

## WHAT IS A *KALIMA*? 'ASTARĀBĀDĪ'S ANSWER\*

Pierre Larcher

### INTRODUCTION

The choice of the present subject undoubtedly derives from the questions aroused in the occasion of this first colloquium of Arabic linguistics. However, the question is not new to me. In fact, it dates back to my days as a PhD student, when I acquainted myself with the writings of a currently well known, at that time less if not ill-known, Arab grammarian. I am obviously referring to Raḍī ad-dīn al-'Astarābādī the author of the *Šarḥ al-Kāfiya* and of the *Šarḥ aš-Šāfiya*, two great commentaries on two compendiums written by another author, Ibn Ḥājjib, namely *al-Kāfiya fi n-naḥw* and *aš-Šāfiya fi š-šarf*. The first deals with syntax and the second with morphology. The first constitutes the core of the present study. I will quote it from the Istanbul edition of 1310H, for a long time the only one available (there exist at least three recent editions today). I will not translate *kalima*, but limit myself to reminding that this word is both the singulative (*ism al-waḥda*) of the collective *kalim* and the singular of the plural *kalimāt*.

#### 1. THE *KALIMA*: NEITHER A WORD, NOR A MORPHEME

At the beginning of the *Šarḥ al-Kāfiya*, Raḍī d-dīn al-'Astarābādī begins with the definition of the *kalām* given by Ibn Ḥājjib in the *Kāfiya*:

An utterance is everything that contains two *kalima* as a result of the predicative connection. This only occurs in the case of two nouns or in the case of a verb and a noun.<sup>1</sup>

which he comments in these terms:

---

\* I kindly thank Professor Giuliano Lancioni for his invitation, which obliged me to resume an old work.

<sup>1</sup> 'Astarābādī (*ŠKāfiya*:1.7): *al-kalām mā taḍammāna kalimatayn bil-'isnād wa-lā yata'attā dālīka 'illā fi smayni 'aw fi fi'l wa-ism.*

By “which contains two *kalima*”, we mean that the utterance is composed by these two or that these constitute its two parts (or elements).<sup>2</sup>

by adding (I will only quote what is relevant to the scope of this paper):

[...] and the two parts of the utterance can be either explicit as in *Zaydun qā'imun* and *qāma Zaydun*, or implicit, [...] or one of the two can be implicit but not the other. The implicit one can be either the verb [...] or the subject, as in *Zaydun qāma*.<sup>3</sup>

It is clear that for Raḍī d-dīn al-'Astarābādī both:

- (1) *Zaydun qā'imun*  
 Zayd standing  
 'Zayd is standing'

and:

- (2) *qāma Zaydun*  
 stood Zayd  
 'Zayd stood'

are utterances (*kalām*), each of them constituted of two *kalima*, i.e. two nouns in (1), a verb and a noun in (2).

However, at the beginning of the third and last part of his work, devoted to the particle (*ḥarf*: this constitutes the third of the three kinds of *kalima* recognised by the Arabic grammatical tradition, after the noun and the verb) he gives as an example:

- (3) = (1) *Zaydun qā'imun*

and:

- (4) *qad qāma Zaydun*

<sup>2</sup> In 'Astarābādī (*ŠKāfiya*: 1.7): *wa-na'nī bi-taḍammunihi l-kalimatayn tarakubbahu minhumā 'aw kawnahumā juz' ayyhi*.

<sup>3</sup> *Wa-juz'ā l-kalām yakūnāni malfūḍayn ka-Zaydun qā'imun wa-qāma Zaydun wa-muqadarrayn [...] 'aw 'aḥaduhumā muqaddar dūna l-'āxar wa-huwa 'immā fi'l [...] 'aw fā'il kamā fi Zaydun qāma*.

indicating that “each of the two abovementioned utterances is composed of four *kalimāt*”,<sup>4</sup> i.e. the *tanwīn* (-n) which is added to *Zayd* and *qā'im* in (3), the particle *qad* which comes before the verb, and the *tanwīn* which is added to *Zayd* in (4).

At this point, a first conclusion can be drawn: a *kalima* is not a word. There would actually be two words in (1) and in (3), two in (2) but three in (4). A *kalima* is not even a morpheme. If by morpheme we mean a minimal meaningful unit, we would at least count six morphemes in (1) and (3), i.e.:

(1') = (3') *Zayd-u-n qā'im-u-n*

and at least four in (2), i.e.:

(2') *qāma Zayd-u-n*

and five in (4), i.e.:

(4') *qad qāma Zayd-u-n*

where, besides the *tanwīn*, we would count as a morpheme the vowel -u- which is the mark of the nominative case. As for Western scholars of Arabic, most of them would see in *qā'imun* and *qāma* the association of two morphemes: a lexical one (to which correspond the “consonantal root”  $\sqrt{qwm}$  and the meaning ‘to stand up’), and a grammatical one (to which correspond the patterns *fā'il* and *fā'ala* and the meanings ‘active participle’ and ‘3rd person m.s. of the perfect’). This would bring the total number of morphemes to seven in (1) and (3), five in (2) and six in (4)...

If a *kalima* is neither a word nor a morpheme, what is it then? The answer to this question is uncertain, because we are faced with two possibilities:

1. either Raḍī d-dīn al-'Astarābādī contradicts himself
2. or he does not contradict himself, in which case we have to find an interpretation of *kalima* compatible with the double counting of *kalimāt* in (1) that the author does, respectively two and four.

<sup>4</sup> 2.319: *kull wāḥid min al-kalāmayn al-maḍkūrayn murakkab min 'arba' kalimāt.*

2. WHAT IF, DESPITE EVERYTHING, THE *KALIMA* WAS A MORPHEME?2.1. *Morphemes Belonging to the Non-Concatenative Morphology*

If a *kalima* is not a word, it *might* be a morpheme. This comes directly from the definition of *kalima* given by Ibn Ḥāḥib as “an expression instituted for a single meaning”.<sup>5</sup> One expression, one meaning: what we have here seems to be a *kalima* as a minimal meaningful unit, i.e. what modern Western—mainly American—linguistics calls a morpheme. Unfortunately, Ḥāḥib’s commentary of the terms “instituted for a single meaning” does not confirm this interpretation:

He means by that the meaning, of which one part is not signified by any part of the expression, whether this meaning has a part (like the meaning of *ḍaraba* which means the *maṣḍar* and the tense) or whether it does not, like the meaning of *ḍarb*.<sup>6</sup>

In other words, an expression can be single, as *ḍaraba* and *ḍarb*, each counting for one and only one *kalima*, but its meaning can be either single, as in *ḍarb*, or multiple, as in *ḍaraba*, which adds to the meaning of *ḍarb* that of the past tense...

Nevertheless, at this point, one could perfectly maintain that *kalima* is a morpheme, in the sense of the American distributionalist linguistics, which “conceives the meaningful unit as a segment of the speech chain, vehicle of a meaning which is exterior to it” (Ducrot & Schaeffer 1995: 360). In such a conception, it is possible that “an unanalysable phonic element can bear many clearly distinct meanings”. This very fact leads to distinguish between morpheme as an abstract unit, and morph as a concrete segment and consequently, to introduce the concept of *allomorph*, when to the same abstract morpheme correspond several concrete morphs, and of *portmanteau morph*, when to the one and same morph correspond several morphemes. This would be the case of *ḍaraba*, counted as **one** *kalima*, even if we recognise **two** meanings in it.

Is *ḍaraba* really an “unanalysable phonic element”? Ḥāḥib goes back to the verb in the past tense. He writes:

<sup>5</sup> In Ḥāḥib (ŠKāfiya: 1.2): *lafḍ mawḍūʿ li-māʿnā mufrad*.

<sup>6</sup> Ḥāḥib (ŠKāfiya: 1.3): *ya nī bihi al-māʿnā allaḍī lā yadullu juzʿ lafḍihi ʿalā juzʿihi sawāʿ kāna li-ḍālika al-māʿnā juzʿ naḥw maʿnā ḍaraba ad-dāll ʿalā l-maṣḍar waz-zaman ʿaw lā juzʿ lahu ka-māʿnā ḍarb*.

As for the verb in the past tense, like *ḍaraba*, this is an interesting point: in fact, it is undoubtedly a *kalima*, though the process is signified by its ordered consonants and the assertion of this process in the past is signified by its form, which affects its consonants. The form is a part of the expression, since it represents the number of the consonants, including the whole set of the short vowels and the lack of short vowels; this set has been the object of a specific institution. Vowels are a part of what is uttered. Thus the verb in the past tense is a *kalima* made up of two parts, each indicating a part of its meaning.<sup>7</sup>

It is then clear that for 'Astarābādī, in *ḍaraba* there are not only two meanings, but to each of these meanings corresponds a part of the expression: the "ordered consonants", i.e. the consonantal root  $\sqrt{ḍrb}$  for *ḍarb* and the whole represented by the number of vocalised or non-vocalised consonants, i.e. the pattern (*wazn*) R1aR2aR3 for the past tense.<sup>8</sup>

'Astarābādī adds that what is true for *ḍaraba* is also true for the plural 'usud 'some lions' as opposed to 'asad 'a lion' and for the diminutive ['usayd 'a lion cub'],<sup>9</sup> for the plurals *masājid* 'some mosques' and *rijāl* 'some men', for *ḍārib* 'striking', *maḍrūb* 'stricken' and *miḍrab* 'a striking tool':

because the diminutive, the plural, the agent, the patient and the tool in the abovementioned examples are signified by the short vowels occurring with the added consonant and/or the long vowel.<sup>10</sup>

He concludes:

It is not correct to claim here that the incidental form is a *kalima* which, as a result of the combination, has become as a part of another *kalima* [...]. These *kalim* constitute therefore a relevant objection, unless one limits the explanation of "complex expression" and says "it is something a part of the expression of which indicates a part of its meaning". In such an expression,

<sup>7</sup> 'Astarābādī (*ŠKāfiya*: 1.5–6): 'ammā l-fi' l-al-māḍī nahw ḍaraba fa-fihī naḍar li-'annahū kalima bi-lā xilāf ma'a 'anna al-ḥadaṭ madlūl ḥurūfihī al-murattaba wal-'ixbār 'an ḥuṣūl ḍālika l-ḥadaṭ fi z-zaman al-māḍī madlūl waznihi aṭ-ṭāri' 'alā ḥurūfihī wal-wazn juz' al-laḥḍ 'id huwa' 'ibāra 'an 'adad al-ḥurūf ma'a majmū' as-sakanāt al-mawḍū'a waḍ'an mu'ayyanan wal-ḥarakāt mim mā yutalaffaḍ bihi fa-huwa 'iḍan kalima murakkaba min juz'ayn yadullu kullu wāḥid minhumā 'alā juz' ma'anāhu.

<sup>8</sup> At first sight, 'Astarābādī's analysis is not so different from that of the Western scholars of Arabic. Actually, it deviates from it on a crucial point: the lexical meaning 'to strike' is not directly related to the consonantal root  $\sqrt{ḍrb}$ . It is only indirectly attached to it since it represents, in the derived word, (here the perfect *ḍaraba*), the base it derives from, i.e. the *maṣdar ḍarb*.

<sup>9</sup> 'Astarābādī does not quote the word itself.

<sup>10</sup> 'Astarābādī (*ŠKāfiya*: 1.6): li-'anna d-dāll 'alā ma'nā at-taṣṣūr wal-jam' wal-fā'il wal-maf'ūl wal-'āla fi l-'amṭila al-maḍkūra al-ḥarakāt aṭ-ṭāri'a ma'a al-ḥarf az-zā'id.

one of these two parts immediately follows the other, whereas in the above-mentioned *kalima*, the two parts are heard together.<sup>11</sup>

In other words, 'Astarābādī recognises that a word like *ḍaraba*, just like other words of the same kind, is not only semantically, but also morphologically complex. According to modern Western linguistics, one can recognise in it not only two morphemes, but also two morphs; according to the Arabic grammatical tradition, one should—given the definition of the *kalima*—see in it two *kalima* and not one. However, the Arabic grammatical tradition has always seen in it only one *kalima*. In order to avoid the contradiction, 'Astarābādī proposes to consider as complex only the expression a part of which *immediately follows* (*muta' aqqib*) the other. This confirms that a *kalima*, though not being a word, a morpheme or a morph, is a segment of the utterance. The criterion of the *consecution* allows us to exclude from the definition of *kalima* the *bound morphemes*, belonging to what today is called *non-concatenative morphology*. It is already clear that this criterion does not allow to exclude the morphemes belonging to the concatenative morphology, i.e. those which 'Astarābādī refers to in the passage above in brackets ([...]) that we quote here:

as we claimed concerning the previous *kalimāt* and as it would be correct to claim concerning the inflectional vowels.<sup>12</sup>

## 2.2. *The morphemes belonging to concatenative morphology*

Here is what the “abovementioned *kalimāt*” were:

If one says: when you say *muslimāni*, *muslimūna*, *baṣrī* and all the imperfect verbs, a part of the expression of each of them indicates a part of their meaning; the *wāw* marks the plural, the *'alif* marks the dual, the *yā'* marks the *nisba* and the prefixes of the imperfect mark at the same time the imperfect and the person. The same can be said for the *tā'* of feminisation in *qā'ima*, the *tanwīn* and the definite article, and the two *'alif* of feminisation. The expression of each of them [i.e. of these *kalimāt*], just like their meaning, is therefore necessarily complex. There are then two *kalima*, not one.<sup>13</sup>

<sup>11</sup> *Wa-lā yaṣiḥḥu 'an naddā'iyā hāhunā 'anna l-wazn aṭ-ṭārī' kalima šārat bit-tarkīb ka-juz' kalima [...] fal-i' tirāḍ bi-hādīhi l-kalim i' tirāḍ wārid 'illā 'an nuqayyida tafsīr al-lafḍ al-murakkab fa-naqūla huwa mā yadullu juz' uhu 'alā juz' mā nāhu wāḥid al-juz' ayni muta' aqqib lil-'āxar wa-fi hādīhi l-kalima al-maḍkūra al-juz' āni masmū' āni mā 'an*

<sup>12</sup> *Kamā dda' aynā fi l-kalim al-mutaqaddima wa-kamā yaṣiḥḥu 'an yuddā' ā fi l-ḥarakāt al-'irābiyya.*

<sup>13</sup> 'Astarābādī (*ŠKāfiya*: 1.5): *'in qīla 'inna fi qawlīka muslimāni wa-muslimūna wa-baṣriyyi wa-jamī' al-'af'āl al-muḍārī' a juz' lafḍ kull wāḥid minhā yadullu 'alā juz' mā nāhu 'id al-wāw*

All the expressions mentioned by 'Astarābādī constitute a sequence of two segments. We shall refer here to his own examples and when he does not give any, we shall refer to those given by his commentator 'Alī b. Muḥammad, known as as-Sayyid aš-Šarīf al-Jurjānī, in the "margin" (*Ḥāšiya*):

1. *muslimāni* 'two Muslims', analysable as *muslim* + *ā(ni)*, where *ā* is the mark of the dual, but also a mark of declension, in this case the nominative, as opposed to *muslimayni*, accusative-genitive;
2. *muslimūna* 'Muslims', analysable as *muslim* + *ū(na)*, where *ū* is the mark of the masculine plural, but also a mark of declension, in this case the nominative, as opposed to *muslimūna*, accusative-genitive ;
3. *baṣriyy*, analysable as *baṣr* (stem of *al-Baṣra*) + *yy*, which marks the 'relation' (*nisba*), i.e. 'from Basra, Basrian';
4. the imperfect prefixes: 'Astarābādī does not give any example, but his commentator gives '*adribu* 'I (will) strike'. It is clear that '*adribu*, *taḍribu* 'you (will) strike', *yaḍribu* 'he (will) strike'... can be analysed as prefixes '*a-*, *ta-*, *ya-* added to a stem *ḍrib*<sup>14</sup> and that these prefixes, as mentioned by 'Astarābādī, mean both the imperfect and the 'state of the agent' (*ḥāl al-fā'il*). His commentator is more detailed, for he speaks on the one hand of 'future or present' (*al-'istiqbāl 'aw l-ḥāl*), on the other of 'speaker, addressee, masculine, for example' (*at-takallum wal-xiṭāb wat-taḍkīr maṭalan*);
5. *qā'ima* 'standing' analysable as *qā'im* + *at-*, where *at-* is a mark of the feminine;
6. the two *tā'* of "feminisation" are exemplified by his commentator (Šarīf Jurjānī *Ḥāšiya* in *ŠKāfiya*: 1. *istidrakāt* 5) through *ḥublā* 'pregnant' and *ḥamrā* 'red'. It is worthwhile noting that the segmentation of *ḥublā* into *ḥubl* + *ā* and of *ḥamrā* into *ḥamr* + *ā* leaves a segment that coincides neither with a word, nor with a stem: the word is *ḥabal* 'pregnancy' in the first case and *ḥumra* 'redness' in the second. We could see here some intermediary (in synchrony) or transitional (in diachrony)

*tadullu 'alā l-jam'iyya wal-'alif' alā t-tatniya wal-yā' 'alā n-nisba wa-ḥurūf al-muḍārā'a 'alā ma'nā fi l-muḍārī' wa-'alā ḥāl al-fā'il 'ayḍan wa-kaḍā tā' at-ta'nūt fi qā'ima wat-tanwīn wa-lām at-ta'rīf wa-'alifā at-ta'nūt fa-yajibu 'an yakūna lafīḍ kull wāḥid mūnhā murakkaban wa-kaḍā l-ma'nā fa-lā yakūnu kalima bal kalimatayn.*

<sup>14</sup> Since the vowels of both the prefix and the stem can be subject to apophony, for example the passive *yudrab* 'he is /will be stricken', the prefix is exactly '(V), t(V), y(V) and the stem is *ḍrVb*...

forms between concatenative morphology and non-concatenative morphology.

We have obviously noted that 'Astarābādī includes in this list the definite article and the *tanwīn*. The definite article is prefixed to the noun, e.g. *ar-rajul* 'the man', and the *tanwīn* is suffixed to the noun, which bears the declension vowel: *Zaydu-n*, *qā'imu-n* etc. 'Astarābādī concludes:

the answer is that all I have mentioned constitutes two *kalima* which have become, as a result of the strong amalgam, like one *kalima*: the compound expression has been given the desinential inflection of the *kalima*, and this because the affixes are not independent; the same can be said of the inflectional vowels.<sup>15</sup>

In his answer, 'Astarābādī, as a good logician, draws the conclusion necessarily resulting from Ibn Ḥājjib's definition of the *kalima* as a minimal meaningful unit. Such a definition leads to recognise as a *kalima* all morphemes belonging to the concatenative morphology, i.e. not only those he generically labels as *hurūf muttaṣila*, but also the declension vowels.

'Astarābādī's originality clearly appears when comparing his text to Šarīf Jurjānī's critical commentary in margin of the *ŠKāfiya*:

It is clear that the *tanwīn* and the article belong to the particles and that both have been counted amongst them: each of the two is therefore immediately one *kalima*. Thus, an expression such as *ar-rajul* is made of two—not one—*kalima*.<sup>16</sup>

Further on he adds:

as for the 'alif of dualisation, the *wāw* of the plural, the *yā* of the *nisba*, the *tā* of vocalised feminisation and the two 'alif of feminisation, we have already said that they belong to the formatives, which are added to the *kalim*.<sup>17</sup>

Šarīf Jurjānī's commentary shows that 'Astarābādī groups together under the name of *hurūf muttaṣila* (translated here as 'affixes') what Šarīf Jurjānī himself distinguishes as *hurūf al-ma'ānī* on one side and *hurūf al-mabānī*

<sup>15</sup> *Faḷ-jawāb 'anna jamī' mā ḍakartu kalimatāni šaratā min šiddat al-'imtizāj kalima wāḥida fa-'u'riba al-murakkab 'i'rāb al-kalima wa-ḍālika li-'adam istiqlāl al-hurūf al-muttaṣila fi l-kalim al-maḍkūra wa-ka-ḍālika al-ḥarakāt al-'i'rābiyya.*

<sup>16</sup> *Lā xafā' fi 'anna t-tanwīn wa-lām at-ta'rīf min hurūf al-ma'ānī wa-qad 'addūhumā fihā fa-kull wāḥida minhumā kalima 'alā ḥiyālihā [sic: read 'alā ḥālihā] fa-naḥw ar-rajul kalimatāni lā kalima wāḥida.*

<sup>17</sup> *'Ammā 'alif at-taṭniya wa-wāw al-jamī' wa-yā' an-nisba wa-tā' at-ta'niṭ al-mutaḥarrika fa-qad qila 'innahā min hurūf al-mabānī zīdat fi l-kalim.*

on the other. The first term is the plural of *ḥarf ma'nā* (litt. 'semantic articulation') which, in the Arabic grammatical terminology, designates the particle, as opposed to *ḥarf lafḍ*, which designates the 'phonic articulation' (consonants and long vowels). The second term is the plural of *ḥarf mabnā*, which designates the elements taking part in word formation (for this reason we have translated it as 'formative'). If 'Astarābādī can operate such a grouping, it is clearly because the *ḥurūf al-mabānī* are as meaningful as the *ḥurūf al-mā'ānī*. Therefore, 'Astarābādī introduces a double conceptual and terminological innovation. Conceptually, he finds the idea of morpheme in its most general sense: this concept groups together the morphemes belonging to non-concatenative morphology ('root' and 'pattern') as well as those belonging to concatenative morphology. In this last case, the concept groups together the morphemes belonging to derivation as well as inflection, and, in the case of inflection, those belonging to purely morphological inflection as well as those belonging to desinential inflection (at the intersection between morphology and syntax). Terminologically, he has a term for "affix", even if this term is less general than his concept of morpheme: in fact, it excludes the inflectional vowels, though they are recognised as morphemes affixed to the word.<sup>18</sup> This is explained by a terminological reason, i.e. the fact that in the Arabic grammatical terminology *ḥarf* and *ḥaraka* are in opposition.

This blurring of the boundary between *ḥurūf al-mā'ānī* and *ḥurūf al-mabānī* is confirmed by the fact that 'Astarābādī transfers the definition of the *ḥurūf al-mā'ānī* as 'expression signifying [not in itself but] in another' (*lafḍ dāll 'alā ma'nā fi ḡayrihi*) to the *ḥurūf al-mabānī*. As he writes: "the *ḥarf* creates its meaning in an expression other than itself, which either precedes it, as in *baṣriyy*, or follows it, as in *ar-rajul*".<sup>19</sup>

Nevertheless, we should not attribute to 'Astarābādī what he does not say. He simply reacts as a logician to Ibn Ḥājjib's definition of the *kalima*. This definition, of an intensional type, has undesirable effects in extension: literally, it would end up considering all morphemes as *kalima*. However,

<sup>18</sup> Cf. Šarīf Jurjānī's commentary to the sentence "the same can be said of inflectional vowels" ('Astarābādī *ŠKāfiya*: 1. *istidrakāt* on p. 5,6–7): "he means that they are also *kalimāt* in themselves, belonging to the category of *ḥurūf*" (*ya'nī 'annahā 'ayḍan kalimāt bi-ra'sihā min qabīl al-ḥurūf*): in this context, Šarīf Jurjānī is obviously referring to 'Astarābādī's *ḥurūf mutaṣṣila*, all the more so as he adds: "but as a result of the strong amalgam between inflectional vowel and something else, he counted the whole as one *kalima*" (*lākin li-šiddat imtizāj al-ḥaraka bi-ḡayrihā 'adda l-majmū' kalima wāḥida*).

<sup>19</sup> 'Astarābādī (*ŠKāfiya*: 1.9): *fal-ḥarf mūjīd li-mā'nāhu fi lafḍ ḡayrihi 'immā muqaddam 'alayhi kamā fi naḥw baṣriyy 'aw mu'axxar 'anhu kamā fi ar-rajul*.

not every morpheme is a *kalima*; moreover, some sets of two morphemes are considered as one *kalima* . . . In order to avoid these undesirable effects, there is no other solution in logic, than to restrict the definition (*qayd*, pl. *quyūd*). A first restriction, the non-consecution of two elements of a set, allows him to exclude from the category of *kalima* the bound morphemes belonging to the non-concatenative morphology. A second restriction, i.e. the incidence of desinential inflection, allows him then to exclude from the category of *kalima* the bound morphemes belonging to the concatenative morphology.

However, in his commentary to the sentence “the compound expression has been given the desinential inflection of the *kalima*”, Šarīf Jurjānī distinguishes between four cases:

this is evident in expressions such as *bašriyy*, *qā'ima*, *ḥublā* and *ḥamrā'*, because the desinential inflection, at the end of the compound expression, affects a part which originally has no right to it. As for the noun affected by the *tanwīn*, this *tanwīn* occurs after the inflectional vowel affecting the first part of the noun. In the case of the dual and the plural, if we take the marks of each of them as the desinential inflection itself, since they replace the inflectional vowels, then there is no desinential inflection affecting the compound expression, but only the first part of it. The desinential inflection, in expressions such as *ar-rajulu* and *'aḍribu*, actually affects the second part of the expression, which has right to it, not the set constituted by the second and the first part: think about it!<sup>20</sup>

Šarīf Jurjānī's commentary has a double merit. It shows the double function of the inflectional vowels in 'Astarābādī: if they are not *kalimāt*, but parts of *kalimāt*, they are at the same time boundaries of *kalimāt*. Moreover, it stresses the inadequacy of 'Astarābādī's criterion: if it allows considering the *tanwīn* as a *kalima*, it no longer allows considering the definite article as a *kalima*! This pernicious effect bears also a paradox: the Arab grammatical tradition has always considered the article as a *kalima*. On the contrary, as Levin (1986; 2007) reminds us this is not the case for the *tanwīn*: a grammarian such as Sībawayhi (d. 177/793?) does not consider

<sup>20</sup> 'Astarābādī (ŠKāfiya, 1.istadrākāt 5): *hadā fi naḥw bašriyy wa-qā'ima wa-ḥublā wa-ḥamrā' dāhir li-'anna l-'irāb fi 'āxir al-murakkab 'alā juz' lā yastahiqquhu 'ašlan wa-'ammā l-munawwan fat-tanwīn fihi ba'd ḥarakat al-'irāb 'alā l-juz' al-'awwal wa-fi l-muṭannā wal-majmū' in ju'ila l-'alāma naḥs al-'irāb qā'ima maqām al-ḥarakāt fa-lā 'irāb lil-murakkab bal lil-juz' al-'awwal wal-'irāb fi naḥw ar-rajulu wa-'aḍribu 'innamā huwa lil-juz' at-tānī alladī yastahiqquhu lā lil-majmū' al-murakkab minhu wa-min al-juz' al-'awwal fa-ta'ammal.*

it as a *kalima*, but as a part of *kalima*, to be included among the *zawā'id* 'added elements'.<sup>21</sup>

So we are now back to the starting point. In the sequence *ar-rajulu* as well as in the sequence *Zaydun*, 'Astarābādī counts the same number of morphemes, i.e. three: *al + rajul + u*; *Zayd + u + n*. In the two sequences, he alternatively counts one *kalima* (*al-rajulu*; *Zaydun*) or two (*al + rajul*; *Zayd + n*), where the inflectional vowel is never counted as a *kalima*, but only as a part of a *kalima*: *al + rajul-u*; *Zayd-u + n*. This confirms the negative response that we have given to the initial question: *kalima* is neither a word, nor a morpheme. In both cases, there is only one word, in which case the Arab grammarians do not use the term *kalima* but *lafḍa*, as Šarīf Jurjānī reminds us in margin of the *ŠKāfiya*: "but, because of the amalgam between the two [i.e. the article and the noun], we apply to them the name of *lafḍa*".<sup>22</sup> In any case, this allows us to provide a positive answer to the same question.

### 3. THE KALIMA, A CONSTITUENT OF THE UTTERANCE

There is only one possible interpretation of the *kalima* compatible with this double counting. It is not a semantic one, even if the *kalima* is a meaningful unit. It is not a morphological one, even if the *kalima* is an "expression", i.e. a concrete segment of the utterance, with a significant exception, that of the "hidden pronoun". It is then a syntactic one: a *kalima* is a "part of the constructed discourse", i.e. a constituent of the utterance, which in its turn can be divided into constituents.

The example *ar-rajulu*, compared to *rajulun* 'a man', shows the principle of segmentation of the utterance into constituents and of its constituents into constituents: this cannot but be distribution.<sup>23</sup> If *ar-rajulu*

<sup>21</sup> See also Levin (this volume)

<sup>22</sup> *Wa-lākin li-šiddat al-'imtizāj baynahumā yuṭlaq 'alayhimā al-lafḍa*. Cf. 'Astarābādī himself, (*ŠKāfiya*: 1.5.3–4): "an expression such as *qālā* ['they both said'] and *qālū* ['they said'] is, just like *artā* [a kind of tree] and *burqu'* ['veil, burkah'], one and only one word. The same applies to all that is uttered in one and only one time, though each of the first two is made of two *kalima*, as opposed to the last two" (*miṭl qawlīka qālā wa-qālū ka-artā wa-burqu' lafḍa wāḥida wa-kaḍā kull mā yutalaffaḍ bihi marra wāḥida ma'a anna kull wāḥid min al-awwalayn kalimatān bi-xilāf at-ṭāniyayn*).

<sup>23</sup> On the syntactic and, more specifically, distributional basis of word classification, see Owens (1989).

and *rajulun* count each as *one kalima*, it is because they have the same distribution. Both can be the subject of a verbal sentence such as:

- (5) *qāma r-rajulu*  
 stood ART-man.N  
 ‘the man stood’
- (6) *qāma rajulu-n*  
 man.N-TANWĪN  
 ‘a man stood’

On this point ‘Astarābādī does not innovate: I refer the reader to Zamaxšārī’s *Mufaṣṣal* in which he defines the *kalām* as “what is composed of two *kalima*, one of which is predicated of the other” adding that “this occurs only in the case of two nouns [...] or of a verb and a noun”.<sup>24</sup> Yet, he gives as an example of the first case *Zaydun ‘axūka* ‘Zayd is your brother’ and *Bišrun šāḥibuka* ‘Bišr is your companion’. Zamaxšārī counts indeed two *kalimāt*, which are nouns, in this utterance, but the second one is divided in its turn into two nouns, a substantive and a suffixed pronoun. Conclusion: *ism* is used here not as a noun in the sense of word class, but in the sense of a noun phrase.

If *ar-rajulu* and *rajulun* count each as *two kalimāt*, it is because the definite article and the *tanwīn* are in complementary distribution. Because of its mutual exclusion with the article, we can see in the *tanwīn* one of the determiners of the noun. However, there is some uncertainty amongst grammarians, on the meaning of this determiner, whereas there is none on the meaning of the article (“particle of definition”). According to Zamaxšārī (*Mufaṣṣal*: 328), the *tanwīn* of *rajulun* is the same as that of *Zaydun*. Since Zayd is a proper noun, this cannot be a mark of indefiniteness: it is only a mark of “declinability” (*makāna* or *tamakkun*). According to ‘Astarābādī (*ŠKāfiya*: 1.13), on the contrary, “the *tanwīn*, in *rajulun*, is a mark of indefiniteness, too” (*at-tanwīn fī rajulun yufīdu t-tankīr ‘ayḍan*). This uncertainty is easily explainable if we think that the *tanwīn* can be pronounced only if the inflectional vowel is pronounced, too. Yet, in the orthoepy of Classical Arabic, there is a case in which the *tanwīn* is not pronounced: the pause. The noun determination system changes whether the distribution is funded on the noun’s inflected (*ar-rajulu* vs *rajulun*) or

<sup>24</sup> Zamaxšārī (*Mufaṣṣal* 6): *al-kalām huwa l-murakkab min kalimatayn ‘usnidat ‘iḥdāhumā ‘ilā l-‘uxrā, and wa-dāka lā yata’attā ‘illā fī smayn [...] ‘aw fī fi’l wa-ism.*

pausal (*ar-rajul* vs *rajul*) forms: in the first case, we can oppose a mark of definition (*al-*) to a mark of indefiniteness (*-n*); in the second, we can oppose only a positive mark of definition to a negative mark of indefiniteness (*al-* vs  $-\emptyset$ ). Actually, since the inflected and pausal forms of the noun coexist, the *tanwīn* of *rajulun* can hardly be anything else than a purely redundant mark of indefiniteness. For this reason, the status of the *tanwīn* hesitates between *kalima* and part of *kalima*.

I am of course well aware that what has been said has to do with a theory of the sentence that, from the 4th/10th century on, has been strongly influenced by the Greek logic: at that time, the term *ʿisnād* acquired the meaning of ‘predication’. Each utterance (*kalām*) is therefore defined as a sentence (*jumla*), i.e. a set (literal meaning of the term) of two elements (*juzʿ*) in a predicative connection. One can easily imagine that at times some recalcitrant structures are introduced by force into the model... In any case, the number and the hierarchy of the *kalimāt* will be further justified with reference to this model: the noun is what can be either subject or predicate in a sentence; the verb is what can be predicate but not subject in a sentence; the particle is what can be neither subject nor predicate in a sentence.<sup>25</sup>

Another influence of the logic can be seen in the distinction, among the *kalimāt*, between the noun and the verb as “expression signifying [each one] in itself” (*lafẓ dāll ʿalā mā nā fi nafsihi*), as opposed to the particle which is “an expression signifying in an [expression] other than itself”.<sup>26</sup> This distinction immediately reminds us of the difference done by ancient logic between categorematic and syncategorematic terms, as presented by Priscian (*Institutiones grammaticae* 2.4.15):

partes orationis sunt duae nomen et verbum, quia haec solae et jam per se conjunctae plenam faciunt orationem; alias autem partes syncategoremata, hoc est consignantia, appellant.

Lalande’s *Vocabulaire technique et critique de la philosophie* (1972) made this quotation in order to support his definition of the term categorematic as “when speaking of words, those which have a meaning in themselves”. Actually, this simply repeats in a semantic way what categorematic does in a syntactic way, i.e. which can constitute by itself the predicate (gr. *katēgorēma*) in a clause.

<sup>25</sup> Cf. ʿAstarābādī (6, ll. 8–12; 8, ll. 31–32). For a critical commentary, cf. Guillaume (1988, though he does not quote ʿAstarābādī). Cf. also Owens (1989).

<sup>26</sup> Cf. Guillaume (1988) for more references.

The same applies here: the logical or semantic vocabulary must not hide the syntactic reality. How to understand that the suffixes of the perfect are categorised as *kalima*, but not the prefixes of the imperfect, when we know that they are both recognised as bound morphemes playing the same semantic role (they mark firstly the person, secondly the gender and/or the number)? How to understand that some suffixes of the perfect and the imperfect, recognised as marks of gender and/or number, as classified as personal pronouns (*ḍamā'ir*) and so as *kalima*, but not the *-t* of *ḍaraba*?<sup>27</sup> Finally, how to understand the existence of the implicit *kalima* which is represented by the “hidden pronoun” (*aḍ-ḍamīr al-mustatir*)? There is no answer to these questions unless one introduces in one way or another the concept of distribution. This is evident in the case of the perfect: 1st and 2nd person suffixes have the same distribution as the 3rd person lexically full subject:

- |                     |                 |                        |
|---------------------|-----------------|------------------------|
| (7) <i>ḍarab-tu</i> | <i>ḍarab-ta</i> | <i>ḍarab(a) Zaydun</i> |
| stroke-I            | stroke-you      | stroke Zayd            |
| 'I stroke'          | 'you stroke'    | 'Zayd stroke'          |

On the contrary, the fact that in the 3rd person of the imperfect the prefix and the lexically full subject occur together leads grammarians not to see in the 1st and 2nd persons some SV structures, but rather to rebuild V[S] structures by analogy with the perfect (considered as a basic structure), i.e.:

- |                              |                          |                        |
|------------------------------|--------------------------|------------------------|
| (8) <i>yaḍribu Zaydun</i>    | <i>taḍribu</i> [ʾanta]   | ʾaḍribu [ʾanā]         |
| he-strike Zayd               | you-strike [you]         | I-strike [I]           |
| 'Zayd strikes (will strike)' | 'you strike/will strike' | 'I strike/will strike' |

Similarly, what makes necessary to see in *Zaydun qāma*, not a *bound sentence* subject/verb, but a *segmented sentence* topic/comment,<sup>28</sup> is the fact

<sup>27</sup> Called “unvowelled *tā'* of feminisation” (*tā' at-ta'nūt as-sākina*) and counted by Zamaxšarī (*Mufaṣṣal*, 328) as a *ḥarf* “the occurrence of which aims at indicating immediately that the subject is feminine” (*duxūluhā lil-'idān min 'awwal al-amr 'anna l-fā'il mu'annaṭ*). The *ḥarf* is undoubtedly a *kalima*, but it is not a major constituent.

<sup>28</sup> We use here the particularly accurate terminology employed by the Swiss linguist Charles Bally (1865–1947: Bally 1965).

that we find in the same context for example, *Zaydun qāma 'abūhu*. This allows again for an analogical reconstruction, i.e.:

- (9) *Zaydun qāma [huwa]*    *Zaydun qāma 'abūhu*  
       Zayd stood [he]            Zayd stood father-he  
       'as for Zayd, he stood'    'as for Zayd, his father stood'

### CONCLUSION

This explains why 'Astarābādī's writings have delighted me for over thirty years. Who else offers us a double analysis of an expression such as 'aḍribu, once as a word (*lafḍa*) and once as a sentence (*jumla*)? As a word, 'aḍribu can be analysed as a sequence of three segments 'a + ḍrib + u, to which correspond at least than five meanings: 1st person singular (marked by the prefix ['']) of the imperfect (marked by the prefixation itself to the stem of the verb) indicative (marked by the vowel -u) of the active voice (marked by the vocalisation of the prefix and of the stem). However, as a sentence, 'aḍribu must be analysed as a sequence of two constituents (*kalima*), one explicit (*maḥḥud bihā*), the verb, and one implicit (*muqaddara*), the "hidden pronoun", subject of the verb.<sup>29</sup> This double analysis allows us to understand not only what a *kalima* is not, but also what it is.

### REFERENCES

#### 1. Primary Sources

'Astarābādī. *ŠKāfiya*. Šarḥ Kāfiyat Ibn al-Ḥāḥib, 2 vols. Istanbul: Maṭba'at aš-šarika aš-šihāfiyya al-ʿuṣmāniyya, 1275H<sup>1</sup> and 1310H<sup>2</sup>. [Reprint Bayrūt: Dār al-kutub al-ʿilmiyya, n.d.].

Šarīf Jurjānī. *Hāšiya*. Al-Ḥāšiya 'alā šarḥ al-Kāfiya, cf. 'Astarābādī.

Zamaxšarī. *Mufaššal*. Al-Mufaššal fi ʿilm al-ʿarabiyya. Bayrūt: Dār al-Jil, n.d.

<sup>29</sup> Pointed out by Šarīf Jurjānī in margin of the *ŠKāfiya* (1.5.below): "the same applies to the prefixes of the imperfect: the *hamza* of 'aḍribu is not a constituent, but is in fact, together with what follows it, one and only one constituent, and the hidden pronoun is another constituent" (*wa-kaḍā fi ḥurūf al-muḍāra'a fal-hamza fi 'aḍribu laysat kalima bal hiya ma'a mā ba'da-hā kalima wāḥida haqīqatan waḍ-ḍamīr al-mustatir kalima 'uxrā*).

2. *Secondary Literature*

- Bally, Charles. 1965. *Linguistique générale et linguistique française*, 4th ed., reviewed and corrected. Berne: Francke.
- Ducrot, Oswald and Jean-Marie Schaeffer. 1995. *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris: Seuil.
- Guillaume, Jean-Patrick. 1988. "Le discours tout entier est nom, verbe et particule. Elaboration et constitution de la théorie des parties du discours dans la tradition grammaticale arabe". *Langages* 92, 25–36.
- Lalande, André. 1972. *Vocabulaire technique et critique de la philosophie*, 11th ed. Paris: Presses Universitaires de France.
- Levin, Aryeh. 1986. "The medieval Arabic term *kalima* and the modern linguistic term morpheme: similarities and differences". *Studies in Islamic History and Civilization in Honour of Professor David Ayalon*, ed. M. Sharon. Jerusalem: Cana & Leiden: Brill, 423–446.
- . 2007. "Kalima". *Encyclopedia of Arabic Language and Linguistics*. Leiden: Brill. 2.545–548.
- Owens, Jonathan. 1989. "The syntactic basis of Arabic word classification". *Arabica* 36, 211–234.