

1

Introduction

Zygmunt Frajzyngier and Erin Shay

1.1 The goal of the present work

Afroasiatic is the fourth largest linguistic phylum in the world, with about 375 living languages spoken by some 300 million speakers (www.ethnologue.org). In the view of the contributors to this volume, this number may well be an overestimation. For the Chadic family alone Ethnologue estimates over 190 languages, while most linguists working in the area estimate the number to be between 140 and 160 languages. The differences boil down to decisions regarding what is a language and what is a dialect. Given the absence of clear-cut criteria for this distinction we leave the question of the actual number of languages open.

This book provides the first-ever typological survey of each of the language families belonging to the Afroasiatic phylum as well as a typological outline of the entire phylum.

The book is addressed to a general linguistic audience, some of whom may be unfamiliar with Afroasiatic linguistics, as well as to linguists who have worked on Afroasiatic languages and would like information about languages from other branches and about the characteristics of the whole phylum.

The approach taken in this book is typological rather than historical, taking for granted the existence of the Afroasiatic family, as confirmed by comparative historical studies. This is because a number of phenomena in contemporary languages can best be explained by the internal structure of the grammatical systems rather than by their origin in an ancestral proto-language. There exist a fair number of studies, chief among them Diakonoff (1988), that take a historical rather than a typological approach to selected issues.

We adhere to the common usage in referring to Afroasiatic as a phylum, rather than a family, on the grounds of the remoteness of the relationships among its various branches (Diakonoff 1988). We also take it as given that the phylum is composed of six language families (from west to east), namely Berber, Chadic, Egyptian, Cushitic, Omotic, and Semitic, even though the internal structure of the phylum remains somewhat

controversial and further internal subgroupings within the phylum cannot be ruled out (for recent proposals based on different criteria, see Diakonoff (1998), Zaborski (2005a), and Ehret (2005)).

The Afroasiatic languages are an important object of study, not only because of their widespread usage but also because of their great typological diversity. It has been claimed that the Chadic family alone is more typologically diverse than the entire Indo-European language family (Diakonoff 1988). The Afroasiatic phylum exhibits great variation with respect to traditional typological criteria such as the position of the predicate in the clause, the structure of the noun phrase, and the structure of the verb. As a result, the question ‘What is a typical Afroasiatic language?’ cannot, at this stage, be answered. The material in this book is intended, among other things, to document this typological diversity.

The Afroasiatic languages are also an important object of study because the languages and the cultures they embody have been instrumental in shaping Judeo-Christian culture, Islamic culture, and much of what has come to be referred to as Western civilization. Speakers of Semitic languages, which belong to the Afroasiatic phylum, developed the alphabetic writing system which, with numerous modifications, is now used in thousands of languages throughout the world. The development of the alphabetical writing system may have been facilitated by the underlying structure of verbal roots and derived nominal forms in Semitic languages, where the consonantal structure alone conveyed a great deal of semantic information.

While some Afroasiatic languages are widely spoken and robust, many languages of the phylum are endangered and may disappear within a few generations. Although a decline in the use of a given language is usually a result of various social forces, sometimes including speakers’ choice, such a loss also means the loss of the most complex intellectual product of those who speak the language. We hope that this book will be a stimulus and a useful tool for scholars to undertake the task of working on hitherto-undescribed or under-described languages.

Although the focus of the book is typological, individual chapters also provide information on the history of the language family; geographical distribution; historical writing systems, if any; and, in some instances, diachronic changes within the family. While the information included in a single volume cannot be exhaustive, we hope that it will serve as a starting point for a more extensive and intensive typological, and eventually historical, study of the families composing the Afroasiatic phylum. The book includes a bibliography of sources and materials for further reading. Since the scholarship on many languages is new, and since there is no agreed-upon standard for many language names, the spelling of language names on the maps and in various chapters may differ.

1.2 History of the recognition of the phylum

The term 'Afroasiatic' was coined by Delafosse in 1914 (cited in Newman 1980) and was reintroduced by Greenberg in 1960. The term captures the fact that this is the only phylum whose member families include languages spoken in Africa and languages spoken in Asia. The phylum has also been called Hamito-Semitic (since F. Müller 1876), Semito-Hamitic (chiefly in the older Russian sources), Afrasian (Diakonoff 1988), Erythraean (Tucker and Bryan 1966; Tucker 1975), and Lisramic (Hodge 1972). The term 'Afrasian' is an Anglicization of the Russian *afrazijskije*, a variation on the term 'Afroasiatic'. The term 'Erythraean' refers to a core geographical area of the family. The term 'Lisramic' is based on the Proto-Asiatic **lis* 'language' and the Proto-Egyptian **rāmāč* 'people'. Appellations for the phylum have been the object of vigorous discussion among linguists, and a special session of the Hamito-Semitic conference held in London (Bynon and Bynon 1975) was devoted to naming the phylum. In contemporary writing by various scholars, the most frequently used terms are 'Afroasiatic', 'Hamito-Semitic', and 'Semito-Hamitic' (see later sections concerning the history of the phylum).

1.3 Evidence for genetic relationships within the phylum

The typical evidence for genetic relationships within the phylum includes numerous comparative word lists showing etymologies across the Afroasiatic families. Some of these lists propose sound correspondences, while others simply provide the presumed cognates. The first of such comparisons was Marcel Cohen (1947), followed by Greenberg (1963), Hodge (1966, 1967), and a number of more recent studies dealing with the whole phylum, two or more families of the phylum, or a single family within the phylum. These include Skinner (1984); Belova *et al.* (1994–7), representing the work of Diakonoff's team; Orel and Stolbova (1995); Naït-Zerrad (1998); Takács (2005 and other works); Ehret (1995); Dolgopolsky (1999); Militarev and Kogan (2000); and Rössler's and Vycichl's numerous studies on Egyptian–Semitic relations. Militarev (2000) uses glottochronology as a means of calculating when the phylum split into various families. The largest of the comparative studies are Belova *et al.* (1994–7), Orel and Stolbova (1995), and Ehret (1995). The larger etymological studies have been criticized for the choice of items taken for comparison and often for the validity of postulated cognates. The cumulative effect of these studies, that of reconfirming the genetic unity of the phylum, is not in doubt.

Another piece of evidence for the genetic relationship of Afroasiatic languages comes from morphology. Across the phylum there are morphemes with similar phonological structures and similar functions. Many of these have long been known for their

occurrence in Egyptian, Semitic, Berber, and Cushitic languages, but it was Greenberg (1963) who demonstrated that the same morphemes also occur in various Chadic languages. Hodge (1969a, first presented in 1965) describes the evolution (what we would call today the ‘grammaticalization’) of determiners in Afroasiatic languages in a way that leaves no doubt as to the genetic relationship of the phylum. David Cohen (2005: 17ff.) provides extensive evidence for the alternation between *a* and non-*a* vowels in the verbal systems of Afroasiatic languages. This is another piece of evidence for the genetic relationship of the six families.

Sasse (1984a) and Blažek (2006) are devoted to the study of case and (mainly locative) prepositions. These studies are driven by the aims of historical linguistics and do not deal with functions of the reconstructed elements within the grammatical systems of the various languages.

Some linguists have claimed that there is a genetic relationship between Afroasiatic and Indo-European languages. Hodge, who called the proposed super-phylum ‘Lislahk’, argued for this relationship in a number of publications (Hodge 1978, 1979, 1981). Proponents of Nostratic theory (Dolgopolsky 1998) include Afroasiatic as a member of the Nostratic family. The Nostratic hypothesis is highly controversial and has very few supporters among specialists in Afroasiatic languages.

Debate as to the internal division within the phylum involves the status of Omotic as a separate family and the question of whether there may be further subdivision within the phylum. With respect to Omotic, the question is whether it is a separate family or whether it should be incorporated within the Cushitic family. The history of Omotic as a family within the Afroasiatic phylum is described in detail in chapter 7 of this volume.

1.4 A snapshot of the history of scholarship

The awareness of relationships among languages within Afroasiatic goes back at least to the ninth century, when Judah ben Quraysh of Morocco, a physician to the emir of Fez, wrote of lexical and phonological similarities between Berber and the Semitic languages Biblical Hebrew, Aramaic, and Arabic (Becker 1984, reviewed by Wansbrough 1986). Hayward (2000) reports that the French orientalist Postel (1538) also pointed out resemblances among Hebrew, Arabic, and Aramaic. To these languages, Ludolf (1702) added Amharic and Ge’ez (Hayward 2000). In 1781, von Schläözer gave the grouping the name ‘Semitic’, based on the biblical Sem, son of Noah (Genesis 5:32). Müller (1876) followed the pattern in naming the Hamitic branch, assumed at that time to consist of Egyptian and Berber. Müller also created the term ‘Hamito-Semitic’ for the larger language family, reflecting the assumption that the phylum could be split into two branches, the Hamitic languages and the Semitic languages. The selection of languages in Meinhof’s 1912 *Die Sprachen der Hamiten* was based on a mixture of

anthropological and linguistic typological criteria and included languages that are now not part of the Afroasiatic family.

Marcel Cohen (1924) was the first to reject a division of the phylum into Hamitic and Semitic branches. He stated emphatically that there is no trait shared by the Libico-Berber (now Berber), Cushitic, and Egyptian languages that would group them together and set them apart from the Semitic languages. However, he retained the term 'Hamito-Semitic' as a purely conventional label. A few linguists still interpret the term in the sense in which it was originally coined, as implying two branches within the phylum. Diakonoff (1998) points out that the use of the term 'Hamito-Semitic' by Orel and Stolbova (1995) wrongly implies a division into two branches. The term 'Hamito-Semitic' is still used in French, Italian, Russian, German, and English writings.

According to Sasse (1981a: 132), some of the languages now classified as Cushitic, such as Beja, Somali, Galla, and Harari, were considered, as of the mid nineteenth century, to belong to the same family as Egyptian, Semitic, and Berber. Sasse cites Lepsius (1844), Beke (1845), d'Abbadie (1845), and Lottner (1860–1) as among those who speculated about the existence of the larger linguistic family. We may add to this list Burton (1856), who stated that 'the Harari appears, like the Galla, the Dankali, and the Somali, its sisters, to be a Semitic graft inserted into an indigenous stock' (Burton 1987 (1856): 153).

The Chadic family was the last to be added to the phylum. Marcel Cohen (1924) did not include Chadic languages in his study of the Hamito-Semitic languages. In Cohen (1947), a comparative study of 500 lexical items, he does include Hausa along with Egyptian, Semitic, Berber, and Cushitic. Although the inclusion of Hausa in the comparative study may be construed as a tentative inclusion of Chadic within the phylum, Cohen does not mention Chadic as a family. As late as 1970 he was reluctant to recognize Chadic as a member of the Afroasiatic phylum: 'Si le tchadien doit réellement nous être adjoint (je crois qu'il doit nous être adjoint mais non pas incorporé) la question se pose aussi: comment le tchadien s'est-il formé?' (Cohen 1970 (1937): 24). The unequivocal inclusion of Chadic in the Afroasiatic phylum is due to Greenberg (1950b).

Fleming (1969, 1974) proposed placing a group of languages previously classified as West Cushitic in a separate branch of Afroasiatic, which he called the Omotic branch. This separation has been accepted by some but rejected by others.

Inclusion of Omotic in the Cushitic family, and by implication in the Afroasiatic phylum, is strongly supported by Cerulli's (1951) study of Kafa (once considered part of West Cushitic, now classified as Omotic). Cerulli provides numerous regular phonological and morphological correspondences between Kafa and the Cushitic languages of the western Sidamo province as well as the Central Cushitic Agau. As

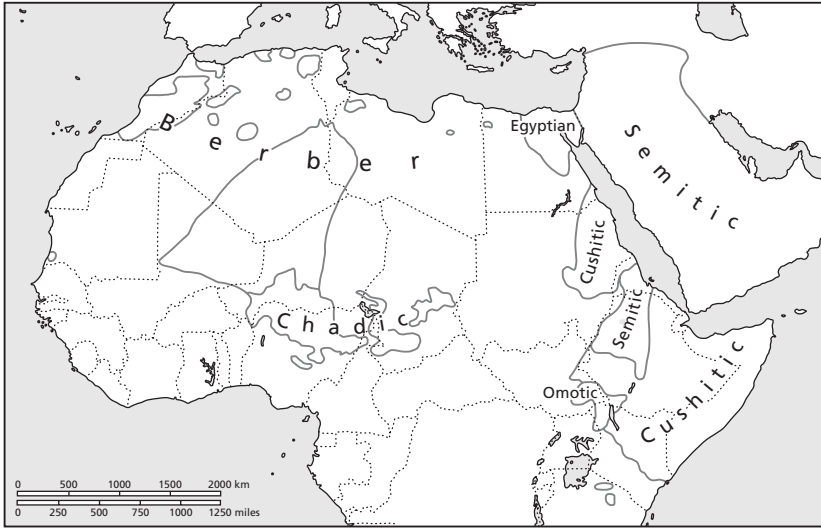
described in chapter 7 in the present volume, the position of these languages within the Afroasiatic phylum remains one of the most controversial issues in Afroasiatic classification. The discipline still lacks a systematic study of regular sound correspondences and of common retentions and innovations that would allow subclassification.

For succinct histories of the concept of the Afroasiatic phylum, the reader is referred to Newman (1980), Diakonoff (1988), and Hayward (2000).

1.5 Geographical range of the Afroasiatic phylum

Afroasiatic languages are spoken in Northern Africa, Central Africa, the Horn of Africa, the Arabian Peninsula, and even in Central Asia (Arabic). Berber languages are spoken in isolated pockets in Mauritania and in Morocco, Algeria, Mali, Niger, Libya, Tunisia, Burkina Faso, as well as in the Siwa oasis in Egypt. The Chadic family, the largest of the phylum, comprises between 140 and 160 languages (estimates vary) spoken in northern Nigeria, southern Niger, northern Cameroon, and southern Chad Republic. Hausa, the Chadic language with the greatest number of speakers, is a vehicular language in West Africa and the official language of Nigeria, and there are pockets of Hausa speakers to be found over large areas of West and Central Africa. Egyptian was the language of ancient Egypt, and its descendant, Coptic, remains the liturgical language of the Coptic Church in Egypt. Cushitic languages are spoken in Eritrea, Ethiopia, Somalia, and northern Kenya, Sudan (Beja) and in isolated pockets in Tanzania. Some Cushitic languages are official in different federal regions of Ethiopia. Somali, also a member of the Cushitic family, is the official language of the Somali Republic. Omotic languages are spoken in southwest Ethiopia. Among the Semitic languages, Hebrew is one of the official languages of Israel, Amharic is one of the official languages of Ethiopia, and Tigre and Tigrinya are the official languages of Eritrea. Arabic, spoken throughout North Africa, the Arabian Peninsula, and in areas outside of Africa, is the official language, or one of the official languages, of Mauritania, Morocco, Algeria, Tunisia, Libya, Egypt, and Chad in Africa, and of Israel, Lebanon, Jordan, Syria, Iraq, Saudi Arabia, United Arab Emirates, Oman, Bahrain, Qatar, and Yemen in the Arabian Peninsula, and in other countries with significant Arabic diaspora. For a map of dialects of Arabic, see Kaye and Rosenhouse (1997: 264). Maltese (Semitic) is one of the official languages of Malta. Despite the wide geographic range of the families of the phylum, and the fact that a number of those languages are the official languages in various countries, many Afroasiatic languages are threatened with extinction because they are spoken by a small number of people in economically and politically unstable environments.

All languages of the Afroasiatic phylum have had extended contact with other Afroasiatic languages and with languages belonging to other families. The Semitic languages



Map 1.1 *Afroasiatic phylum*

of Ethiopia (Ethiosemitic) have been in contact with Cushitic and Omotic languages, and some of the Ethiosemitic languages share a number of typological features with Cushitic and Omotic languages that they do not share with other Afroasiatic languages. Akkadian, an East Semitic language, has been in contact with Sumerian, and North-Eastern Neo-Aramaic has been in extensive contact with Kurdish. Berber languages have been in contact with Arabic and Chadic languages and with Nilo-Saharan languages. Chadic languages have been in contact with Niger-Congo, Nilo-Saharan, Semitic, and Berber languages. Such contacts have no doubt induced changes in Afroasiatic languages. The clause-final position of the verb in Amharic and Tigrinya (Ethiosemitic) languages, for example, is attributed to contact with Cushitic and Omotic languages, and the clause-final position of the verb in Akkadian is attributed to contact with Sumerian. Given the absence of systematic phonological, morphological, and syntactic reconstructions, in most instances we are unable to state categorically which typological features are due to language contact.

1.6 State of the art in Afroasiatic scholarship

A few Afroasiatic languages, including Egyptian, Hebrew, Arabic, and Aramaic, have been the objects of study for more than 200 years, and the amount of literature on these languages is very large indeed. As might be expected, the longer the history of scholarship on a given language or family, the more publications are available. Yet even

for language families with the longest scholarly tradition there are fundamental gaps in the scholarship. Izre'el (2002) and Khan (2002) acknowledge the fact that traditional Semitic scholarship seldom dealt with syntax. Even such fundamental components of grammar as the aspectual and tense systems in Semitic languages remain poorly studied (Izre'el 2002). For the state of the art in Semitic studies, see Izre'el (2002) and chapter 4 of the present volume.

Tosco (1994a) and (1994b) represent typological studies of the syntax of East Cushitic languages. Dolgopolsky (1973) is a massive reconstruction of lexical roots in Cushitic languages. The languages classified now as Omotic are included in Dolgopolsky's study as West Cushitic. Zaborski (1975) is a study of the verbal forms in Cushitic languages. Bender (2000) is a comparative study of Omotic morphology. For a compendium of literature on individual families, the reader is referred to the bibliography at the end of the volume.

Very little scholarship has been devoted to the typology of the Afroasiatic phylum as a whole. The most recent and most complete surveys of Afroasiatic languages are Diakonoff (1988) and Petráček (1989). The latter, a two-volume publication in Czech, is a textbook for students of Semitic, Egyptian, and African languages. The work is conceived as an exhaustive review of the literature on the Afroasiatic phylum and individual families; the history of classification; theoretical issues in genetic classification; sociolinguistic issues; and areal linguistics. The list of references, which ends with 1987, takes up eighty pages. Only ten pages of the two-volume work are devoted to typology, the focus of the present volume.

With the exception of Egyptian and Semitic, there have been no typological studies of any single branch of the Afroasiatic phylum. For most of the branches, and for the phylum itself, there have been no comparative studies of phonological processes, of the syntax of simple or complex sentences, of semantic categories encoded, of reference systems, or of any of the other functional domains that comprise a complete grammar of a language.

Hodge (1971, a reprint of Hodge 1970) is a collection of essays on various Afroasiatic families. These essays, some written by the most eminent scholars of the time, are very brief. Most list the important references for the family surveyed and provide some information about the phonological system and bits and pieces of morphology. None of the essays offers a picture of the grammatical system of any of the families surveyed. Since Hodge (1971), there have been many collections of papers published on Afroasiatic linguistics, e.g. Perrot *et al.* (1981), Lecarme *et al.* (2000), Zaborski (2001), Bender *et al.* (2003), Lecarme (2003), Fronzaroli and Marassini (2005), Lonnet and Mettouchi (2005), and Mettouchi and Lonnet (2006). Studies in these volumes are devoted to single topics in individual languages and do not pretend to offer a survey of any of the families, much less of the entire phylum. The only

work that attempts to look at a large number of issues in Afroasiatic languages remains Diakonoff (1988).

1.6.1 Phonological reconstruction

Most of the work conducted so far on Proto-Afroasiatic reconstruction has dealt with the sound inventory and the lexicon. Diakonoff and his collaborators have worked for many years on reconstructing the Afroasiatic consonants, vowels, tones, and vocabulary. Their results have been published in instalments, first in Russian and then in English (Diakonoff *et al.* 1992; Belova *et al.* 1993; Belova *et al.* 1994–97). Diakonoff (1988) reconstructed a consonantal system involving four manners of articulation for obstruents: voiceless, emphatic, and voiced stops and voiceless continuants. He postulated labial, dental, and palatalized fricatives and affricates, and labial, velar, labialized-velar, post-velar, labialized post-velar, pharyngeal, and laryngeal places of articulation. Diakonoff also postulates two nasals, *m* and *n*; two liquids, *r* and *l*; and palatal and labial glides. He does not include the prenasalized stop *mb*, posited in Greenberg (1965) as a Proto-Afroasiatic phoneme.

Orel and Stolbova (1995) is an attempt to reconstruct 2,672 lexical items. The book is subtitled ‘Materials for a Reconstruction’. The authors postulate a Proto-Afroasiatic consonantal system consisting of voiced, voiceless, and emphatic stops and fricatives. They also postulate seven places of articulation: labial, dental, lateral, velar, post-velar, pharyngeal, and laryngeal. Unlike Diakonoff, Orel and Stolbova do not postulate the labialized velar consonants *k^w*, *g^w*, and *q^w*. They postulate a six-vowel system consisting of *i*, *ü*, *e*, *a*, *o*, and *u*. The Orel and Stolbova reconstruction focuses on lexical items rather than on the phonological system. It does not deal with constraints on syllable structure or with phonological processes such as vowel or consonant harmony.

Ehret (1995) reconstructs about forty consonantal phonemes, with three manners of articulation (voiceless, voiced, and emphatic). As places of articulation he posits labial, dental, velar and labiovelar; alveolar and palatal; nasal; and laryngeal. He also postulates glides, an *r*, and four laterals: *l*, *dl* (corresponding to the *ḫ* of Jungrathmayr and Shimizu’s Proto-Chadic reconstruction (1981)), *tl’*, and *ʃ*. He retains the labial consonants postulated by Greenberg (1958) but explicitly rejects the notion, advocated in Greenberg (1965), of prenasalized stops in Proto-Afroasiatic. Ehret further postulates, albeit tentatively, the existence of tone in Proto-Afroasiatic, basing his conclusion on the analysis of tones in Ngizim (West Chadic) and Mocha (Omotic) and on the fact that tones are attested in Cushitic, Omotic, and Chadic languages. Unlike Orel and Stolbova, Ehret takes phonological constraints into consideration. He takes the emergence of phonological constraints as evidence for innovations that are the basis of his subclassification.

With respect to the consonantal system, most reconstructions agree that Proto-Afroasiatic had three series of obstruents and that the only continuants were voiceless. Note, however, Ehret's reconstruction of the voiced lateral continuant. Diakonoff and his associates postulate a series of labiovelar consonants, while Orel and Stolbova (1995) claim that labial velars derive from 'velar consonants followed by the sequence *au*'. While such sequences often result in a labiovelar and eventually a labial stop (Frajzyngier 1989a), there are also labial velar consonants that cannot be explained as deriving from the sequence velar-*a-u*. The evidence that some labial velar stops may, in fact, be underlying is provided by languages where such stops occur in word-final position and where there is no rule of final *a* deletion. This is the case in Hdi (Central Chadic), e.g., the noun *màrkw* 'wife' (Frajzyngier with Shay 2002).

The posited reconstructions show much greater variation with respect to the number of vowels in Proto-Afroasiatic. Diakonoff (1980) considers the possibility that there were only two vowels, the low vowel *a* and a high vowel realized as *i*, *u* or a central vowel. Ehret reconstructs five short and five long vowels for Proto-Afroasiatic: *a*, *aa*, *e*, *ee*, *i*, *ii*, *o*, *oo*, *u*, and *uu*. For a critique of reconstructions in Orel and Stolbova (1995), see Diakonoff (1998), and for a critique of methodologies in Orel and Stolbova (1995) and Ehret (1995), see Ratcliffe (n.d.).

1.6.2 Reconstruction of morphology and syntax

Little work has been done on reconstructing the Proto-Afroasiatic morphological system, and most of this work has been devoted to the morphology of the verb (Diakonoff 1988). Greenberg (1955) described the use of *a*, in the place of a different vowel, as the marker of nominal plurality, and also described the suffix *-en* as a marker of plurality. Greenberg (1952) and (1953) dealt with verbal forms involving gemination, which he analysed as coding present tense. Greenberg (1960) dealt with gender and number agreement, citing *t* as the Afroasiatic marker of feminine gender and citing a contrast between *k* as the marker of masculine gender and *t* as the feminine marker. Greenberg also stresses similarities among pronominal systems, in particular possessive pronouns, among various families of the phylum. Greenberg (1963), which presented a revised classification of African languages, posited a number of Proto-Afroasiatic grammatical morphemes, including pronouns; the causative *-s* (although subsequent studies have shown that this suffix does not occur in Chadic (Frajzyngier 1985a)); the prefix *m-* as a marker of place names, instrument, and agent; and the internal *a*-plural. The internal *a*-plurals occurring in all Afroasiatic languages constitute a strong morphological argument in favour of a genetic relationship among Afroasiatic languages. Greenberg also postulates a variety of other morphemes, all of which have been shown by subsequent studies to be the product of grammaticalization processes whose original source was the same as that posited