

PRONOUN REFERENCE IN GERMAN

by

L. BRANDWOOD

THE PROBLEM AND ITS CAUSE

THE problem arises simply from the fact that German regularly attributes the masculine and feminine as well as the neuter gender to inanimate objects, whereas English rarely departs in this respect from a uniformly neuter representation. Conversely, an animate being may be denoted as neuter, but since in this case the same is true for the most part of English, the problem is encountered here to a much less degree. Thus a neuter pronoun in German referring to animals or even a human being, if it is a child, may readily be translated into the equivalent neuter pronoun in English; it is only when a somewhat later stage of life is reached (e.g. *das Mädchen*) that English cannot follow the German usage.

In short, therefore, a German masculine or feminine pronoun must sometimes be converted in English to a neuter pronoun, a neuter pronoun to a masculine or feminine pronoun. In order to determine whether any conversion is necessary, the antecedent of the pronoun must be identified, which for reasons that will become apparent through the illustrations presented is not always a simple matter from the point of view of mechanical translation.

THE MATERIAL INVESTIGATED

The present study was undertaken to determine how far purely syntactical considerations would solve the problem of pronoun reference. For this purpose it was decided to examine scientific texts in which persons were but infrequently mentioned. While this had the advantage of excluding the human/non-human criterion and so of making a more purely syntactical solution possible in theory, it was also automatically rendered defective. For since the human/non-human distinction was the only problem, and since a complete solution of this was not possible by consideration of syntax

alone, the means of discovering criteria for promoting the solution further were lost through this exclusion.

The investigation was limited to personal pronouns, the problem being most acute in their case, as they are generally capable of wider separation from their antecedent than the other pronoun forms. Several further restrictions also applied:

1. Since on occasions there is difficulty in determining whether it is functioning as a pronoun proper or merely as an introductory particle, the neuter form *es* was not considered for the sake of simplicity in compiling material.
2. The genitive forms were omitted as being more closely related in their use to the possessive pronouns, with which they might be treated more satisfactorily.
3. By contrast, although no problem exists in translating the plural forms, the nominative and accusative *sie* were included in order to increase the amount of material available from the texts examined. Without them this would have been reduced by about 38%. Needless to say, they are just as useful as the other forms for ascertaining the rules of pronoun reference.

The forms of the personal pronoun examined, therefore, were: *er*, *ihn*, *ihm* (masculine and neuter), *sie* (feminine singular and plural of all genders, nominative and accusative), *ihr*.

The texts, which were from books on the various sciences - physics, chemistry, biology, mathematics, geology, zoology, etc., amounted to approximately 200,000 words and yielded the following:

nominative pronouns (*er*, *sie*) - 416

accusative pronouns (*ihn*, *sie*) - 74

dative pronouns (*ihm*, *ihr*) - 14

accusative pronouns preceded by a preposition - 10

dative pronouns preceded by a preposition - 30

As a result of their low occurrence no attempt was made to proceed further with the cases other than the nominative, since whatever system they might have been reduced to would not have been of any great applicability. Even the 400 odd instances of the nominative pronoun are far from being sufficient material for a final analysis. Some facts, however, have become clear from the present study, defective though it be, and these are accordingly given in the following account. It must be left to future investigations to remedy these defects through a larger material, selected more appropriately from texts in which persons as well as things find frequent mention.

PRONOUN CATEGORIES

The pronouns could be divided into three main categories:

- A. those occurring in a clause which was not the first in its sentence, in which case the reference was usually, but not invariably, to some substantive in a preceding clause of this sentence;
- B. those occurring in the first clause of a sentence, the last clause of the preceding sentence being independent. In this case the reference was usually to a substantive in this last clause of the preceding sentence;
- C. those occurring in the first clause of a sentence, the last clause of the preceding sentence being a dependent one, in which case the reference could be to this dependent or a preceding independent clause.

Within these categories it would have been possible to classify the various types of pronoun reference in several ways: the one actually adopted was as follows:

CLASS A. I. INCLUDED CLAUSES

A.I. (a) If the pronoun occurred in a dependent clause (other than a relative clause) included within another clause, it referred to the nearest congruent substantive in the preceding part of the thus divided clause, though excluding any substantive in parenthesis:

- e.g. *Wird ein Glied eines solchen Systems durch einen "Treffer",
wie er in der Strahlengenetik vorkommt,
* ionisiert,
so muss sich die Ladung durch Verschiebung der innermolekularen
Elektronen und der zwischenmolekularen Protonen über einen Teil
des Systems ausbreiten.

This was true of all the 15 instances which occurred in the texts examined.

It should be noted that where the part of the clause preceding the included clause was divided yet again by another included clause, the pronoun was not to be referred to any substantive in this previous included clause;

- e.g. **Der Saft* der Bauchspeicheldrüse,
in dem das Enzym Trypsin enthalten ist,
* spaltet,
wenn er nicht zuvor mit dem Darm in Berührung kam,
* keine höheren, eigentlichen Proteine.

* In the German examples, where a clause is divided by an included clause, then the two halves of the divided clause will each be preceded by an asterisk.

A.I. (b) If the included clause containing the pronoun was a relative clause, the only method of identifying the antecedent seemed to be to examine all the substantives (except that referred to by the relative pronoun) in the preceding part of the divided clause in the hope that only one would show congruence in gender and number. Any methods of resolving the problem, if more than one should be congruent, did not come to light from the material at hand, inasmuch as in the 5 instances of this type the congruence was always unique:

e.g. *Eine Mischung von *Sauerstoff und Wasserstoff* im Verhältnis, in welchem sie im Wasser enthalten sind,
* heisst Knallgas.

A.I. (c) If the pronoun occurred in the second of two included clauses, both being of the same type and dependent on co-ordinate elements, it referred to the subject of the first Included clause:

e.g. *Mit der Erkenntnis,
dass *die Sonne* den Hauptkörper und den Mittelpunkt unseres Systems bildet,
* und der naheliegenden Vermutung,
dass sie uns, von einer sehr grossen Entfernung aus gesehen, wahrscheinlich als ein Stern erscheinen würde,
* gewann diese Sphäre Fixarum Stellarum plötzlich ein ganz neues Interesse.

This was the only instance in the texts examined.

A.I.(d) In 3 instances of the type A.I. (a) the substantive referred to was itself a pronoun, with the result that its antecedent, either in the same or a preceding sentence, had to be identifiable to make a correct translation of the following pronoun possible:

e.g. Wenn es auch möglich ist,
alle oben zusammengestellten kolloidalen Systeme herzustellen,
so wurden doch *Suspensionen und Emulsionen* am genauesten studiert;
*gerade *diese* werden auch,
wenn sie im oben definierten Dispersionsgrad aufgeteilt sind,
* kolloidale Lösungen genannt.

In the 3 instances concerned this identification would in fact have been made in accordance with the relevant rule in each case.

CLASS A. II. PARENTHETICAL CLAUSES

A. II. (a) If the pronoun occurred in a parenthetical clause, it referred to the subject - provided this was congruent - of the clause in which or after which the parenthesis stood:

e. g. *Während

wie unten noch begründet wird,

* *die klassische Dispersion* praktisch keine wichtige Rolle spielt

(sie tritt nur in einer höheren Näherung in Erscheinung),

tritt eine solche im Gas mit chemisch reagierenden Molekülen auf.

There were 12 such instances including 5 where the substantive referred to was itself a pronoun, and with each of these the further antecedent was correctly identifiable.

It should be remarked that, apart from clauses enclosed in brackets, those introduced by d.h. were also regarded as parenthetical.

e.g. *Die Linse* weist für alle diese Elektronen das gleiche

Brechungsvermögen auf,

d.h. sie ist achromatisch.

A.II. (b) There was a further instance of this type, in which the pronoun could not from lack of congruence refer to the subject of the clause after which the parenthesis stood, and in fact referred ultimately to the subject of the preceding sentence:

e.g. *Die Kristalle* vieler Hydrate mit hohem Wasserdampfdruck haben das Bestreben,

*Wasser zu verlieren,

wenn sie der Luft ausgesetzt sind,

* und zerfallen in Pulver.

Man nennt diesen Vorgang Verwitterung (sie verwittern).

In such cases the pronoun might be treated as falling under Class B or C, whichever was relevant. In the present instance the pronoun would be referred to its correct antecedent under rule B. II.a (q. v.)

CLASS A. III. CO-ORDINATE CLAUSES

A.III. (a) If the pronoun occurred in the second of two co-ordinate clauses, it referred to the subject of the first, provided this was congruent. In the case of independent clauses both those connected specifically by a conjunction (e.g. und, und zwar, oder, oder aber, aber, denn,

jedoch, auch, dann, teils - teils, einerseits - andererseits) and those left unconnected were regarded as co-ordinate:

- e.g. (a) *Einteilung* nach der Wertigkeit hat gewisse Aussichten,
aber sie bringt Substanzen von diametral
entgegengesetzten Eigenschaften zusammen.
(b) *Manche feste Stoffe und Flüssigkeiten* ziehen Wasser an,
ohne sich jedoch notwendigerweise darin aufzulösen,
sie heissen ebenfalls hygroskopisch.

There were 34^φ instances where the pronoun in the second co-ordinate clause, being congruent with the subject of the first, referred to this, 1 instance where, though congruent, it did not:

- e.g. Der Endpunkt der Titration wird durch eine kleine, zu der
Lösung zugegebene Menge *eines Indikators* angezeigt,
er ändert die Farbe,
je nachdem die Lösung sauer oder alkalisch 1st.

It may be noted that in this case the reference was to the agent in a passive construction, that is to say, what would have been the subject of the corresponding active construction.

A.III. (b) In 1 instance the pronoun was congruent not with the subject of the first co-ordinate clause but with its dependent genitive, to which it referred.

- i.e. Die Einfachheit *der Gesetze* der chemischen Vereinigung führt zu der
Überzeugung,
dass sie auf Grund einer einfachen Vorstellung über den elementaren
Aufbau der Materie erklärt werden können,
und tatsächlich lassen sie sich aus der Atarihypothese leicht ableiten.

A.III. (c) In 2 further instances the pronoun, being congruent neither with the subject nor its dependent genitive, referred to the direct object of the first co-ordinate clause:

- i.e. Die Kraftwerke vom Rheinfall bei Schaffhausen bis zur
niedersächsischen Tiefebene geben *ihre Energie* in das
grosse westdeutsche Hochspannungsnetz,
und dem Bedarf entsprechend, wird sie den Verbrauchern
zugeleitet.

^φIn 4 of those the substantive referred to was also a pronoun, but in each case further reference to the correct, antecedent was practicable under the appropriate rules.

A.III.(Id) Finally there was 1 instance of this type which fell under neither (a), (b), or (c), referring instead to a substantive in the preceding sentence:

- i.e. Wenn *eine reine Substanz* in zwei oder mehrere Substanzen von niedrigerem Gewicht zerlegt werden kann, so nennt man sie eine Verbindung; ergeben sich stets kompliziertere Substanzen bei ihrer Vereinigung mit anderen Stoffen, und spaltet sie sich nie in einfachere Substanzen, so heisst sie ein Element.

If a pronoun were thus found to refer outside its own sentence, it could be treated in the same way as the exception in A.II. (a) above, that is as coming under Class B or C. The present instance in fact would be referred to its antecedent by rule B. II. (a).

CLASS A. IV. SUBSTANTIVAL CLAUSES

There were comparatively few examples of a nominative personal pronoun in this type of clause, so that little could be learned about the nature of the reference here. Those kinds occurring were as follows:

1. *clause (with or without dass) = direct object of preceding clause*

(a) The pronoun referred to the subject of the preceding clause (1 instance):

- i.e. *Der Verfasser* sagt im Vorwort, er beanspruche nicht, irgend etwas wesentlich Neues zu bieten.

(b) The pronoun was not congruent with the subject of the preceding clause and referred instead to its dependent genitive (1 instance):

- i.e. Oft zeigt schon die Prüfung *eines Körpers* mit dem blossen Auge, dass er aus mehr als einem Bestandteil zusammengesetzt ist.

(c) The pronoun referred to the substantive in the initial prepositional phrase of the preceding clause; in 1 instance the preposition was *für*,

- i.e. Für *diese Bahnen* können wir aber auf die gleiche Weise wie auf Seite 84 zeigen, dass sie sich alle in einem anderen Punkte P^1 der Ebene $x = X_0$ schneiden müssen.

in 5 instances *von* (2 of the substantives in this case being a relative pronoun)

- e.g. *Man vermischt die fragliche Substanz *mit einer anderen*,
von *der* man vermutet,
dass sie mit ihr identisch ist,
* und bestimmt den "Mischschmelzpunkt."

(d) The pronoun was not congruent with any substantive in the preceding clause, containing as this did only the impersonal *man*, and referred instead to the direct or prepositional object of the clause (conditional) before this:

- e.g. Fragt man am Schluss *nach den therapeutischen Anwendungen* von
Chinonen,
so wird man einsehen,
dass sie noch sehr begrenzt sind.

2. clause (with *dass*) = subject of preceding clause

(a) In 3 instances there was no substantive in the preceding clause for the pronoun to refer to, and it referred in fact to a substantive in the clause before this - in 2 of the instances to a substantive in a prepositional phrase,

- e.g. Bei der Analyse der Absorptionskurven gehen wir zunächst
von *der nuclearen Komponente* aus,
da empirisch bekannt ist,
dass sie den Hauptanteil der Sterne liefert.

in 1 to the direct object,

- i.e. Kristallisiert man *diese Formen*,
so zeigt sich,
dass sie zueinander spiegelbildlich sind.

(b) In 1 instance the pronoun was not congruent with the substantive in the preceding clause; it referred to the substantive in a prepositional phrase in the preceding sentence:

- i.e. Auf die Frage nach *den Ursachen* des warmen Tertiärklimas
soll hier nicht näher eingegangen werden.
Mir scheint jedenfalls,
dass sie trotz der Überlegungen von Kerner und Brooks
nicht nur in der veränderten Land- und Meerverteilung
und ähnlichen irdischen Faktoren gesucht werden.

A resemblance between this and the example under A. IV 1 (d) above may be observed: in both cases the construction is passive or virtually passive

(*man*) and impersonal, with the result that the substantive after *nach* acquires the Importance of the subject function.

CLASS A. V. WHERE THE PRONOUN OCCURRED (EXCLUDING THOSE TYPES OF CLAUSE ALREADY MENTIONED) IN

1. an independent clause preceded by one dependent clause or by two, if co-ordinate,
2. a dependent clause preceded by an independent clause,
3. a dependent clause preceded by another dependent clause, the main clause following later.
- and provided that in each case the clause preceding that containing the pronoun possessed neither direct object nor predicative substantive.

If this preceding clause consisted of two co-ordinate predicates with a common subject, reference was made to the construction of the second predicate to decide whether the clause contained a direct object or predicative substantive. ϕ

A. V. (a) If the pronoun was congruent with the subject of the preceding clause, it referred to this:

e.g. (a) Obwohl *die zentrische Formel* in mancher Richtung unbefriedigend ist,

erklärt sie das Fehlen von Isometrie in den Ortho-Di-Substitutionsprodukten.

(b)*Zerfällt zum Beispiel *ein Stoff* nur, wenn er am Katalysator adsorbiert ist,

*und ist bei normalen Bedingungen schon die Sättigungsgrenze der Adsorption erreicht,

so wird eine Erhöhung des Druckes keine Vermehrung der Zahl der zerfallenden Moleküle bedingen.

ϕ The reason for the exclusion in No. 1 above of two preceding dependent clauses, unless co-ordinate, is the difficulty encountered with such instances or determining to which of the two the pronoun in the independent clause refers. Similarly in No.3 only those instances are taken into account in which the main clause comes after the dependent clause containing the pronoun, since in those instances where the main clause precedes the two dependent clauses the pronoun may refer either to a substantive in the preceding dependent clause or to one in the main clause before this.

This was true of 39 instances^φ not true of 4. These were

e.g. (a) Ein Gang ist meist *durch einen Fehler* in der Messanordnung bedingt,
auch wenn er innerhalb der Fehlergrenzen liegt.

(b) Die Rolle der "Zwischenstrahlung" kann natürlich auch *von der Primärstrahlung* übernommen werden, indem sie ihre Qualität oder ihre Energie ändert.

(c)*Der *Saft* der Bauchspeicheldrüse, in dem das Enzym Trypsin enthalten ist,
* spaltet,
wenn er nicht zuvor mit dem Darm in Berührung kam,
* keine höheren, eigentlichen Proteine (Eiweissabbauprodukte);
+ erst wenn ein Extrakt der Darmschleimhaut, in dem die Enterokinase enthalten ist,
+ dazugegeben wird,
spaltet er auch höhere Eiweisskörper.

(d) Die Namen *neu entdeckter Elemente* endigen im allgemeinen auf -um,
wenn sie Metalle sind.

In the first two the pronoun, though congruent, with the subject, referred to the agent of the preceding clause⁺, in the third to the subject of the preceding sentence, and in the fourth to the subject's dependent genitive (see (b)).

A.V.(b) If the pronoun was congruent not with the subject of the preceding clause, but with the dependent genitive or *von* substantive (= dependent genitive) of the subject, it referred to the latter (3 instances).

e.g. Die biologische Bedeutung *der Klopfsignale* dürfte auf dem Gebiet der Fortpflanzung liegen,
indem sie eines der Mittel darstellen,
durch welche das Männchen sich als solches gelten lässt.

^φ 6 of these having a pronoun as the subject of the preceding clause, each with its antecedent determinable under the relevant rule.

+ c.f. the exception in A.III (a).

There were 10 instances in which the pronoun did not refer to the subject of the preceding clause or its dependent substantive, having no congruence with either. These were of the following types:

A. V. (c) The pronoun referred to the agent of the preceding clause (2 instances); ^φ

A.V. (d) The pronoun referred to a dative object after *entsprechen* (1 instance);

A, V. (e) The pronoun referred to a substantive after the impersonal construction *es handelt sich um...* (1 instance); ⁺

A. F. (f) The pronoun referred to a substantive in a prepositional phrase (2 instances):

e.g. *Bei den wirklichen Gasen ist diese aber merkbar und zwar um so mehr, je näher sie ihrem Kondensationszustand sind.*

It may be noted that in each of these 6 instances the pronoun could refer for reasons of congruence to only one substantive in the preceding clause. In the other 4 instances the pronoun could not for the same reasons refer to any substantive in the preceding clause. They were as follows:

A.V. (g) The pronoun referred to the subject of the preceding Independent clause but one:

i.e. *Diese Partikel fallen ungefähr entlang den magnetischen Feldlinien auf die Sonne zurück, und es besteht eine merkliche Wahrscheinlichkeit, dass sie durch Kerntreffer Neutronen erzeugen.*

A.V. (h) The pronoun referred to the subject of a clause Included within the preceding clause:

i.e. **Die Leichtigkeit, mit der sich diese *Stoffe* ineinander umwandeln lassen, * spricht dafür dass sie alle nur Modifikationen der einen Substanz, Wasser, sind.*

A.V. (i) In the last 2 instances the pronoun referred to the subject (pronominal) of the last clause in the preceding sentence:

e.g. *Man besass solche Tafeln schon im Altertum, und zwar ist uns als erste die Sehnentafel des *Ptolemäus* überliefert,*

^φ c. f. the exceptions mentioned in A. III. (a) and A.V. (a)

⁺ c. f. the remarks on the prepositional phrases in A.IV.2(b)

die er für astronomische Zwecke ums Jahr 150 n.Chr.
zusammengestellt hat.
Sie befindet sich in seinem Werke "Megale syntaxis,"
in dem er auch das nach ihm genannte Weltsystem entwickelt.

CLASS A. VI. *WIE* CLAUSES (NOT INCLUDED)

In the 4 instances where a pronoun occurred in a *wie* clause after a clause containing a direct object, it referred either to this:

e.g. An erster Stelle habe ich *diese Gruppe* zu nennen,
weil sie sich vor Erfindung der Logarithmen entwickelt haben.

or to its dependent substantive:

e.g. Diese Analogie lässt uns vermuten,
dass es möglich sein wird,
auch mit Elektronenstrahlen die gleiche Art von *Abbildung* zu
erzielen,
wie sie auf dem lichteoptischen Gebiete seit langem bekannt ist.

CLASS A. VII. WHERE THE PRONOUN OCCURRED (EXCLUDING THOSE TYPES OF CLAUSE ALREADY MENTIONED) IN

1. an independent clause preceded by one dependent clause or by two, if co-ordinate,
2. a dependent clause preceded by an independent clause,
3. a dependent clause preceded by another dependent clause, the main clause following later
- and provided that in each case the clause preceding that containing the pronoun possessed either a direct object or a predicative substantive.

A.VII. (a) If the pronoun was congruent with the subject, but not with the direct object or predicative substantive of the preceding clause, it referred to the subject. There were 11 such instances:

e.g. Diese *Potentialdifferenz* kann in einem lokalisierten Teil des Nervs
die anliegenden Nerventeile reizen,
wenn sie stark genug ist.

In 2 of the 11 Instances the subject referred to was itself a pronoun; both would have been referred to their correct antecedent by the relevant rule in each case.

A. VII. (b) If the pronoun was congruent not with the subject of the preceding clause, but with the direct object or predicative substantive, it referred to the latter:

e.g. Wenn man *Saccharose*, mit verdünnten Säuren erwärmt, wandelt sie sich in gleiche Teile Glukose and Fructose um.

Though true of 13 instances, this clearly could not apply in the following:

e.g. Die obige Reaktion kann man ebenso gut erklären unter der Annahme, dass *die Moleküle* von Wasserstoff und Chlor die gleiche gerade Anzahl von Atomen enthalten.

Es gibt jedoch gewichtige physikalische und chemische Beweise dafür, dass sie zweiatomig sind.

A. VII. (c) If the pronoun was congruent with both the subject and the direct object or predicative substantive of the preceding clause, it referred to the subject:

e.g. Zum Beispiel können *NH₄Cl und FeCl₃* Mischkristalle bilden, obgleich sie nicht isomorph und bestimmt nicht chemisch ähnlich sind.

This was true of 6 instances, not true of 1:

i.e. Wenn die Schauer des 3. Maximums *zwei geladene Mesonen* enthalten, sollten sie sich in der Nebelkammer unter geeigneten Bedingungen zeigen.

It should be observed, however, that in this case the pronoun reference was genuinely ambiguous. Only through the subsequent context did it become clear that the reference was to the object and not the subject.

CLASS A. VIII. DEPENDENT CLAUSE AFTER DEPENDENT AND MAIN CLAUSES

The pronoun occurred in a dependent clause following another dependent clause, but the main clause of the sentence came earlier than both instead of later as in A.V. The possibility existed, therefore, that the pronoun would sometimes refer to a substantive in the main clause or an earlier dependent clause rather than to one in the immediately preceding dependent clause.

A.VIII.(a) This in fact happened in all 3 instances where the pronoun stood in a relative clause:

e.g. Vielleicht liegt dies daran, dass *die Verfasser* nicht auf die Korrektur achteten,

*welche nach der Theorie des Skineffektes den geraden Draht von der Spule,
die sie tatsächlich zu ihren Messungen benutzten,
* unterscheidet.

In each case the pronoun referred to the subject (or its dependent substantive) of the next but one clause preceding.

A. VIII. (b) Of the other 10 instances of this type all but one referred to some substantive in the dependent clause immediately preceding, the exception being:

Die Kristalle vieler Hydrate mit hohem Wasserdampfdruck: haben das Bestreben,
Wasser zu verlieren,
wenn sie der Luft ausgesetzt sind.

where congruence Indications pointed away from the Immediately preceding clause to the one preceding this. The other 9 followed the rules of A.VII.

A.VIII.(c) If the pronoun was congruent either with the subject of the preceding clause or its dependent substantive (but not with both), it referred to this. There were 8 such instances, in 3 of them the subject referred to being also a pronoun. Reference of this to its antecedent would have been difficult in one case without consideration of meaning, or rather the human/non-human distinction:

i.e. Wir danken *den Herren* für die Übersendung von Separaten,
welche denselben Tagesgang aus Koinzidenzregistrierungen
statistisch erfassten,
indem sie die Methode der Stichtage auf den Tagesgang anwandten.
Sie wählten....

A.VIII. (d) If the pronoun was congruent with the direct object and not with the subject (or its dependent substantive) of the preceding clause, it referred to this (1 instance).

CLASS B. I.

The pronoun stood in the first clause of a sentence, the last clause of the preceding sentence being independent and containing neither direct object nor predicative substantive.

B.I. (a) If congruent with the subject of the last clause in the preceding sentence, the pronoun referred to this:

e.g. *Sauerstoff* lässt sich leicht durch Erhitzen von vielen sauerstoffhaltigen Verbindungen darstellen, so aus Quecksilberoxyd,

aus dem er zuerst von Priestley gewonnen wurde,
* weiter aus Kaliumchlorat und aus Kaliumpermanganat.

Für Handelszwecke wird er durch Wasserelektrolyse oder aus der Luft
fabrikmässig hergestellt.

What may have become apparent from previous illustrations may be noted explicitly, namely that when only part of a clause comes at the end of a sentence (as in the above), the whole clause to which this part belongs is nevertheless to be regarded as the final clause for the purposes of pronoun reference.

The above rule applied to 65 instances, 9 of these having a pronoun for the subject referred to. In 1 of the latter this pronoun would not have been referred to its correct antecedent by the relevant rule. Apart from this there were 3 instances to which the above rule did not apply, in one this was due to the reference being to the agent in a passive construction instead of to the subject^φ. In another the pronoun, though congruent with the subject of the last clause of the preceding sentence, referred not to this but to a subject in the sentence before:

i.e. *Die Opisthobranchiata* haben infolge Rückdrehung meist keine
Chiastoneurie mehr
und die eine erhaltene Kieme liegt meist hinter dem Herzen.
Die Schale ist oft rückgebildet,
die Geschlechtsorgane sind zwittrig.
Sie leben im Meer, einige auch freischwimmend.

In the third the pronoun referred to a substantive in a prepositional phrase, or to its apposition. This was in the nature of a predicative substantive in this case:

i.e. An manchen Stellen sind die Lymphkanäle *zu stecknadelkopf - bis
erbsengrossen Knoten* erweitert, den sog. Lymphdrüsen.
Sie bilden die weissen Blutkörperchen,
die hier den Giften und Bakterien entgentreten.

B.I. (b) If congruent not with the subject, but with its dependent genitive or *yon* substantive, the pronoun referred to the latter. In such cases the subject was generally a "mass" or other abstract noun;

e.g. *Anzahl, Gruppe, Menge, Art, Ordnung, Bestimmung,
Bedeutung, Prüfung, Einfachheit.*

^φc. f. the exceptions under A.III. (a) and A.V. (a)

This was true of 5 instances but not of a further 1, where the reference was again to the agent in a passive construction - or rather to the dependent substantive of the agent:

- i.e. Weitgehende Aufklärung dieser Gesetzmässigkeiten und Vorgänge, insbesondere in quantitativer Hinsicht, wurde durch die Untersuchung *der Atomspektren* der Elemente erhalten.
Sie haben den Charakter von diskontinuierlichen Linienspektren.

B.I.(c) In 1 instance the pronoun referred to the dative substantive (= subject) after the impersonal construction *es gelingt...*

B.I. (d) In another instance there was no sentence preceding that in which the pronoun stood, only the chapter title (VORPRÜFUNGEN), to which it in fact referred.

B.I.(e) In 5 instances the pronoun was congruent neither with the subject of the last clause in the preceding sentence nor its dependent substantive. It referred to a substantive in a prepositional phrase, though there was not in each case only one with which it was congruent. Where it was congruent with more than one, however, it referred to that in an "unattached" prepositional phrase.^φ

- e.g. Es lässt sich auch hier das Potential in der Nähe des Achsenpunktes Z_0 in *eine Reihe* nach y und nach $\Delta z = z - z_0$ entwickeln;
sie lautet $V(z, y) = \dots$

B.I. (f) In 5 instances the pronoun could not, from lack of congruence, refer to any substantive in the preceding sentence. It referred instead to the subject (or its dependent substantive) of the sentence before this. Had there been any substantive in the last clause of the preceding sentence congruent with the pronoun, those instances would have formed exceptions to rule B.I. (a) like the second mentioned there.

B.I. (g) In 1 instance the pronoun referred to a substantive in a prepositional phrase, as in B.I. (e), but with this difference, that the reference was to the plural component of a singular compound noun, thereby creating some difficulty in establishing congruence:

- e.g. Danach ist für diese *Nucleonengruppe* die Multiplizität je Stern in Blei höher als in Luft.

Sie sind aber nicht mehr zur Auslösung von Sternen fähig.

CLASS B. II

The pronoun occurred in the first clause of a sentence. The last clause of the preceding sentence was independent and contained a direct object.

^φ i.e. unattached to a substantive.

B.II. (a) If the pronoun was congruent with the subject of the last clause in the preceding sentence, it referred to this:

e.g. *Die Schauer* des 3. Maximums enthalten also wahrscheinlich zwei geladene Teilchen;
daneben können sie auch eine gewisse Zahl n_0 von neutralen Teilchen enthalten.

There were 31 such instances, the subject in 4 being itself a pronoun. The antecedent of this in turn would have been identified by the relevant rule in each case.

B.II.(b) If the pronoun was congruent not with the subject, but with the direct object, it referred to the latter.

e.g. Der umgekehrte Vorgang, *die Spaltung* von Estern durch Wasser bezeichnet man mit Verseifung:
sie ist eine Hydrolyse.

Though this applied to 11 instances (in one of which the direct object was a pronoun, but likewise causing no difficulty in reference to its antecedent), there was 1 to which it did not:

i.e. Die thixotrope Regelatinerungsgeschwindigkeit kann man oftmals durch *eine geeignete leichte Bewegung* erhöhen;
Sie ruft eine mechanische Koagulation der Teilchen hervor und heisst Rheopexie.

This may be compared to the exceptions in A. III. (a), A. V. (a), B.I. (a) and (b), since the reference here is to what is really the agent in the virtually passive *man* construction.

B.II. (c) If the pronoun was congruent with neither the subject nor the direct object, but with the dependent substantive of the subject, it referred to this (1 instance).

B.II.(d) In 2 instances the pronoun referred to the dependent genitive or *von* substantive of the direct object. In one it was congruent with neither the subject, its dependent substantive, nor the direct object; in the other it was also congruent with the direct object, to which it would be referred by rule (b) on the previous page. The direct object in this case, however, was *Art*, and it might be asked whether a pronoun ever refers to such a noun when it has a dependent substantive and not rather to this substantive.

B.II. (e) In 2 instances the pronoun referred to a substantive in apposition to the direct object, not being congruent either with the subject or the direct object:

e.g. Das Element geht eine wohldefinierte Verbindung mit Wasserstoff ein,
den Schwefelwasserstoff.

Er ist ein unangenehm riechendes Gas.

B.II.(f) In 1 instance the pronoun referred to the indirect object in the preceding sentence, again lacking congruence with both subject and direct object:

i.e. Die OH-Gruppe gibt *den Phenolen* einen saueren Charakter;
Sie können mit Basen unter Salzbildung reagieren.

B.II. (g) In 1 instance the pronoun referred to the substantive governed by the impersonal construction *man hat es zu tun mit...*^φ

B.II.(h) In 1 instance the reference was to a substantive in a prepositional phrase (other than *von*) dependent on the direct object. There was no congruence with the latter or the subject of the sentence:

i.e. Wir mustern nun aus diesen Gesichtspunkten das experimentelle
Material *über die Schauer* des 2. Max.

Nach Ziff. 7 und 10 werden sie durch geladene Teilchen ausgelöst,
deren.....

B.II. (i) In 1 instance the pronoun could not refer to any substantive in the preceding sentence by reason of lack of congruence; its reference was in fact to the direct object of the sentence before this. There was, however, in the intermediate sentence a possessive pronoun referring to this direct object:

i.e. Für Gebirgshöhen haben Dallaporta, Puppi und Merlin
einen schwachen nuclearen Übergangseffekt auch für
Sterne in Blei errechnet.

Sein Vorhandensein in 2950 m Höhe können wir unseren
Messungen nicht mit Sicherheit entnehmen.

Er kann in Kurve d (Abb. 2) enthalten sein.

CLASS B. III

The pronoun stood in the first clause of its sentence, while the last clause of the preceding sentence was independent and contained a predicative substantive.

^φ c.f. A. IV. 2. (b).

B-III. (a) If the pronoun was congruent with the subject of the last clause in the preceding sentence, it referred to this:

e.g. Der angegebene Wert gilt allerdings nur bei kleinen Geschwindigkeiten, denn *die Masse des Elektrons* ist keine konstante, sondern eine von dessen Geschwindigkeit abhängige Größe; sie wächst mit der Geschwindigkeit und wird bei Annäherung derselben an die Lichtgeschwindigkeit unendlich gross.

There were 32 such instances, 3 of these having a pronoun also for the subject referred to. In each case this would in turn have been referred to its antecedent by the appropriate rule.

B-III. (b) If congruent not with the subject, but with the predicative substantive, the pronoun referred to the latter:

e.g...
$$\text{Das Verhältnis } \frac{\text{Lichtgeschwindigkeit im Vakuum}}{\text{Lichtgeschwindigkeit im Medium}} = n \text{ heisst}$$

der Brechungsindex einer Substanz.

Er ist von der Wellenlänge des Lichtes abhängig.

This was valid for 3 instances; not, however, for the following:

e.g. Die oben erwähnten Reaktionen sind die Grundlage des sogenannten Kontaktverfahrens für die Fabrikation der Schwefelsäure.

Sie wird auch nach dem Bleikammerverfahren technisch hergestellt

Where the reference was to the dependent genitive of the prepositional phrase, dependent on the dependent genitive of the predicative substantive.

B-III. (c) In 1 instance the pronoun referred to the dependent genitive of the predicative substantive, being congruent neither with the latter nor with the subject of the sentence.

CLASS C.

The pronoun stood in the first clause of a Sentence, and the last clause of the preceding sentence was a dependent clause.

This class, comprising about 10% of all the instances, was the least amenable to syntactic treatment for the solution of pronoun reference. The difficulty was to determine not only to which substantive but also to which clause the pronoun referred, since although the reference was more frequently (73%) to the main than to a dependent clause in the preceding sentence, the ratio was not sufficiently one-sided to permit those instances where reference was to the latter to be ignored. On the basis of analogy from the previous classes, however, it was to be expected that

reference would be primarily to the subject of a clause, and reliance on this for identifying the antecedent's clause as well as the antecedent itself produced the following results.

CLASS C.I.

The pronoun referred to a substantive in the main clause of the preceding sentence.

C.I. (a) Of all the subjects of the clauses in the preceding sentence the pronoun was congruent only with that of the main clause, or if also with a pronoun which was the subject of its clause, then this too referred to the subject of the main clause:

e.g. (a) *Die qualitative Analyse* einer chemischen Verbindung gibt Auskunft darüber,

welche Bestandteile: Ionen, Atome oder Atomgruppen, in ihr enthalten sind.

Sie zerfällt in drei Teile:

(b) *Freier Schwefel* findet sich in vulkanischen Gegenden,

wo er oft mittels überhitztem Wasserdampf geschmolzen und von Erde befreit wird.

Er kommt als Stangenschwefel oder in Pulverform als Schwefelblüte in den Handel.

There were 21 such instances. In 5 of them the subject of the main clause was itself a pronoun, but in each case the antecedent of this would already have been identified by the relevant rule.

C.I.(b) The pronoun was congruent either with no subject of a clause or only with a relative pronoun subject, it was, however, congruent with the dependent genitive or *von* substantive of the subject of the main clause, and if also with a relative pronoun subject, this referred to the former (2 instances).

C.I. (c) In 2 instances, where the pronoun was congruent - as i

(a) - with the subject of the main clause and a relative pronoun subject, the latter did not refer as there to the subject, but to the direct object of the main clause. In one of the instances the pronoun referred to the subject. In the other to the direct object of the main clause:

i.e. (a) *Die Parabelmethode* gestattet eine Abwandlung, welche sich von der Annahme "r klein gegen I" freimacht. Sie wird deshalb überall dort mit Vorteil verwendet, wo die Elektronenbahnen Umkehrpunkte aufweisen.

(b) Die dazwischenliegenden Elemente habe *restliche Elektronen*, die nicht zu vollständig abgeschlossen, stabilen Gruppen vereinigt sind.

Sie üben Kraftwirkungen nach aussen aus und sind Einflüssen von aussen zugänglich;

C. I. (d) In 1 instance the pronoun was congruent not with the subject of the main clause, but with those of two following co-ordinated relative clauses, the relative pronoun itself being the subject in the second. It was to this, and through it to the direct object of the main clause, that the pronoun referred.

Neuerdings hat man *Elemente* gefunden, deren Kernladung und Masse zwar gleich sind, die aber eine verschiedene Anordnung der Protonen und Neutronen im Kern besitzen.

Sie heissen Kernisomere ...

It may be noted that here the direct object is really the subject of a virtually passive construction^φ and that therefore the instance approximates to the C.I. (a) type.

C.I. (e) The pronoun was not congruent with the subject, dependent substantive of the subject, or direct object of any clause in the preceding sentence. It referred to the dependent genitive of the direct object in the main clause (1 instance):

i.e. Im Ausgleich dieser Ladung hinter der Erde vermuten Chapman und Ferraro das Zustandekommen des *Ringstroms*, dessen Natur und Aufbau keineswegs geklärt sind. Er scheint so hohe Stabilität anzunehmen, dass...

C.I. (f) Not congruent with the subject of the main clause the pronoun referred to the relative pronoun subject of the intermediate clause, and through this to the predicative substantive of the main clause (1 instance):

i.e. $K_e = -eE$ ist *die Kraft*, welche das elektrische Feld auf das Elektron ausübt. Sie ist der elektrischen Feldstärke entgegengerichtet.

C. I. (g) Though not congruent with the subject of the main clause, the pronoun was congruent with the subject(s) of the intervening clause(s). These subjects, being pronouns, referred in one instance to the indirect, in the other to the direct (dative) object of the main clause:

^φ c.f. B. II. (b).

i.e. *Die klare Erkenntnis,
 dass es sich hier wirklich um ein ganz neues astronomisches Problem handelte,
 * und die ersten Ansätze zur Vorbereitung seiner Lösung verdankt die Wissenschaft *William Herschel*,
der zu Ende des 18. Jahrhunderts als erster mit Hilfe des von ihm erbauten grössten Spiegelteleskops seiner Zeit den bewussten Versuch unternahm,
 die Gestalt des unsere Sonne umgebenden Sternsystems auf Grund von Beobachtungstatsachen zu ermitteln.
 Hierbei bedient er sich einer Reihe von an verschiedenen Stellen ausgeführten Sternabzählungen.

and

Wir danken *den Herren* für die Übersendung von Separaten, *welche* denselben Tagesgang aus Koinzidenzregistrierungen statistisch erfassten, indem sie. die Methode der Stichtage auf den Tagesgang anwandten. Sie wählten.....

C.I. (h) The pronoun was not congruent with the subject of the main clause, but was congruent with a relative pronoun subject, this referring to the agent in the main clause (1 instance):

i.e. Eine andere Methode zur Auffindung der Wertigkeit von Elementen wurde durch *Mitscherlich* eingeführt,
der eine Ähnlichkeit zwischen den Kristallen analoger Verbindungen entdeckte.
 Diese Erscheinung ernannte er Isomorphie.

C.I. (i) The pronoun was not congruent with the subject of any clause, the reference being in fact to a substantive in a prepositional phrase in the main clause (1 instance):

i.e. In diesen Arbeiten von *F. Tödt* finden sich auch Angaben, die geeignet scheinen,
 jene Frage nach dem Mechanismus der Korrosion durch einen unmittelbaren Befund aufzuklären.
Er teilt nämlich mit, dass ...

C.I. (j) The pronoun was congruent both with the subject of the main clause and with that of the intervening dependent (relative) clause. It referred, however, to the prepositional relative pronoun, and through this to the direct object of the main clause (1 instance)

i.e. Die photochemische Chlorknallgasreaktion besitzt
eine "Induktionsperiode" zu Beginn der Bestrahlung,
während der keine Reaktion stattfindet.
Sie ist durch die Anwesenheit von Verunreinigungen bedingt.

CLASS C. II

The pronoun referred to a substantive in a dependent clause of the preceding sentence.

C.II. (a) Of all the subjects of the clauses in the preceding sentence the pronoun was congruent only with that of a dependent clause, or if also with a relative pronoun subject, this referred to the former:

e.g. Wer heute in Westdeutschland eine Lampe anknipst,
kann nicht mehr sagen,
woher die *Energie* stammt,
die sie zum Glühen bringt.
Sie kann ebensogut aus dem österreichischen Wasserkraftwerk
Bludenz wie aus den grossen Moorkraftwerken an der
holländischen Grenze kommen.

There were 7 such Instances.

C.II. (b) Not congruent with the subject of any clause nor with a dependent substantive of the subject in the main clause, the pronoun referred to the dependent *von* substantive (or its apposition) of the subject in the intermediate dependent clause (1 Instance):

i.e. Es ist ein unangenehm riechendes Gas,
von dem sich eine Reihe *von Salzen, die Sulfide,*
ableiten;
sie sind meist in Wasser unlöslich ...

C.II. (c) The pronoun was not congruent with the subject (or its dependent substantive) of any clause, nor with the direct object of the main clause, but referred to that of the dependent clause (1 instance):

i.e. Jedenfalls haben wir ein Interesse daran,
die sphärische Aberration nach Möglichkeit zu verringern;
ganz beheben lässt sie sich ... Im Gegensatz zur Lichtoptik nicht.

C. II. (d) The pronoun was not congruent with the subject of the main clause, congruent with the relative pronoun subject of the intermediate clause, which in turn referred to the direct object of the main clause. According to the rules of C.I. the pronoun would have been referred in such a case to this relative pronoun - and thereby to the direct object of the main clause, whereas in fact its reference was to the direct object of the relative clause (1 instance):

i.e. Die zweite Gruppe enthält die energieärmeren
Nucleonen $S-107 < E < S \cdot 10^8$ eV,
die noch Sterne auszulösen vermögen und in der photographischen
Platte *die sogenannten "grauen" Spuren* bilden.
Soweit sie aus mehrspurigen Sternen hervorgehen,
ist ihre Winkel Verteilung weitgehend isotrop.

C.II.(e) Though not congruent with the subject of the main clause, the pronoun was congruent with the subject of the three succeeding dependent clauses, this in one being a relative pronoun referring to the previous subject. It was to this relative and its antecedent that the pronoun referred (1 instance)

i.e. Die Beobachtung lehrt,
*dass *parallele Lichtstrahlen*,
die zur Hauptachse parallel auffallen,
* durch die Brechung in der Konkavlinse divergent
gemacht werden, und zwar so,
dass ihre Rückwärtsverlängerungen durch einen Punkt der
Hauptachse gehen;
sie scheinen also von diesem Punkt aus zu divergieren.

C.II.(f) The pronoun was congruent with the subject of the main clause and with the subjects of the two succeeding dependent clauses, the latter of these being a relative pronoun referring to the former. As in C. II. (e), the pronoun referred to this relative pronoun and its antecedent (1 instance):

Diese Kurven sind dadurch charakterisiert,
dass an ganz bestimmten Stellen des Spektralbereichs *selektive*
Absorptionsbanden mit ausgeprägten Maxima liegen,
die durch ihre spektrale Lage, ihre Höhe und ihre
Halbwertsbreite gekennzeichnet sind.
Sie sind bestimmten Atomen oder Atomgruppen in
Molekül, sogenannten Chromophoren, einzuordnen.

CONCLUSION

The nature of the syntactical Items referred to and the number of pronouns out of a total of 392^φ referring to each was as follows (column A.):

	A	B	C	D	E
Subject	296	283	10	2	1
Direct object	38	33	2		3
Substantive in prepositional phrase	22	18	2		2
Dependent substantive of subject	16	14	1		1
Agent	7	2			5
Dependent substantive of direct object	4	3			1
Predicative substantive	3	3			
Apposition of direct object	2	2			
Indirect object	1	1			
Dependent substantive of agent	1	0			1
Dependent substantive of prepositional phrase	1	0			1
Dependent substantive of predicative substantive	1	1			

Correct reference, however, was not possible in all cases; Under B are given the number of instances in which the pronoun was congruent with and referred to an antecedent in its expected location (i.e. the immediately preceding clause or sentence) according to the relevant rule.

In some cases the relevant rule failed to find a substantive congruent with the pronoun in question in the expected location, but a switch to the next relevant rule (c. f. for example A. II.b) succeeded in identifying the correct antecedent. Such Instances are shown under column C and were of the following types:

Subject A. II. b, A. III. d, A. V. g, A. V.h, A. V.i, B. I.f-4.

Direct object A. IV.2.a, B.II.1.

Substantive in prepositional phrase A. IV. 2. a, A. IV. 2.b.

Dependent substantive of subject B. I. f.

^φ The 24 instances of class A.I. are excluded, the pronoun there referring without exception to the nearest congruent noun before the Included clause.

On the other hand there were one or two instances of this nature where the relevant rule did find a congruent substantive in the normal location, with the result that the pronoun was referred to the wrong antecedent, the true antecedent being in an abnormal (i.e. earlier) location. These are given under column D. They were of the following nature:

A. V. a. and B.I.a. - the pronoun was referred to the subject of the immediately preceding clause or sentence, but actually referred to that of the sentence before this.

Finally there were occasions where a particular rule, adapted to the majority of instances, met with an exception and made an error in reference. These are listed under E alongside the syntactical item to which they ought to have been referred. In fact, most of them were referred - as would be expected - to either the subject or the direct object of the clause. The types were as follows:

Incorrectly referred to the subject; actually referred to:-

the direct object: A.VII.c, C.I.j, C.II.d.
the agent: A. III. a, A.V.a. (2), B.I.a.
a substantive in a prepositional phrase: B.I. a, B. I.g.
the dependent genitive of the subject: A.V.a.

Incorrectly referred to the direct object; actually referred to: -

the subject (of the preceding sentence): A.VII.b.
the agent: B. II.b.
the dependent substantive of the direct object: B. II. d.

Incorrectly referred to the predicative substantive; actually referred to:-

the dependent genitive of a prepositional phrase dependent on the dependent genitive of the predicative substantive: B.III.b.

Incorrectly referred to the dependent substantive of the subject; actually referred to:-

the dependent substantive of the agent: B. I.b.

The number of errors, then, produced by the above in connection with the texts concerned would have been 19; that is, those listed under D and E plus two others, the antecedent of which though correctly identified was itself a pronoun, and the antecedent of this would not have been identified correctly. This means something over a 95% correct solution. Naturally, being adapted to a particular set of instances, the rules produce a better result than would be obtained, if they were applied to an entirely fresh set.

Nevertheless, syntactical methods may always be expected to solve a fairly high percentage of pronoun references - or at least so one would assume from an experiment carried out in the course of compiling the present data. For after half of the instances had been collected, the rules were formulated on the same lines as the above to cover as many of them as possible. Then, when the other half had been collected, these rules were applied to them to see in how many cases they would solve the pronoun reference. The result was 83%.

It would seem reasonable, therefore, to assume that with a more detailed and reliable set of rules, such as might be expected from further investigation of the problem, a solution in the region of 90% would be likely.

A complete solution of the problem by syntactical considerations alone is not possible, since examples exist which, though identical in syntactical structure, nevertheless differ in their pronoun reference. In such cases the reference rule can only be chosen on a basis of attempting to cover the greatest possible number of instances, thereby at the same time necessitating that in the "exceptions" the pronoun will be referred to the wrong antecedent. That the translation also will be wrong does not automatically follow. This, as was remarked at the beginning, will arise only from confusion of the human/non-human distinction in a context mentioning persons as well as other animate and inanimate things.

The possibility of a more complete solution to the problem lies in a combination of syntactical methods with a system of word classification. The latter has been found in the course of past investigations to be essential for solving the residue of most problems after application of syntactical methods, often an extremely detailed classification appears to be required, but in the present instance a very broad one would suffice, such as a division of the words in the dictionary, especially verbs and adjectives, according to whether they represented or were characteristic of a human animal, a non-human animal, an animate object, or an inanimate object, of course, some would be characteristic of more than one, and then perhaps more so of one than another; such things would have to be taken into account.

In some instances it would then be possible to ascertain whether a pronoun represented a human or non-human being or thing simply from the immediate context (e.g. the nature of the verb, a predicative adjective or substantive, or a prepositional phrase). For the majority, however, the classification would need to be used in conjunction with syntactical rules, although these would have a somewhat different function than before. Instead of making a final decision about the identity of the pronoun's antecedent they would serve merely to indicate the various possible antecedents from the point of view of congruence and usage. The appropriateness of each antecedent in the pronoun's context would then be tested. If this

scheme were used, the only instance where an error could occur would be where there were two (or more) possible antecedents, one human, one non-human, and each was found to fit equally well into the context of the pronoun. How often such instances occur - if at all - without also being ambiguous in the German is a matter for further investigation. They would obviously be least likely when the other antecedent besides the human one belonged to the class of inanimate object, most likely when it belonged to that of non-human animal. If instances did occur where the context provided no criterion for distinguishing between the possible antecedents, the syntactical rules could be used as a final resort in the original manner described above with a 9:1 chance in general of selecting the correct antecedent.

A much simpler and better solution from the point of view of economy in the translation program would be to re-adopt in principle the idea of the pre-editor, though putting the responsibilities on the writer of the article or book rather than on an editor in the usual sense. In short, prescriptive grammar. With pronoun reference, as with most problems in mechanical translation, 90% of the instances follow a regular pattern. Why should not the other 10% do so too? In most cases a small alteration would have sufficed, as for example the conversion of a passive to the corresponding active construction (c. f. Class A. III.a.):

Der Endpunkt der Titration wird durch eine kleine, zu
der Lösung zugegebene Menge *eines Indikators* engezeigt,
er ändert die Farbe,
je nachdem die Lösung sauer oder alkalisch 1st.

becoming

Eine kleine, zu der Lösung zugegebene Menge *eines*
Indikators zeigt den Endpunkt der Titration an,
er ändert die Farbe,

which would have been correctly solved under A. III.b.

Once aware of the existence of such difficulties for mechanical translation an author could quite easily avoid raising them. The only question mark against such prescriptive grammar concerns the number of rules an author might be expected to memorise and apply to his compositions. If they were to grow so numerous as to cause considerable inconvenience in writing, a professional pre-editor would become a necessity, and if - as some believe - such extraneous assistance to mechanical translation is to be avoided, then it would be advisable to resort to the prescriptive grammar solution only with the most intractable problems.

ACKNOWLEDGEMENT

This work was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U.S. Navy (office of Naval Research); and in part by the National Science Foundation.

DISCUSSION ON THE PAPER BY DR. L. BRANDWOOD

MR. R. H. RICHENS: Dr. Brandwood's descriptive syntactical analysis of selected German sentences will be of value, not only to those working on the mechanical translation of German, but also to those working on other languages where analogous problems arise. However, the results we have been given raise in an acute form the problem of the relative importance of syntactic and semantic analysis in machine translation. Dr. Brandwood has shown that the antecedents of the German pronouns that he has studied cannot be located by any simple syntactic procedure; it was necessary to devise a considerable number of probability rules, most of which have exceptions with appreciable frequency. The human translator does not trouble with such rules since the antecedent is located by semantic congruence and it is important therefore to integrate both syntactic and semantic analytic procedures into any machine translation programme with pretensions to generality. The methods being investigated by the Cambridge Language Research Group for semantic analysis are the thesaurus method and a technique involving three types of analysis of interlingual formulae. Dr. Brandwood tentatively resurrects the notion of the pre-edited text. I cannot follow him here, since, leaving various impracticabilities aside, the interest of machine translation resides largely in attacking rather than minimizing the problems inherent in existing natural languages.

MR. E. A. NEWMAN: I thought it worth while to mention what seemed to me an example of the very great importance of context in translation. We saw an example of translation this morning which might have been called English to English, English to Russian, Russian to Russian, Russian to English, English to English, and we noted, many of us, I think, that the English to Russian and Russian to English was very easy, but the English to English and the Russian to Russian showed signs of having a lot of difficulty. This does rather illustrate that context is of very great importance.

PROF. Y. BAR-HILLEL: We heard this afternoon two examples of investigations into machine translation which in a certain sense lay at opposite ends of this research. None of these investigations, I would say, is representative of much of what is going on in machine translation. But it is, of course, not my task to fill in the gap.

Though the investigations we listened to are then not typical of MT research in general, they are indicative of a certain disappointment in the

achievements of MT in the last years. This disappointment seems to me a necessary reaction to the original sin of having staked the aims too high. Occasionally, however, the despair has yielded strange fruit. From the fact that we still don't have a working programme for translation between one pair of languages that would be both fully-automatic and high-quality, some drew the conclusion that one must now aim for fully-automatic high-quality translation between any pair of languages. Others drew the conclusion that one has to go into extremely minute investigations of minute occurrences. This is indeed necessary so long as fully-automatic high-quality translation is still the aim. Nothing less than what Mr. Brandwood has given us will suffice for solving one of the million problems which one could cite in this connection. But clearly what Dr. Brandwood did is not even enough to solve the problem he set himself. He went with extreme care over certain samples in order to derive from their analysis rules that would fit these samples hoping that they would fit in general. I don't think there is any good reason for this hope. Nobody who knows German will find any difficulty in concocting counter-examples to any of the rules which Dr. Brandwood derived from his analysis. Dr. Brandwood himself apparently made no such attempt, misled by a rather frequently found "empirical approach" which distrusts the rules found in the grammar books - probably because they were given by normatively-minded grammarians - and hopes to arrive at better rules from analysis of a large enough corpus.

I regard this approach as enormously wasteful. To throw away insights of the best grammarians (only because you were unable to derive from these insights a working method of MT and had therefore to find a scapegoat on whom to put the blame, instead of concluding that your original aim was too high!) and to start from scratch seems to me extremely bad policy. But, worse than that. The refusal to put the "derived" rules to test by cooking up counter-examples, and the concomitant decision to wait instead for a counter-example to come up by itself in some future sample, is a caricature of the empirical approach and definitely contrary to the standard scientific method which, if only possible, dispenses with observation in favour of experimentation.

The decisive question whether fully-automatic, high-quality translation is a reasonable aim is not one whose answer should be dependent on one's temperament. I am convinced that this question must and can be decided on a purely scientific level. I think I can prove that this aim is totally unrealistic on the basis of machinery available or under construction today, and of programming techniques in use or in prospect today. I would not dare to say anything about machines of which we do not even have blueprints, and of programming techniques of which nobody knows even the principles today.

I do not think that Chomsky's recently obtained deep insights into the structure of languages (*ref. 1*) are at all helpful to machine translation,

Their major impact on MT is rather to show that mechanical syntactical analysis is incomparably more difficult than linguists had believed before. (For an elaboration of this evaluation, see my papers, *refs. 2 and 3*).

Let me finish my comments with a challenge. I defy anybody to give me any *non-ad-hoc* rules which will provide for the correct machine translation from English into French of both the following two sentences: "The pen is in the box" and "The box is in the pen". Notice that within a suitable longer context - and I leave it to you to imagine the two contexts - the translation will be unique for anybody who knows sufficiently well English and French. Now, would someone volunteer to translate the sentences into French?

DR. L. MEHL (CHAIRMAN): "*La plume est dans la boite*" and "*La boite est dans la plume*". (Laughter).

You see the inconvenience of having a chairman who does not speak English perfectly! He translates like a machine.

PROF. BAR-HILLEL: Now this is exactly what I expect a machine to come up with, if it must come up with a unique translation. Dr. Mehl gave us these translations because he worked for a minute like a machine and because he might not know the other meanings of 'pen' not being a native English speaker.

DR. L. KOWARSKI: I think your example is probably even worse than you expect, because a possible French word for 'pen' in the second sense is 'box'. "*La boite est dans le box.*"

PROF. BAR-HILLEL: Well, I am surprised. I indeed did not know that the correct French rendering of 'pen' in the second sentence is 'box'. Now, why is it the case that you, Dr. Kowarski, who knows English and French well had no difficulty in finding the correct renderings? Knowledge of the two languages is not enough! You had to use, in addition, your knowledge of the world acquired in many years of learning, schooling, and observation. You knew from experience and not only from your knowledge of the languages, that the entitles denoted by '*plume*' are generally smaller than the entitles denoted by '*boite*' whereas the entitles denoted by 'box' (in French) are usually larger. It was this empirical knowledge of the relative sizes of certain objects, a knowledge which it is not very difficult for a

REFERENCES

1. CHOMSKY, K. Syntactical Structures, *s'Gravenhage* (1957).
2. BAR-HILLEL, Prof. Y. Decision method for structure in ordinary language. *Logique et Analyse*, 1959, **2**.
3. BAR-HILLEL, Prof. Y. Some linguistic obstacles to machine translation. *Proc. of the Second International Congress on Cybernetics, Namur* (1959).

human to acquire, I admit, which was a necessary condition for the uniqueness of the translation. I know of no method, surely not the Thesaurus method advocated by the Cambridge group, which would enable a machine to come up with the correct translation and not be completely *ad hoc*. It is pointless to believe, that a machine will have built into it the knowledge of all size relationships of all the objects in the world. A machine, however, that will be able to *acquire* this knowledge will surely be of a degree of complexity of a human being, and the point of having machine translation would be lost altogether.

If my challenge is left unanswered, then the aim of high-quality, fully-automatic translation must be considerably lowered. I admit that thereby the great attractiveness of the problem will be reduced, for me and for other people. However, those people who are interested in MT not as a purely intellectual problem but as something of great practical value will do best if they reduce their aim to either fully-automatic but low-quality or else to high-quality but only partially-automatic translation.

DR. A. M. UTTLEY: I have never done anything on machine translation, but seeing this problem I would like to write something more on the board. "The spring is in". Now I might ask the mechanical translators to translate that. I think their reply would be that that is not fair.

PROF. Y. BAR-HILLEL: Is that a full-stop?

DR. UTTLEY: No. They would say: "That is not fair, let me have a little more context. I do not know whether this is a spring in a lock, or a spring in a garden". I would say the same thing of the first two sentences and that given a little more context, a translation could be made mechanically by non ad hoc methods.

MR. P. E. TRIER: I believe the author has set himself a very difficult problem in the codification of these pronoun references for machine translation, since even he himself has given several ambiguous or wrong references. On page 320 at the top: "Ein Gang ist meist durch einen Fehler in der Messanordnung bedingt". That would refer at least as probably to the word "Gang" as to the word "Fehler."

In the middle of page 322: "Diese Analogie lässt uns vermuten"- "sie" refers to the "Art" and not "Abbildung".

Page 323- " when he talks about the 'maximum' he does refer himself to the ambiguity which is quite unresolvable within the sentence that is given.

Lastly on page 324, about the middle "Wir danken den Herren für die Übersendung von Separaten, welche denselben Tagesgang aus Koinzidenz-registrierungen statistisch erfassten." Here "welche" must refer to

"Separaten" and not to "Herren." With these genuine difficulties, I believe the machine is going to have quite a problem.

PROF. Y. BAR-HILLEL: No, it is not the missing context that does not allow the machine to translate correctly my two sample sentences. Even had I embedded the two sentences in two suitable paragraphs, the machine would still have been unable to decide that in the one case 'pen' has to be rendered by '*plume*' and in the other by '*box*', except by an *ad hoc* procedure. However, as soon as someone will come up with a procedure, according to which the machine will be able to resolve the ambiguity, I commit myself to concoct in fifteen seconds a simple, nothing special, counter-example for which this procedure will fall. (This is the reason why I use the term '*ad hoc*'.)

By the way, there exists to this day no procedure which takes into account larger-than-sentence contexts. But though such contexts mean a new order of magnitude of complexity, I am quite ready to agree that this extension will have to be tackled if one wishes to increase the machine part in the whole translation process.

MR. R. H. RICHENS: Professor Bar Hillel is misinformed in supposing that no procedures exist for dealing with ambiguities of the type he mentions. Furthermore, in the procedure in which I am most interested, the paragraph and not the sentence is the field for semantic comparison. Briefly, this procedure involves translation of all word segments into the range of interlingual semantic formulae which each potentially represents. Two principal operations are then performed on these expressions: (1) testing the semantic congruence between the terms of dyadic relations and the relations themselves; and (2) production of a semantic spectrum for the entire paragraph and using this for making decisions for such ambiguities as remain to be resolved. This procedure is fallible but gives a correct solution in an as yet unknown number of cases; it is mathematically definite. Prof. Bar Hillel quoted two sentences devoid of context. They would not be expected to occur in the same paragraph except in a paper dealing with difficulties in machine translation. If they were embedded in likely contexts, they would not appear to offer insuperable difficulties to existing semantic analytic procedures.

PROF. Y. BAR-HILLEL: As a matter of fact, I discussed my sample sentences with other members of the Cambridge Group, Miss Masterman, Mr. G. Parker-Rhodes and Mr. Needham, less than two weeks ago in Washington, and they were unable to present a procedure that would provide for a correct translation of both sentences. By the way, I did not claim at all that both sentences have to appear together in some far-fetched story. I am quite

satisfied with dealing with either sentence at a time. I still claim that a machine will have no way of deciding from the given sentence, in many if not all contexts of paragraph length, that 'pen' means *'plume'* rather than *'box'*. I presented you with two sentences only for pedagogical reasons. The members of the Cambridge Group in Washington sat for some time to cook up a Thesaurus which would resolve the ambiguity, without success. And I think that Miss Masterman at least was quite ready to admit at the end that a Thesaurus would not resolve all semantical ambiguities. I do not know what the other members of the Group believe in this respect.

DR. L. MEHL: If we consider legal questions, it is possible that in a certain passage of such a text, the same word can be taken to have a different meaning. Your example is quite obvious here, but in juridical text, for example, ambiguous words are relatively frequent, and the automatic translation, of juridical texts would be very difficult. On the question concerning the translation of "pen" you say "box" in French. I could have said "enclos" or "parc" just as well as "box" because the French "box" means "garage" or "parking". But for horses "box" is correct.

DR. KOWARSKI: By the way, there is another translation of "The box is in the pen" in French, it would be "Le combat a lieu au pénitencier."

MR. J. F. COALES: I would like to make one point that seems to me relevant, that both these sentences are essentially ambiguous, ambiguous in English. You can take "pen" and "box" in either way and the sentence makes sense in a certain context; which to choose is really a question of probability. You cannot translate an imprecise sentence into a precise sentence, and I suspect such a case should be treated on a probability basis, which I think the Thesaurus method does in fact do. It must take into account the size of the pen versus the box, and the question of whether it can be small or large then has some effect on how you translate it.

DR. L. MEHL: It is clear in some of these examples that the probability of the appearance of these two sentences in the same passage is highly unlikely, but in fact I have experience of juridical texts. Having written a book concerning tax questions, I made an alphabetic index for the book, and I noticed that certain words appearing in the same place in the alphabetic index because they were identical in form were totally different in meaning in the same book. I think it confirms your own observations concerning the difficulty of automatic translation.

DR. L. BRANDWOOD (In reply): Mr. Trier has kindly pointed out several "errors". He is correct in saying that on page 320 the pronoun reference could be to *Gang*

just as easily as to *Fehler* - in which case the example would not be an exception, but would comply with rule A.V.a.

He is not correct in saying that on page 322 *sie* refers to *Art* instead of *Abbildung*. This may be most easily seen by subtracting each of the two words in turn from the sentence, thus

- (1) auch mit Elektronenstrahlen die gleiche *Abbildung* zu erzielen, wie sie auf dem lichtoptischen Gebiete seit langem bekannt ist. conveys practically the same information as the original sentence, whereas
- (2) auch mit Elektronenstrahlen die gleiche *Art* zu erzielen, wie sie auf dem lichtoptischen Gebiete seit langem bekannt ist. has an entirely different sense.

Mr. Trier is also incorrect when he declares that on page 324 the pronoun must refer to *separaten* not to *Herren*. That the *welche* does in fact refer to *Herren* is indicated by the following *indem* clause, which must have a common subject with *erfassten*, which it explains, and one could hardly say of *Separaten* that they *die Methode* *anwandten*. Mr. Trier is apparently not acquainted with the fact that in German a relative pronoun may refer to a syntactical unit other than that which it immediately follows.

Prof. Bar-Hillel's discourse is at first sight a plausible attempt - for reasons one can but guess at - to discredit all the work done on mechanical translation in this country. On further consideration it can be seen that there is little logic and even less fact to substantiate his remarks. These may be taken in the order in which they were made.

First there is the statement about my investigating samples in order to find rules that would fit these samples in the hope that they would also fit in general. "I don't think", Prof. Bar-Hillel says, "there is any good reason for this hope". Apparently he has not read or chooses to ignore the fact mentioned on page 337 of the paper, that the set of rules derived in this way covered over 80% of the instances in a fresh sample. Prof. Bar-Hillel's *belief* to the contrary is of no great value. As for the subsequent "Nobody who knows German will find any difficulty in concocting counter-examples to any of the rules which Dr. Brandwood derived from his analysis", we might ask who would wish to deny this seeing that the investigation was described solely for the purpose of illustrating work in progress and no claim was made to have arrived at a complete solution to the problem. Even so, any such "concocted" counter-examples would fall under the 10-20% not accounted for by those rules with the widest application, and would therefore represent what would be - on the existing evidence - the less common grammatical constructions.

Prof. Bar-Hillel is quite right in assuming that I myself made no attempt to "concoct" counter-examples. It is disputable, however, whether he is also correct in saying that I was "misled by a rather frequently found 'empirical approach' which distrusts the rules found in the grammar books -". If a complete system of a language's grammar is to be obtained, then it is

essential to use the method of examining texts for the simple reason that no individual is acquainted with the whole, merely with a part of his native tongue; he uses not all the available vocabulary and grammatical constructions but only a personal selection. Consequently he cannot "concoct" all the possible exceptions to a grammatical rule derived from the observation of a number of instances in a sample of text. It will probably be generally agreed that in reading one not infrequently comes across a grammatical construction which strikes one as completely different from one's own usage. A rough-illustration might be provided by reference to one of Prof. Bar-Hillel's statements, though the unusualness in this instance is due to incorrect rather than just variant English constructions: "I defy anybody to give me any *non-ad-hoc* rules which will provide for the correct machine translation from English into French of both the following two sentences: 'The pen is in the box' and 'The box is in the pen'. Notice that within a suitable longer context the translation will be unique for anybody who knows sufficiently well English and French." Although the pleonastic "both" and "two" in the first sentence and the position of the adverb between the verb and its object in the second are new to me - I should never have been able to "concoct" them - the sentences exist and it is perhaps conceivable that someone might wish to translate them, together with the rest of the discussion, into German. In this event allowance would have to be made for such constructions in the programme. One would not want to find such phrases in the output as *beiden der beiden Sätze*.

Furthermore, to say - as Prof. Bar-Hillel does - that the empirical method "distrusts" the rules found in grammar books is ridiculous. The truth is that, while the mechanical translator would dearly like to use the rules formulated by traditional grammars, he cannot, either because they are not sufficiently precise for his purpose or because they are non-existent for the problem with which he is dealing. Let Prof. Bar-Hillel try to find rules or "insights" bearing on the problem of pronoun reference even in the larger German grammars; he will discover that they are conspicuous by their almost complete absence, why else would an eminent professor of German philology Professor B. Ulvestad of Bergen University, be investigating this problem at the present time?

Prof. Bar-Hillel's emphasis on the sin of proceeding beyond *traditional* grammar or of employing the empirical approach, "contrary to the *standard* scientific method", and his insistence, based on a consideration of *available* electronic computers, that "fully automatic high-quality translation is an unreasonable aim," make it abundantly clear that with his inability to see beyond the present state of progress Prof. Bar-Hillel is the eternal pessimistic conservative, the scientist who before Galileo's observations thought it impossible that the earth should travel round the sun, the military expert who, with the Maginot Line in mind, asserted

that France was impregnable - until Hitler's paratroopers overran it in a matter of weeks, the man who before the Wright brothers gave their demonstration was incredulous of the possibility of flight by machines heavier than air and who even afterwards could see no further than his nose and scoffed at the notion that a later development of this same machine might cross the Atlantic, not to mention reach the moon.

As regards Prof. Bar-Hillel's challenge he is not taking much of a risk in "defying" anyone to produce at such short notice non-ad-hoc rules to translate his two sentences correctly. The problem of multimeaning is one if not *the* most difficult problem of mechanical translation, and it would be foolish to deny that the translation of the type of sentence quote by Prof. Bar-Hillel causes some difficulty, on the other hand it would be equally foolish to deny, as he does, that there is any possible *non-ad-hoc* solution. He is mistaken, moreover, in thinking that an "empirical knowledge of the relative sizes of certain objects" is a "necessary condition for the uniqueness of the translation". This knowledge is in fact derivable from the sentences themselves, since the subject of the verb "to be in" must be smaller in size than the object governed by the verb, at least when both represent material things and can be spoken of as having size. Knowledge of relative size alone is not sufficient to permit a unique translation here, since - as Mr. Coales observes - both sentences are completely ambiguous as they stand. Thus "the pen is in the box" might be translated as

- (a) la plume est dans la boîte.
- (b) le box est dans la boîte.

Since in (b) *le box* could refer to this particular piece in a child's toy model of a farm. Similarly "the box is in the pen" could be either of

- (a) la boîte est dans le box.
- (b) la boîte est dans la plume.

where (b) would refer to a box small enough to be concealed in a hollow pen, such a box in fact as could have held the poison which the famous Greek orator Demosthenes was said to have carried around with him in his pen, and by which he took his life on the occasion of his impending arrest by the soldiers of Alexander the Great.

The uniqueness of meaning, then, comes only from the context in which the particular sentence is found, and this being so I see no reason why such sentences should not be capable of being dealt with by a Thesaurus method - or in many cases even by the much simpler method outlined in the book "Mechanical Resolution of Linguistic Problems (1958)" pp. 94-95 by Booth, Brandwood and Cleave especially in view of the fact that in scientific texts, the material for which mechanical translation was primarily intended, ambiguity is by no means as common as Bar-Hillel's preoccupation with it would suggest, and can be reduced still further by the use of idioglossaries.