日本語など，自然言語とは独立に存在するという前提に立っている．これにつきましては，既に相当議論されていて，そのようなものはないとの結論が出ている．言語は民族の文化を担っているので，民族の自然の分節，見方に依存している．ランダムハウスでmoustacheを引いてみると，原文のlipを「唇」と訳しますと，おおよそ「上唇に生える毛」ということになりますが，これは日本語の唇と，英語のlipの分節の違いに起因している．このような食い違いが語彙の多くの部分で生じるので，CETAでは中間言語方式をやめて，言語ごとに異なる部分を認めたトランスファ方式を考えだしたわけです．これは翻訳の質と正確さに多大な影響を与えます．翻訳は基本的には二か国語間で行うもので，原言語側の表現と最も近い目標言語側での表現は何かを考えながら行う．その意味で，中間言語という雑挟物を間に挟むのは面白くない．中間言語方式は目標言語を意識しないのに対して，トランスファ方式では

むしろそれが翻訳には必要であると考える．

ただ最近新しい立場から中間言語が言われています．それを仮にここでは言語共通という言葉を使いますと，言語独立な語彙は諦めて，現存するすべての言語の語彙の和集合を，つまり，すべての言語に現れるものを中間言語の語彙と考える．したがって，この新しい中間言語方式では，原言語の概念表現から目標言語の概念表現にトランスファする過程が，辞書の記述中に直接書かれていて，トランスファ方式のように突出した過程になっていない，それが特色だと思います（図3参照）．

石　崎　現在のトランスファ方式で，どういうものをトランスファしているかという点が重要です．トランスファ方式は構文トランスファとか意味トランスファとか言っています．

天　野　それはすごく誤解があります．グルノーブル大学で書かれた図1では，解析をとことんやると中間言語方式になる．中間言語方式はそこまでやっているので翻訳の品質が良いと誤解されていますが，現在の技術はそこまでいっていない．翻訳技術の観点からみますと，中間言語方式もトランスファ方式もさほど差がないと思います．

中　村　両方式について少し別の観点から補足したいと思います．両方式を議論するときの前提として，語彙，シンタックス，セマンティックスの三つを区別して議論しないといけないと思います．語彙の問題についてはさらに専門用語と一般用語とを分けて議論すべきでしょう．一般用語，例えば体の部分の名称については，言語共通の語彙を設定することは非常に難しい．天野さんの moustache はその一例です．そのほかに動詞としての knee の意味は「膝で押す」の意味ですから，日本語では目的語まで含んだ表現になります．語彙の対応はこのように複雑ですから，その処理方法に，中間言語方式とトランスファ方式とで差が出てくる．それに対して専門用語では，例えば「コンピュータ」は各言語でそれほど大きく異なるはずがないので，両方式ともさほど大きな差が出ない．

次に文の意味解析を行って，「この文のこの語が行為主体で，この語が行為の対象である」という中間表現を得ている場合には，いずれの方式を採用しても大きな差が出てこない．ただし，表層レベルの翻訳のテクニックに依存するようなこと，すなわち典型的な訳し方などを考えると，

それを中間言語でうまく表現できるかどうか疑問が残ります．

語彙の問題

石　崎　中村さんが考えている語彙についてもう少し詳しく説明していただけないでしょうか．

中　村　機械翻訳の方式を比較する前提として，語彙とシンタックスとを区別して議論しようということです．トランスファ方式であれば，原言語の表層の単語の意味を原言語の観点から区別できる範囲で区別し，これを内部表現の語彙にします．中間言語方式であれば，この区別は理想的には全言語共通の区別になると思います．英語の head は「首から上の部分」という意味で，英語ではそれ以上意味の区別をしません．日本語の「首を切り落とす」も「顔を隠す」もいずれも head を使う．トランスファ方式では，head の意味として「体の部位」，「場所」，「頭脳」といった区別は英語の側でしますが，「体の部位」を更に細分することはしない．細分化はトランスファの段階で行うことになります．

内　田　中間言語方式とトランスファ方式で用いる語彙を比較するとき，中間表現を構成する語彙が何なのかをもう少しはっきりさせないといけないと思う．中間言語方式を採用していると言っている人でも，中間言語の語彙については必ずしもコンセンサスが得られていないのではないでしょうか．私は，各言語が共通に持っている概念の集合と，各言語が個別に持っている概念とを足し合わせたものが，中間言語の語彙であると考えます．それに対して，言語独立な基本概念（プリミティブな概念）を集めたものが中間言語の語彙であるとする考え方もあります．

天　野　言語共通と言語独立とをごちゃ混ぜにして議論をしてはいけないと思います．言語独立というのは，例えば人体を現存の言語とは独立に，どう分節するかを考えることです．人体の部分の切り分け方にはいくらでもあるわけで，言語独立の問題はまさにここにあるといってよいのではないでしょうか．

司　会　中村さんの例で言えば，体の部位については，言語によっていろいろ指す場所が異なる．ただこれは，言語による差というより文化による差ではないでしょうか．その点はどうでしょうか．

天　野　言語ごとに自然の分節が異なるということだと思います．

内　田　言語における差というより，言語が持っている概念の差があるっていうことですよね．色を3色にとらえる言語は，色についての概念は三つしかな

い．7色にとらえる言語は，色についての概念が七つしかない．言語ごとに独立に概念があるわけです．それがたまたまある言語の「赤」という概念と，別の言語の red という概念とが似ていてそれらが対応しているから，「赤」を red と訳すことで翻訳が行われる．

中　村　今の分節の議論は，機械翻訳だけの問題ではなく，翻訳可能性一般の議論でしょう．これは翻訳論の研究としては面白いと思いますが，実用的あるいは工学的な立場からの機械翻訳の研究としてはあまり有益ではないと思います．ソビエトには「訴訟」という概念がないから，「訴訟」を翻訳することは不可能だという議論もありますが，それを今言っても生産的ではないでしょう．

内　田　私はそういうことを言ったのではなくて，翻訳するとき目標言語に原言語の概念がなくても，目標言語の側でそれに近い概念を選んで翻訳すればよいということで，言語ごとに概念のセットが違うということはあり得るし，あり得ても翻訳は可能だということを言いたかったのです．

梶　中間言語の語彙を，対象とする言語の持つ概念の和集合として定義したとき，本質的にはトランスファ方式と同じことになるのではないかと思います．中間言語方式の場合は，原言語の語から概念，概念から目標言語の語という2段階の写像を行うわけですが，概念を示すシンボルが，結局は原言語あるいは目標言語の語（あるいは語の並び）になるので，原言語の語と目標言語の語との間に対応関係をつけておくことが必要になります．多言語間翻訳の場合，特定の言語の語彙を中間言語の語彙とすれば，その言語と他の言語との間の対応づけをすればよいというメリットはありますが…．その場合，原言語と目標言語の対訳辞書を使う場合と比べて翻訳の質はどうかという疑問があります．

プリミティブな概念と中間言語

村　木　私は，我々が持っている語彙は文化・文明の所産であり，生活様式の異なりを反映していると考えます．ある言語に存在する概念が他の言語ではなかったりずれたりしています．両言語に共通する語，例えば専門用語などは，それらと1対1に対応する中間言語の語を考えればよい．一方，他の言語にはない概念は，言語に依存しない共通なプリミティブな概念を組み合わせてパラフレーズすることになります．パラフレーズしたものが万人に納得がいくものかどうか，それを近似的な対応づけと呼ぶかどうかは，中間言語の問題とは別の問題のように思えます．このようにして中間言語の語彙が作られます．どちらにしても，いったん決めた中間言語の語彙は，知識の量が増えるにつれて成長しますが，中間言語そのものは標準的な記述言語であります．

天　野　中間言語をプリミティブなものを導入して扱うと，こんな問題が起きます．例えば「豆腐」という日本語を考え，これがある言語にない概念であるとしますね．この「豆腐」を「大豆を煮てすりつぶし，湯に溶かした中ににがりを入れて蛋白を固めたもの」などという中間言語によるパラフレーズ（中間表現）は，英語でもスワヒリ語でも，バスク語でも，異なる辞書開発者が作って同じになるとは思えない．更に「豆腐が好き」を「大豆で煮てすりつぶし，湯に溶かした中ににがりを入れて蛋白を固めたものが好き」と訳しても，それは翻訳にはならない．

内　田　村木さんの言う中間言語は，あらゆる概念を記述しうるプリミティブな概念があり，それにより中間言語が構成されるということだと思います．それに対して私の言っている中間言語は，すべての言語の持つ概念の和集合をとってきて，それらの間の関係をつけたものが中間言語の語彙を構成しますから，必ずしもプリミティブなものの存在を前提にしない．概念間の関係づけはトランスファ方式に近いと言えますが，トランスファ方式では，解析の目標，生成の出発点である中間表現として，表層の語そのものを使っていることが多いのに対して，我々の主張する中間言語方式では，概念の世界を表す中間表現を経由して解析と生成を行っています．つまり，このような中間言語方式ですと，表層より一歩踏み込んだ表現を翻訳の対象とすることになります．

天　野　例えば日本語では兄，姉，弟，妹という人間関係があります．一方，兄弟・姉妹というのもある．ところが英語では brother と sister の二つしかな

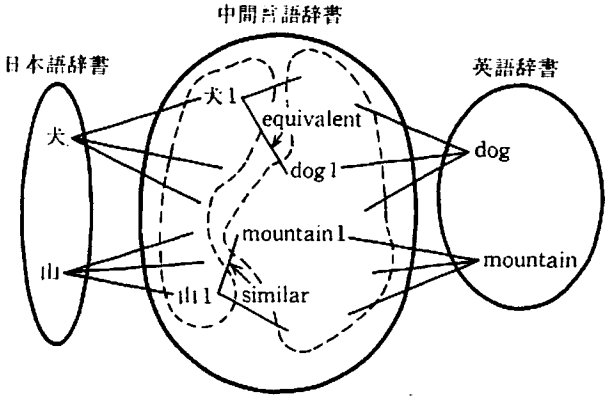

図 3　EDRの中間言語辞書

い．もし中間言語の設計者が，年齢という相を忘れていたとすると，兄，弟というものを表しようがなくなる．ありとあらゆることを表現可能なプリミティブが，言語独立に考えられるだろうかということなんです．

村　木　新しい概念に気がついたらもちろんそれは組み込めばよい．

司　会　そろそろどのようにして中間言語を設計するか，またそれは可能かという問題に入っているようです．

石　崎　中間言語の設計は次のようにするのがよいと考えます．翻訳の対象となる言語をいくつか用意します．その中の概念を全部洗いだしてきて，妹であろうが younger sister であろうが洗いだしてきて，それを全部統合して概念記述を作り，そこから出発して体系化していくわけです．このようにして体系化された概念の世界が，中間言語の語彙の体系になります．妹と sister との間にダイレクトな対応がつかない場合には，我々は概念展開法と呼ぶ方法を使って，sister に対して年齢の上下関係という視点を用いて妹と対応させ，それを新しい概念として付け加えます．したがって言語独立という話にはならない．

我々の研究所で進めている方法では，最初に翻訳対象言語を設定します．中国語，タイ語，英語，日本語でもなんでもいいのですけれど．最初は分野を限定することになるかもしれませんが，とにかく出てきた概念をまず集める．それでどのくらい体系化できるかが問題ですが，例えば，Schank の言っていたプリミティブは体系化した概念体系の最も上位に現れる．したがって，それだけで全部であると考えるのは土台無理な話です．実際には個々の言語の持つ語に近い下位概念が沢山あって，それらの間の対応づけを考えることが重要になります．

司　会　石崎さんの話の中に出てきた Schank の理論についてちょっと補足させて下さい．Schank は，動詞に関する概念は，極めて少数のプリミティブ（彼はこれをプリミティブアクトと呼んでいる）の組合せで構成されると考えていたのでしたね．

内　田　もう一度．石崎さんの方式で，兄とか姉がどうなるのかを聞かせて下さい．

石　崎　家族関係に関する概念体系ですが，兄弟・姉妹という概念が英語の brother, sister に対応します．日本語を考えますとそれは更に分かれるのですが，それらを含めたトータルな概念体系を考えます．そこでは上位に brother や sister があります．一方，兄・弟・妹・姉に対応した概念もありますが，例えば，兄弟と兄・弟の間は年齢差という概念でつながれる．妹と sister との関係は sister のところに書くことになります．その方式はいろいろ考えられますが，sister のところに分類の視点として aging を用います．

内　田　村木さんの方法では，兄とか弟とかいう概念は，性別，年齢，それから兄弟関係という三つのプリミティブから構成される．こうしたプリミティブがあるから言語のコミュニケーションが成り立つのだし，中間言語もあり得るという立場だと思いますが，それと石崎さんの考え方とは少し違う気がしますが．

石　崎　それはほとんど同じですね．兄弟姉妹のうちで，性別，年齢の上下関係を分類の視点や制約記述に用います．

中　村　今まで私は中間言語の語彙を構成する語は，一つのシンボルで表現されていると思っていたのですが，構造を持っているわけですね．例えば，先ほどの knee のうちの「膝で押す」という概念に対応する中間言語の語は「押す」の道具格が「膝」であるというネットワーク構造で表現されるわけでしょうか．

石　崎　いいえ，少し違います．knee の中の「膝で押す」という概念に対応するシンボルは中間言語の語彙の中に存在し，そのシンボルに対応する概念辞書の中に，「体の部分である膝で何かを押す」などといった詳細が記述されます．

司　会　中間言語方式を考えている研究者の中でも，プリミティブを導入してその組合せで概念を表現すべきだとする立場と，そうすべきではないとする立場があるように思いますが，いかがでしょうか．

内　田　プリミティブを導入すると，そのプリミティブの組合せで表現できる概念が，現在の研究レベルでは限られてしまうのではないでしょうか．つまりプリミティブの組合せで考えた途端に，情報が落ちてしまう恐れがある．私は，原言語の文を解析して中間言語による中間表現を抽出したとき，たとえ抽出された情報が目標言語で全部表し得なくても，抽出した中間言語の表現の中には解析した文の情報が完全に保持されている必要があると思います．別の目標言語では全部表しうる可能性がありますから．我々の方式は図3のところで説明しましたように，対象言語の概念の和集合をとり，言語対ごとに概念間の関係をつけていく方式で概念体系を作ります．それが中間言語の語彙であると主張していますか

ら，必ずしもプリミティブなものを前提にしていません．

村　木　私たちは，プリミティブなもので中間言語の語（概念）を構成する立場をとっていますが，その動機を述べさせて下さい．プリミティブなものをどれだけ用意するかという問題ですが，それは相当数あります．最初の出発点は，とにかく2言語間の翻訳を可能とする中間言語が設定できれば，その他の言語も考慮した多言語間翻訳システムが経済的に開発できると考えました．それを実証するにはまだ時間がかかりそうですが，とにかく，2言語間の翻訳システムで抽出された中間言語の語彙に含まれる語の数がかなり多いとしたら，その語彙を用いて，ある言語になくて他の言語にある概念を表すことができるはずだと考えました．言い換えると先の中間言語の語彙がプリミティブの相当な部分をカバーするものとして機能すると考えたわけです．

この考え方でやりますと，語の持つ表層的な違い，例えば日本語モーダルが助動詞として現れるのに対して，英語では動詞や副詞であったりするといった問題をどう扱うかなど，いろいろ面白い問題が出てきました．

中間表現の構造

司　会　中間言語の語彙の話をしてきたわけですが，中間言語の語彙を用いて表現した中間表現の構造についての問題はありませんか．

梶　ここにいる皆さんが頭に描いている中間言語は，深層格関係を中心とした意味ネットワーク表現を記述する言語だと思うのですが，そこで問題になる言語間のギャップについて述べたいと思います．意味ネットワーク表現は，中間表現の要件をかなり満たしていると思いますが，一つの文に対する中間表現が一つに定まらないという問題があります．中間言語方式の利点の一つは，原言語解析と目標言語生成の独立性であり，開発が独立にできるということもその中に含まれると思いますが，同じ意味を表す中間言語による表現が一意に決まらないと，独立に作った解析部と生成部とを組み合わせても良い翻訳文が得られない可能性があります．

簡単な例を挙げますと，日英翻訳で，自動詞・他動詞変換というのがあります．日本語では「AによりBが向上した」と言いますが，英語では「AがBを向上させた」と言います．日本語的な発想で考えた中間言語表現では，恐らくAが自動詞的概念「向上する」の原因になります．一方，英語的な発想では，Aが他動詞的概念「向上させる」の動作主になります．このようなギャップを吸収して解析の目標，生成の出発点を合わせないとうまく翻訳きません．

言語の定義としては，語彙，シンタックス，セマンティックスを決めれば十分なわけですが，中間言語を機械翻訳システムの設計あるいはインプリメントの方法論として考えますと，中立的な表現の基準を決めることが重要で難しい問題です．例えば，通産省の「近隣諸国間の機械翻訳プロジェクト」では，最初，各国語の対訳例文をそれぞれ中間言語で表現して，どのあたりが中立的な表現かということを決める作業をしましたが，例文対応ではなく，一般的な判定基準が必要なのです．個々の言語に依存した中間言語表現を比較して，妥協点を探る作業が避けられないのであれば，中間言語方式の利点は半減します．そうした部分は，意味トランスファとして残すほうがよい．意味トランスファは，中間言語上のパラフレーズとしてみることもできますが，どのパラフレーズが必要かは言語対に依存するので，トランスファ方式であると考えられます．

内　田　自動詞・他動詞変換で梶さんのおっしゃることは間違っています．英語でも，Aは動作主になるわけではなく原因になります．言い換えますと，そうなるような英語の解析を行わなくてはならないということです．むしろ解析の問題だと思います．これは生成についても言えます．生成では目標言語を意識しますから，原因となっているAを，目標言語（英語）の主語としたとき，動詞として何を選ぶかを生成の段階で決めることになります．したがって，梶さんのおっしゃった問題は中間言語方式の問題ではないと言えます．

天　野　梶さんの言われることも，内田さんの言われることもどちらもわかるような気がします．梶さんの言いたかったことは，特定の言語を意識すると，その言語では原因を動作主としたほうが自然な表現であることがあります．トランスファ方式では，目標言語を意識してトランスファするので，このような問題に対処しやすい．中間言語方式では，目標言語を意識しないので対処しにくいということだと思います．

司　会　正直言って中間言語方式は，今開発中で，これからも研究を進めなければならないと思います．梶さんの提起した問題は，中間言語方式の問題というより，むしろ今の技術で，中間言語方式で考えている中間表現を原言語の文から抽出できるか，あるいはその中間表現からうまく目標言語の文を生成で

きるかということだと理解すべきだと思います．現実にそこに問題があるとしたら，もう少し現実路線をとったほうがよい，それがトランスファ方式であるということでしょうか．それでも，中間言語方式の良さを追求すべきであるとする研究者が一方におり，両者の間でさまざまな議論がなされてきたと思います．この座談会は両方式の決着をつけるのが目的ではありません．むしろ問題点を整理し，将来の方向を見定めたいということです．この辺りで話題を変えましょうか．

石　崎　その前に，梶さんから「近隣諸国間の機械翻訳プロジェクト」についてのコメントがありましたので，それに関連して一言述べさせて下さい．我々の研究所でもこのプロジェクトに参加していますが，プロジェクトを推進している（財）国際情報化協力センターに中間言語ワーキンググループがあります．そこでは中間言語で中間表現を表すために，概念間の関係を表す関係子とか，概念の性質を記述するための属性子を多言語を参照して設定しています．記述形式を統一すれば，言語固有の概念が出てきてそれを記述したとき，共通に理解し合えるだろうということで作業を進めているわけです．したがいまして，さまざまな言語を見ながら中間言語を設計しているわけでして，言語独立に設計しているわけではありません．

エスペラント語や英語は中間表現として適切か

司　会　中間言語の表現としてエスペラント語を選んだらどうなりますか．

石　崎　エスペラント語は，状況意味論で言う1個1個の単語がさまざまな意味を担うことができて，情報伝達の観点からは効率が良いと言われています．これは英語であろうと日本語であろうと自然言語一般に言える性質です．このことは，これらの言語で表現されたものは意味的曖昧性を内在しているということを意味しています．このような意味的曖昧性は，人間なら文脈に依存した情報でかなり解消することができますが，コンピュータにとってこれは困難な問題になります．したがって，中間言語の表現としてエスペラント語などを採用しますと，この曖昧性のために，生成側に問題が出てきます．

内　田　解析段階では極力曖昧性を解消したいわけですから，せっかく解消した曖昧性が，エスペラント語による中間表現に直したとたん曖昧性を持ってしまうのでは，何のために解析で曖昧性を解消したのかわからなくなります．これは，中間言語方式であろうとトランスファ方式であろうと，共通して言えることだと思います．

天　野　解析能力の問題はありますが，理想を言えばとにかく曖昧性はできるだけ解消したいわけです．

司　会　解析段階で，せっかく解消した曖昧性が，中間表現に直したとたん曖昧性を持つ可能性があるので，エスペラント語を中間表現とするのは適切でないという結論ですね．

中間表現と曖昧性

司　会　ある語の意味が，後続文が出てきて初めて決まるという文はかなりあると思います．後続文が未処理の段階で，ある語の意味が未定な状況では，その中間表現はどうなるのでしょうか．

内　田　曖昧性を含んだ言語現象は私の中間表現では残念ながらとらえられません．それだけでなく，文を読んだとき，その読みの音が読み手に何かを連想させる場合も中間表現で表しにくい．音をどう中間言語による中間表現で表すか，これはなかなか難しい．

村　木　原言語の持つ意味的曖昧性を目標言語の側に移したい，移せるということがわかるためには，既に目標言語の知識を解析段階で使う必要が出てくるかもしれません．中間言語方式では解析と生成とを独立に設計してあるので，これは具合が悪く問題が確かにあります．

中　村　トランスファ方式ですと，このような場合，概念とか意味とかを経由せず直接目標言語の適切な語を選択する道が残されていますので，問題が少ないと思います．人間の翻訳者が，非常にうまい単語を選んで何とか翻訳できるような例であれば，トランスファ方式でそれを実現することができると思います．

両方式の合流の可能性

司　会　トランスファ方式と中間言語方式は，将来互いに接近する可能性があるのか，それとも別の道を歩むことになるのか，この点はどうでしょうか．

天　野　各言語の持つ概念の和集合で中間言語の語彙を構成する立場の中間言語方式であれば，梶さんも指摘していましたが，トランスファ方式と中間言語方式とは既に合流していると言えます．つまり，中間言語の語彙を作るとき相手言語を取り込み，意識しているでしょう．この場合，各言語の持つ概念で共通なものが相当あるに違いない．新しい言語を取

り入れるときには，その言語に固有な概念だけを付加すればよいわけです．一方，トランスファ方式ですと，言語対ごとに辞書作りをするので，多くの言語の概念間に共通部分があっても，それに気がつかない．何度もその部分を重複して作らざるを得ない．トランスファ方式で多言語間翻訳を行うときには，そうしたデメリットがあります．そこで，トランスファ方式で使う辞書を，初めに述べた方式で作ることも考えられます．そうすることにより，多言語間翻訳をトランスファ方式で行う場合にコストが安く済むことになります．

ただプリミティブを設定する立場の中間言語で非常に気になるのは，例えばA言語の語を，マクロなラベルではなくプリミティブな概念の組合せで表現しようとする．言い換えると，語彙項目の分解を行うことになる．語彙項目を分解したものから文を生成することは，全く不可能と言わないまでも，困難だろうと思います．多くの場合，語彙項目を分解したものの直訳的な訳文を生成せざるを得なくなり翻訳の質の低下を招くと思います．

内　田　私も天野さんの意見と同じで，中間言語方式とトランスファ方式の合流はあり得る，というよりむしろ合流してトランスファ方式は消えてなくなるのではないかと考えています．例えば，天野さんが指摘した語彙を設定するとき，トランスファ方式では共通部分が共通部分として生かせないというデメリットは大きい．何度も繰り返しますが，我々の方式ですと各言語の持つ概念の和集合で中間言語の語彙を設定し，原言語固有の概念があれば，それを持たない目標言語に翻訳するとき，何らかのパスにより目標言語の側の概念（の組）を選択するので，これはある種のトランスファと言えるかもしれない．そういう意味で両者の合流は既になされていると思います．私が最初アトラスと呼ぶシステムを発表したときには，それを意味トランスファ方式と呼んでいました．

石　崎　私どもの研究所では，文脈理解の研究をしています．それを利用したコントラストと呼ぶ翻訳システムを作成し，それを文脈情報変換方式と呼んでいます．文脈情報まで使いますと，意味トランスファよりもう少し深い解析をしていることになりますから，中間言語方式に近いと考えています．

中　村　私は今実用システムを開発しないでいいはずですから（笑い），その前提では，ほぼ内田さんに近い考えを持っています．つまり純粋なトランスファ方式は将来なくなるのではないかと思います．特に辞書の部分に関して，電子化辞書研究所（EDR）で開発されている辞書を利用するのが，中間言語方式であるとすれば，トランスファ方式が中間言語方式に吸収されることになります．

図3は，私がイメージしたEDRの単語辞書の内容です．EDRの辞書では，例えば「犬」という表層語の概念を「犬1」，「犬2」のように原言語のレベルで区別し概念記号を与えます．もちろん純粋に原言語だけで区別するのではなく，他の言語も考慮しますが，すべての言語で共通になるかどうかは考えません．その上で，区別された概念記号の間をequivalent（同義），similar（類義）というリンクで接続して翻訳に利用すると私は理解しています．そこで図3の点線のように，日本語や英語といった表層の言語に対応した中間言語辞書内の概念記号の集合をイメージすることができますし，この段階ではトランスファ用の辞書に非常に近いものになります．

このとき，理想を言えば，図の点線で表された「ある言語の概念」という壁はないと考えますが，実際，辞書を作成していく段階では，この点線を設定せざるを得ないと思います．私の希望としてはこの壁を徐々に取り外すように辞書を作っていただきたいわけです．つまり，equivalentというリンクでつながれた概念記号を一つの概念記号にマージしてしまう．その方法として，手作業で行うのも正攻法でしょうが，司会の田中さんが提案しているような，機械的にマージすることにも興味を持っています．専門用語であれば，簡単にマージできるのではないですか．

内　田　専門用語についてはequivalentというリンクを使わず，始めから直接概念記号が設定できるので，マージする必要はありません．マージする必要があるのは一般用語のほうです．

中　村　マージする努力を積み重ねていけば，村木さんの主張されている中間言語辞書も実現できるかもしれない．辞書に関しては，実用面，研究面でこのような考え方になっていくのではないでしょうか．もっともこのようにして出来上がった辞書は，単語の意味を日英といった言語ペアに対して必要以上に細かく分解してしまっている可能性がある．そこで出来上がった中間言語辞書から逆にトランスファ辞書を生成して利用するというのはどうでしょうか．つまり，複数のトランスファ辞書をマージして中間言語辞書を作り，そこから再びトランスファ辞書を生成する．

司　会　なぜわざわざトランスファ用の辞書を生成し

なければならないのですか.

天　野　nか国語のシステムなど要らないというユーザに，nか国語の辞書の入った大きなシステムを供給する必要はないからです．辞書はメモリを食って不経済です．実際には，翻訳需要の90%以上は英日システムですし.

中　村　ヨーロッパ共同体のように，一つの文を一気に複数の言語に翻訳しなければならないというような場面は別にして，普通は特定の言語対での翻訳のニーズが多いと思いますが，その場合にはトランスファ方式のほうが作りやすいのではないでしょうか.

梶　コンピュータの負荷というか処理効率の面で，トランスファ方式が有利な点もあります．中間言語方式はある意味で理想的な方式ですが，厳密な意味での中間言語方式は難しいですから，トランスファ方式を併用するのが現実的だと思います．その意味でトランスファ方式はなくならないと思います.

天　野　2か国語を意識することは翻訳では絶対必要だと思う．中間言語を設定すると，そこで情報が失われてしまう可能性がある．英語から日本語へ翻訳しようとするとき，わざわざ別の言語（中間言語）に翻訳する必要はない．そうした途端，翻訳の正確さと品質が損なわれてしまうと思います.

村　木　トランスファ方式だと具合が良くて，中間言語方式だと具合が悪い，あるいはその逆のこともあるかもしれません．翻訳方式そのものもまだ研究段階にあるといってよいでしょう．私も7000を越えると言われる言語すべてに当たったわけではないのですが，現在の中間言語方式の技術ですべて解決できると信じてはいますが，それ以上のものではありません．ただ我々は，トランスファ方式（図2参照）と異なりVの溝（図1参照）に変換の橋を掛けていない，少数の言語しかまだ扱っていないのですが，橋を掛ける必要は全く感じていないのです．橋を掛ける気ならいつでも掛けられますが，そうする必要を感じていません.

内　田　現在中間言語についてはコンセンサスの得られたものは存在しない．しかし，中間言語方式のメリットを最大限生かせる方式をもっと研究する必要があるのではないでしょうか.

司　会　そうすると，中間言語方式は理想であり不可能に近いという立場をとるのではなく，その理想に向かう努力を今後積み重ねる必要があるということになりますね．一方，トランスファ方式にはトランスファ方式の利点もある．特定の言語対を考えたとき，翻訳システムを開発しやすいという利点です．したがって，トランスファという考え自体も消えずに残るという考えもあることになりますね.

機械翻訳の今後

司　会　最後に機械翻訳の今後についてお願いします.

村　木　機械翻訳は変わると思います．いつになっても100%正確な翻訳ができるとは思いませんが，少なくとも機械翻訳の基本機能が今のかな漢字変換と同様，コンピュータに組み込まれていくと思います．変な文をコンピュータに入力すれば，コンピュータがそれを指摘してくれて，それに従って文を作ると翻訳結果も同時に手に入る，そういう時代が来ると思う.

内　田　機械翻訳は現在のレベルでも実用化されています．ただ現在のレベルは誰が使っても軽く使えるというものではない．100%翻訳することもできません．したがいまして，原言語と目標言語の双方を知っていないと使えないという面がある．専門家しか使えないシステムと言えるかもしれません．100%翻訳するのは100年たっても200年たっても無理かもしれない．ただ性能向上に向けて努力しなければなりません．そのためには文脈処理の研究が今後重要になると思います．それと共に，実用化の立場からは，今ある機械翻訳システムの使い方を研究することも必要だと思います.

中　村　実用面でなく研究面で考えて，やはり私は機械翻訳を研究テーマとして続けていきたいと思っています．自然言語処理の研究テーマとしては，言語知識ベースの作成，文脈・談話処理や文章生成に興味を持っています．これらを単独で考えていると，何を目標にしているのかわからなくなることが多いのですが，機械翻訳を対象にすると物事が具体的に考えられるわけです．機械翻訳そのものについては，文脈構造を表現することを考えた場合，単文のときとは違って中間言語方式的な手法のほうが考えやすいかもしれない．もっとも，日本語と英語で文章作成の発想，すなわち話題の展開の仕方が異なるので，やはり，文脈構造トランスファが必要になるかもしれません.

天　野　私の立場は，トランスファ方式は目先のものでなく，翻訳の本質を突いたものであるということです．語彙における言語普遍なものはない．なぜなら言語は各民族固有の文化を担い，背景にしているわけで，自然界をどう分節するかで各言語の語彙が決まってしまうからです．現状の自然言語処理技術

のレベルでは，文の分析レベルが十分でないため，中間言語方式のデメリットも目立たないといえます．私見ですが，両方式の比較検討も大切ですが，一つの言語の言語学的な分析をもっと深めていく必要があります．

梶　現在の機械翻訳システムを使いこなすための技術開発の必要性については，内田さんと同感です．制限言語の普及，機械翻訳しやすい文を作るのを支援するツールの実現に力を入れる必要があります．一方，次世代の翻訳方式も研究していく必要があります．中間言語について言いますと，今日の議論は，原言語や目標言語の構造に基礎を置いた中間言語，つまり構造束縛型の機械翻訳の枠組の中での中間言語であったわけです．コミュニケーション内容の表現手段としての中間言語は本来，原言語や目標言語の構造とは関係ないものであり，今までとは全く違ったレベルの中間言語を考えることも将来の課題であると思います．

石　崎　中間言語方式はトランスファ方式と比べて，中間表現の情報が一部欠落してしまうのではないかという意見があったと思いますが，私はむしろ逆ではないかと考えています．中間言語方式の定義にもよりますが，中間言語方式では概念辞書を扱います．概念辞書にどれほどの量の情報が書き込まれているか，それをどのくらいうまく使えるかによって，情報が落ちる落ちないが決まると思います．自然言語の解析と生成で十分使うことができる概念辞書を着実に構築し，それにより解析の結果，情報が落ちないような方策を考え，それで中間言語方式の翻訳システムを作りたいと思っています．そうした方向の研究が必要だと思います．1960年代に考えた中間言語方式がうまくいかなかったのは当時の技術的な制約と構想に欠陥があったためだと考えます．今後，自然言語処理の基礎的な技術を十分発展させることが，機械翻訳の将来にとって必要であると思います．

司　会　「中間言語方式をめぐって」ということで長い間議論してきましたが，この辺でひとまず座談会を打ち切りたいと思います．中間言語方式のさまざまな問題がこの座談会を通じて明らかになったと思います．中間言語方式のイメージもただ一つに決まっているわけではないということもわかりました．これからは，もしかするとグルノーブル大学の中間言語方式とは異なる中間言語方式があるということを外に向かって発言する必要があるかもしれません．トランスファ方式がベストな方式であるとする立場もありました．両方式の利害得失は，長期の立場に立つか，短期の立場に立つかで変わるでしょうが，それは読者の皆さんの判断に委ねたいと思います．途中激論もありましたが，この問題をめぐって長時間議論したのはこれが初めてではないかと思います．出席者の皆様には本当にありがとうございました．