

# e国際共通語研究会

2011年3月28日

特定非営利活動法人セマンティックコンピューティング研究開発機構 (ISeC)

## e国際共通語研究会の開催

ISeC として、 e 国際共通言語研究会をスタートさせる。要点は、以下である。

- (1) 期間は、2 年間とし、成果を書籍や論文にまとめる。
- (2) 年寄りにとっては楽しめるもの、若い人には新研究領域の開拓となるものとする。
- (3) 実際に集るのは、2 ヶ月に一度程度とし、Web 上に議論を集約する仕組みを設ける。
- (4) 研究会代表は、石塚理事長、横井理事、内田理事の 3 名による共同代表とする。
- (5) ISeC 関係者と ISeC シンポジウム (2010 年 11 月 6 日) 参加者に呼びかけを行う。

## 開催趣旨

### (1) なぜ e 国際共通語なのか

バベルの塔の崩壊 (旧約聖書、創世記) 以来、言葉の壁の克服は、人類の永遠の課題であった。国際共通語 (Common Language) への希求は、絶えることがなかった。アプローチには、既存言語 (自然言語) による共通語と新規言語 (人工言語) による共通語とがあった。

既存言語 (自然言語) による共通語には、

ラテン語 : 近代ヨーロッパ知識人の共通語

フランス語 : 外交用共通語

ピジン言語 (ピジン英語、ピジン仏語等) : 商売のための共通語

国連公用語

欧州連合公用語

などが使われてきた。そして、最近では、インターネットによる諸活動のグローバル化によって、英語が世界共通語となりつつある。ヨーロッパ域内のビジネス公用語としての英語、韓国人や中国人の英語能力の高まり、日本企業における英語の社内公用語化 (楽天、ユニクロ等) などの動きがある。日本企業の社内公用語化は、'No English, No job.'とまでに謳われるようになった。

新規言語 (人工言語) による共通語については、Wikipedia 「人工言語」の項に以下のよ

うな事例が挙げられている。

1950年以前に発表された言語：

イド語、インターリングア、エスペラント、オクツィデンタル（後にインターリングに改称）、ソルレソル、ノヴィアル、ベーシック英語、ヴォラピュク

1950年以降に作られた言語：

3CL、地球語、トキポナ、ノシロ語、ユーロパント、ログラン、ロジバン、BABM、世界語 74・07

しかしながら、エスペラントを始めとして、人工言語による人が理解する国際共通語はことごとく普及に失敗した。国際共通語は、人が理解するものではなく、コンピュータに理解させるべきものである。人は、それぞれの母語を用い、母語と共通語間の翻訳・通訳はコンピュータが自動的に行う。ただし、現状の機械翻訳技術・自動通訳技術を前提にすると、国際共通語は、制限自然言語となる。

コンピュータが理解する国際共通語をe国際共通語と名付ける。e国際共通語は、プログラミング言語などのいわゆるコンピュータ言語ではない。コンピュータのための自然言語、コンピュータの母語である。人間のための自然言語、すなわち、人間の母語は、国や地域の特性を反映しながら人間活動や社会活動の微細にいたるまでの表現能力や伝達能力が求められる。一方、コンピュータの母語、e国際共通語には、まずは、そのような微細な表現能力や伝達能力は必要無く、適切に制限された自然言語でよい。

人々は、制限された母語ではあるが、e国際共通語の介在を意識することなく、それぞれの母語で全世界とコミュニケーションできる。情報サービスや情報システムの人間と係る部分には、e国際共通語が用いられることになる。すべてのサービスやシステムは、一気にグローバルイゼーションを達成することができる。ローカライゼーションの手間隙は必要なくなる。

なお、このe国際共通語と基本的な趣旨を同じくする具体的な活動がある。国連大学の関連機関が推進する UNL (Universal Networking Language) である。UNL は、CWL (Common Web Language) とも呼ばれている。e国際共通語の研究は、UNLの多くを参考に進められることになる。

## (2) e国際共通語の言語特性

e国際共通語の言語特性を列挙する。

### (1) 概念を表現するネットワーク構造言語

人間の母語としての言語は、音（音声）と概念（意味）とを結びつけるための記号の体系である。言語の発生史からも、また、幼児の言語獲得過程からも明らかなように、言語としては、音声言語が基本である。そして、人間の聴覚器官や発声器官の逐次性から、言語の基本構造は、本質的に線状構造である。しかしながら、コンピュータは、直接的に話す・聞くこと、読み・書きすることはしない。したがって、

e 国際共通語は、概念レベルを表現する記号の体系となる。そして、概念の構造は、入れ子構造のネットワーク構造によって近似される。なお、人間の母語と異なり、表現の経済性という制約は緩くなる。そこで、e 国際共通語は、概念レベルでの自然言語の意味記述言語という側面をもつことになる。

(2) 人が使う母語との相互自動翻訳・自動通訳

人間が使う各種の母語の共通の仕組みを抽出することになる。そのためには、最近の対照言語学の成果が利用できる。まずは、日本語、英語、中国語、これら 3 言語の共通の仕組みから始めることになる。e 国際共通語と人間の各種母語とを結びつけるためには、現状の機械翻訳・自動通訳技術の活用が必要である。翻訳技術に関しては、最近の新動向の技術の利用も必要であるが、かつての中間言語方式や中間言語トランスファ方式技術の利用が肝要である。e 国際共通語と人間の各種母語との間の翻訳・通訳は、完全自動でなければならない。したがって、人間の母語に対する効果的な制限言語の仕組みが必要である。しかしながら、機械翻訳のために試みられたかつての制限言語は、制限言語のジレンマから逃れることができなかった。このジレンマの適切な解決法が必要である。とりあえずの方法として、マインドマップやコンセプトマッピングのような図式言語の利用もひとつの方法であろう。

(3) コンピュータ言語やコンピュータ指向表現形式との相互変換

e 国際共通語には、知識表現言語や意味記述言語や各種論理式などのコンピュータ言語やコンピュータ指向表現形式との間に相互変換の仕組みが必要である。この仕組みによって、各種推論系との連携処理が可能になる。そして、高度な意味処理や知識処理や推論処理などが実現されることになる。そして、この仕組みのためには、e 国際共通語は、コンピュータ向きの表現メディアと概念レベルで連携できるものでなければならない。その連携の基盤として、概念記述言語 CDL (Concept Description Language) を設定することになる。

(4) 他の表現メディアとの連携表現・連携処理

コンテンツとしての情報表現には、様々な表現メディアが組み合わされて使用される。自然言語は、あくまでも、そのような表現メディアのひとつである。実際には、他の記号系メディア（数式、論理式等）、ビジュアル系メディア（図、表、ダイアグラム等）、イメージ系メディア（画像、映像、音像等）などが組み合わされて使用される。通常、自然言語は、これらの表現メディアを連携させる汎用メディアとしての役割を担う。したがって、e 国際共通語は、これらの各種表現メディアを概念レベルで連携させる汎用概念メディアの役割を担わなければならない。そのためのメディア連携の基盤として、概念記述言語 CDL を設定することになる。

(5) 自身の可視化・可読化に図式言語と制限英語

システムの実装・管理のためには、e 国際共通語自身の可視化、可読化が必要になる。そのためには、国際共通語となりつつある英語を使用することになる。図式言語と

制限英語の適切な組み合わせが有効であろう。

### (3) e国際共通語の言語仕様とシステム仕様

研究会の主題は、e国際共通語の言語仕様とシステム仕様を策定することである。ここでは、言語仕様の要点だけを掻い摘んでおく。

- ① 語彙は内容語と関係語に分かれる。
- ② 語義は単義である。多義ではない。
- ③ 内容語はノードに、関係語はアークに対応付ける。
- ④ 内容語の表現特性は、ノードの属性・属性値によって表す。
- ⑤ 言語の構造は、複合ノードの入れ子構造によって表現する。

文章構造、段・連段構造、文・連文構造、単文・複文構造、単純語・複合語構造

修飾・被修飾の関係、係り受け関係

修飾の範囲（スコープ）

照応・参照関係

等々

- ⑥ 共通常識としての概念体系を持つ。
- ⑦ 表現の経済性は緩くなる。

### (4) e国際共通語の役割

e国際共通語の役割は、グローバル情報サービス基盤となることである。すなわち、①ローライゼーション問題を解消するための情報サービス・情報システムのグローバルライゼーション基盤となること、②共有オントロジーの定義記述や共有情報のメタデータ記述のための情報のグローバル共有基盤となること、③危機管理マニュアルや環境保全マニュアルや共通規約・規定などの情報のグローバル管理基盤となることである。

### (5) テーマ設定の意義

e国際共通語研究というテーマ設定の第一の意義は、大きな物語を提示することである。日本のICT（研究開発・ビジネス・技術政策）は、実務的な事柄に囚われ、そこに大きな物語が無くなった。原因のひとつは、かつてのバブル期に語らえた大きな物語の試みがことごとく成功とは言い難いこととなったことである。そして、競争原理の導入によって無秩序な成果主義がはびこり、細かな目先の事柄に終始することになったことである。そして、大きな物語を語れるのは、米国ICT企業のみとなった。IBMの「Smarter Planet」然り、Googleの「世界の全情報を検索可能に」然りである。日本の得意はまとまること、システム化することである。本テーマは、この日本の得意に沿った大きな物語のひとつである。

テーマ設定の第二の意義は、ICT技術の革新に向けた展望を提示することである。技術は、個別技術の時代から総合技術の時代となり、技術開発とビジネス化とユーザ視点との一体

化が求められるようになった。そのような技術動向に即した総合的な革新的ICT技術の目標設定としては、e国際共通語や産業日本語を中核とする日本言語システム、ヒト型知能ロボット、日本型情報基盤構築などを挙げるができる。

テーマ設定の第三の意義は、ICT 関連学会の再統合に向け、さらに、ICT 業界の再編成に向けた方向性を提示することである。そして、そこには、それらの実務を担う若きリーダー達への期待がある。

テーマ設定の第四の意義は、言語研究の革新に向けた潮流を提示することにある。分析科学の側面に加え言語研究に合成科学の側面をもたせ、言葉をデザインするという新しい言語研究のアプローチを提示することである。