

# *Compositionality in N + N compounds in Jordanian Arabic and English*

ABDEL RAHMAN MITIB ALTAKHAINEH 

*The University of Jordan, Amman, Jordan*

[a.altakhaineh@ju.edu.jo](mailto:a.altakhaineh@ju.edu.jo)

---

## **Abstract**

This article aims to examine to what extent English and Jordanian Arabic (JA) have the same classification of N + N compounds based on their degree of compositionality. It also attempts to propose a universally applicable classification of compositionality in N + N compounds. I suggest a modified version of the degree of compositionality based on previous classifications by Fernando (1996), Dirven and Verspoor (1998), and Kavka (2009). The new classification is based on the semantic contribution of the head and the non-head to the meaning of the whole compound. After I have applied the new scale to the JA data, I argue that English and JA have compounds that exhibit the four degrees of compositionality; namely completely compositional, semi-compositional, semi non-compositional and completely non-compositional. The article concludes with some recommendations for future research.

**Keywords:** Morphology, compounding, compositionality, Jordanian Arabic, English

## **Résumé**

Cet article vise à examiner dans quelle mesure l'anglais et l'arabe jordanien (JA) ont la même classification des composés N + N, selon leur degré de compositionnalité. Il tente également de proposer une classification universellement applicable de la compositionnalité des composés N+N. Je propose une version modifiée du degré de compositionnalité basée sur les classifications précédentes de Fernando (1996), Dirven et Verspoor (1998), et Kavka (2009). La nouvelle classification se fonde sur la contribution sémantique de la tête et de la non-tête à la signification de l'ensemble du composé. Après avoir appliqué la nouvelle échelle aux données de l'arabe jordanien, je soutiens que l'anglais et l'arabe jordanien ont des composés qui présentent les quatre degrés de compositionnalité, à savoir complètement compositionnel, semi-compositionnel, semi non-compositionnel et complètement non-compositionnel. L'article se termine par quelques recommandations pour les recherches futures.

**Mots clés:** Morphologie, composition, compositionnalité, Arabe jordanien, Anglais

---

I wish to thank the three anonymous reviewers and the editor-in-chief for their useful comments on an earlier draft of the paper.

## 1. INTRODUCTION

Jordanian Arabic (henceforth JA),<sup>1</sup> which is a variety of colloquial Arabic, is a Semitic language spoken in Jordan by approximately 10 million people. JA exhibits different types of morphological processes that result in new words. However, one of the predominant morphological processes in Arabic is compounding. JA exhibits two constructions which are identical in form but very different in terms of their semantics. These constructions are referred to as Synthetic Genitive Constructions (henceforth SGCs) (Altakhaineh 2019) and are normally comprised of two nouns, but may alternatively be composed of an adjective and a noun. The first constituent in this construction is always indefinite, while the latter can either be definite or indefinite (Fassi Fehri 2012: 156). With regards to the semantics of these two constructions, one is based on a relationship of possession (i.e., phrase), while the other is clearly not (i.e., compound). In the latter, there is a concatenation of two nouns with a meaning that is usually predictable and compositional. However, the study of compositionality in Arabic compounds has not been given full attention in the literature on Arabic language and linguistics. This study aims to provide a systematic and analytical description of compositionality in N + N compounds in JA and English, and consequently suggests a modified classification based on the previous proposals of Fernando (1996), Dirven and Verspoor (1998), and Kavka (2009). The current study seeks answers to the following questions:

1. What is the best classification of compounds in JA and English, based on their degree of compositionality?
2. To what extent do English and JA compounds share the same degree of compositionality?

This study is organised as follows: Section 2 defines a compound and identifies headedness in English and JA compounds. Section 3 defines compositionality and identifies its classifications for English compounds. Section 4 provides an overview of the methods of data collection and analysis employed in the current study. Section 5 suggests a modified classification of compositionality for compounding in English and JA. Finally, Section 6 summarises the main findings of the study and concludes with some recommendations for further research.

## 2. WHAT IS A COMPOUND?

Scholars have provided different definitions for compounds (Marchand 1960: 11, Fabb 1998: 66, Olsen 2000: 280, Carstairs-McCarthy 2002: 59, Ralli 2013: 10). These definitions agree that compounds consist of two or more words, stems, or roots which are combined to form a new complex word, for example *school bus*, *classroom* or *egghead*. Two main types of compounds exist in the relevant literature, based on the presence (endocentric) or absence (exocentric) of a semantic head. An

---

<sup>1</sup>Abbreviations used: JA : Jordanian Arabic; MSA: Modern Standard Arabic; RHHR: Right-Hand Head Rule; SGC: Synthetic Genitive Constructions.

endocentric compound is semantically headed since it denotes a hyponym of the head element in the compound, that is, the semantic head is inside the compound (Lieber 2005: 178, Bauer 2009: 348), for example *sailboat*, whereas an exocentric compound is not semantically headed.<sup>2</sup> It does not denote a hyponym of either of its elements; that is, the semantic head is outside the compound, for example *scarecrow* (Fabb 1998: 66-67, Booij 2002: 81, Haspelmath 2002: 87-89).

In English, endocentric compounds are normally right-headed based on the Right-Hand Head Rule (RHHR), first suggested by Williams (1981: 248), who states that “in morphology, we define the head of a morphologically complex word to be the right-hand member of that word.” For example, a *bookshop* is a *shop*. Conversely, Arabic tends to place the head element in endocentric compounds in the initial/left position, as illustrated in (1) (Altakhaineh 2016: 14).

- (1) minjaar l-xafab IS A minjaar  
 saw the-wood IS A saw  
 ‘the wood saw’

The compound *minjaar l-xafab* “the wood saw” is a hyponym of *minjaar* “saw”. To make the compound plural, the word *minjaar* ‘saw’ is pluralised rather than *l-xafab* ‘the wood’. Therefore, the head of the compound is *minjaar* “saw” (Altakhaineh 2016). It has been argued that the position of the head is a parameter; that is, it can be either the right or the left element of a construction in a language (Selkirk 1982).

Another difference between English and JA is the form of the compound. As mentioned previously, compounds in JA are similar to phrases in form, that is, both types of compounds are SGCs. Some criteria have been proposed to distinguish between the two constructions. Specifically, Altakhaineh (2019: 33) suggests that a JA compound consists of at least two adjacent words, where the non-head is normally non-referential. In other words, adjacency and referentiality can be used to draw the line between these two constructions: the two elements of compounds are inseparable (3), while those of phrases are separable (2).

- (2) ?awaaʕii kull t<sup>s</sup>-t<sup>s</sup>ullaab (phrase)  
 clothes all the-students  
 ‘the clothes of all students’
- (3) ʕaruus (\*baʕd<sup>s</sup>/kull) l-bihaar (compound)  
 bride (\*some/all) the-seas  
 lit. the bride of (\*some/all) sea(\*s)

Example (2) shows that the elements of a phrase in JA are separable in the sense that an element can be inserted between them (i.e., *kull* ‘all’), while those of a compounds in (3) are adjacent or inseparable. Regarding the property of referentiality, the non-head of phrases is referential (4), while that of compounds is not (5):

<sup>2</sup>Note that a semantic definition is only one part of defining exocentricity and headedness; the former can also be determined on formal grounds while the latter, morphologically speaking, cannot be defined in purely semantic terms.

- (4) beet haaða z-zalameh (phrase)  
 house this the-man.M  
 ‘this man’s house’
- (5) raaʔid (\*haaða) l-fadʕaaʔ (compound)  
 pioneer (\*this.M) the-space.M  
 ‘the pioneer of (\*this) space’  
 (Intended meaning of the compound is ‘astronaut’)

Example (4) shows that the non-head of phrases in JA is referential, as evidenced by the use of the demonstrative *haaða* ‘this’. In contrast, in (5), the demonstrative cannot be used with the non-head of the compound because it is non-referential.

The semantics of these two constructions is also different. The elements of phrases denote a relationship of possession, whereas those of compounds do not. This relationship can be shown using the word *tabaʕ* ‘of/for’ which can act as a possessive marker in JA. This marker appears in phrases but not in compounds, especially non-compositional ones, as in (6–7):

- (6) s-siyyarah tabʕat z-zalameh (phrase)  
 the-car.FSG for.FSG the-man.MSG  
 ‘the man’s car’
- (7) \*f-jaʕar tabaʕ l-banaat (compound)  
 the-hair for the-girl  
 lit. the hair (\*for) the girls  
 (intended meaning ‘the candy floss’)

Other criteria were found to be useful in differentiating phrases from non-compositional compounds, but not from compositional ones, as follows:

- (8) galam wu mastʕarit l-walad (phrase)  
 pen and ruler the-boy  
 ‘the boy’s pen and ruler’
- (9) findʒaan wu ʔibriig f-jaay (compositional compound)  
 cup and pot the-tea  
 ‘the tea cup and pot’

Examples (8–9) show that the head of phrases, such as *galam* ‘pen’ in (8), and the head of compositional compounds, such as *findʒaan* ‘cup’ in (9) can be coordinated using the conjunction *wu* ‘and’ with other nouns, without any syntactic or semantic problems. Conversely, the head of non-compositional compounds cannot be coordinated with other nouns. In addition, coordination can be used as a criterion to differentiate between compositional compounds, on the one hand, and semi-compositional and non-compositional compounds, on the other. The head of semi-compositional and non-compositional compounds cannot be coordinated with other nouns. This can be seen in the following examples:

- (10) ʕarus (\*wu ʕariis) l-baħir (non-compositional compound)  
 bride (\*and bridegroom) the-sea  
 ‘the female (\*and male) merpeople’  
 lit. the/a bride (\*and bridegroom) of the sea

- (11) *faʃar* (\*wu dʒadaayil) l-banaat  
 hair (\*and braids) the-girls  
 ‘the candyfloss (\*and braids)’  
 lit. the hair (\*and braids) of the girls (non-compositional compound)
- (12) *hajar* (\*wu tʃuub) z-zaawiyah  
 stone (\*and bricks) the-corner  
 ‘the cornerstone (\*and bricks)’  
 lit. the stone (\*and bricks) of the corner (semi-compositional compound)
- (13) *baab* (\*wu fubbaak) l-ʃaamood  
 door (\*and window) the-pillar  
 ‘one of the doors of Al-Aqsa Mosque’  
 lit. the door (\*and window) of the pillar (semi-compositional compound)

Examples (10) and (11) show that the heads of non-compositional compounds, such as *ʃaruus* ‘bride’ in (10) and *faʃar* ‘hair’ in (11) cannot be coordinated with other nouns without resulting in an impossible reading. Similarly, examples (12–13) demonstrate that the heads of semi-compositional compounds, such as *hajar* ‘stone’ in (12) and *baab* ‘door’ in (13) cannot be coordinated with other nouns. Thus, examples (10–13) demonstrate that the possibility/impossibility of head coordination in N + N compounds can be used to distinguish between compositional compounds, on the one hand, and semi-compositional and non-compositional compounds, on the other. In addition, coordination of the non-heads is impossible with both semi-compositional and non-compositional compounds:

- (14) *ʃaruus* l-bahir (\*wu n-nahir)  
 bride the-sea (\*and the-river)  
 ‘the mermaid of the sea (\*and the river)’  
 lit. ‘the bride of the sea (\*and river)’ (non-compositional compound)
- (15) *faʃar* l-banaat (\*wu l-wlaad)  
 hair the-girls (\*and the-boys)  
 ‘the candyfloss (\*and the boys)’  
 lit. the hair of the girls’ (\*and the boys) (non-compositional compound)
- (16) *hajar* z-zaawiyah (\*wu l-muθallaθ)  
 stone the-corner (\*and the-triangle)  
 ‘the cornerstone (\*and triangle)’  
 lit. the cornerstone (\*and triangle) (semi-compositional compound)
- (17) *baab* l-ʃaamood (\*wu l-ʃaff)  
 door the-pillar (\*and the-line)  
 ‘one of the doors of Al-Aqsa Mosque’  
 lit. the door of the pillar (\*and the line) (semi-compositional compound)

Examples (14) and (15) demonstrate that if the non-head of non-compositional compounds, such as *lbahir* ‘the sea’ in (14) and *lbanaat* ‘the girls’ in (15), is coordinated, the compound loses its non-compositional reading. The same applies to examples (16–17) where the coordination of the non-head results in an anomalous reading.

### 3. COMPOSITIONALITY AND ITS TYPES

Neef (2009: 394) states that a linguistic expression is considered compositional if its meaning is derived from the meanings of its parts and the way it is structured. For example, the English compound *school bus* is compositional because its meaning is determined by its components; *school* and *bus*. Neef (2009: 395) points out that the compositional meaning of a compound with the constituents AB is ‘B that has something to do with A’. That is, every compositional compound which consists of two elements can be interpreted in an associative way, especially N+N compounds. For instance, the German compound *Fisch•frau*, lit. fish•woman ‘is a woman that has something to do with fish’ (Neef 2009).

In English, three levels of semantic compositionality in compounds can be distinguished. The fact that there are levels within compositionality has been acknowledged by Fernando (1996: 36), who examined the levels of compositionality in idioms. In particular, Fernando argues that in addition to a pure idiom, which is completely non-compositional, there is another type of idiom, which he refers to as semi-idiom. The latter refers to a sequence that has one or more literal elements and at least one element that has a non-literal sub-sense. The semantics of a semi-idiom is not as complex as that of a pure idiom, since its meaning is partially transparent. For instance, one can infer from the idiom “to promise someone the moon” that something is being promised.<sup>3</sup> Similarly, Dirven and Verspoor (1998: 60) argue that compounds can be placed on a cline of transparency, which includes transparent compounds, partially transparent compounds, and non-transparent compounds. Examples of the three classes are *apple tree*, *high street*, and *red tape*, respectively. The three levels of compositionality are presented in Table 1.

In the first level, the compounds are completely compositional in the sense that the meaning of the whole compound is the total sum of its parts. For instance, the compound *bookshop* is a shop that has something to do with books. In the second level, the meaning of the compound is not completely the total sum of its parts, but the head makes a clear contribution to the meaning of the whole compound. For instance, the compound *high street* is a street, even though it is not necessarily high. The compounds in the third level are completely non-compositional in the sense that the meaning of the whole compound cannot be derived from the sum of its parts. For example, *egghead* refers to neither a *head* nor an *egg*. Its meaning, ‘intellectual’, is not related to both elements combined together.

Semantically speaking, by applying the ‘IS A’ condition, which was suggested by Allen (1978: 11), it seems that English compounds are usually semantically headed, but there are some that are headless. This principle is normally used to differentiate between endocentric and exocentric compounds. Allen’s ‘IS A’ condition is given in (18).

- (18) In a compound [ [ ]A [ ]B ]C      C ‘IS A’ B

<sup>3</sup>The difference between phrasal idioms and compounds is that in the former, the elements are not normally adjacent, whereas the elements in the latter are adjacent (see section 2).

	Levels of compositionality	Examples
1	Completely compositional	bookshop, houseboat, darkroom, physics teacher.
2	Semi-compositional	high street, blackboard, cathouse, small talk.
3	Completely non-compositional	egghead, white-collar, airhead, faint-hearted, bluestocking

**Table 1:** Levels of compositionality in English by Dirven and Verspoor (1998: 60)

This can be seen in the endocentric compounds in examples (19) and (20):

(19) A school bus IS A bus

(20) A book shop IS A shop

This principle can be used to show that *faintheart* and *airhead* are exocentric compounds, as in (21) and (22):

(21) A faintheart IS NOT A heart

(22) An airhead IS NOT A head

Bauer (1998: 67) suggests that non-compositional compounds are listed in the dictionary, whereas syntactic constructs such as phrases are not, although he points out that this can be considered a lexicographical rather than a linguistic criterion. In particular, Bauer states that “many linguists seize one aspect of listedness – namely idiomaticity – and use that as a criterion for compound status”. Examples would be a compound like *high street*. Later, Kavka (2009: 33) argues that compositionality is the most important criterion that distinguishes compounds from free combinations, claiming that, like idiomatic expressions, compounds are non-compositional. Kavka (2009: 33) suggests that “their status will be understood more readily if they are viewed as parts of concrete, contextually defined utterances”.

On the other hand, Lieber (2005: 376) points out that compounding in many languages is highly productive and new compounds are very often compositional in meaning, especially when context is taken into account. In other words, it is easy to dismiss this non-compositional criterion for compoundhood, at least in languages like English; the more productive the process of compounding in a language, the less likely individual compounds are to be lexicalised or listed (Lieber and Štekauer 2009: 7). The same applies to phrases: since *old hand* and *green fingers* are non-compositional, whereas *beautiful house*, *long journey* and *tall man* are compositional. Therefore, compositionality is not a reliable criterion for distinguishing compounds from phrases in English. In other languages, things may of course be different. For instance, Borer (2009: 491-2) shows that compositionality in Hebrew is a reliable criterion for distinguishing between compounds on the one hand (e.g., *beyt (ha)sefer*, lit. house (the) book ‘(the) school’) and R-constructs (i.e., possessive constructs such as *beyt (ha)mora*, lit. house (the) teacher ((the) teacher’s house)) and M-constructs (i.e., modification constructs such as *kos (ha)mic*, lit. glass (the) juice ‘(the) juice glass’), on the other. Borer suggests that the meaning of productive

syntactic constructs (i.e., R-constructs and M-constructs) is completely predictable from their components, while the meaning of compounds in Hebrew is not predictable from the individual N components; hence, it is non-compositional. In JA, Altakhaineh (2019: 33) suggests that the non-compositionality criterion can only identify non-compositional compounds; it cannot differentiate between P-constructs (phrases), on the one hand, and compositional compounds, on the other. Therefore, it is partially applicable. Borer's M-construct, whose meaning is totally predicted from its parts, is a compositional compound in Arabic. Thus, in Arabic, R-constructs are phrases, while M-constructs are compositional compounds. What Borer refers to as compounds in Hebrew, which are entirely non-compositional, are also found in Arabic. This suggests that while compositionality can be used to pinpoint compounds in Hebrew, it cannot be used as a criterion in Arabic, since both compositional and non-compositional compounds can be found.

In relation to degrees of semantic compositionality, Bauer (1983: 56) argues that compounds that exhibit different degrees of compositionality are attested in the literature.<sup>4</sup> For example, unlike *understand*, which has been completely lexicalised, compounds like *playboy* and *bedstead* still have some relation to the meaning of their parts; they are only partially compositional (Bauer 1983: 56–57). According to Marelli and Luzzatti (2012: 653) and Ji et al. (2011), semantic transparency or degrees of semantic compositionality is all about how well the combination of the two elements of the compound, rather than the two individual words, determines the meaning of the whole compound. Along these lines, semantic compositionality can be said to lie on a synchronic and diachronic continuum, with full compositionality at one end and no compositionality at the other end, and with various degrees of compositionality in between (Fernando 1996, Dirven and Verspoor 1998).

Similarly, Kavka (2009) argues that compositionality can be regarded as a scalar phenomenon, in which multiword expressions are viewed as fully compositional, such as *shoot a bird* and *red ink*; semi-compositional, such as *shoot a film* and *red carpet*; or non-compositional, such as *shoot the breeze* and *blue blood*. With respect to the difference between fully compositional and semi-compositional, Kavka (2009) states that *playground* is fully compositional, since it literally means a ground to play on, while *life boat* does not mean ‘\*a living boat’, rather ‘a boat used for saving lives’. Similar examples are *bulldog*, *horse-fly*, *stone-fish*, etc. He also indicates that the sequence of elements in the majority of non-compositional compounds is fixed, as in *lazybones* (not \**boneslazy*). With regard to their lexical flexibility, modifications are always external, that is, they modify the whole

---

<sup>4</sup>There is consensus concerning the concept of semantic transparency: it is usually considered to mean semantic compositionality. In fact, in a number of cases, compositionality, in relation to compounds, has a similar definition to semantic transparency. Roelofs and Baayen (2002: 132) indicate that “A morphologically complex word is semantically transparent if its meaning is compositional”. According to Girju et al. (2005: 488), “the meaning of compositional compounds can be successfully derived from the meaning of the noun constituents”. The distinction between semantic compositionality and semantic transparency will not be discussed any further here.

compound, rather than one or the other element, as in *an intolerable lazybones*, the word *intolerable* describes the whole compound.

According to Kavka (2009), the meaning of a multiword expression may become fixed through time, which reduces the degree of compositionality, resulting in it having a figurative meaning. This may suggest that the expression moves along the scale of compositionality. Historically, compositionality can be regarded as a cline on which multiword expressions are situated. Expressions which are referred to as idiomatic are on the non-compositional side of the cline.

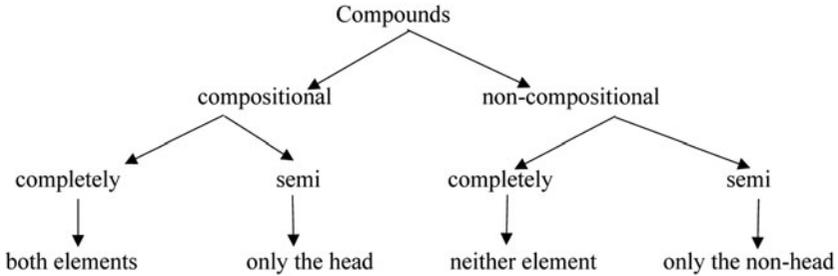
Fernando (1996), and Dirven and Verspoor (1998)'s classifications, as well as Kavka's (2009) proposal, which are quite similar, can also be applied to compounds in JA. If we also take into account that compositionality can be regarded as a scalar phenomenon with fully compositional at one end, and fully non-compositional at the other end, with some intermediate degrees in between, and if such degrees of compositionality can be based on the semantic contribution of the head and the non-head to the meaning of the whole compound, then a new, slightly modified, classification can be proposed for both English and JA. That is, if only the head of the compound contributes to the meaning of the whole compound, then it is semi-compositional, whereas if only the non-head of the compound contributes to the meaning of the whole compound, then it is semi non-compositional. This is illustrated in Figure 1 below.

Examples of each type of compound described in Figure 1 from English are provided below:

- |                  |   |
|------------------|---|
| (23) swansong    | (completely non-compositional compound) |
| (24) scarecrow   | (semi-non-compositional compound)       |
| (25) bookshop    | (completely compositional compound)     |
| (26) high street | (semi-compositional compound)           |

Examples (23) and (24) show two levels of non-compositionality. The meaning of the compound *swansong* 'farewell performance' in (23) is derived from neither *swan* nor *song*. That is, *swansong* IS NOT A *song*, which means that the compound *swansong* is semantically headless and does not denote a hyponym of its head *song*. The compound *scarecrow* 'an object to scare birds' in (24) is also semantically headless since *scarecrow* IS NOT A *crow*, which means the compound does not denote a hyponym of its head. However, the meaning of the non-head, *scare*, seems to contribute to the whole meaning of the whole compound.

In addition, two levels of compositionality can be depicted in English compounds in (25) and (26). In (25), the compound *bookshop* is endocentric, that is, semantically headed, since *bookshop* IS A *shop*, suggesting that the compound denotes a hyponym of its head. The meaning of the compound *bookshop* 'a shop where books are sold' is a sum of its two parts *book* and *shop*; thus it is completely compositional. In contrast, *high street* in (26) is also endocentric since *high street* IS A *street*, which means that the compound is a hyponym of its head. However, *high street* is not necessary *high*, which means that only the head contributes to the meaning of the compound, making it semi-compositional.



**Figure 1:** Levels of compositionality vs. non-compositionality in compounds

#### 4. METHODOLOGY

The analysis in the current study is focused mainly on JA compounds in comparison with English compounds. For the purpose of this study, the examples that are analysed were collected from various resources, including: (1) some research articles on JA compounding such as Altakhaineh (2016, 2017, 2019) and Zibin and Altakhaineh (2018); (2) some social media websites, such as Facebook and Twitter; and (3) my intuition as a native speaker of Jordanian Arabic (see Altakhaineh 2016). I collected around 60 examples of JA compounds. These compounds were classified into four degrees of compositionality, namely, completely compositional, semi-compositional, semi non-compositional and completely non-compositional, based on the judgement of 40 native speakers of JA. The 40 informants recruited in the current study are residents of Amman (the capital of Jordan) and are part of the same social network. They are students majoring in linguistics at the University of Jordan. Their mean age is 21. They are all native speakers of JA with working knowledge of Modern Standard Arabic (henceforth, MSA). I informed the participants that taking part in the study was voluntary and they could withdraw at any time during the experiment. Each participant was given a compositionality judgement task in which he/she was asked to classify 60 JA compounds as: 1) completely compositional; 2) semi-compositional; 3) semi non-compositional; and 4) completely non-compositional. To ensure that the participants understood the task, the compounds were all contextualised, and I explained to the informants the notion of compositionality and the four above degrees with illustrative examples. The informants were also asked to explain their choice briefly. Only the compounds that were agreed upon by 80% of the informants (50 out of 60 examples) were included in the study.

The current study adopted a corpus-driven approach to data analysis. According to Tognini-Bonelli (2001: 84), more insightful results can be yielded when the researcher adopts a corpus-driven approach in comparison with a corpus-based approach. This is because in the corpus-driven approach, the researcher analyses all the data in the corpus, which provides more accurate results. In corpus-driven studies, the researcher usually does not have preconceived hypotheses which he/

she intends to test based on the corpus. Conversely, in this approach the researcher starts with the corpus itself (Tognini-Bonelli 2001: 84). Nevertheless, some researchers (e.g., Zibin 2018) have argued that in reality the majority of corpus-driven studies do start with some preconceptions about what they are going to find in the corpus, which means that these studies at least have an initial prediction with regards to the findings. Still, in these kinds of studies, initial predictions are examined by analysing the entire corpus. (Deignan 2008: 282). That is, since the corpus size is small, the researcher can manually analyse the entire corpus and even if he/she has preconceived notions, these can be examined thoroughly. In large-size corpora, the researcher is unable to manually examine the entire corpus. Rather, an automated software-based search is normally performed. Studies which adopted a corpus-driven approach indicated that more accurate results are elicited when the researcher analyses all the data in the corpus (e.g., Moon 1998, Zibin 2018). Hence, in this study, I analysed all the instances of compounds collected from different sources and made judgements regarding their compositionality, as well as consulting with the native speaker informants who took part in the study.

## 5. COMPOSITIONALITY IN JA

This section classifies the JA compounds into four levels, based on their compositionality. For JA, examples of each type of compound described in Figure 1 are provided below (see Altakhaineh 2019 for more examples of N + N compounds in JA):

### 5.1 Completely non-compositional compounds (neither element)

Examples (27–29) were judged by 90% of the informants to be completely non-compositional, since the meaning of compound is not derived from either of its elements.

- (27) *yazl/fafar*    *l-banaat*    (completely non-compositional compound)  
 spinning/hair    the-girls  
 ‘the candyfloss’  
 lit. the girls’ hair
- (28) *lifb*    *fyaal*    (completely non-compositional compound)  
 playing    children  
 ‘something which is insignificant’  
 lit. children’s play
- (29) *?abuu*    *l-hool*    (completely non-compositional compound)  
 father    the-terror  
 ‘the sphinx’  
 lit. the father of terror

In example (27), *fafar lbanaat* ‘candyfloss’ IS NOT *fafar* ‘hair’. Semantically, it is neither related to *fafar* ‘hair’ nor to *lbanaat* ‘the girls’. The same applies to examples (28–29) where the meaning of the compound *lifb fyaal* ‘something which is insignificant’ in (28) is neither *lifb* ‘playing’ nor *fyaal* ‘children’ and the meaning of *?abuu lhool* ‘the sphinx’ in (29) is not related to either of its constituents. This

suggests that when there is no contribution of the meaning of either the head or the non-head to the meaning of the entire compound, this compound can be regarded as completely non-compositional. This also suggests that this type of compound is semantically headless since the compounds *faʕar lbanaat* ‘candyfloss’, *liʕb ʕyaal* ‘something which is insignificant’, and *?abuu lhool* ‘the sphinx’ do not denote hyponyms of either element, that is, the semantic head is outside the compound.

## 5.2 Semi-non-compositional compounds (only the non-head)

The following are examples of semi-non-compositional compounds in JA.

- (30) ʕarus            l-baħir    (semi-non-compositional compound)  
 bride            the-sea  
 ‘the mermaid’  
 lit. the sea bride

- (31) raaʔid            l-fadʕaaʔ    (semi-non-compositional compound)  
 pioneer            the-space  
 ‘astronaut’

lit. pioneer of the space

- (32) faras            n-nahir    (semi-non-compositional compound)  
 horse            the-river  
 ‘the hippopotamus’  
 lit. the river horse

Comparing examples (27–29) with (30–32) shows that two levels of non-compositionality can be distinguished in JA compounds. Specifically, JA compounds can be completely non-compositional when the meaning of the whole compound (e.g., candyfloss) is not related to the meaning of both elements combined. Conversely, *ʕarus lbahir* ‘mermaid’, in example (30), IS NOT A *ʕarus* ‘bride’. However, *ʕarus lbahir* ‘mermaid’ is a mythical creature that lives in *lbahir* ‘the sea’. Thus, at least one element of the compound, in this case the non-head, contributes to the meaning of the whole compound. Hence, it is semi-non-compositional. The same can be noted in example (31) in which the non-head *lfadʕaaʔ* ‘the space’ contributes to the meaning of the compound *raaʔid lfadʕaaʔ* ‘astronaut’ since he/she works in space despite not being a pioneer. Similarity, in (32) *faras nnahir* ‘the hippopotamus’, the non-head *nnahir* ‘the river’ contributes to the meaning of the compound, since hippos usually inhabit rivers. These examples were judged by 80% of the informants as semi-non compositional. It can be observed that both completely non-compositional and semi-non-compositional compounds are exocentric. In particular, both types are semantically headless, since they do not denote hyponyms of either of their elements. The interesting point here is that even when compounds are semantically headless (i.e., exocentric), in semi-non-compositional compounds, the non-head still contributes to the meaning of the compound even if it does not denote a hyponym of either of its elements. This may suggest that the lack of a semantic head does not necessarily imply that the meaning of the compound is not derivable from the meaning of its non-head.

It can also be noted that some compounds, such as *ʕaruus lbahir* ‘the mermaid’, are based upon metaphors. The issue of whether metaphoric/metonymic compounds are endocentric or exocentric has been subject to a great deal of debate in the relevant literature. For some researchers (e.g., Guevara and Scalise 2009), compounds based on metaphor and/or metonymy are regarded as exocentric. Bauer et al. (2013: 478) suggest that bahuvrihis<sup>5</sup> are exocentric, but leave open the possibility of analysing them as endocentric because of their figurative reading. They indicate that bahuvrihis of this type are productive in English, for example *air head*, *bone head*, *acid head*, *cheesehead*, *butterhead*, *bottle head* and *redhead*, which are all metonymic in nature. Bauer et al. (2013: 478) conclude that the main difference between bahuvrihis and endocentric attributive compounds is that the referents of bahuvrihi compounds are characterized figuratively, normally being metaphorical or metonymic. Therefore, analyzing them as regular endocentric compounds with a metonymic or metaphorical interpretation of the head noun is possible (Bauer et al. 2013: 478-9). Andreou and Ralli (2015) argued that nominal bahuvrihis should be analyzed as endocentric compounds based on metonymy. Other researchers (e.g., Benczes, 2006) proposed the umbrella term “creative compound” for any metaphorical and/or metonymical compound, and discarded the traditional “endocentric” and “exocentric” labels. It can be observed that whether compounds based upon metaphors are endocentric or exocentric is non-conclusive; yet, based on the analysis of JA compounds in this study, the compounds based upon metaphor usually lack a semantic head, and as such can be regarded as exocentric.

### 5.3 Completely compositional compounds (both elements)

Examples (33–35) are of completely compositional compounds.

- (33) ʕasʕiir      t-tuffaah      (completely compositional compound)  
juice          the-apple  
‘the apple juice’
- (34) findʒaan      f-faay          (completely compositional compound)  
cup            the-tea  
‘the tea cup’
- (35) muʕallim      l-fiizyaaʔ      (completely compositional compound)  
teacher        the-physics  
‘the physics teacher’

---

<sup>5</sup>Bahuvrihi is a hyponym of an implicit or unexpressed semantic head. For instance, the bahuvrihi *hardhat* does not denote a special kind of *hat*, but refers to an individual who uses, possesses or is characterised by that kind of hat. Its semantic head is not explicitly expressed, rather it is implicitly understood, as being ‘person/one’. Čermák (1997: 13) notes that as a consequence, it is not possible to analyse a bahuvrihi compound into its immediate elements; it is solely interpretable as predicated of an unrealised ‘third party’. The lack of a semantic head and the external nature of their reference means that bahuvrihis are structurally exocentric (Čermák 1997).

Two levels of compositionality can be detected in compounds in JA. For example, *ʕasʕiir ttuffaah* ‘the apple juice’, in example (33), IS A *ʕasʕiir* ‘juice’. Semantically, *ʕasʕiir ttuffaah* is juice made from apples. Hence, this compound is completely compositional. Similarly, *findʒaan ffaay* in example (34) IS A *findʒaan* ‘cup’ and both elements contribute to the meaning of the whole compound. The same applies to example (35). These examples are also endocentric, which means that they are semantically headed and the compound denotes a hyponym of its head. That is, *ʕasʕiir ttuffaah* ‘the apple juice’ is a hyponym of ‘juice’, *findʒaan ffaay* ‘the tea cup’ is a hyponym of ‘cup’ and *muʕallim lfiiɣyaaʔ* ‘the physics teacher’ is a hyponym of ‘teacher’.

#### 5.4 Semi-compositional compounds (only the head)

The following are examples of semi-compositional compounds.

- (36) *hajar z-zaawiyah*<sup>6</sup> (semi-compositional compound)  
 stone the-corner  
 ‘the cornerstone’
- (37) *baab l-ʕaamood* (semi-compositional compound)  
 door the-pillar  
 ‘one of the doors of Al-Aqsa Mosque’  
 lit. the door of the pillar
- (38) *ʔaklit l-muluuk* (semi-compositional compound)  
 meal the-kings  
 ‘a meal of white beans’  
 lit. the meal of kings

The compound *hajar zzaawiyah* ‘the cornerstone’, in example (36), IS A *hajar* ‘stone’, which is used as a foundation in any building. In other words, it is not part of the second element *zzaawiyah* ‘the corner’; the relationship between the elements is determinative rather than associative (see Adams 1973, Olsen 2000: 279). Thus, this compound is semi-compositional, because only the head contributes to the meaning of the whole compound (see Figure 1). The same applies to example (37) *baab lʕaamood* which IS A *baab* ‘door’ but is not part of the second element *lʕaamood* ‘the pillar’ and thus was judged by 90% of the informants as semi-compositional. Example (38) *ʔaklit lmuluuk* IS A *ʔaklih* ‘meal’ but is not related to kings, at least not contemporarily. It can be observed that both completely-compositional and semi-compositional compounds are endocentric, as they are semantically headed and they denote hyponyms of their heads. This can be seen as the exact opposite of completely non-compositional and semi-non-compositional compounds, since the latter are both exocentric; and while in non-compositional compounds neither elements contribute to the meaning of the compound, in semi-non-compositional ones, only the non-head does. When the compound is semantically headless, that is, exocentric, neither element, or only the non-head, can contribute to the meaning of the

<sup>6</sup>Note that the compound *hajar zzaawiyah* can be used metaphorically to refer to ‘something that is crucial in any process (political, social, economic, etc.)’.

compound, whereas when the compound is semantically headed, that is, endocentric, either both elements, or only the head, contributes to the meaning of the compound. It can be observed that there is an interplay between the notions of endocentricity and exocentricity, on the one hand, and semantic compositionality, on the other.

Note that many examples of non-compositional compounds in JA may have been borrowed from other languages, such as English, and integrated into JA at some point in history. JA does not have one-to-one equivalents for words like ‘mermaid’, that is, *hūuriyyat lbahr* ‘beautiful woman in the sea’.<sup>7</sup> Therefore, JA uses two words to describe one lexical item. This can explain why N + N constructs are very productive. Note, also, that the outcome of this integration is, in many cases, non-compositional.

Out of the total 50 examples judged by the 40 informants based on their level of compositionality, four compounds were found to be completely non-compositional, seven semi-non compositional, 33 completely compositional, and six semi-compositional. Based on the data collected in the current study, JA compositional compounds were more frequent than non-compositional ones, and by extension endocentric compounds were more frequent than exocentric ones. This tentative result may have implications for the productivity of these types of compounds. That is, compositional compounds in JA, both completely and semi, could be more productive than non-compositional ones, both completely and semi (see Lieber 2005: 376).

## 6. CONCLUSION

The examination of compounds in JA and English has revealed that there is an interplay between the notions of endocentricity vs. exocentricity, on the one hand, and semantic compositionality on the other. That is, if the compound is endocentric, that is, semantically headed based on Allen’s (1978) IS A principle, both of its elements, or only the head, contributes to the meaning of the compound. Conversely, if the compound is exocentric, that is, semantically headless, then neither element, or only the non-head, contributes to the meaning of the compound. In particular, there are four degrees of compositionality based on how the head and/or the non-head of a compound contributes to the meaning of the whole compound. It has been indicated that if the meaning of the whole compound: (1) is derived from both the head and the non-head, then it is completely compositional; (2) is not derived from either the head or the non-head, then it is completely non-compositional; (3) is only derived from the head, then it is semi-compositional; and finally (4) is only derived from the non-head, then it is semi non-compositional. This new taxonomy based on Fernando (1996), Dirven and Verspoor (1998), and Kavka’s (2009) classifications helps morphologists have a better understanding of the internal structure of compounds in English and JA. It also explains the role of the semantic head in coining new compounds in English and JA. It was proposed that this study may have implications pertaining to the productivity of compounds in JA, based on their degree of compositionality, that is, compositional compounds (both

<sup>7</sup>*hūuriyyah* is the most beautiful young woman with a fair skin found in heaven.

completely and semi) could be more productive than non-compositional ones (both completely and semi). Thus, it is suggested that a study to measure the productivity of the four types of compositionality is needed. Finally, it is recommended that this taxonomy should be applied to other languages to verify its reliability cross-linguistically.

## REFERENCES

- Adams, Valerie. 1973. *An introduction to Modern English word-formation*. London: Longman.
- Allen, Margaret Reece. 1978. *Morphological investigations in English*. PhD dissertation: University of Connecticut.
- Altakhaineh, Abdel Rahman Mitib. 2016. Headedness in Arabic compounds within the synthetic genitive construction. *SAGE Open* 6(4): 1–16. <doi.org/10.1177/2158244016674514>
- Altakhaineh, Abdel Rahman Mitib. 2017. Arabic compounds within the cross-linguistic compound taxonomy of Scalise and Bisetto (2009). *Lingue e Linguaggio* 16(1): 101–118. <doi.org/10.1418/87002>
- Altakhaineh, Abdel Rahman Mitib. 2019. Identifying N+ N compounding in Modern Standard Arabic and Jordanian Arabic. *Studia Linguistica* 73(1): 1–36. <doi.org/10.1111/stul.12087>
- Andreou, Marios, and Angela Ralli. 2015. Form and meaning of bahuvrihi compounds: Evidence from Modern Greek and its dialects. In *Semantics of complex words*, ed. Laurie Bauer, Lívia Körtvélyessy and Pavol Štekauer, 163–185. Dordrecht: Springer
- Bauer, Laurie. 1983. *English word formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 1998. When is a sequence of two nouns a compound in English? *English Language and Linguistics* 2: 65–86.
- Bauer, Laurie. 2009. Typology of compounds. In *The Oxford handbook of compounding*, ed. Rochelle Lieber and Pavol Štekauer, 343–356. Oxford: Oxford University Press.
- Bauer, Laurie, Rochelle Lieber, and Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Benczes, Réka. 2006. Creative compounding in English: the semantics of metaphorical and metonymical noun–noun combinations. Philadelphia: John Benjamins.
- Booij, Geert. 2002. ‘Constructional idioms, morphology, and the Dutch lexicon. *Journal of Germanic Linguistics* 14(4): 301–327.
- Borer, Hagit. 2009. Afro-Asiatic, Semitic: Hebrew. In *The Oxford handbook of compounding*, ed. Rochelle Lieber and Pavol Štekauer, 386–399. Oxford: Oxford University Press.
- Carstairs-McCarthy, Andrew. 2002. *An introduction to English morphology*. Edinburgh: Edinburgh University Press.
- Čermák, Jan. 1997. A prow in foam: The Old English bahuvrihi compound as a poetic device. *Prague Studies in English*, 22, 13–31.
- Deignan, Alice. 2008. Corpus linguistics and metaphor. In *Cambridge handbook of metaphor and thought*, ed. Raymond W. Gibbs, 280–294. Cambridge: Cambridge University Press.
- Dirven, René, and Marjolyn Verspoor. 1998. *Cognitive exploration of language and linguistics*. Amsterdam: John Benjamins.
- Fabb, Nigel. 1998. Compounding. In *The Handbook of Morphology*, ed. Arnold M. Zwicky and Andrew Spencer, 66–83. Oxford: Blackwell.

- Fassi Fehri, Abdelkader. 2012. *Key features and parameters in Arabic grammar*. Amsterdam: John Benjamins Publishing.
- Fernando, Chitra. 1996. *Idioms and idiomacity*. Oxford: Oxford University Press.
- Girju, Roxana, Dan Moldovan, Marta Tatu, and Daniel Antohe. 2005. On the semantics of noun compounds. *Computer speech and language* 19(4): 479–496. <doi.org/10.1016/j.csl.2005.02.006>
- Guevara, Emiliano, and Sergio Scalise. 2009. Searching for universals in compounding. In *Universals in language today*, ed. Sergio Scalise, Elisabetta Magni and Antonietta Bisetto, 101–28. Dordrecht: Springer.
- Haspelmath, Martin. 2002. *Understanding morphology*. London: Hodder
- Ji, Hongbo, Christine L. Gagnéri, and Thomas L. Spalding. 2011. Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language* 65(4): 406–430.
- Kavka, Stanislav. 2009. Compounds and idiomatology. In *The Oxford handbook of compounding*, ed. Rochelle Lieber and Pavol Štekauer, 19–33. Oxford: Oxford University Press.
- Lieber, Rochelle. 2005. English word-formation processes. In *Handbook of word-formation*, ed. Pavol Štekauer and Rochelle Lieber, 375–428. Dordrecht: Springer.
- Lieber, Rochelle, and Pavol Štekauer. 2009. Introduction: Status and definition of compounding. In *The Oxford handbook of compounding*, ed. Rochelle Lieber and Pavol Štekauer, 3–18. Oxford: Oxford University Press.
- Marchand, Hans. 1960. *The categories and types of present-day English word-formation: A synchronic-diachronic approach*. Wiesbaden: Otto Harrassowitz.
- Marelli, Marco, and Claudio Luzzatti. 2012. Frequency effects in the processing of Italian nominal compounds: Modulation of headedness and semantic transparency. *Journal of Memory and Language* 66(4): 644–664. <doi.org/10.1016/j.jml.2012.01.003>
- Moon, Rosamund. 1998. *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Oxford University Press.
- Neef, Martin. 2009. Indo-European, Germanic: German. In *The Oxford handbook of compounding*, ed. Rochelle Lieber and Pavol Štekauer, 386–399. Oxford: Oxford University Press.
- Olsen, Susan. 2000. Compounding and stress in English: A closer look at the boundary between morphology and syntax. *Linguistische Berichte* 181: 55–69.
- Ralli, Angela. 2013. *Compounding in Modern Greek*. London: Springer Science and Business Media.
- Roelofs, Ardi, and Harald Baayen. 2002. Morphology by itself in planning the production of spoken words. *Psychonomic Bulletin and Review* 9(1): 132–138. <doi.org/10.3758/BF03196269>
- Selkirk, Elisabeth. 1982. *The syntax of words*. *Linguistic Inquiry Monographs* 7. Massachusetts: MIT Press.
- Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- Williams, Edwin. 1981. On the notion ‘lexically related’ and head of a word. *Linguistic Inquiry* 12(2): 245–274.
- Zibin, Aseel, and Abdel Rahman Mitib Altakhaineh. 2018. An analysis of Arabic metaphorical and/or metonymical compounds: A cognitive linguistic approach. *Metaphor and the Social World* 8(1): 100–133. <doi.org/10.1075/msw.16023.zib>
- Zibin, Aseel. 2018. The effect of the Arab Spring on the use of metaphor and metonymy in Jordanian economic discourse: A cognitive approach. *Review of Cognitive Linguistics* 16(1): 254–298. <doi.org/10.1075/rcl.00010.zib>