# ON SEVERAL TYPES OF LINGUISTIC ISOMORPHISM

By

Hiroshi AMAKI

(Received November 18, 1968)

In the present paper, we shall investigate the diverse types of linguistic isomorphism. For the purpose, before preceding, we shall explain the notations and terminology which are used in the Book [2] by S. Marcus.

Let $\Gamma$ be a given finite set — vocabulary. Elements of $\Gamma$ will be called **words**. We denote by T the free semi-group generated by $\Gamma$. By definition, the elements of T will be finite strings (briefly, strings). Now let $\Phi$ be a subset of T. Then we shall call the strings which belong to $\Phi$ **marked strings**. If P is a given partition of $\Gamma$, each set of P will be called a P-cell, and we denote by $P(a)$ a P-cell containing the word $a$. Moreover we notice that for two distinct words $a$, $b$ we have either $P(a) = P(b)$ or $P(a) \cap P(b) = \phi$ (where $\phi$ is the empty set). Now consider a triple $\{\Gamma, P, \Phi\}$, and we shall call such a triple a **language with paradigmatic structure** (briefly, **language**). Let $x \varepsilon \Gamma$, and $y \varepsilon \Gamma$. We shall say that $x$ **dominates** $y$, and we shall write $x \rightarrow y$, if for each pair of strings $p$ and $q$ such that the string $pxq$ is marked, then the string $pyq$ is also marked. Thus for any $x \varepsilon \Gamma$, $y \varepsilon \Gamma$, the set $S(x) = \{y\,;\, x \rightarrow y \text{ and } y \rightarrow x\}$ determines a partition of $\Gamma$ into disjoint sets and such a partition is called a **family** S. Furthermore, the unit partition of T is, by definition that partition for which $E(x) = \{x\}$ when $x \varepsilon \Gamma$.

The above partitions S and E are the most useful in the study of linguistics. The notion of marked strings and family has been introduced by O. Kouladgina in his paper [1]. The notions of domination and of family have been studied in detail by S. Marcus in his papers (for example, [3]). A finite sequence $P(x_1)$, $P(x_2)$, $\cdots$ ,$P(x_n)$ of the cells of a partition P of $\Gamma$, is called a P-**structure**, and we shall say that this P-structure is marked if there exists a marked string $y_1 y_2 \cdots y_n$ such that $y_1 \varepsilon P(x_1)$, $y_2 \varepsilon P(x_2)$, $\cdots$ , $y_n \varepsilon P(x_n)$. Let $P(x)$ and $P(y)$ be two cells of P. Then we shall say that $P(x)$ and $P(y)$ are P-**equivalent** and we shall write $P(x) \leftrightarrow P(y)$, if for each pair of P-structures $P_1$ and $P_2$, the P-structure $P_1\, P(x)\, P_2$ and $P_1\, P(y)\, P_2$ are either both marked or both unmarked. Let us consider a language $\{\Gamma, P, \Phi\}$. Put, for each $x \varepsilon \Gamma$,

$$P'(x) = \bigcup_{P(x) \leftrightarrow P(y)} P(y)$$

Then the set $P'(x)$ determines a partition of $\Gamma$ into disjoint sets. The partition $P'$ is called the **derivative** of the partition P.

Now let us consider two languages $L_1 = \{\Gamma_1, P, \Phi_1\}$ and $L_2 = \{\Gamma_2, P, \Phi_2\}$.

In the Book [2] by S. Marcus already referred to, various types of isomorphism of $L_1$ and $L_2$ are introduced as follows :

P $\Phi$-isomorphism : there exists a 1 : 1 mapping f of $\Gamma_1$ onto $\Gamma_2$, such that $P_2(f(x)) = f(P_1(x))$ for each $x \in \Gamma_1$ and such that the string $f(x_1)\, f(x_2) \cdots f(x_n) \in \Phi_2$ if and only if the string $x_1\, x_2 \cdots x_n \in \Phi_1$ $(x_i \in \Gamma_1,\ 1 \leqq i \leqq n)$.

P′ S-isomorphism : there exists a 1 : 1 mapping $g$ of $\Gamma_1$ onto $\Gamma_2$, such that $P'_2(g(x)) = g(P'_1(x))$ and $S_2(g(x)) = g(S_1(x))$ for each $x \in \Gamma_1$, where $S_1$ and $S_2$ are the partitions into families in $L_1$ and $L_2$, respectively.

PP′ S-isomorphism : there exists a 1 : 1 mapping $h$ of $\Gamma_1$ onto $\Gamma_2$, such that $P_2(h(x)) = h(P_1(x))$, $P_2'(h(x)) = h(P_1'(x))$ and $S_2(h(x)) = h(S_1(x))$ for any $x \in \Gamma_1$.

Now we shall also define the new types of isomorphism, as follows.

P′ $\Phi$-isomorphism : there exists a 1 : 1 mapping $r$ of $\Gamma_1$ onto $\Gamma_2$, such that $P_2'(r(x)) = r(P_1'(x))$ for each $x \in \Gamma_1$ and such that the string $r(x_1)\, r(x_2) \cdots (x_n) \in \Phi_2$ if and only if the string $x_1\, x_2 \cdots x_n \in \Phi_1$ $(x_i \in \Gamma_1,\ 1 \leqq i \leqq n)$.

PP′ $\Phi$-isomorphism : there exists a 1 : 1 mapping $v$ of $\Gamma_1$ onto $\Gamma_2$, $P_2(v(x)) = v(P_1(x))$, $P_2'(v(x)) = v(P_1'(x))$ for any $x \in \Gamma_1$, and such that the string $v(x_1)\, v(x_2) \cdots v(x_n) \in \Phi_2$ if and only if the string $x_1\, x_2 \cdots x_n \in \Phi_1$ $(x_i \in \Gamma_1,\ 1 \leqq i \leqq n)$.

Thus we shall have some results on the above described types of isomorphism of $L_1$ and $L_2$. In the first place, we have the following proposition.

Proposition 1. If two languages $L_1$ and $L_2$ are P′ $\Phi$-isomorphic, they are also P′ S-isomorphic.

Proof. Let $y \in S_1(x)$. By definition, for any pair of strings $p$ and $q$ we have either $pxq \in \Phi_1$, $pyq \in \Phi_1$ or $pxq \notin \Phi_1$, $pyq \notin \Phi_1$. Hence, by hypothese, there exists a 1 : 1 mapping $r$ of $\Gamma_1$ onto $\Gamma_2$ such that $P'_2(r(x)) = r(P_1'(x))$ and such that the string $pxq \in \Phi_1$, if and only if the string $r(p)\, r(y)\, r(q) \in \Phi_2$. The latter means $r(y) \in S_2(r(x))$. Thus all of the required conditions of P′ S-isomorphic languages $L_1$ and $L_2$ are fulfilled.

Proposition 2. There exist two P′ S-isomorphic languages $L_1$ and $L_2$ which are not P′ $\Phi$-isomorphic.

Proof. Let $\Gamma_1 = \{a, b, c\} = \Gamma_2$, $P_1 = E = P_2$, $\Phi_1 = \{ab,\ ac\}$, and $\Phi_2 = \{aab,\ aac\}$. Then we have $S_1(a) = P_1'(a) = \{a\} = S_2(a) = P_2'(a)$, $S_1(b) = P'(b) = \{b, c\} = S_2'(b) = P_2'(b)$. By taking as $\varphi$ the identical mapping of $\Gamma_1$, it follows that for any $x \in \Gamma_1$.

$$\varphi(P_1'(x)) = P_2'(\varphi(x)) \text{ and } \varphi(S_1(x)) = S_2(\varphi(x)).$$

Hence $L_1$ and $L_2$ are P′ S-isomorphic. On the other hand, $L_1$ and $L_2$ are not P′ $\Phi$-isomorphic, since the length of each string of $L_1$ is equal to 2, whereas the length of each string of $L_2$ is equal to 3.

Now we shall have the following propositions.

Proposition 3. There exist tow $P'$ $\Phi$-isomorphic languages $L_1$ and $L_2$ which are not $P$ $\Phi$-isomorphic.

Proof. Let $\Gamma_1 = \{a, b, c, d\} = \Gamma_2$, $P_1 = E$, $\Phi_1 = \{aa, ab, ba, cc, cd, dc\} = \Phi_2$, $P_2(a) = \{a, b\}$, $P_2(c) = \{c, d\}$. Then we have $P_1'(a) = \{a, b\}$, $P'(c) = \{c, d\}$, $P_2'(a) = \{a, b\}$, $P_2'(c) = \{c, d\}$. By taking as $\varphi$ the identical mapping of $\Gamma_1$, it follows that for each $x \in \Gamma_1$ $P_2'(\varphi(x)) = \varphi(P_1'(x))$ and the string $x_1 x_2 \in \Phi_1$ if and only if $\varphi(x_1)\varphi(x_2) \in \Phi_2$ for any $x_1, x_2$.

Hence $L_1$ and $L_2$ are $P'$ $\Phi$-isomorphic. On the other hand, $L_1$ and $L_2$ are not $P$ $\Phi$-isomorphic, since $P_1$ is the unit partition E, whereas $P_2 \neq P_1$.

Proposition 4. There exist two $P$ $\Phi$-isomorphic languages $L_1$ and $L_2$ which are not $P'$ $\Phi$-isomorphic.

Proof. Let $\Gamma_1 = \{a, b, c, d\} = \Gamma_2$, $P_1 = E = P_2$, $\Phi_1 = \{ab, ac, c, d\}$ and $\Phi_2 = \{ba, bd, d, c\}$, Then we have $P_1'(a) = \{a\}$, $P'(b) = \{b, c\}$, $P'(d) = \{d\}$, $P_2'(a) = \{a, d\}$, $P_2'(b) = \{b\}$, $P_2'(c) = \{c\}$. Now define a 1 : 1 mapping $\Psi$ of $\Gamma_1$ onto $\Gamma_2$ as follows : $\Psi(a) = b$, $\Psi(b) = a$, $\Psi(c) = d$, $\Psi(d) = c$. Thus using the mapping $\Psi$, we see easily that $L_1$ and $L_2$ are $P$ $\Phi$-isomorphic, but these languages are not $P'$ $\Phi$-isomorphic, since $\Psi(P_1'(x)) = P_2'(\Psi(x))$ for any $x \in \Gamma_1$.

By V. A. Uspenskii. a language is said to be **adequate** if we have $S(x) \subseteq P'(x)$ for any $x \in \Gamma$ (see, [5]). Then we obtain the following proposition.

Proposition 5. If $L_1$ and $L_2$ are $P'$ $\Phi$-isomorphic and $L_1$ is adequate, then $L_2$ is also adequate.

Proof. If $L_1$ and $L_2$ are $P'$ $\Phi$-isomorphic, in view of proposition 1, these languages are $P'$ S-isomorphic. Moreover, since $L_1$ is adequate, by Proposition 57 of [2], L is also adequate.

By S. Marcus (see [4]), a language is said to be **completely adequate**, if for any two words $x$ and $y$ such that $x$ dominates $y$ we have $y \in P'(x)$.

Proposition 6. If $L_1$ and $L_2$ are $P'$ $\Phi$-isomorphic and $L_1$ is completely adequate, $L_2$ is also completely adequate.

Proof. Since $L_1$ is completely adequate, for any pair of the strings $p$ and $q$ we have $y \in P_1'(x)$, $pxq \in \Phi_1$ and $pyq \in \Phi_1$. By hypothese, there exists a 1 : 1 mapping $r$ of $\Gamma_1$ onto $\Gamma_2$ such that $P_2'(r(x)) = r(P_1'(x))$ and such that $r(p) r(x) r(q) \in \Phi_2$ if and only if $pxq \in \Phi_1$. Hence we have $r(y) \in r(P_1'(x)) = P_2'(r(x))$, $r(p) r(x) r(q) \in \Phi_2$, $r(p) r(y) r(q) \in \Phi_2$. That is $L_2$ is completely adequate.

Now we shall have the following statements.

Proposition 7. If $L_1$ and $L_2$ are $PP'$ $\Phi$-isomorphic, these languages are also $PP'$ S-isomorphic.

Proof. This proof follows immediately from both the proof of proposition 1 and the definitions of PP′ $\Phi$-isomorphism and PP′ S-isomorphism.

Proposition 8. There exist two PP′ S-isomorphic languages which are not PP′ $\Phi$-isomorphic.

Proof. Let $\Gamma_1 = \{a, b, c, d\} = \Gamma_2$, $P_1(a) = \{a\} = P_2(a)$, $P_1(b) = \{b\} = P_2(b)$, $P_1(c) = \{c, d\} = P_2(c)$, $\Phi_1 = \{ac, bc, ad, bd\}$ and $\Phi_2 = \{acd, bcd, cab, dab\}$. Then we have $P_1'(a) = \{a, b\} = S_1(a)$, $P'(c) = \{c, d\} = S_1(c)$, $P_2'(a) = \{a, b\} = S_2(a)$ and $P_2'(c) = \{c, d\} = S_2(c)$. Taking for $h$ the identical mapping of $\Gamma_1$, it is easy to see that $L_1$ and $L_2$ are PP′ S-isomorphic. However, these languages are not PP′ $\Phi$-isomorphic, since the length of each string of $L_1$ is equal to 2 whereas the length of each string of $L_2$ is equal to 3.

Proposition 9. If $L_1$ and $L_2$ are PP′ $\Phi$-isomorphic, these languages are also P $\Phi$-isomorphic.

Proof. This proof follows immediately from the definitions of PP′ $\Phi$-isomorphism and P $\Phi$-isomorphism.

Proposition 10. There exist two P $\Phi$-isomorphic languages $L_1$ and $L_2$ which are not PP′ $\Phi$-isomorphic.

Proof. Let $\Gamma_1 = \{a, b, c\}$, $\Gamma_2 = \{x, y, z\}$, $P_1 = E = P_2$, $\Phi_1 = \{ab, ac, c\}$ and $\Phi_2 = \{xy, xz, z\}$. Then we have $P_1'(a) = \{a\}$, $P_1'(b) = \{b, c\}$, $P_2'(x) = \{x, z\}$, and $P_2'(b) = \{y\}$. Now define a 1 : 1 mapping $\varphi$ of $\Gamma_1$ onto $\Gamma_2$ as follows : $\varphi(a) = y$, $\varphi(b) = x$, $\varphi(c) = z$. Thus using the mapping $\varphi$, it is easy to see that $L_1$ and $L_2$ are P $\Phi$-isomorphic, but these languages are not PP′ $\Phi$-isomorphic, since $\varphi(P_1'(x)) \neq P_2'(\varphi(x))$ for any $x \varepsilon \Gamma_1$.

Proposition 11, If $L_1$ and $L_2$ are PP′ $\Phi$-isomorphic, theses languages are also P′ $\Phi$-isomorphic.

Proof. This proof follows immediately from the definitions of PP′ $\Phi$-isomorphism and P′ $\Phi$-isomorphism.

Proposition 12. There exist two P′ $\Phi$-isomorphic languages $L_1$ and $L_2$ which are not PP′ $\Phi$-isomorphic.

Proof. Let $\Gamma_1 = \{a, b, c\} = \Gamma_2$, $P_1 = E$, $P_2(a) = \{a\}$, $P_2(b) = \{b, c\}$ and $\Phi_1 = \{ab, ac, aa\} = \Phi_2$. Then we have $P_1'(a) = \{a\}$, $P_1'(b) = \{b, c\}$, $P_2'(a) = \{a\}$ and $P_2'(b) = \{b, c\}$. Now by taking as $\varphi$ the identical mapping of $\Gamma_1$, it follows that $L_1$ and $L_2$ are P $\Phi$-isomorphic. However, these languages are not PP′ $\Phi$-isomorphic, since $P_1$ is the unit partition E, whereas $P_2 \neq P_1$.

## References


[1] Kulagina, O S. ; On one method of defining grammatical notions on the basis of set theory (in Russian). Probl. Kibernetiki 1, 203−214.

[2] Marcus, S. ; Algebraic Linguistics ; Analytical Models, 1967, Academic Press, New York and London.

[3] Marcus, S, ; Sur un modèle logique de la catégorie grammaticale élémentaire. I, Revue de math. pures et appl., Acad. R. P. R., 1962, 7, 1, 91−107.

[4] Marcus, S. ; Sur un modèle logique de la catégorie grammaticale élémentaire II. Zeitschr. f. math. Logik und Grundlagen d. Math. Bd. 8. S. 323−329 (1962).

[5] Uspenskii, V. A. ; On defining the part of speech in the set-theoretic system of language (in Russian). Byull. Bed. Probl. Masinngo Perevoda. No. 5, (1957).