Phonological Data & Analysis Volume 4, Article 2: 1–21
(2022) https://doi.org/10.3765/pda.v4art2.61

Received 05 May 2021; revised 21 November 2021; accepted 20 February 2022.

Phonological effects on word order: AB constructions in Hebrew

Noa Handelsman[a] & Outi Bat-El[b*]

[a]Tel Aviv University – handelsman@mail.tau.ac.il
[b]Tel Aviv University – obatel@tauex.tau.ac.il

The literature on what we call AB constructions (freezes, irreversible binomials), such as *odds and ends* and *copy paste*, attributes the fixed word order to both phonological and nonphonological, mostly semantic constraints. However, some researchers attribute a prominent role to phonology, while others view semantics as the major contributor to word order of AB constructions. In this paper we evaluate the role of phonology in Hebrew AB constructions with reference to a harmonic grammar with weighted constraints, where constraint weight is calculated on the basis of its effect in our corpus. The grammar reveals that semantic constraints weigh more than phonological constraints in both the cumulative weight and the average weight. Nevertheless, phonology affects a great number of data items, in particular those where semantic constraints are mute. We thus conclude that although syntax and semantics are responsible for word order, phonology determines word order when the other modules do not have a say.

Keywords: binomials; fixed word order; phonology–semantics interaction; weighted constraints; Hebrew

1 Introduction: AB constructions

Word order is usually assigned by syntax. The phrases *queens chase kings in the park* and *kings chase queens in the park* are semantically different, as syntax assigns agent–patient relations according to the noun's position within the phrase. However, the phrases *queens and kings walk in the park* and *kings and queens walk in the park* are semantically identical. Despite the semantic identity, *kings and queens* is used much more often than *queens and kings*, to the extent that the order between the elements may be considered irreversible. Similarly, Malkiel (1959) notes with regard to *odds and ends* that "… an inversion of the two kernels – **ends and odds* – would be barely understandable to listeners caught by surprise" (p. 113).

There are several terms referring to this type of constructions, including 'fixed order coordinates' (Abraham 1950), 'irreversible binomials' (Malkiel 1959), and 'freezes' (Cooper and Ross 1975), all denoting fixed word order, and 'lexical pairs' (Kaye 2009), which does not. However, scholars agree that irreversibility is scalar; for example, 5.7% of the 2,720 binomials in Gustafsson (1976) were reversible. For this reason, Mollin (2014) proposed an (ir)reversibility scale, whereby each lexical pair is assigned an irreversibility score. In Mollin's corpus-based study, *king and queen* is positioned at one edge of the scale with the highest irreversibility score (100), together with *law and order* and many others, while *beliefs and*

* We thank Evan Cohen and the participants of TAU Phonology Circle for helpful comments and suggestions. All disclaimers apply.

values is at the other edge with the lowest score (50). In-between, one can find *fish and chips* (99.05), *salt and pepper* (82.45), and *family and friends* (76.62).

In this paper we use the general term AB constructions. We include under this term binomials (1a,b), which include a linker, and copulative compounds (hereafter: copulatives), also known as *dvandva* compounds (1c).

(1) AB constructions

a. *and*-constructions *odds and ends*
b. *or*-constructions *heads or tails*
c. copulatives *copy paste*

There is an extensive literature on AB constructions (see reviews on English AB constructions in Mollin 2014 as well as Kopaczyk and Sauer 2017), beginning with Sanskrit dvandva compounds (e.g. *mātā-pitarau* 'parents; lit. mother-father') in Pāṇini's grammatical treatise *Aṣṭādhyāyi*. There are many studies on English AB constructions (Jespersen 1905, Malkiel 1959, Cooper and Ross 1975, Gustafsson 1976, Gil 1989, McDonald, Bock and Kelly 1993, Benor and Levy 2006, Copestake and Herbelot 2011, Lohman 2014, Mollin 2014, Renner 2014, Kopaczyk and Sauer 2017), sometimes with a comparative orientation (Abraham 1950). Studies on other languages are limited to one or two for each language; e.g. Ancient Hebrew (Kaddari 1966, Avishur 1976), Basque (Jacobsen 1982), German (Lambrecht 1984), Hungarian (Pordány 1986), Arabic (Gorgis and Al Tamimi 2005, Kaye 2009), Italian (Masini 2006), Korean and Japanese (Kwon and Masuda 2019). In this paper we add Modern Hebrew to this list, with a comparative glimpse at the end.

Our study focuses on the principles that determine the preferred and sometimes irreversible order between the elements, and in particular on the interaction among them. There is a consensus with respect to the wide range of factors affecting word order in AB constructions – phonological and semantic, as well as pragmatic and psycholinguistic. However, some researchers attribute a prominent role to phonology (Renner 2014), or more specifically prosodic phonology (Jespersen 1905, Gil 1989), while others view semantics as the major contributor to word order of AB constructions, where semantics is often a cover term for all non-phonological principles (Abraham 1950, Cooper and Ross 1975, McDonald, Bock and Kelly 1993, Benor and Levy 2006).

In addition to the linguistic factors, various studies suggest a frequency-based principle, stating that the first element in an AB construction is more frequent than the second (Kaddari 1966, Wright, Hay and Bent 2005). Fenk-Oczlon (1989) argues that frequency alone is sufficient because it overlaps, at least partially, with other constraints (see §4 below), such as SHORT-LONG (frequent words are often shorter than infrequent ones) and PROTOTYPE FIRST (prototypical words are often frequent). This follows form Zipf's (1935/1965) principle of least effort, which connects between structural complexity and the frequency of its use. Although most linguists agree that frequency plays a role in language knowledge, the extent of its effect in the construction of grammar is controversial (see studies in Gülzow and Gagarina 2007).

We do not deny the potential role of frequency, but find it unlikely to play a major role in our model. First, Fenk-Oczlon (1989) acknowledges that frequency has a larger impact when the constructions are reversible, while the AB constructions in our database were irreversible, or with low degree of reversibility. Fenk-Oczlon (1989) also claims that frequency overlaps, at least partially, with other constraints (e.g. PROTOTYPE FIRST, SHORT-LONG) which are addressed in our paper. Moreover, Pinker and Birdsong (1979) managed to eliminate the effect of frequency in their experimental study by using nonce-words, showing that SHORT-LONG is active without the effect of frequency. [1]

[1] In addition, there are methodological reasons for not addressing frequency. First, the AB constructions in our database were coined during different time periods, and therefore a corpus-based frequency check of each element in a construction would be misleading; for example, both words in *kro u xtov* 'reading and writing' are rarely used today,

We present here a study of Hebrew AB constructions, aiming to shed light on the role of phonology in determining word order. We show that although the most prominent constraints in the grammar are semantic, the power of phonology emerges to a great extent. More specifically, the effect of phonology emerges when semantics is indifferent, i.e., in data items where there are no relevant semantic constraints; apparently, there are quite a few such data items. This state of affairs, where phonology does the job of other modules when they are mute, is introduced in Golston (1995), where prosody selects one out of two syntactically well-formed structures.

Our account of the interaction of semantics and phonology in the grammar of Hebrew AB constructions is couched within the framework of Optimality Theory (OT; Prince and Smolensky 2004 [1993]), which not only provides a solid infrastructure for constraint interaction, but also allows the interaction of constraints from different modules of grammar. We use Harmonic Grammar with weighted constraints (Legendre, Miyata and Smolensky 1990, Smolensky and Legendre 2006, Pater 2009, 2016), where the weight of the constraints is calculated on the basis of its violation/satisfaction in our corpus.

We continue this introduction with a brief review of studies on phonology and word order (§2) and then move on to our methodology (§3). First, we present our corpus (§3.1), and then discuss two tests that were essential to set the constraints: prototypicality test (§3.2) and syllabification test (§3.3). We then introduce all the constraints that were found to be relevant to our corpus (§4), distinguishing between non-phonological (§4.1) and phonological (§4.2) constraints. With the constraints at hand, we turn to the grammar of Hebrew AB constructions (§5), starting with the basic tenets of Harmonic Grammar and its unique properties compared to classic OT. We then introduce the model we used to assign constraint weight and provide the list of the relevant constraints and their weight (§5.1). At this point, we highlight the role of phonology in the grammar of Hebrew AB constructions, pointing out the conditions under which phonology prevails (§5.2), and continue with evaluating our grammar against new data items (§5.3). At the end, we provide a summary of constraint interaction, again highlighting the role of phonology (§5.4). Finally, we raise two questions that require further study (§6) and then conclude (§7).

2 Phonology and word order

The impact of phonology on word order is reflected in a general *end-weight tendency* (Wasow 1997; attributed to Behaghel 1909), whereby long constituents are pushed towards the end of the sentence. As in many cases of phonology–syntax interface (Inkelas and Zec 1990), the impact of phonology, in particular prosody, often arises when the syntax is mute (Golston 1995), although there is often a default structure.

The most well-known phenomenon of the end-weight tendency is Heavy NP shift, whereby a long direct object lands at the end of the phrase rather than next to the verb (Hawkins 1983, Siewierska 1988, Stallings and Macdonald 2011). As exemplified below, the direct object usually precedes the indirect object (2a), unless the direct object is long, in which case it is shifted to the end of the sentence (2b); crucially, there is no NP shift when the direct object is short (2a).

(2) Heavy NP shift

a. Direct object precedes indirect object – default

i. Jack sent the book to Jill
ii. ?Jack sent to Jill the book

---

and so are the second word in *zug o peret* 'even or odds' and the first in moen ve niman 'sender and addressee'. Second, Hebrew has many homonyms and homographs, and therefore a frequency check of a written corpus would include many false results. For example, the written Hebrew word פרט corresponds to *péret* 'odd number' (as in *zug o peret* 'even or odds'), *prat* 'detail'/'except of', *parát* 'to cash a bill into smaller notes'/'to strum', *perét* 'to detail', and *porát* 'to be detailed'.)

b. Indirect object precedes direct object – when direct object is long

i. I $_{\text{VP}}$[introduced $_{\text{PP}}$[to Mary] $_{\text{NP}}$[some friends that John had brought to the party]]
ii. $^{?}$ I $_{\text{VP}}$[introduced $_{\text{NP}}$[some friends that John had brought to the party] $_{\text{PP}}$[to Mary]]

The function of Heavy NP shift is parsing facilitation (Hawkins 1990, 1994) and utterance planning (Wasow 1997), where the measuring units of length are phonological (segments or syllables, though some argue for grammatical length, where the measuring units are words (McDonald, Bock, and Kelly 1993)). Word length is, of course, one of several non-syntactic factors determining word order, as shown in Bresnan et al.'s (2007) study of English dative constructions (e.g. … *gave books to all the children* vs. … *gave all the children books*); here, the length of the recipient interacts with other factors such as givenness, definiteness and animacy.

The selection between English *s*-genitive and *of*-genitive (e.g. (e.g. *the man's computer* vs. *the computer of the man*) is shown to be affected by phonology, particularly rhythm and stress pattern. Shih et al. (2015) argue that in addition to semantic, pragmatic and processing factors, the Principle of Rhythmic Alternation (Selkirk 1984) is a reliable predictor of the selection of genitive construction in their corpus of spoken English.

The effect of phonology on word order is also found in Tagalog. In their corpus-based study of adjective–noun phrases, Shih and Zuraw (2017) found that the free variation in the order of adjectives and nouns in Tagalog is not at all free, but rather imposed by segmental and prosodic constraints. For example, the default adjective–noun order in *sawíng báyan* 'unfortunate country/people' leads to stress clash, and therefore the reverse order *báyang sawíʔ* arises.

The topic of the present study is the effect of phonology on word order in AB constructions. While in Tagalog there is a default word order that phonology can adjust, in AB constructions there is no default order – word order is determined by phonological and semantic constraints. For example, in *art and design* the short element precedes the long one (phonology), and in *right and wrong* the positive element precedes the negative (semantics). Cooper and Ross (1975) propose the cover principle *me first*, stating that "first conjuncts refer to those factors which describe the prototypical speaker" (p. 67). This principle includes various semantic subprinciples, many of which adopted from earlier studies (Abraham 1950, Malkiel 1959); for example, positive–negative, light–dark, male–female, animate–inanimate, near–far, present– future, top–bottom (where X–Y means X comes before Y).[2] These principles interact with various phonological principles, including short–long, which falls under the above-mentioned end-weight tendency.

There is a consensus among the researchers with regard to the interaction of phonological and semantic principles in determining the order in AB constructions. However, it is not yet clear which of these two modules is the major contributor. While some researchers claim that semantic factors prevail (McDonald, Bock and Kelly 1993, Benor and Levy 2006, Mollin 2012), others assert that phonology, in particular prosody is the major contributor to word order in AB constructions (Gil 1989, Wright, Hay and Bent 2005, Renner 2014).

In this paper, we present a study of Hebrew AB constructions, aiming to shed light on the role of phonology in determining word order. We show that although some semantic constraints are rarely violated (e.g. MALE-FEMALE), the power of phonology may emerge to the extent that in some data items phonology is the major contributor.

[2] It must be said out loud that the correlations implied from *Me First* (e.g. positive–negative, light–dark, male–female) are, at best, inappropriate.

3 Methodology

In this section we provide details on the selection of the data items on the basis of the irreversibility score (§3.1), and introduce two preliminary tests which were conducted in order to determine how to set some of the constraints: the prototypicality test (§3.2) and the syllabification test (§3.3).

3.1 The data

In order to determine which AB constructions are in fact irreversible, we adopted Mollin's (2012) notion of (ir)reversibility scale, whereby AB constructions are assigned with a reversibility score. Our corpus, which was initially comprised of 95 Hebrew AB constructions, was drawn from written material and native speakers' intuition. For each of the 95 items, the frequency of both possible orders was indicated on the basis of google search results. Then, the reversibility score was calculated with Mollin's (2012) formula *[freq/(freq+revfreq)]×100*, where *freq* corresponds with the frozen order, and *revfreq* corresponds with the reverse order. For example, *mukdám o meuχáʁ* 'early or late' had 19,700 hits on a google search, while the reverse *meuχaʁ o mukdam* had 3220 hits; the irreversibility score given to this binomial was thus 86 ([19,700/(19,700+3220)]×100). The distribution of the 95 data items on the scale is given in Table 1.

Table 1: Irreversibility score

score		AB constructions	
100	Absolute	10	11%
85-99	Strong	56	59%
65-84	Moderate	18	19%
51-64	Weak	5	5%
0-50	Reversible	6	6%

For the purpose of our study, we eliminated the 6 reversible data items (score 0-50) and analyzed a total of 89 items (see Appendix A) with an average irreversibility score of 90. The data items are distributed as follows:

(2) The corpus of Hebrew AB constructions (n =89)

a. Binomials

And constructions 54 (e.g. *mélaχ ve-pílpel* 'salt and pepper')

Or constructions 7 (e.g. *paχót o-jotéʁ* 'less or more')

b. Copulatives 28 (e.g. *aaték adbék* 'copy paste')

In addition to the 89 items which were used to calculate the weight of the constraints, we examined 17 and-constructions consisting of personal names (listed in Table 4). Since all but two of the relevant constraints in binomial names are phonological, these 17 items allowed us to evaluate the model with emphasis on the phonological constraints (§5.3).

3.2 Prototypicality test

As will be elaborated in §4, the grammar presented in this paper consists of phonological and non-phonological constraints. One of the non-phonological constraints is the PROTOTYPE FIRST, which requires the first word in the AB construction be more prototypical than the second one. In order to determine which of our AB constructions satisfies this constraint, a prototypicality test was conducted.

Participants: 30 native speakers of Hebrew (age 20-40), with no previous knowledge of the study, participated in the test.

Procedure: The participants were given a written form with a list of all the AB constructions, and were asked to choose, within each construction, the element that is more central in the world (the term prototypical was not used). The participants were requested to be as objective as possible, disregarding their own personal preferences. In order to control the effect of word order on the participants' choice, 15 participants took the test where the order of constructions was AB, and another 15 participants took the test where the order of constructions was BA.

Results: Prototypicality was assigned to an item only when a statistically significant majority of the participants agreed upon it ($p<0.05$; a binomial test). For example, in item (a) in Table 2, A was significantly more central than B, and in item (b) it was the other way around; in item (c) there was no significant difference in the centrality of the two elements.

Table 2: Prototypicality test results

a.	klavím ve ganavím	A	klavím	'dogs'	21	p = .013
		B	ganavím	'thieves'	8	
b.	báal ve iʃá	A	báal	'husband'	9	p = .044
		B	iʃá	'wife'	19	
c.	áχ ve aχót	A	áχ	'brother'	12	p = .285
		B	aχót	'sister'	16	

Accordingly, the constraint PROTOTYPE FIRST was considered satisfied in (a), violated in (b), and not relevant in (c).

3.3 Syllabification test

We identified two cases of potential variation in syllabification, which could affect our model, since some of the phonological constraints refer to syllable structure.

The first case appears in constructions where a word with an initial consonant cluster is preceded by a vowel-final word (e.g. *neʃiká stiʁá* 'kiss slap') or a linker (e.g. *óχel ve ʃtijá* 'food and drinks'). There are two ways to syllabify such a sequence: (a) morphologically-sensitive syllabification, whereby the left edge of the word is aligned with a left edge of a syllable, thus preserving the cluster – *…ve.ʃtijá;* (b) morphologically-insensitive syllabification, whereby the two consonants of the input cluster are heterosyllabic – *…veʃ.tijá.* This variation is relevant to the constraint ALIGNLEFT(IC, PHPH), which requires the input initial cluster to reside at the edge of the AB construction (see §4.2). ALIGNLEFT is violated in morphologically-sensitive syllabification, where the cluster is at the middle of the AB construction – *…ve.ʃtijá*, but not in the morphologically-insensitive syllabification, where there is no cluster – *…veʃ.tijá*.

The second case of variation is found with the linker *ve* 'and', which has a normative allomorph *u* before labials and consonant clusters (Amir Coffin and Bolozky 2005:246).[3] The variation between *u* and *ve* is relevant to the constraint *VV, which prohibits two adjacent vowels; e.g. *kʁó u χtóv* 'reading and writing' violates *VV but *kʁó ve χtóv* does not. The *u* allomorph is limited to normative Hebrew and rarely used in casual speech. However, most speakers have knowledge of the allomorph, and therefore we had to eliminate this potential variation.

Participants: 10 Hebrew native speakers (ages of 22-35), with no previous knowledge of the study, participated in the test.

Material: All participants were presented with a list of 28 AB constructions from the corpus. The data items met the following conditions (i) an initial cluster in one of the two words (or final cluster, relevant to

[3] There is actually a third allomorph, *va*, but since we do not have constraints that refer to vowel quality, the difference between *ve* and *va* is not relevant.

only one data item) or (ii) the conjunctive marker appeared in the environment of *u* in normative Hebrew. That is, all items were subject to resyllabification.

Procedure: The participants were asked to read each item out loud 4 times: Two times in each order – AB vs BA, and within each order one time in a normal pace and another in a slow pace; the latter pace potentially provides syllable boundaries. In the slow pace mode, participants were requested to read slowly, following an audio example, but were not requested to syllabify the words. All participants performed 2 tryout readings prior to the test. All 270 readings were transcribed and examined for (i) resyllabification and (ii) the u allomorph of *ve*.

Results: Resyllabification occurred in only 2 items (out of 270 readings), in both the normal and the slow reading paces. Hence, we concluded that no resyllabification occurs and initial clusters remain intact, preserving morphology-sensitive syllabification.

The allomorph *u* appeared conclusively in two constructions only: *χomá u migdál* 'wall and tower' and *kró u χtóv* 'reading and writing'. Within these two AB constructions, the constraint *VV is violated (*χomá u migdál* and *kró u χtóv*) and hence its weight was calculated accordingly.

Relevant to *VV is our transcription method with regard to glottals. Since they are rarely produced in casual speech, orthographic glottals were not transcribed, giving rise to vowel initial words (e.g. *ód ve adáʁ* 'majesty and glory', *óʁ ve a͡tsamót* 'skin and bones'). Note that a literate speaker "knows" that *ód ve adáʁ* is actually *hód ve hadáʁ*, since the initial glottal fricative has a corresponding letter in the alphabet. However, the decision to eliminate glottals from the phonetic script is based on their poor representation in speech, to the extent that their phonemic status is questionable. Faust (2005) argues that a glottal in Modern Hebrew is actually the vowel /a/ thus supporting our glottal-less transcription.

4 The relevant constraints

In this section, we present the constraints relevant to the grammar of word order in the AB constructions in our corpus. There are two measures of relevancy – relevant to the corpus and relevant to the data item. A constraint is considered relevant to the corpus if the corpus includes one or more data items whose order cannot be accounted for without this constraint. A constraint is considered relevant to a data item if it does not obtain the same result in the two possible orders, AB and BA. All the constraints but one (NATIVE-LOAN; see §4.1) are drawn from the literature on AB constructions and/or the general literature on universal constraints. We assume that the constraints are universal; for some there is evidence in Pinker and Birdsong's (1979) experimental study and for others there is evidence in general linguistic studies.

As our main interest here is the role of phonology in determining word order, the constraints are divided into phonological and non-phonological constraints, where the latter ones are mostly semantic. Some of the constraints are presented in a A–B scheme, which indicates that A appears before B.

4.1 Non-phonological constraints

Below is the list of non-phonological constraints that were found relevant to the data items in our corpus.

TEMPORAL SEQUENCING (Tai 1985), also referred to as 'iconic sequencing constraint' (Malkiel 1959): A appears before B within the real-world referents (e.g. *avót ve baním* 'fathers and sons').

MALE-FEMALE (Cooper and Ross 1975): A is male and B is female (e.g. *áχ ve aχót* 'brother and sister').

Note that this constraint specifically refers to male and female, and not to masculine and feminine, as the latter include grammatical gender. Thus, the constraint is relevant to *aχ ve aχót* 'brother and sister' but not to *léχem ve χemá* 'bread and butter', where *léχem* is grammatically masculine and *χemá* is grammatically feminine.

POSITIVE NEGATIVE (Cooper and Ross 1975): A is objectively more positive than B (e.g. *kén o ló* 'yes or no').

PROTOTYPE FIRST (Kelly, Bock and Keil 1986): A is more prototypical than B within the same semantic field (e.g. *mélaχ ve plpel* 'salt and pepper').

NATIVE LOAN: A is a native word and B is a loanword (e.g. *taʁbúve spóʁt* 'culture and sports').

We limit loanwords to those exhibiting salient non-native features (Schwarzwald 1998). For example, *spóʁt* is a loanword since there are no Hebrew words with both initial and final clusters, but *salát* 'salad' is not considered a loanword because it is similar to native words like *ʃalát* 'remote control' and *gamád* 'dwarf'. [4]

4.2 Phonological constraints

Seven phonological constraints were found relevant to our corpus, most of which fall under syllable structure optimization (Benor and Levy 2006, Shih and Zuraw 2017). The constraints are defined below, accompanied with examples of the two possible orders, indicating with > that the first order is the most common one, and sometimes irreversible.

SHORT LONG (Cooper and Ross 1975): A is shorter than B in syllabic and/or segmental terms (e.g. *smól jamín* 'left right' < *jamín smól*).

In most studies, the units considered for the evaluation of SHORT-LONG are syllables, and in a few they are segments (Sobkowiak 1993). In most cases, the two measuring units converge (e.g. *χút ve máχat* 'thread and needle'), in some only the syllable is relevant (e.g. *lév ʁeá* 'heart lung'), and in yet others only the segments are relevant (e.g. *buʃá ve χeʁpá* 'shame and disgrace'). There is no data item in our corpus that respects SHORT-LONG with reference to one measure and violates it with reference to another (hypothetical *t͡ʃíps reá* 'chips lung'). Therefore, we combined the two measuring units into one constraint.

ALIGNL (IC, PHPH) (ALIGNL; McCarthy and Prince 1993, Benor and Levy 2006): A word initial cluster (IC) is aligned with the left edge of the phonological phrase (e.g. *gvinót ve jáin* 'cheeses and wine' > *jáin ve gvinót*).[5]

ALIGNL disfavors word initial clusters in the middle of the AB constructions, pushing the cluster to the left edge.[6] Recall from §3.3 that word boundary aligns with syllable boundary, and thus, an initial cluster in the second element persists when preceded by a vowel-final element (e.g. *ó.χel . ve.ʃti.já* 'food and drinks').

Cooper and Ross (1975) proposed a constraint with the opposite effect, giving priority to fewer consonants at the edges. However, based on experimental data from English and French, Pinker and Birdsong (1979) restate this constraint, which like ALIGNL, gives priority to complexity at the margins. This effect is universally more plausible since the left edge is known for its unique status in adult languages (Nelson 2003); it resists weakening caused by phonological alternation (Beckman 1998), and in some cases it also induces strengthening (Bat-El 2014).

*VV (Casali 1997, Benor and Levy 2006, Shih and Zuraw 2017): A sequence of two adjacent vowels is prohibited (e.g. *emét o χová* 'truth or dare' < *χová o emét*).

Vowel hiatus in Hebrew often persists (e.g. *ʃoéʁ* 'gate keeper', *maéʁ* 'fast', *jaavód* 'to work 3.MS.SG.FUT), more so in careful than in causal speech. However, there are derived environments where the constraint is active, and the VV sequence is resolved via vowel deletion (e.g. /asa-u/ → *asú* 'to do 3.PL.PAST') or glide epenthesis (e.g. /dati-a/ → *datijá* 'religious FM.SG'). We expect *VV to be active in AB constructions, not only because of the derived environment, but also because AB constructions reside in the periphery of the lexicon, where the effect of universal markedness constraints tends to emerge (McCarthy and Prince 1994, Bat-El 2005).

*CLASH (Liberman and Prince 1977, Hayes 1995, Kager 1993, 1999): A sequence of two adjacent stressed syllables is prohibited (e.g. *béten gáv* 'stomach back' > *gáv béten*).

[4] In Hebrew, word similarity is based on configurations (Bat-El 2011), which consist of prosodic structure, vocalic pattern and affixes (if any). For example, *mataná* 'gift' and *maxalá* 'disease' are similar, and so are *kélev* 'dog' and *béten* 'abdomen'.

[5] Note that we consider an AB construction a phonological phrase (PhPh) since the two words in the phrase preserve their primary stress. We also assume that the domain of ALIGNL is both the prosodic word and the phonological word.

[6] A parallel constraint, ALIGNR, disfavors word final clusters. For our purposes, only the left edge is relevant since there are very few word final clusters in Hebrew (Asherov and Bat-El 2019).

The effect of *CLASH emerges only in copulative compounds; in the other types of AB constructions there is an unstressed linker between the two elements.

*LAPSE (Selkirk 1984, McDonald, Bock and Kelly 1993, Benor and Levy 2006): A sequence of two or more adjacent unstressed syllables is prohibited (e.g. *milím ve láxan* 'lyrics and melody' > *láxan ve milím*).

There is no secondary stress in Hebrew (Becker 2002, Cohen, Silber-Varod and Amir 2018) and thus all syllables that do not carry main stress, including the syllable of the linker, are unstressed. *LAPSE violations were counted in derived environment only, i.e. when the sequence of two unstressed syllables was at a word boundary. For example, there is one violation of *LAPSE in *ón ve ʃiltón* 'wealth and authorities', and two violations in *óxel ve ʃtijá* 'food and drinks'. In *nisím ve niflaót* 'miracles and wonders' there is a sequence of three unstressed syllables, as in *óxel ve ʃtijá*, but only two of the unstressed syllables is at a word boundary and therefore *nisím ve niflaót* has only one *LAPSE violation.

THE OBLIGATORY CONTOUR PRINCIPLE (OCP; Goldsmith 1976): Adjacent segment sharing place of articulation are prohibited (e.g. *χamút͡s matók* 'sour sweet' > *matók* χamút͡s).

5 The grammar of AB constructions

The formal analysis of Hebrew AB constructions presented in this section adopts the framework of Harmonic Grammar (HG; Legendre, Miyata and Smolensky 1990, Smolensky and Legendre 2006, Pater 2009, 2016), which differs from classic OT (Prince and Smolensky 2004[1993]) in the constraint scheme; in classic OT the constraints are strictly ranked while in HG they are weighted.

Consider the schematic tableaux in (3) below. In (3a), A is the optimal candidate because its competitor B violates the dominant constraint CON1 and thus eliminated. The fact that candidate A violates more constraints than candidate B is irrelevant, as the strict ranking is the significant factor. This OT tableau contrasts with the HG tableau in (3b), as in the latter candidate B is optimal. Here we see the gang effect (Pater 2009); although CON1 weighs more than CON2 and CON3 independently, the cumulative weight value of the latter two allows them to gang up against CON1.

(3) Ranked vs. weighted constraints

a. OT – A is optimal

	CON1	CON2	CON3
a. ☞ Candidate A		*	*
b. Candidate B	*!	*	**

b. HG – B is optimal

	CON1 [1.5]	CON2 [1]	CON3 [1]	$\mathscr{H}$
a. Candidate A		-1	-1	-2
b. ☞ Candidate B	-1			-1.5

In HG tableaux, the weight of constraints appears below their names, and a violation is indicated by the negative integer -1. The cumulative value of violations for each candidate is on the rightmost column, under $\mathscr{H}$, where the winning candidate is the one with the cumulative value closest to 0.

5.1 The weight of the constraints

Weight value is often assigned with arbitrary integers (Pater 2009, 2016), but here we drew the weight values from the corpus. Each constraint in the two lists in §4 was assigned a numerical weight according to the formula X:Y, where X is the number of times the constraint was satisfied in the data items, and Y is the number of times the constraint was relevant, i.e. satisfied and/or violated. A constraint is irrelevant when its environment is not met (e.g. the environment of MALE-FEMALE is not met in *mélaχ ve pílpel* 'salt and pepper'), or when both AB and BA equally fare with respect to a constraint (e.g. *VV is equally violated in *áχ ve aχót* 'brother and sister' and *aχót ve áχ*). Constraint weight could be any number between 0 (never satisfied in our corpus) to 1 (never violated in our corpus). For example, SHORT-LONG is satisfied by 53 data items and violated by 18, and its weight value is thus 0.75 (=53:(53+18)). Similarly, 0.83 (=38:(38+6)) is the weight value of PROTOTYPE FIRST, and 0.41 (=13:(13+19)) of *LAPSE. We used 89 data items to assign weight, all with two elements, with or without a linker (see the list in Appendix A).

We started the analysis with a wide range of constraints from the literature on AB constructions, and gradually eliminated those that are not relevant to the data or do not add the efficiency of the grammar (i.e. the results are identical with and without them). The constraints that were eventually found relevant for our data items are provided in Table 3 below (see §4 for definitions), with their weights at the rightmost column (from the highest to the lowest value).

Table 3: Constraint weight (phonological constraints are shaded)

Constraint	Satisfactions	Violations	Weight
MALE-FEMALE	5	0	1.00
TEMPORAL ICONICITY	16	1	0.94
POSITIVE-NEGATIVE	12	2	0.86
PROTOTYPE FIRST	38	6	0.83
*CLASH	8	2	0.80
SHORT-LONG	53	18	0.75
ALIGN LEFT	7	4	0.64
NATIVE LOAN	3	2	0.60
OCP	12	12	0.50
*VV	9	10	0.47
*LAPSE	13	19	0.41

There are 11 active constraints in the grammar of Hebrew AB constructions – 5 non-phonological, mostly semantic, with an overall weight of 4.23, and 6 phonological with an overall weight of 3.57. At the top of the list are four semantic constraints – MALE-FEMALE, TEMPORAL-ICONICITY, POSITIVE-NEGATIVE and PROTOTYPE FIRST, where the first one is never violated in this corpus. These details alone support the studies that attribute word order in AB constructions primarily to semantic/pragmatic factors (Cooper and Ross 1975, Benor and Levy 2006, Mollin 2012). However, as we show in §5.2 below, the effect of phonology emerges under several circumstances.

5.2 Where does phonology prevail?

As noted, there is a consensus in the literature that both phonological and non-phonological factors play a role in determining word order in AB constructions, and this is also true for Hebrew AB constructions. As shown below, there are cases where semantics prevails (4a) and others phonology prevails (4b).

(4) Constraint interaction

a. Semantics prevails: *χatúl ve aχbáʁ* ‘cat and mouse’

	Prototype First [0.83]	*VV [0.47]	ℋ
a. ☞ χatúl ve aχbáʁ		-1	-0.47
b. aχbáʁ ve χatúl	-1		-0.83

b. Phonology prevails: *báʁ misadá* ‘bar-restaurant’

	*Clash [0.80]	Short-Long [0.75]	Native-Loan [0.60]	ℋ
a. ☞ báʁ misadá			-1	-0.60
b. misadá báʁ	-1	-1		-1.55

Here we focus on the role of phonology in determining word order in Hebrew AB constructions, and thus consider the cases where phonology prevails. We start with the cases where only phonological constraints are relevant, either without constraint interaction (5a), or with interaction between them (5b).

(5) Only phonological constraints are active

a. All constraints are satisfied: *χút ve máχat* ‘thread and needle’

	Short-Long [0.75]	*Lapse [0.41]	ℋ
a. ☞ χút ve máχat			0
b. máχat ve χút	-1	-1	-1.16

b. Constraint interaction: *béten gav* ‘stomach back’

	*Clash [0.80]	Short-Long [0.75]	OCP [0.50]	ℋ
a. ☞ béten gav		-1		-0.75
b. gáv béten	-1		-1	-1.3

Despite the fact that the top four constraints on the weight list are semantic constraints (Table 3), there are two conditions under which phonology prevails in its interaction with semantics. The first condition is when the weight of the semantic constraint is relatively low (6a), and the second is a gang effect (see §5), whereby the cumulative weight of the phonological constraints outweighs the weight of the semantic constraint (6b).

(6) Phonology – semantics interaction

a. *báʁ misadá* ‘bar restaurant’

	*Clash [0.80]	Short-Long [0.75]	Native-Loan [0.60]	ℋ
a. ☞ báʁ misadá			-1	-0.60
b. misadá báʁ	-1	-1		-1.55

b. *bejt͡sá ve taʁnególet* 'egg and chicken' – gang effect

	PROTOTYPE FIRST [0.83]	SHORT-LONG [0.75]	*LAPSE [0.41]	H
a. ☞ bejt͡sá ve taʁnególet	-1		-1	-1.24
b. taʁnególet ve bejt͡sá		-1	-2	-1.57

To summarize, phonology plays a role in determining word order in three cases: in data items where only phonological constraints are relevant (5); in data items where the relevant non-phonological constraint is outweighed by phonological constraints (6a); and in data items exhibiting gang effect (6b).

5.3 Focusing on the phonological constraints: Binomial names

In order to further evaluate the role of phonology in determining word order in Hebrew AB constructions, we examined binomial names, where only two non-phonological constraints are potentially active – MALE-FEMALE and NATIVE-LOAN; all other relevant constraints are phonological. The corpus, presented in Table 4 below, consists of 17 binomial names, including celebrities and characters known as a duo (cf. English *Tom and Jerry, Romeo and Juliette, Will and Grace*; Wright, Hay and Bent 2005, Tachihara and Goldberg 2020).

Table 4: Binomial personal names

M – male, F – female, N – native, L – loan

	M-M (n=8)	F-F (n=2)	M-F (n=6)	F-M (n=1)
N-N	ʃáj ve dʁóʁ tál ve aviád ási ve gúʁi davíd ve golját	t͡síli ve gíli óʁna ve éla	dát͡s ve dát͡sa ílan ve ilanít amnón ve tamáʁ adám ve χavá ámi ve támi ʃimʃón ve dlilá	
L-L	lóndon ve kíʁʃenbaum gʁájnik ve álteʁman tájχeʁ ve zaʁχóvit͡ʃ			
N-L	éli ve maʁiáno			ofíʁa ve béʁkovit͡ʃ

The data items in Table 4 are grouped with reference to the two non-phonological constraints, thus allowing to highlight those that involve only phonological constraints (shaded). Overall, 77% (13/17) binomial names were accounted for by our grammar (cf. 83% (74/89); see (13) below). The four data items that were not accounted for were *t͡síli ve gíli*, where none of the constraints were relevant; *óʁna ve éla*, where the only relevant constraints – SHORT-LONG – is violated (with reference to the number of segments); and *ʃimʃón ve dlilá* and *tál ve aviád*, where the expected gang effect is not borne out.

Note that there are two binomial names under the M-F column in Table 4, where MALE-FEMALE converges with SHORT-LONG: *dát͡s ve dát͡sa* and *ílan ve ilanít*. This is due to the feminine suffixes *-a* and *-it*, which often appear in feminine names (e.g. masculine *ron* vs. feminine *róna* or *ronít*). This is also true for *áχ ve aχot* 'brother and sister' from the general data, but not for *bén o bát* 'boy or girl' or *báal ve iʃá* 'husband and wife'. That is, MALE-FEMALE is independently required.

In 7 out of the 8 binomial names of mixed gender, the constraint MALE-FEMALE determines the order. However, there is one – *ofíʁa ve béʁkovit͡ʃ* – where the female name appears first. As shown in (7) below, this is due to a gang effect, whereby the low-valued phonological constraints gang against the high-valued semantic constraint.

(7) A gang effect in binomial names: *ofíʁa ve béʁkovit͡ʃ*

	MALE-FEMALE [1]	SHORT-LONG [0.75]	NATIVE-LOAN [0.60]	*VV [0.47]	*LAPSE [0.41]	ℋ
a. ☞ ofíʁa ve béʁkovit͡ʃ	-1				-1	-1.41
b. béʁkovit͡ʃ ve ofíʁa		-1	-1	-1	-2	-2.64

5.4 Summary of constraint interaction

We developed the grammar of Hebrew AB constructions on the basis of 89 data items – 61 binomials and 28 copulatives, and examined it against 17 binomial names. However, as summarized in Table 5 below, different data items provided different types of information.

Table 5: Distribution of constraint interaction

	Type of constraint interaction	General	Names	Total	
a.	**No interaction – all constraints are respected**	**30**	**6**	**36**	**(33.9%)**
	□ Only phonological constraints	10	3		
	□ Only semantic constraints	5	0		
	□ Phonological and semantic constraints	15	3		
b.	**Interaction**	**44**	**7**	**51**	**(48.1%)**
	□ Only phonological constraints	12	3		
	□ Only semantic constraints	0	0		
	□ Phonological and semantic constraints	32	4		
c.	**Not accounted for**	**15**	**4**	**19**	**(17.9%)**
	□ No relevant constraint	1	1		
	□ The loser is the actual form	4	3		
	Total	**89**	**17**	**106**	

In 33.9% of the data items there was no constraint interaction, i.e. all the constraints were satisfied (Table 5, a). Crucial for our focus of this study is that in 13 of these 37 data items, only phonological constraints were relevant, as opposed to 5 where only semantic constraints were relevant. That is, there are more cases where phonology alone determines word order in AB constructions. The same is true for the cases where there is constraint interaction (Table 5, b), which constitute 48.1% of the data items. In 15 out of the 51 data items, the interaction is among phonological constraints only, as opposed to zero data items where only semantic constraints are relevant. In addition to these 28 data items where only phonological constraints are active, there are 3 data items where phonology prevailed in its interaction with semantics (see §5.2). Thus, phonology determines word order in 31 data items, which constitute 35% of the 87 data items that our grammar accounts for.

Our analysis accounts for 82% of the data (Table 5), which means that there were 19 data items that cannot be explained. In 2 there are no relevant constraints (e.g. *t͡síli ve gíli*) and in 17 the loser is the actual form.

(8) The loser is the actual form (☞ marks the optimal candidate and ☺ the actual form)

a. *óʁ ve a͡tsamót* 'skin and bones'

	PROTOTYPE FIRST [0.83]	SHORT-LONG [0.75]	*LAPSE [0.41]	ℋ
a. ☺ óʁ ve a͡tsamót	-1		-1	-1.24
b. ☞ a͡tsamót ve óʁ		-1		-0.75

b. *χúmus tχína* 'hummus tahini'

	PROTOTYPE FIRST [0.83]	ALIGN LEFT [0.64]	OCP [0.50]	ℋ
a. ☺ χúmus tχína		-1	-1	-1.14
b. ☞ tχína χúmus	-1			-0.83

Note that in (8a), also a traditional OT analysis would fail, since the actual form violates the dominating constraint PROTOTYPE FIRST. In (8b), however, a traditional OT analysis would derive the actual form, because the actual form respects the dominating constraint PROTOTYPE FIRST, while our harmonic OT analysis predicts a gang effect. Only 4 out the 17 data items where the loser is the actual form are correctly derived by a classic OT analysis. This result does not give advantage to classic OT since an examination of the data items predicted by our analysis reveals that 6 data items would not be predicted within a classic OT analysis. That is, the two analytical approaches have (almost) similar results.

Finally, it is important to highlight the correlation between the percentage of unaccounted forms and the degree of irreversibility (see Table 1 for the irreversibility scale). As shown in Table 6 below, the higher the degree of irreversibility the lesser the percentage of unaccounted forms.

Table 6: Irreversibility score and unaccounted forms

	Total	Unaccounted	
Absolute	10	–	0%
Strong	56	6	11%
Moderate	18	5	28%
Weak	5	3	60%

The correlation in Table 6 thus shows that our model is better suited for AB constructions with a low degree of reversibility, i.e. for irreversible AB constructions.

6 Questions for future study

In the course of our research, we have encountered two questions that require further study, in particular with reference to other languages; one question relates to ABC constructions (§6.1) and the other to the universal status of the grammar of AB constructions (§6.2).

6.1 ABC constructions

We found 5 ABC constructions in Hebrew – 3 and-constructions and 2 copulatives.

(9) ABC constructions

a. áf ózen gaʁón ‘nose ear throat’
b. dát géza ve mín ‘religion race and gender’
c. éven nijáʁ ve mispaʁáim ‘rock paper and scissors’
d. séks samím ve ʁókenʁol ‘sex drugs and rock-and-roll’
e. χúmus t͡ʃíps salát ‘hummus chips salad’

We expected the grammar developed on the basis of constructions with two elements to account for constructions with three elements as well, taking into consideration the six possible orders. However, our grammar accounted for only one of the five in (9), much below the ~82% success rate encountered with the general corpus (§5.2) and the binomial names corpus (§5.3).

Below we provide two tableaux, one where the ABC construction is derived by the proposed grammar (10a) and one where it does not (10b). There are three (out of six) candidates in each tableau, the winner and the two runner ups. The optimal candidate is marked with ☞ and the actual form with ☺.

(10) ABC constructions – success and failure

a. Derived by the grammar: *éven nijaʁ ve mispaʁáim* ‘rock paper and scissors’

	*CLASH [0.80]	SHORT-LONG [0.75]	*VV [0.47]	*LAPSE [0.41]	OCP [0.50]	ℋ
☞☺ a. éven nijáʁ ve mispaʁáim				-2	-1	-1.32
b. nijáʁ éven ve mispaʁáim	-1	-1		-2		-2.37
c. nijáʁ mispaʁáim ve éven		-2	-1	-1		-2.38

b. Not derived by the grammar: *af ózen gaʁón* ‘nose ear throat’

	*CLASH [0.80]	SHORT-LONG [0.75]	*LAPSE [0.41]	ℋ
☞ a. ózen áf gaʁón		-1		-0.75
b. áf gaʁón ózen	-1			-0.80
☺ c. áf ózen gaʁón	-1		-1	-1.21

There are no sufficient data to understand the failure of the proposed grammar, but we have some thoughts that should be perused with more data, probably experimental. ABC constructions, like compounds, have three potential syntactic structures – [ABC], [A[BC]] and [[AB]C]. It is possible that different types of constructions, for examples, trinomials (e.g. *dát géza ve mín*) vs. triplets (e.g. *áf ózen gaʁón*), would have different syntactic structures, and it is also possible that some of the constraints are restricted to apply only within a constituent and not across constituents; e.g. within BC in [A[BC]] but not across AB.

6.2 Is the grammar of AB constructions universal?

Cross-linguistic variation lies in the different constraint rankings each language adopts, or in the different relative constraint weights it assigns. However, specific phenomena found in different languages may, at least partially, share a grammar. For example, the ranking WORD=FOOT » MAX characterizes hypocoristics in many languages, thus yielding disyllabic forms (e.g. English *ǽlɪks*, Hebrew *smádi*, Spanish *kónθe*; Bat-El 2005). It is thus reasonable to ask whether the grammar of AB constructions is universal.

Based on Abraham’s (1950) comparative study, this line of research does not look promising. For example, several semantically identical AB constructions in German and English display a reverse order despite the similarities in prosodic structure (e.g. *Butter und Brot* vs. *bread and butter*, *Wasser und Brot* vs.

bread and water). More examples are given in the table below, which provides a comparative sample of parallel AB constructions in English, Hebrew, and Arabic.[7] In Table 7, a, the order is identical in all three languages; and in the rest of the examples there is an odd one out (shaded) – in Arabic (b), Hebrew (c), and English (d).

Table 7: A comparative sample

	English	Hebrew	Arabic
a.	life and death	χaím ve mávet 'life and death'	alħajá:t wa lmáwt 'life and death'
b.	day and night	jóm ve lájla 'day and night'	le:l w inaha:r 'night and day'
c.	sweet and sour	χamút͡s matók 'sour sweet'	ħiliw u ħa:mið 'sweet and bitter'
d.	bride and groom	χatán ve kalá 'groom and bride'	ʔilʕarí:s w ilʕarú:s 'the groom and the bride'

Compared to English and Hebrew, Arabic displays a reverse order in 'day and night' (b), which could probably be attributed to SHORT-LONG; if this is the case, we have to say that SHORT-LONG outranks (or weighs more than) POSITIVE-NEGATIVE. However, the ranking SHORT-LONG » POSITIVE-NEGATIVE does not hold for Arabic *alħajá:t wa lmáwt* 'life and death' (a).

Despite such potential conflicting evidence (which may diminish when more constraints are taken into consideration and in a harmonic grammar), it is worth perusing the question as to whether the grammar of AB constructions is universal. This may involve language-specific effects, phonological and perhaps also cultural, but the core grammar could be universal. [8]

7 Concluding remarks

Most studies agree that the grammar determining word order in AB constructions consists of the interaction of phonological and non-phonological, mostly semantic constraints. However, there is a disagreement with regard to the prominent module in this interaction – is it phonology or semantics? Our study suggests that there is no quantitative primary module, but rather different spaces. We reached this conclusion on the basis of a grammar we developed for Hebrew AB constructions (§5) – a harmonic grammar with weighted constraints, where weight was determined on the basis of the function of the constraint in the corpus. Our grammar predicts 83% (74/89) of the main corpus (§5.2) and 76% (13/17) of the binomial names (§5.3).

Some of the forms that are incompatible with our grammar may reflect the influence of English (e.g. *ʃaχóʁ laván* 'black white', *sakín ve mazlég* 'knife and fork', *gviʁotáj ve ʁabotáj* 'ladies and gentlemen') and others, as noted in §5.4, are rather low on the scale of irreversibility (e.g. *sakín ve mazlég* 'knife and fork').

We found that the semantic constraints are much heavier than the phonological constraints; the top four constraints were semantic, and the cumulative weight of the semantic constraints was higher than that of the phonological constraints. Thus, in the interaction between phonology and semantics, the semantic constraints often determined word order in most cases. This alone may suggest that phonology is only second to semantics when it comes to word order in AB constructions (Benor and Levy 2006, Mollin 2012, Shih et al. 2015).

However, the prominent role of phonology emerges when semantics (and, of course, also syntax) is mute, and this happens quite often in our corpus. In addition to the two cases of phonological gang effect

[7] The Arabic data were drawn from Gorgis and Al Tamimi (2005) and Kaye (2009). However, due to dialectal variation within and across in these studies, we adjusted all forms to the dialect of Palestinian Arabic spoken in the center of Israel. We thank Suma Samara for the help with these data.

[8] An example of a cultural effect can be drawn from Kaye (2009), who attributes the order in *albá:rid wa lħa:rr* 'cold and hot' to POSITIVE-NEGATIVE, since cold is better than hot if one lives in the dessert.

((6) and (7)), there were 28 data items where only phonological constraints were relevant, as opposed to 5 data items where only semantic constraints were relevant (Table 5).

The module responsible for word order is not phonology, and we thus do not expect phonology to determine word order; nevertheless, it seems that it does. Phonology selects the order AB when both syntax and semantics equally accept both AB and BA.

Appendix A

DATA ITEMS

Binomials (n=61)			*Reversibility score*
1.	á ve dá	'this and that'	93
2.	avót u baním	'fathers and sons'	98
3.	áχ ve aχót	'brother and sister'	91
4.	báal ve iʃá	'husband and wife'	73
5.	beád ve néged	'pros and cons'	99
6.	bejt͡sá ve taʁnególet	'egg and chicken'	96
7.	bén o bát	'boy or girl'	90
8.	bʁaχót ve iχulím	'blessings and greetings'	92
9.	buʃá ve χeʁpá	'shame and disgrace'	99
10.	dáf ve ét	'paper and pen'	77
11.	dáf ve ipaʁón	'paper and pencil'	84
12.	dín ve χeʃbón	'report (lit. law and bill')	99
13.	emét o χová	'truth or duty (lit. truth or dare')	100
14.	gvinót ve jáin	'cheeses and wine'	65
15.	it͡súv ve χitúv	'body shaping (lit. design and carving')	81
16.	jóm ve lájla	'day and night'	87
17.	kafé ve maafé	'coffee and pastry'	96
18.	kalkalá ve niúl	'economics and management'	62
19.	kán ve aχʃáv	'here and now'	99
20.	kéfel ve χilúk	'multiplication and division'	97
21.	kén o ló	'yes or no'	96
22.	kéʃev ve ʁikúz	'attention and focus'	99
23.	klavím ve ganavím	'dogs and thieves'	100
24.	kʁó u χtóv	'reading and writing'	96
25.	léχem ve χemá	'bread and butter'	92
26.	madá ve teχnológja	'science and technology'	93
27.	masá ve matán	'negotiation'	100
28.	mélaχ ve pílpel	'salt and pepper'	91
29.	milím ve láχan	'lyrics and melody'	72
30.	moén ve nimán	'sender and addressee'	93
31.	mukdám o meuχáʁ	'sooner or later'	86
32.	nisím ve niflaót	'miracles and wonders'	99
33.	ód ve adáʁ	'majesty and glory'	99
34.	ómoim ve lésbijot	'homosexuals and lesbians'	78
35.	ón ve ʃiltón	'wealth and authorities'	93
36.	óʁ ve at͡samót	'skin and bones'	53
37.	ovéʁ ve ʃáv	'checking account (lit. pass and back)'	99
38.	óχel ve ʃtijá	'food and drinks'	95
39.	páam o paamáim	'once or twice'	100
40.	paχót o jotéʁ	'less or more'	82
41.	pípi ve káki	'pee and poop'	81
42.	pó ve ʃám	'here and there'	100
43.	soʁág ve baʁíaχ	'grate and bolt'	100

44.	sulamót ve neχaʃím	'ladders and snakes'	73
45.	ʃedím ve ʁuχút	'demons and ghosts'	62
46.	ʃení ve χamiʃí	'Monday and Thursday'	98
47.	ʃnít͡sel ve piʁé	'schnitzel and mashed potatoes'	58
48.	táam ve ʁéaχ	'taste and smell'	95
49.	taʁbút ve spóʁt	'culture and sports'	91
50.	zúg o péʁet	'evens and odds'	96
51.	χád ve χalák	'clear cut (lit. sharp and smooth)'	93
52.	χaím ve mávet	'life and death'	83
53.	χáj ve kajám	'alive and existing'	100
54.	χalifá ve anivá	'suit and tie'	95
55.	χatúl ve aχbáʁ	'cat and mouse'	100
56.	χét͡s ve kéʃet	'bow and arrow'	95
57.	χibúʁ ve χisúʁ	'addition and subtraction'	97
58.	χók ve sédeʁ	'law and order'	98
59.	χomá ve migdál	'wall and tower'	100
60.	χupá ve kiduʃín	'Huppah and matrimony'	98
61.	χút ve máχat	'thread and needle'	83

	Copulatives (n=28)		*Reversibility score*
62.	aaték adbék	'copy paste'	99
63.	báʁ misadá	'bar restaurant'	52
64.	béten gáv	'stomach back'	90
65.	d͡ʒín tónik	'gin tonic'	89
66.	gomʁím holχím	'finish go'	97
67.	íks igúl	'tic-tac-toe (lit. ex circle)'	97
68.	jám jabaʃá	'sea land'	95
69.	kadíma aχóʁa	'forward backward'	92
70.	káχ tén	'take give'	89
71.	lemála lemáta	'up down'	97
72.	lév ʁeá	'heart lung'	99
73.	limonít luíza	'lemon grass verbena'	71
74.	mélaχ máim	'salt water'	93
75.	mizʁáχ maaʁáv	'east west'	88
76.	nagáta nasáta	'touch drive'	100
77.	neshiká stiʁá	'kiss slap'	76
78.	pinúj binúj	'evict build'	99
79.	plús mínus	'plus minus'	97
80.	ʁámle lód	'Ramla Lod'	81
81.	smól jamín	'left right'	79
82.	solél boné	'construction company'	98
83.	ʃaχóʁ laván	'black white'	90
84.	ʃóko banána	'chocolate banana'	82
85.	ʃóko vaníl	'chocolate vanilla'	87
86.	tsfon mizʁáχ	'north east'	98
87.	tút banána	'strawberry banana'	92
88.	χamút͡s matók	'sour sweet'	98
89.	χúmus tχína	'hummus tahini'	80

References

Abraham, Richard D. 1950. Fixed order of coordinates: A study of comparative lexicography. *The Modern Language Journal* 34(4). 276–287. https://doi.org/10.2307/318913.

Amir Coffin, Edna & Shmuel Bolozky. 2005. *A reference grammar of Modern Hebrew.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511811081.

Asherov, Daniel & Outi Bat-El. 2019. Syllable structure and complex onsets in Modern Hebrew. *Brill's Journal of Afroasiatic Languages and Linguistics* 11(1). 69–95. https://doi.org/10.1163/18776930-01101007.

Avishur, Yitzhak. 1976. Pairs of parallel words coupled in close proximity in Jewish Aramaic – in Biblical Aramaic and in the Aramaic translations. *Beit Mikra* 21. 247–262. [in Hebrew]

Bhandare, V.V. 1995. Structural and semantic aspects of the dvandva compound. *Annals of the Bhandarkar Oriental Research Institute* 76. 89–96.

Becker, Michael. 2002. *Hebrew stress: Can you hear those trochees?* Tel-Aviv: Tel-Aviv University MA thesis.

Benor, Sarah Bunin & Roger Levy. 2006. The chicken or the egg? A probabilistic analysis of English binomials. *Language* 82. 233–278. https://doi.org/10.1353/lan.2006.0077.

Bat-El, Outi. 1995. On the apparent ambiguity of the schwa symbol in Tiberian Hebrew. *Langues Orientales Anciennes Philologie et Linguistique* 5–6. 79–96.

Bat-El, Outi. 2005. The emergence of the trochaic foot in Hebrew hypocoristics. *Phonology* 22. 1–29. https://doi.org/10.1017/s0952675705000515.

Bat-El, Outi. 2011. Semitic templates. In Marc van Oostendrop, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *Blackwell companion to phonology*, 2586–2608. Malden, MA: Wiley-Blackwell. https://doi.org/10.1002/9781444335262.wbctp0108.

Bat-El, Outi. 2014. Staying away from the weak left edge: A strengthening strategy. In Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi & Nicola Lampitelli (eds.)*, The form of structure, the structure of form: Essays in honor of Jean Lowenstamm*, 193–207. Amsterdam: John Benjamins. https://doi.org/10.1075/lfab.12.15bat.

Beckman, Jill N. 1998. *Positional faithfulness*. Amherst, MA: University of Massachusetts dissertation.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: The Royal Netherlands Academy of Arts and Sciences.

Casali, Roderic F. 1997. Vowel elision in hiatus contexts: Which vowel goes? *Language* 73. 493–533. https://doi.org/10.2307/415882.

Cohen, Evan G., Vered Silber-Varod & Noam Amir. 2018. The acoustics of primary and secondary stress in Modern Hebrew. *Brill's Annual of Afroasiatic Languages and Linguistics* 10. 1–19. https://doi.org/10.1163/18776930-01001001.

Cooper, William E. & John R. Ross. 1975. World order. *Chicago Linguistic Society (CLS)* 11(2). 63–111.

Copestake, Ann & Aurélie Herbelot. 2011. Exciting and interesting: Issues in the generation of binomials. In Anja Belz, Roger Evans, Albert Gatt & Kristina Striegnitz (eds.), *Proceedings of the UCNLG+Eval: Language Generation and Evaluation Workshop,* 45–53. Edinburgh: Association for Computational Linguistics. https://aclanthology.org/W11-2707.

Faust, Noam. 2005. *The fate of gutturals in Modern Hebrew*. Tel-Aviv: Tel-Aviv University MA thesis.

Fenk-Oczlon, Gertraud. 1989. Word frequency and word order in freezes. *Linguistics* 27. 517–556. https://doi.org/10.1515/ling.1989.27.3.517.

Gil, David. 1989. Freezes, prosodic theory, and the modularity of grammar. *Folia Linguistica* 23. 375–386.

Goldsmith, John. 1976. *Autosegmental phonology.* Cambridge, MA: MIT dissertation.

Golston, Chris. 1995. Syntax outranks phonology. *Phonology* 12(3). 343–368. https://doi.org/10.1017/s0952675700002554.

Gorgis, Dinha T. & Yasser Al Tamimi. 2005. Binomials in Iraqi and Jordanian Arabic. *Journal of Language and Linguistics* 4(2). 135–151.

Gülzow, Insa & Natalia Gagarina (eds.). 2007. *Frequency effects in language acquisition defining the limits of frequency as an explanatory concept*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110977905.

Gustafsson, Marita. 1976. The frequency and "frozenness" of some English binomials. *Neuphilologische Mitteilungen* 77(4). 623–637.

Hawkins, John A. 1983. *Word order universals.* New York: Academic Press.

Hawkins, John A. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21(2). 223–261. https://www.jstor.org/stable/4178670 .

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Hayes, Bruce. 1995. *Metrical structure theory: Principles and case studies*. Chicago: The University of Chicago Press.

Inklas, Sharon & Draga Zec (eds). 1990. *The phonology-syntax connection*. Chicago: The University of Chicago Press. https://doi.org/10.3765/bls.v8i0.2025.

Jacobsen Jr, William H. 1982. Basque copulative compounds: A problem in irreversible binomials. *Berkeley Linguistics Society (BLS)* 8. 384–397. https://doi.org/10.3765/bls.v8i0.2025.

Jespersen, Otto. 1912. *Growth and structure of the English language.* Leipzig: B.G. Teubner.

Kaddari, Tzvi. 1966. Noun composites ("dvandva" type compounds) in Biblical Hebrew. *Lĕ onénu* 30. 113–135. [in Hebrew]

Kager, René. 1993. Alternatives to the iambic-trochaic law. *Natural Language & Linguistic Theory* 11. 381–432. https://doi.org/10.1007/bf00993165.

Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

Kaye, Alan S. 2009. Cultural ingredients in Arabic lexical pairs (binomials). *Word* 60. 65–78. https://doi.org/10.1080/00437956.2009.11432593.

Kelly, Michael H., Kathryn J. Bock & Frank C. Keil. 1986. Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language* 25. 59–74. https://doi.org/10.1016/0749-596x(86)90021-5.

Kopaczyk, Joanna & Hans Sauer. 2017. Defining and exploring binomials. In Joanna Kopaczyk & Hans Sauer (eds.), *Binomials in the history of English*, 1–23. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316339770.001.

Kwon, Nahyun & Keiko Masuda. 2019. On the ordering of elements in ideophonic echo-words versus prosaic dvandva compounds, with special reference to Korean and Japanese. *Journal of East Asian Linguistics* 28. 29–53. https://doi.org/10.1007/s10831-019-09189-1.

Lambrecht, Knud. 1984. Formulaicity, frame semantics, and pragmatics in German binomial expressions. *Language* 60(4). 753–796. https://doi.org/10.2307/413798.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic grammar – A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations (CU-CS-465-90). *Computer Science Technical Reports* 447. http://scholar.colorado.edu/csci_techreports/447.

Liberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8. 249–336.

Lohmann, Arne. 2014. *English co-ordinate constructions: A processing perspective on constituent order*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139644273.

Malikel, Yakov. 1959. Studies in irreversible binomials. *Lingua* 8. 113–160. https://doi.org/10.1016/0024-3841(59)90018-x.

Masini, Francesca. 2006. Binomial constructions: Inheritance, specification and subregularities. *Lingue e Linguaggio* 2. 207–232.

McCarthy, John J. & Alan Prince. 1993. Generalized alignment. In Gert Booij & Jaap van Marle (eds.), *Yearbook of morphology*, 79–153. Dordrecht: Kluwer.

McDonald Janet L., Kathryn Bock & Michael H. Kelly. 1993. Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology* 25. 188–230. https://doi.org/10.1006/cogp.1993.1005.

Mollin, Sandra. 2012. Revisiting binomial order in English: Ordering constraints and reversibility. *English Language and Linguistics* 16(1). 81–103. https://doi.org/10.1017/s1360674311000293.

Mollin, Sandra. 2014. *The (ir)reversibility of English binomials.* Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/scl.64.

Nelson, Nicole A. 2003. *Asymmetric anchoring*. New Brunswick, NJ: Rutgers dissertation.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33(6). 999–1035. https://doi.org/10.1111/j.1551-6709.2009.01047.x.

Pater, Joe. 2016. Universal grammar with weighted constraints. In John J. McCarthy & Joe Pater (eds.), *Harmonic grammar and harmonic serialism*, 1–46. London: Equinox Press.

Pinker, Steven & David Birdsong. 1979. Speakers' sensitivity to rules of frozen word order. *Journal of Verbal Learning and Verbal Behavior* 18. 497–508. https://doi.org/10.1016/s0022-5371(79)90273-1.

Pordány, Laszlo. 1986. A comparison of some English and Hungarian freezes. In Jacek Fisiak (ed.), *Papers and studies in contrastive linguistics,* 119–127. Washington, DC: Center for Applied Linguistics.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality theory: Constraint interaction in generative grammar.* Oxford: Basil Blackwell. [1993, TR 2, Rutgers University Cognitive Science Center].

Renner, Vincent. 2014. A study of element ordering in English coordinate lexical items. *English Studies* 95(4). 441–458. https://doi.org/10.1080/0013838x.2014.908039.

Schwarzwald, Ora. 1998. Word foreignness in modern Hebrew. *Hebrew Studies* 39. 115–142. https://doi.org/10.1353/hbr.1998.0000.

Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Shih, Stephanie S., Jason Grafmiller, Richard Futrell & Joan Bresnan. 2015. Rhythm's role in genitive construction choice in spoken English. In Ralf Vogel & Ruben van de Vijver (eds.), *Rhythm in cognition and grammar*, 207–234. Berlin: Mouton de Gruyter.

Shih, Stephanie S. & Kie Zuraw. 2017. Phonological conditions on variable adjective and noun word order in Taglog. *Language* 93(4). 317–352. https://doi.org/10.1353/lan.2017.0075.

Siewierska, Anna. 1988. *Word order rules*. London: Croom Helm. https://doi.org/10.2307/415269.

Smolensky, Paul & Géraldine Legendre. 2006. *The harmonic mind: From neural computation to optimality-theoretic grammar* (Cognitive Architecture, Vol. 1). Cambridge, MA: MIT Press.

Sobkowiak, Włodzimierz. 1993. Unmarked before marked as a freezing principle. *Language and Speech* 40(3). 177–187. https://doi.org/10.1007/s10936-010-9163-x.

Stallings, Lynne M. & Maryellen C. MacDonald. 2011. It's not just the "heavy NP": Relative phrase length modulates the production of heavy-NP shift. *Journal of Psycholinguistic Research* 36. 393–414. https://doi.org/10.1177/002383099303600403.

Tachihara, Karina & Adele E. Goldberg. 2020. Cognitive accessibility predicts word order of couples' names in English and Japanese. *Cognitive Linguistics* 31(2). 231–249. https://doi.org/10.1515/cog-2019-0031.

Tai, James H-Y. 1985. Temporal sequence and Chinese word order. In John Haiman (ed.), *Iconicity in syntax*, 24–26. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.6.04tai.

Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9(1). 81–105. https://doi.org/10.1017/s0954394500001800.

Wright, Saundra K., Jennifer Hay & Tessa Bent. 2005. Ladies first? Phonology, frequency, and the naming conspiracy. *Linguistics* 43(3). 531–361. https://doi.org/10.1515/ling.2005.43.3.531.

Zipf, George K. 1935/1965. *The psycho-biology of language: an introduction to dynamic philology*. Cambridge, MA: MIT Press (1935 First edition. Houghton: Mifflin).

Noa Handelsman
School of Education
Tel Aviv University
handelsman@mail.tau.ac.il

Outi Bat-El
Department of Linguistics
Tel Aviv University
obatel@tauex.tau.ac.il