Glossa a journal of general linguistics

Faghiri, Pegah and Pollet Samvelian. 2020. Word order preferences and the effect of phrasal length in SOV languages: evidence from sentence production in Persian. *Glossa: a journal of general linguistics* 5(1): 86. 1–33. DOI: https://doi.org/10.5334/gjgl.1078

RESEARCH

Word order preferences and the effect of phrasal length in SOV languages: evidence from sentence production in Persian

Pegah Faghiri and Pollet Samvelian
Sorbonne Nouvelle, Paris, FR
Corresponding author: Pegah Faghiri (pegah.faghiri@sorbonne-nouvelle.fr)

Heaviness (or phrasal length) has been shown to trigger mirror-image constituent ordering preferences in head-initial and head-final languages (heavy-late vs. heavy-first). These preferences are commonly attributed to a general cognitive pressure for processing efficiency obtained by minimizing the overall head-dependents linear distance – measured as the distance between the verb and the head of its left/right-most complement (Hawkins's Minimizing Domains) or as the sum of the distances between the verb and its complements (Dependency Length Minimization). The alternative language-specific accessibility-based production account, that considers longer constituents to be conceptually more accessible and views heavy-first as a salient-first preference, is dismissed because it implies differential sentence production in SOV and SVO languages. This paper studies the effect of phrasal length in Persian, a flexible SOV language displaying mixed head direction and differential object marking. We investigated the effect of linear distance as well as the effect of conceptual enrichment in two sentence production experiments. Our results provide clear evidence that support DLM while undermining Hawkins's MiD. However, they also show that some length effects cannot be captured by a dependency-distance-minimizing model and the conceptual accessibility hypothesis also needs to be taken into account to explain ordering preferences in Persian. Importantly, our findings indicate that distance minimization has a less strong effect in Persian than previously shown for other SOV languages.

Keywords: Word order preferences; grammatical weight; dependency distance minimization; conceptual accessibility; SOV languages; Persian

1 Introduction

This paper aims to contribute to our understanding of word order universals related to "grammatical weight" in OV vs. VO languages, by presenting data from Persian, an SOV language with mixed head direction. Grammatical weight (or heaviness) refers to the structural complexity and/or the length (number of words) of a constituent in relation to other constituents of a sentence. The first generalization on the matter was originally formulated by Behaghel (1909) and is known as *das Gesetz der wachsenden Glieder* (or Behaghel's [fourth] law) or, in more recent terminology, as the end-weight principle. Roughly speaking, it maintains that when ordering elements with comparable grammatical status, the longer element comes last in the sequence.

In psycholinguistics and cognitive sciences, this phenomenon favored availability-based incremental models of sentence production that assumed a universal "end-weight" (or short-before-long, more commonly used in this literature) ordering preference (e.g. De Smedt (1994)). However, given that the opposite preference, that is, long-before-short ordering, has since been documented for head-final languages such as Japanese (Hawkins

1994; Yamashita & Chang 2001), Hawkins argues that the end-weight can no longer be considered as a valid cross-linguistic generalization for sentence production and should be replaced by "heavy-first" or "heavy-last" depending on the typological type, i.e. OV and VO respectively (Hawkins 2007: 93).[1]

These mirror-image preferences presently seem to be well-established in the literature and are broadly assumed to result from a general/universal principle, according to which minimizing dependency distance facilitates language processing and comprehension (e.g. Hawkins (2007; 2014); Gildea & Temperley (2010); Temperley & Gildea (2018); Futrell et al. 2015)). Dependency distance or the length of a dependency refers to its span, generally measured in number of words intervening between a head and (the head of) its dependent. Dependency distance minimization roughly refers to choosing among alternative orders (e.g. V-NP-PP and V-PP-NP) the one that involves a smaller number of intervening words (see Section 2.2 and 3.2 for details).

A production-oriented account has also been proposed by Yamashita & Chang (2001), who roughly claim that longer constituents are conceptually more accessible and that constituent ordering in the preverbal domain is more sensitive to conceptual factors (hereto referred to as the conceptual-accessibility hypothesis), hence the long-before-short preference in OV languages (see Section 2.2). This account can be criticized as a *post hoc* accommodation, since it does not provide independent evidence supporting the claim that longer constituents are conceptually more accessible. Moreover, it posits differential sentence production principles in SOV and SVO languages (e.g. Tanaka et al. (2011)), while it is generally expected that the latter must be universal.

In this paper, we present sentence production data from Persian, studying the effect of the length in interaction with factors that are related to conceptual accessibility. While previous studies investigated the effect of length in transitive and ditransitive constructions, they do not take into account factors such as definiteness and animacy: 1) In experimental studies, namely on Japanese (Yamashita & Chang 2001) and Basque (Ros et al. 2015), all verbal complements are construed as definite entities; in ditransitive constructions the subject and the IO are human and the DO is inanimate, while in transitive constructions both the subject and the DO are human; 2) Large-scale corpus studies that support a cross-linguistic preference for shorter dependencies measure dependency distance while abstracting over these factors altogether (e.g. Futrell et al. (2015); Haitao Liu et al. (2017)).[2]

Indeed, interactions between the effect of these factors are not expected if we assume that length-based preferences derive from the tendency to minimize the distance between the verb and its complements, or, similarly, to avoid center-embedded constructions (Wasow 2002; 2013), in order to obtain a more optimal parsing. On the other hand, these interactions become relevant when the effect of the phrasal length is explained through its contribution to conceptual accessibility, since it is plausible to posit a hierarchy among factors that are known to enhance the conceptual accessibility of a constituent: definiteness, animacy or semantic role may be more prominent cues to conceptual accessibility

[1] It should be noted that the cross-linguistic generalizations regarding head-final vs. head-initial languages made by Hawkins are called into question by other more recent typological studies involving African languages in particular (see e.g. Dimmendaal (2011: 304–305)).

[2] There are of course a number of detailed corpus studies on different weight-related ordering phenomena in a given language that take into account a bundle of factors, such as animacy, definiteness, givenness, semantic relatedness, alongside constituent weight/length. In particular, in English, a number of studies (corpus-based and/or experimental) have shown that the end-weight preference applies notwithstanding other factors (e.g. Arnold et al. (2000); Gries (2003); Bresnan et al. (2007); Lohse et al. (2004); Melnick (2017), see Melnick (2017) for a review). The point is that studies claiming a cross-linguistic/universal end-weight principle have not investigated the role of weight in interaction with other factors.

than phrasal length.[3] Accordingly, for instance, phrasal length is expected to have a weaker effect with animate/given constituents compared to inanimate/new constituents, because the former are already highly accessible.

A length-based effect corresponding to the long-before-short preference has been documented for Persian by corpus-based (Faghiri & Samvelian 2014) and experimental studies (Faghiri et al. 2014; 2018).[4] However, these studies also showed that this effect is limited to specific cases and depends on other factors, such as the type of arguments involved (Subject/DO vs. DO/PP argument) and differential object marking (see Section 3.3). More specifically, the relative order between the direct object and the prepositional argument varies primarily with the degree of definiteness of the DO. A certain degree of variation triggered by a long-before-short preference was observed for bare and indefinite (unmarked) DOs in corpus as well as experimental data. Interestingly, some of these variations may not be straightforwardly accounted for in terms of dependency distance minimization. Furthermore, for definite (marked) DOs, corpus data showed only a trivial variation regardless of the relative length between the two constituents, and no experiments were conducted. Corpus data did not reveal any length-based effects on the order between the subject and the DO either and neither did a follow-up production experiment in which the length of the DO was manipulated by adding a relative clause.

Another issue is to determine what metric of the head-dependents distance provides more accurate predictions. Interestingly, syntactic properties of Persian allow us to tease apart the two main available measures, that is, Temperley's "Dependency Length Minimization"[5] (Temperley 2007) based on Gibson's "Dependency Locality Theory" (Gibson 2000), and Hawkins's "Minimize Domains" (formerly "Early Immediate Constituent") principle (Hawkins 1994; 2004). Previous studies on Persian mainly considered Hawkins's model and argued that it falls short in its account of Persian data. In this paper, we consider the predictions of both metrics. As we will see in Section 3.2, only Temperley's model correctly predicts a long-before-short length-based preference for the constructions under study in Persian.

We have conducted two complementary experimental studies in order to explore cases that are problematic for a dependency-distance minimization account. Based on our findings, we will argue that both a parsing-oriented account in terms of dependency-distance minimization and a production-oriented account in terms of the conceptual accessibility hypothesis can be advantageously recruited to explain word order preferences in Persian, and arguably in other languages.

In the following sections, we will first provide an overview of the relation between grammatical weight and word order, along with the available accounts of the long-before-short preference in SOV languages. In Section 3, we will introduce Persian and its interest for the study of length-based effects on word order, while discussing the findings of previous studies. We will present our two experimental studies in Section 4 and discuss our results in Section 5.

[3] With respect to grammatical roles, it is worth mentioning that in Japanese and Basque, longer DOs are more likely to shift over an IO than over a subject (Yamashita & Chang 2001; Ros et al. 2015). Note that in both languages the DO follows the IO in canonical sentences.

[4] It should be noted that the tendency for the postposition of heavy constituents after the verb has also been reported in a corpus study by Rasekh-Mahand et al. (2016) that we will discuss in Section 3.3. The latter studies the extraposition of relative clauses (in the postverbal domain), also see FN 29.

[5] Note that some studies (e.g. Futrell et al. (2015)) use dependency length minimization or its acronym DLM in a broad way and regardless of the metric used to measure the linear head-dependents distance. In this paper, we use dependency distance minimization to refer to the general tendency and retain dependency length minimization (and DLM) only to refer to Temperley's model.

2 Background

2.1 Grammatical weight and word order

The study of the role of grammatical weight (or heaviness) on the linear order of the constituents in a sentence is an important research topic in language sciences, including theoretical linguistics, psycholinguistics, as well as linguistic typology. It is an old topic, yet continues to motivate ongoing theoretical debates. Interest in the topic within general syntax goes far back. The first generalization on the matter, known as the "end-weight principle", was originally formulated by Behaghel (1909) as the "law of increasing constituents" (*Gesetz der wachsenden Glieder*). According to Wasow (1997: 82), who presents one of the key contributions to the topic, the term end-weight was first used by Quirk et al. (1972) in their description of the English grammar where they find a "tendency to reserve the final position for the more complex parts of a clause or sentence" (Quirk et al. 1972: 943).

Early accounts of end-weight are formulated in terms of language processing and comprehension (Bever 1970; Kimball 1973; Frazier & Fodor 1978), assuming that postponing heavy elements facilitates processing. Meanwhile, some studies do not find parsing-oriented (or hearer-oriented) accounts entirely convincing and argue that postponing heavy elements serves to facilitate the speaker's utterance planning and production (Wasow 1997; Stallings et al. 1998; Arnold et al. 2000; Wasow 2002; Chang 2009; Stallings & MacDonald 2011). Important early contributions to this view are two experimental works by Stallings et al. (1998) and Wasow (2002), who studied the heavy NP shift in English in a series of sentence production experiments, so as to credit the hypothesis that end-weight cannot only be explained by hearer-oriented accounts.

Indeed, a straightforward account of the end-weight effect can be framed in terms of availability-based incremental models of sentence production (e.g. Garrett (1980); Kempen & Hoenkamp (1987); De Smedt & Kempen (1987); Levelt (1989); Bock & Levelt (1994); De Smedt (1994); Kempen & Harbusch (2003)). In this view, the linear order of constituents reflects the order in which they become available for production, as long as grammar rules do not intervene. Constituents that become available at an earlier point in time can occupy an earlier linear position than constituents emerging later. Phrasal length is one factor modulating availability: everything else being equal, short constituents require less processing time and thus become available for production sooner than longer ones. As a matter of fact, the short-before-long or simply "short-first" preference (more commonly used in the psycholinguistic literature) presents a strong empirical argument for availability-based models of sentence production.

However, this account implies that the end-weight preference should be universal, as it has indeed been explicitly or implicitly assumed in many studies (see Hawkins (2007) for an overview). In other words, since the architecture of the language production system is assumed to be universal and the availability-based explanation is grounded in general principles of cognition, the short-before-long preference is expected to hold in all languages. Yet, as has been underscored by Hawkins's typological survey (Hawkins 1994 and subsequent work), OV languages like Japanese show the mirror image of the preference observed in VO languages such as English, that is, the long-before-short preference.

2.2 Long-before-short preference in OV languages

Two types of hypotheses are available to account for the long-before-short preference in OV languages: 1) parsing-oriented and 2) production-oriented.

1. Dependency-based distance-minimizing accounts motivated by efficient processing, such as the Early Immediate Constituent (EIC) or the Minimize Domains (MiD)

principles (Hawkins 1994; 2004; 2014), the Dependency Locality Theory, hereafter DLT, (Gibson 1998; 2000) or the Dependency Length Minimization tendency, hereafter DLM, (Temperley 2007; Gildea & Temperley 2010; Temperley & Gildea 2018), measure the complexity of parsing on the basis of the (linear) distance between the head verb and its complements (in number of words) and predict mirror-image preferences in head-initial vs. head-final languages.[6] These models, however, differ in the measure they use to rank different alternative sentences.

Hawkins (1990 and subsequent work) proposes a theory of the human parser based on efficient processing of complexity in grammar. In a nutshell, between two equally grammatical constructions, the human parser prefers the one that can be processed with optimal efficiency. To measure complexity, Hawkins takes into account the number of words that are needed to be parsed in order to recognize the syntactic structure of a sentence, or, in other words, all its immediate constituents (IC). The Early Immediate Constituent (EIC) principle,[7] or its more recent version Minimize Domains (MiD),[8] provides a metric to predict word order preferences between alternative/equally grammatical sentences (e.g. V-NP-PP vs. V-PP-NP). This metric depends on the Constituent Recognition Domain (CRD),[9] which measures the size of the processing domain, roughly, for a VP in a VO language, the distance between the verb and the head of its last/rightmost complement (see Section 3.2 for illustrative examples). MiD predicts that, with an equal number of words, a sentence with a smaller CRD should be preferred.

DLT and DLM, on the other hand, calculate the cumulative distance between the verb and the heads of all its dependents, referred to as the dependency length, and predict a preference for the sentence with a shorter dependency length. The two models differ substantially. While Gibson's DLT only takes into account constituents with "new discourse referents" (Gibson 1998: 12), the DLM model includes all words when calculating the dependency length (see Temperley 2007: 303–304).[10]

2. Recall that accessibility-based incremental models of sentence production predict a universal end-weight preference. To account for both long-before-short and short-before-long preferences in a unified manner, that is, via the same model of sentence

[6] A similar account attributes both heavy-first and heavy-last preferences to the tendency to avoid center-embedded complex structures that cause extra difficulty for both production and processing/comprehension (Wasow 2002; 2013).

[7] Hawkins initially defines the EIC principle as "The human parser prefers to maximize the left-to-right IC-to-word ratios of the phrasal nodes that it constructs." (Hawkins 1990: 233). In Hawkins (2004 and subsequent work), this principle is defined as: "The human parser prefers linear orders that minimize CRDs (by maximizing their IC-to-non-IC [or IC-to-word] ratios), in proportion to the minimization difference between competing orders." (Hawkins 2004: 32).

[8] The MiD principle is defined as: "The human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed. The degree of this preference is proportional to the number of relations whose domains can be minimized in competing sequences or structures, and to the extent of the minimization difference in each domain." (Hawkins 2004: 31).

[9] Hawkins defines CRD as follows: "The CRD for a phrasal mother node M consists of all non-terminal and terminal nodes dominated by M on the path from the terminal node that constructs the first IC on the left to the terminal node that constructs the last IC on the right." (Hawkins 2004: 32). Note that Hawkins later relabels CRD to "Phrasal Combination Domain (PCD)" – while keeping "essential aspects of the definition [remain] unchanged" – in order to make it "compatible with both production and comprehension." (Hawkins 2004: 107). Both labels are used in subsequent work.

[10] Temperley (2007) gives the following arguments to support this choice: 1) Gibson himself comments on the fact that this aspect of the DLT is provisional "[i]t is also likely that processing every intervening word, whether introducing a new discourse structure or not, causes some integration cost increment" (Gibson 1998: 13).; 2) If complexity is defined "in terms of processing time or computational effort […] the total complexity of a sentence […] should reflect the total of all integration costs. Since each dependency ultimately contributes to the integration cost of a word (the word on its right end), the total integration cost of a sentence is simply equal to the sum total of all its head-dependent distances." (Temperley 2007: 303–304).

production, Yamashita & Chang (2001) propose a language-specific version of the accessibility-based model, assuming that: a) longer constituents are conceptually more salient, and therefore more accessible, than shorter ones, and b) in flexible languages and in the preverbal domain, conceptual factors have a stronger effect on sentence production than formal factors. Departing from the observation that long-before-short invalidates availability-based models of sentence production, the authors suggest that both of these preferences can receive an accessibility-based account within the framework of incremental production.

Their rationale is that: 1) since both conceptual and form-related factors are shown to influence word order preferences (see Bock 1982), a production system's sensitivity to these factors can be viewed as being language-specific; 2) longer constituents, while being more complex, are lexically richer, hence conceptually more accessible than shorter constituents. Accordingly, they claim that, being more sensitive to conceptual factors, Japanese speakers prefer long-before-short ordering, while English is more sensitive to formal factors and English speakers prefer the opposite ordering. This difference, the authors argue, resides in the fact that: 1) English has a more rigid word order than Japanese; 2) In English weight-based shifts in ordering take place after the verb, and they assume, following Stallings et al. (1998), that "verbs exert strong influence over the phrases that follow them, including their order" (Yamashita & Chang 2001: B54)). More precisely, the authors suggest that "[s]ince accessibility of meaning and form can have different influences in production (Bock 1982), it could be that Japanese speakers are focusing more on conveying meaning (putting enriched material earlier), while English speakers are focusing on sequencing forms (putting easily accessed word earlier)." (Yamashita & Chang 2006: 6).

Each of these two hypotheses (parsing-oriented or production-oriented) tackles the problem from a different angle. Hence, they are not contradictory and can be assumed to hold simultaneously, as long as they are not falsified by the data. Furthermore, in previous studies, the predictions of these hypotheses converge for the majority of data discussed – coming from strictly head-final languages such as Japanese (Hawkins 1994; Yamashita & Chang 2001), Korean (Choi 2007) or Basque (Ros et al. 2015) – and makes it difficult to tease them apart.

Parsing-oriented (distance-minimizing) accounts nevertheless seem to have gained more approval among researchers because these accounts maintain the same mechanism for VO and OV languages. Moreover, a couple of recent large-scale cross-linguistic corpus studies conducted on available dependency tree-banks support the universality of dependency-distance minimizing hypothesis (e.g. Futrell et al. (2015); Haitao Liu et al. (2017)). The latter have also included Persian data as providing support for this hypothesis. However, it is important to note that there is a strong tendency for the postposition of subordinate clauses in Persian. For object clauses, this is a grammatical constraint. But even relative clauses that modify a preverbal NP can be extraposed to the postverbal domain.[11] This can easily tip the overall balance in favor of dependency-distance minimization in such large-scale corpus-based analyses. Note that there are also corpus studies that report contradictory results for mixed type languages (see Yao (2018) for Mandarin Chinese and Zoey Liu (2019) for a large-scale cross-linguistic study of PP ordering across 31 languages

[11] Moreover, the corpus study by Rasekh-Mahand et al. (2016) shows that the grammatical weight is one of the main factors which triggers the extraposition of relative clauses in the postverbal domain. Based on this result, the authors argue that in Persian, unlike other SOV languages, heavy constituents move rightward.

including a number of languages with mixed type patterns, e.g. head-initial PPs appearing before the verb).

Yamashita & Chang's (2001) language-specific production-oriented account, on the other hand, stipulates a difference in the production system of OV and VO languages, to which not all studies subscribe (e.g. Tanaka et al. (2011); Ros et al. (2015); Tachihara & Goldberg (2020)). Moreover, this model can also be called into question as a *post hoc* accommodation of existing models, namely, it does not provide independent underpinning for the claim that longer constituents are conceptually more accessible (but see e.g. Karimi & Ferreira (2016)). The claim that salience plays different roles in typologically different languages needs to be empirically justified.

However, there are also studies that pay credit to this language-specific availability-based model and adhere to the conceptual accessibility hypothesis. For instance, Kempen & Harbusch (2003), in their account of word order scrambling as a consequence of incremental sentence production, maintain the availability-based model but with a more inclusive definition considering conceptual factors in line with Yamashita & Chang (2001) to determine constituents' availability along with formal complexity and heaviness. Moreover, Stallings & MacDonald (2011) reiterate that: 1) it is the relative accessibility of constituents that influences word order during sentence production and 2) several factors other than length and complexity, including lexical-semantic properties, can possibly modulate accessibility.[12] They argue that (relative) length is a strong modulating factor of accessibility in English, while it has a weaker effect in Japanese. They also suggest that how (relative) length affects order in other languages is an empirical question.

It is important to bear in mind that salience is a complex and multi layered notion, and it is crucial to capture its various dimensions and the ways in which they interact. Conceptual accessibility of constituents, for instance, remains a vague and sometimes even confusing notion in the literature, especially with regard to its relation with discourse-related accessibility and the information structure of the sentence (e.g. topics are accessible entities and are generally realized by short/simple constituents such as pronouns).[13]

There however seems to be an agreement on the definition of accessibility in psycholinguistic literature, as "the ease with which the concept associated with a noun phrase (NP) can be retrieved from memory", as well as a consensus that the latter is "one of the most influential factors" in the processing and resolution of ambiguous pronouns (Karimi & Ferreira 2016: 507).[14] Importantly, in this line of studies, the informativity and/or specificity of a referring expression, such as an NP, is considered as a potential factor to enhance its accessibility. The length of an NP directly reflects the amount of information attached to it, and experimental studies (on pronoun ambiguity resolution) suggest that longer (referring) NPs are more accessible by virtue of being more informative – roughly because the additional information can "make it easier to retrieve the associated representation from memory" (Karimi & Ferreira 2016: 520). In this paper, we view conceptual

[12] Indeed, different semantic and pragmatic properties related to conceptual accessibility are known to influence ordering preferences, including animacy (e.g. Bock et al. (1992); Ferreira (1994); Kempen & Harbusch (2004); Branigan et al. (2008)), definiteness (e.g. Grieve & Wales (1973); Bader & Häussler (2010)), givenness (e.g. Bock & Irwin (1980); Arnold et al. (2000); Ferreira & Yoshita (2003)) or salience (e.g. Osgood & Bock (1977); Bock & Warren (1985); Prat-Sala & Branigan (2000)).

[13] Note that pronouns are generally used by speakers, as a choice of referring expression, to refer to highly accessible references, that is, referents that are (supposed to be) already activated (in the memory of listeners/comprehenders) and/or entities that are easily identified in the context.

[14] In other words, it is common in studies on pronoun ambiguity resolution to assess the accessibility of a referring expression by its potential to be identified as the antecedent of a pronoun by listeners/comprehenders (see Karimi & Ferreira (2016) for an overview).

accessibility along the same vein and assume that the informativity and/or specificity of a referring expression enhances its conceptual accessibility.

3 Length-based effects in Persian

3.1 The interest of Persian data

Persian is an SOV and a three-way *pro*-drop (subjects, direct objects and obligatory PP arguments can be omitted) language with flexible word order and mixed head direction. While NP objects precede the verb, clausal objects follow it. Furthermore, NPs and CPs are head-initial[15] and Persian has prepositions instead of postpositions. As such, Persian shares several properties with both VO and OV languages studied so far and makes the comparison promising for our understanding of length-based effects on word order variations.

In addition, Persian exhibits differential object marking (DOM),[16] realized by the enclitic *=rā* and triggered (roughly) by definiteness.[17] Note that Persian does not overtly mark definiteness. Definite NPs can be formed either by different definite determiners, like demonstratives, e.g. *in medād* 'this pencil', or by no overt determination, *medād* 'the pencil'. In Persian, a definite DO is always *rā*-marked (1a).[18] A DO lacking *=rā*, and carrying no determination or quantification, like *medād* or *medād=e qermez* in (1b), will necessarily receive a "bare noun" reading, that is, a nonspecific (existential or a kind-level/generic) reading. Note also that Persian bare nouns are underspecified for number and can consequently yield a mass reading (even if countable). Contrary to bare DOs, indefinite DOs are always specified for number and have an existential reading (1c). Indefiniteness is overtly marked, by the enclitic *=i*, e.g. *medād=i*, the cardinal *yek* e.g. *yek medād*, or the combination of both, e.g. *yek medād=i* 'a pencil'. Naturally, indefinite NPs are also formed by different indefinite quantifiers, e.g. *čand medād* 'a few pencils'.[19] Indefinite DOs can also be *rā*-marked to receive a specific reading, however, here we only use the label "indefinite" to refer to non-*rā*-marked indefinite DOs.

(1) a. Mahsā (in) medād(=e qermez)=rā xarid
Mahsa (this) pencil(=EZ[20] red)=RA buy.PST.3SG
'Mahsa bought (this/) the (red) pencil.'

b. Mahsā medād(=e qermez) xarid
Mahsa pencil(=EZ red) buy.PST.3SG
'Mahsa bought (red) pencils/a (red) pencil.'

[15] In Persian NPs, unbound determiners, quantifiers as well as classifiers precede the head noun and all dependents (adjectives or adjective phrases, PP modifiers, the possessor NP, and the relative clause) follow the head noun.

[16] Coined by Bossong (1985), DOM denotes the property of some languages with overt case-marking of some but not all direct objects depending on semantic and pragmatic features, see also Aissen (2003).

[17] DOM is a well-known feature of Persian, yet the object of ongoing controversial debate with no uncontroversial or straightforward account available in the literature (see Samvelian (2018) for a review). However, going into more detail is beyond the scope of this paper. It is sufficient to bear in mind that, while related to definiteness, DOM is far too complex to be captured by a single binary ±definite feature in Persian.

[18] *Rā* behaves as a phrasal affix and attaches to rightmost element of the NP. Note, however, that in NPs containing a relative clause *rā* generally appears before the latter (ia), and marginally after the relative clause (ib).

(i) a. in medād=rā [ke qermez ast] xarid-am
this pencil=RA that red is buy.PST-1SG

b. in medād [ke qermez ast]=rā xarid-am
this pencil that red is=RA buy.PST-1SG
'I bought this pencil which is red.'

[19] The noun remains in the singular form even when it denotes more than one entity.

[20] The enclitic =(y)e, the Ezafe, links the head noun to its modifiers and to the possessor NP (see Samvelian (2018) for a review).

c. Mahsā yek medād(=e qermez) xarid
Mahsa a pencil(=EZ red) buy.PST.3SG
'Mahsa bought a (red) pencil.'

DOM has important bearings on the canonical order between the NP and PP complements. While, in previously studied OV languages, the canonical order is given as S-IO-DO-V with no nuances or controversies, positing a canonical order in Persian ditransitive constructions is neither straightforward nor uncontroversial. In this paper, we rely on the following generalizations supported by quantitative studies (e.g. Faghiri et al. (2018)): 1) Bare DOs have a strong bias towards PP-NP-V, 2) *Rā*-marked DOs have a strong bias towards the NP-PP-V order, and 3) Indefinite DOs are more versatile with a fair inclination towards the PP-NP-V order.[21] Faghiri et al. (2018) argue that these ordering preferences can be captured by a cline on the basis of the degree of determination/definiteness of the DO that maps into a given-first as well as a salient-first preference (Faghiri et al. 2018: 182–183).

Interestingly, as we will show in the next section, different distance-minimizing accounts discussed in Section 2.2 do not predict the same length-based ordering preferences for Persian, contrary to previously investigated OV languages such as Japanese, Korean or Basque. Hence, the Persian data provide us with cases that make it possible to tease apart between, on one hand, Hawkins's MiD and, on the other, DLT and DLM. We limit our discussion to the predictions of DLM as a less restrictive version of Gibson's DLT and will not discuss DLT separately.[22]

As noted by Temperley (2007), one way to compare MiD and DLT/DLM is to examine mixed-branching constructions, containing both head-final and head-initial constituents, for which the predictions the two models differ (p. 323).[23] Temperley uses a small set of data (87 occurrences) from Turkish provided by Hawkins (1994) to test the predictions of these models but his results are not conclusive (Temperley 2007: 323–324).[24] In the next section, we present the predictions of these models for Persian data following Temperley (2007: 303–304).

3.2 Predictions of distance-minimizing models

In this section, we present the predictions of MiD and DLM for Persian data discussed in this paper. In Section 2.2, we saw that these models both account for length-based ordering preferences in OV and VO languages from a processing point of view by positing that some orders are less complex to process. However, they build on different measures of complexity. While both measures depend on the relative length between the constituents involved, they differ in the way the dependency distance is operationalized.

Recall that MiD depends on the size of the Constituent Recognition Domain (CRD) and takes into account the number of words needed to be parsed in order to obtain all immediate constituents (IC) of the sentence. Between two alternative word orders, MiD votes for

[21] Theoretical studies, however, have grouped indefinite DOs with bare DOs, claiming that the former occur in the same linear position as the latter, which is adjacent to the verb. This has been argued to constitute strong evidence in favor of the hypothesis of two distinct syntactic positions for the DO in Persian, depending on its markedness (see Faghiri & Samvelian (2016) for an overview).

[22] Recall that these two only differ with respect to the words that should be included when measuring the dependency length (all words or only those introducing a new discourse reference).

[23] Temperley (2007 and subsequent work) only tests the predictions of DLM, noting that "it seems fair to regard [them] as tests of [DLT]" as well because despite "the small differences between the two proposals" they are "clearly closely related" (Temperley 2007: 304).

[24] "[T]he data reveal a consistent preference for 'short-long' ordering, contrary to both theories. We should note, however, that the body of data available here is very small (87 cases). While this particular test is inconclusive, it points to a possible way of testing the DLT against the EIC theory, given further data." (Temperley 2007: 323).

the sentence that yields a greater IC-to-Word ratio, that is, for an equal number of ICs, a sentence with a smaller CRD is preferred. This is illustrated by the pair of examples in (2) from Hawkins (2014: 104–105). The CRD and IC-to-Word ratio are calculated for the VP. The latter contains three ICs: Verb, PP1 and PP2. Note that the preposition (or postposition in head-final PPs) is straightforwardly considered to be the constructing category of a PP constituent.[25] In other words, a (left-to-right) linear word-by-word parser recognizes the PP constituent when it parses the preposition.[26]

In (2a), Verb, PP1 and PP2 can be recognized on the basis of five words (italicized in the example). The CRD contains 5 words and hence the sentence has an IC-to-Word ratio of 3/5 (60%). (2b), on the other hand, has an CRD of 9 words and an IC-to-Word ratio of 3/9 (33%). Consequently, (2a), which reflects a short-before-long ordering, is considered to be less complex and easier to process than (2b), and thus should be preferred.

(2) a. The man [$_{VP}$ *waited* [$_{PP_1}$ *for his son*] [$_{PP_2}$ *in* the cold but not unpleasant wind]]
1 2 3 4 5

b. The man [$_{VP}$ *waited* [$_{PP_2}$ *in the cold but not unpleasant wind*] [$_{PP_1}$ *for* his son]]
1 2 3 4 5 6 7 8 9

Hawkins has also proposed a more fine-grained metric that allows the ranking of two sequences with the same CRD score.[27] An aggregate IC-to-Word ratio is calculated for the CRD instead of a simple IC-to-Word ratio: IC-to-Word ratios are calculated at each word in the CRD and the mean of all the IC-to-Word ratios is taken into account (see Hawkins (1990: 233–234)). In (3), IC-to-Word ratios are calculated at each word for the pair of examples in (2) above. Aggregate IC-to-Word ratios are 75.3% and 58.2%, respectively for (3a) and (3b).

(3) a. The man [$_{VP}$ *waited* [$_{PP_1}$ *for his son*] [$_{PP_2}$ *in* the cold but not unpleasant wind]]
$\frac{1}{1}$ $\frac{2}{2}$ $\frac{2}{3}$ $\frac{2}{4}$ $\frac{3}{5}$

b. The man [$_{VP}$ *waited* [$_{PP_2}$ *in the cold but not unpleasant wind*] [$_{PP_1}$ *for* his son]]
$\frac{1}{1}$ $\frac{2}{2}$ $\frac{2}{3}$ $\frac{2}{4}$ $\frac{2}{5}$ $\frac{2}{6}$ $\frac{2}{7}$ $\frac{2}{8}$ $\frac{3}{9}$

DLM, on the other hand, represents complexity of a sentence as the total sum of all the lengths of its dependencies. The length of a dependency is defined as the number of words spanned; a dependency connecting adjacent words is considered to have a length of 1 (Temperley 2007: 303). Between two alternative orders, the one with the shortest dependency-length is preferred. For two minimal-pair sentences that only differ in the order of their constituents such as those in (2), DLM only needs to compare the sum of the lengths of the two dependencies – between the verb and the head of each PP – that differ between the two options. This is schematically illustrated by the pair in (4), adopted from Temperley (2017: 317), where A and B can represent PP1 and PP2 above. In (4a) the dependency lengths are of 1 and 4 words and in (4b) of 1 and 8 words. Consequently,

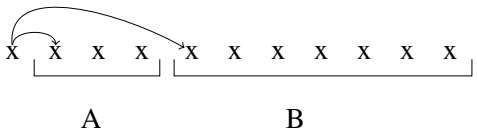
[25] The constructing category of an NP constituent is considered to be the determiner or the head noun (see Hawkins 1994: 60–62). Note that in languages with case marking (e.g. Japanese), Hawkins assumes the case marker to be the constructing category of case-marked NPs (Hawkins 2004: 108).

[26] Hawkins argues that "head categories such as P immediately project to mother nodes such as PP, enabling the parser to construct and recognize them online." (Hawkins 2014: 12).

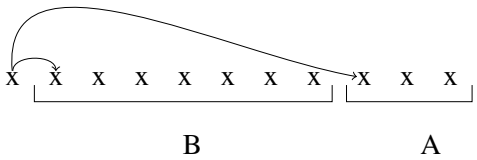
[27] In Hawkins's subsequent work, this proposition is subsumed in a broader principle, namely the Maximize Online Processing (MaOP) principle (see Hawkins 2004: 49–58).

(4a), which reflects a short-before-long ordering, has a shorter total dependency length (5 vs. 9).

(4) a.

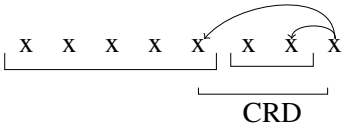


b.

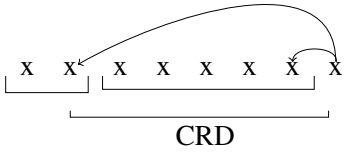


In a consistently head-final language such as Japanese, as illustrated by schematic examples in (5) from Temperley (2007: 323), both measures predict a preference for the long-before-short ordering. (5) presents two alternative orders for a verb final sentence with two preverbal dependents (of 2 and 5 words long each). (5a) has a smaller CRD compared to (5b) as well as a shorter total dependency length. Hence, according to both measures, (5a) is less complex and should be preferred.

(5) a.



b.



Finally, it is worth highlighting that besides the direction of length-based preferences, both models predict that the strength of preference depends on the size of reduction in complexity (or the gain in efficiency), which amounts to the difference between the complexity/efficiency measures of the two alternative orderings, that is, the difference between total dependency lengths in DLM, and the difference between CRDs or IC-to-Word ratios in MiD. This difference is in turn directly related to the relative length of the constituents involved. In other words, the rate of length-based shifts is expected to increase with the relative length of the constituents.

In what follows, we present the predictions of these two models for a number of cases of constituent ordering variation in Persian where the constituents differ in length. Distinct predictions are provided for *rā*-marked and non-*rā*-marked DOs, given that due to their formal difference the latter cannot be treated in the same way. Note that while a non-*rā*-marked DO can only be viewed as a head-initial NP, a *rā*-marked DO can either be viewed as head-initial or head-final depending on whether *rā* is considered as the head (constructing category) of the NP or not.[28] Here, we consider both possibilities. For non-

[28] Recall that in languages that display case-marking, Hawkins assumes the case-marker to be the constructing category of a case-marked NP (Hawkins 2004: 108) – as opposed to the (head) noun or, if applicable, the determiner. *Rā* is a differential object marker and does not appear on all (object) NPs (see Section 3.1). Being comparable to a case-marker, *r*ā can be safely assumed to be the constructing category of the NP. However, it is also reasonable to consider all object NPs on a par with respect to a left-to-right parsing.

rā-marked DOs, we discuss the relative order between the DO and the PP argument. For *rā*-marked DOs, in addition, we discuss the relative order between the DO and the subject.

For illustration, we take a length difference of 4 words between the two constituents and consider the two possibilities where one constituent has a phrasal length of 6 and the other 2 and vice versa. We use Hawkins's more fine-grained metric when relevant. Note that the latter is not taken into account by Temperley (2007).

We calculate the measures in minimal-pair sentences using schematic examples similar to (4) and (5) for simplification. Each sentence contains 3 ICs including the V(erb). We use bold to mark the head (constructing category) of the other constituents. In PPs, this would be the P (reposition), and in NPs, the first/leftmost element of the NP, which can be a determiner or a noun (represented by X that stands for any word). For *rā*-marked NPs, we are also considering the analysis in which *rā* is the constructing category. In short *rā*-marked NPs, the latter would appear at the right-edge. But in the case of long *rā*-marked NPs, which here we assume to contain relative clauses of 4 words length, *rā* appears just before the relative clause.

1) Non-*rā*-marked DOs and *rā*-marked DOs treated as head-initial NPs

In (6) and (7), we consider the relative order between an NP and a PP complement in the preverbal domain. In (6), we consider the case where the PP is longer than the NP and in (7) the reverse.

(6) a. $[[_{PP}\ \mathbf{P}\ X\ X\ X\ X\ X]\ [_{NP}\ \mathbf{X}\ X]\ \mathbf{V}]$ $DL_{NP} = 2, DL_{PP} = 8 : DL_T = 10$
$\frac{1}{1}\ \frac{1}{2}\ \frac{1}{3}\ \frac{1}{4}\ \frac{1}{5}\ \frac{1}{6}\qquad \frac{2}{7}\ \frac{2}{8}\ \frac{3}{9}$ CRD = 9 / IC-to-$W_{agg.ratio}$ = 36.8%

b. $[[_{NP}\ \mathbf{X}\ X]\ [_{PP}\ \mathbf{P}\ X\ X\ X\ X\ X]\ \mathbf{V}]$ $DL_{NP} = 8, DL_{PP} = 6 : DL_T = 14$
$\frac{1}{1}\ \frac{1}{2}\qquad \frac{2}{3}\ \frac{2}{4}\ \frac{2}{5}\ \frac{2}{6}\ \frac{2}{7}\ \frac{2}{8}\ \frac{3}{9}$ CRD = 9 / IC-to-$W_{agg.ratio}$ = 47.4%

(7) a. $[[_{NP}\ \mathbf{X}\ X\ X\ X\ X\ X]\ [_{PP}\ \mathbf{P}\ X]\ \mathbf{V}]$ $DL_{NP} = 8, DL_{PP} = 2 : DL_T = 10$
$\frac{1}{1}\ \frac{1}{2}\ \frac{1}{3}\ \frac{1}{4}\ \frac{1}{5}\ \frac{1}{6}\qquad \frac{2}{7}\ \frac{2}{8}\ \frac{3}{9}$ CRD = 9 / IC-to-$W_{agg.ratio}$ = 36.8%

b. $[[_{NP}\ \mathbf{X}\ X]\ [_{PP}\ \mathbf{P}\ X\ X\ X\ X\ X]\ \mathbf{V}]$ $DL_{NP} = 8, DL_{PP} = 6 : DL_T = 14$
$\frac{1}{1}\ \frac{1}{2}\qquad \frac{2}{3}\ \frac{2}{4}\ \frac{2}{5}\ \frac{2}{6}\ \frac{2}{7}\ \frac{2}{8}\ \frac{3}{9}$ CRD = 9 / IC-to-$W_{agg.ratio}$ = 47.4%

Predictions:
In both (6) and (7), both pairs have the same CRD but (b) has a greater aggregate IC-to-Word ratio. The gain in efficiency is the same (10.6%) in each pair; (a) has a shorter total dependency length. The reduction is the same (4 words) in each pair.

- MiD predicts a preference for (b) that reflects a short-before-long ordering.
- DLM predicts a preference for (a) that reflects a long-before-short ordering.
- The strength of preference depends on the relative length between the two constituents – regardless of the direction of the difference (NP > PP or NP < PP).

In (8), we consider the relative order between the subject and the DO in the preverbal domain. Note that here we provide one pair of examples to illustrate both the case where the subject is longer than the DO and the case where the subject is shorter. In each exam-

ple, NP1 and NP2 can represent respectively the subject and the DO (Subj > DO) and vice versa (Subj < DO).

(8) a. $[[_{NP_1}\ \mathbf{X}\ X\ X\ X\ X\ X]\ [_{NP_2}\ \mathbf{X}\ X]\ \mathbf{V}]$ $DL_{NP_1} = 8, DL_{NP_2} = 2 : DL_T = 10$
$\frac{1}{1}\ \frac{1}{2}\ \frac{1}{3}\ \frac{1}{4}\ \frac{1}{5}\ \frac{1}{6}\qquad\frac{2}{7}\ \frac{2}{8}\ \frac{3}{9}$ CRD = 9 / IC-to-$W_{agg.ratio}$ = 36.8%

b. $[[_{NP_1}\ \mathbf{X}\ X]\ [_{NP_2}\ \mathbf{X}\ X\ X\ X\ X\ X]\ \mathbf{V}]$ $DL_{NP_1} = 8, DL_{NP_2} = 6 : DL_T = 14$
$\frac{1}{1}\ \frac{1}{2}\qquad\frac{2}{3}\ \frac{2}{4}\ \frac{2}{5}\ \frac{2}{6}\ \frac{2}{7}\ \frac{2}{8}\ \frac{3}{9}$ CRD = 9 / IC-to-$W_{agg.ratio}$ = 47.4%

Predictions:
Both sentences have the same CRD but (b) has a greater aggregate IC-to-Word ratio. The gain in efficiency is the same in each case; (a) has a shorter total dependency length. The reduction is the same (4 words) for each case.

- MiD predicts a preference for (b) that reflects a short-before-long ordering.
- DLM predicts a preference for (a) that reflects a long-before-short ordering.
- The strength of preference depends on the relative length between the two constituents – regardless of the direction of length difference (Subj > DO or DO > Subj).

2 *Rā*-marked DOs treated as head-final constituents with *rā* as the head (constructing category).

In (9) and (10), we consider the relative order between an NP and a PP complement in the preverbal domain. In (9), we consider the case where the PP is longer than the NP and in (10) the reverse. Recall that in the latter case the NP contains a relative clause of 4 words that appears after *rā*.

(9) a. $[[_{PP}\ \mathbf{P}\ X\ X\ X\ X\ X]\ [_{NP}\ X\ \mathbf{X}_{=ra}]\ \mathbf{V}]$ $DL_{NP} = 1, DL_{PP} = 8 : DL_T = 9$
1 2 3 4 5 6 7 8 9 CRD = 9 / IC-to-W = $\frac{3}{9}$ = 33.4%

b. $[[_{NP}\ X\ \mathbf{X}_{=ra}]\ [_{PP}\ \mathbf{P}\ X\ X\ X\ X\ X]\ \mathbf{V}]$ $DL_{NP} = 7, DL_{PP} = 6 : DL_T = 13$
1 2 3 4 5 6 7 8 CRD = 8 / IC-to-W = $\frac{3}{8}$ = 37.5%

(10) a. $[[_{NP}\ X\ \mathbf{X}_{=ra}\ X\ X\ X\ X]\ [_{PP}\ \mathbf{P}\ X]\ \mathbf{V}]$ $DL_{NP} = 7, DL_{PP} = 2 : DL_T = 9$
1 2 3 4 5 6 7 8 CRD = 8 / IC-to-W = $\frac{3}{8}$ = 37.5%

b. $[[_{PP}\ \mathbf{P}\ X]\ [_{NP}\ X\ \mathbf{X}_{=ra}\ X\ X\ X\ X]\ \mathbf{V}]$ $DL_{NP} = 5, DL_{PP} = 8 : DL_T = 13$
1 2 3 4 5 6 7 8 9 CRD = 9 / IC-to-W = $\frac{3}{9}$ = 33.4%

Predictions:
Sentence (9b) has a smaller CRD and hence a greater IC-to-Word ratio; Sentence (10a) has a smaller CRD and hence a greater IC-to-Word ratio. The gain in efficiency is the same in each case; Sentences (9a) and (10a) have shorter total dependency lengths. The reduction is the same (4 words) in each pair.

- MiD predicts a preference for NP-PP-V order regardless of the relative length.
- DLM predicts a preference for (a) that reflects a long-before-short ordering.
- The strength of the preference depends only on the relative length between the two constituents – regardless of the direction of length difference.

In (11) and (12), we consider the relative order between the subject and the DO in the preverbal domain. (11) illustrates the case where Subj > DO and (12) the case where Subj < DO.

(11) a. $[[_{NP_1}\ \mathbf{X}\ X\ X\ X\ X\ X]\ [_{NP_2}\ X\ \mathbf{X}_{=ra}]\ \mathbf{V}]$ $DL_{NP_1} = 1$, $DL_{NP_2} = 8$: $DL_T = 9$
1 2 3 4 5 6 7 8 9 CRD = 9 / IC-to-W = $\frac{3}{9}$ = 33.4%

b. $[[_{NP_2}\ X\ \mathbf{X}_{=ra}]\ [_{NP_1}\ \mathbf{X}\ X\ X\ X\ X\ X]\ \mathbf{V}]$ $DL_{NP_1} = 6$, $DL_{NP_2} = 7$: $DL_T = 13$
1 2 3 4 5 6 7 8 CRD = 8 / IC-to-W = $\frac{3}{8}$ = 37.5%

(12) a. $[[_{NP_1}\ \mathbf{X}\ X]\ [_{NP_2}\ X\ \mathbf{X}_{=ra}\ X\ X\ X\ X]\ \mathbf{V}]$ $DL_{NP_1} = 8$, $DL_{NP_2} = 5$: $DL_T = 13$
1 2 3 4 5 6 7 8 9 CRD = 9 / IC-to-W = $\frac{3}{9}$ = 33.4%

b. $[[_{NP_2}\ X\ \mathbf{X}_{=ra}\ X\ X\ X\ X]\ [_{NP_1}\ \mathbf{X}\ X]\ \mathbf{V}]$ $DL_{NP_1} = 2$, $DL_{NP_2} = 7$: $DL_T = 9$
1 2 3 4 5 6 7 8 CRD = 8 / IC-to-W = $\frac{3}{8}$ = 37.5%

Predictions:
Sentence (11b) has a smaller CRD and hence a greater IC-to-Word ratio; Sentence (12b) has a smaller CRD and hence a greater IC-to-Word ratio. The gain in efficiency is the same in each pair; Sentences (11a) and (12b) have shorter total dependency lengths. The reduction is the same (4 words) in each pair.

- MiD predicts a preference for the DO-first order regardless of the relative length.
- DLM predicts a preference for a long-before-short ordering.
- The strength of the preference depends on the relative length between the two constituents – regardless of the direction of length difference.

To sum up: in different pair-wise comparisons examined here, DLM consistently predicts a long-before-short preference that increases with the relative length between the two constituents, regardless of the choice of analysis for *rā*-marked NPs. MiD, on the other hand, predicts a short-before-long preference that increases with the relative length, provided all NPs are treated as head-initial constituents. Otherwise, MiD predicts a preference for putting the *rā*-marked NP first (in both cases of word order variation studied). Note that while the NP = *rā*-PP-V order corresponds to the canonical order in Persian (see Section 3.1), the DO = *rā*-S-V order is non-canonical, given that Persian is an SOV language.

3.3 Previous studies on Persian

A number of quantitative studies have already addressed the issue of grammatical weight in Persian. Rasekh-Mahand et al. (2016) investigate the role of weight in the extraposition of relative clauses in the postverbal domain[29] in a corpus-based study and argue that this extraposition supports Hawkins's MiD and provides evidence that "Persian, a seemingly SOV language behaves typologically as a VO language, in which the heavy constituents shift rightward" (Rasekh-Mahand et al. 2016: 21).

[29] Recall that in Persian, a relative clause modifying a preverbal noun can be extraposed after the verb:

(i) yek mard āmad [ke hame = rā mi-šenāxt]
a man come.PST.3SG that all = RA IPFV-know.PST.3SG
'A man came who knew everybody.'

In line with studies on the heavy NP shift in different SOV and SVO languages, Faghiri and colleagues (Faghiri & Samvelian 2014; Faghiri et al. 2014; 2018) study the effect of weight on the order between the NP and the PP complement (referred to as direct and indirect objects in a broad manner) and report a long-before-short preference in the preverbal domain. However, they also highlight that this effect is limited to specific cases. Importantly, these studies report zero length-based effects in constructions for which other SOV languages such as Basque (Ros et al. 2015) and Japanese (Yamashita & Chang 2001) are shown to have length-based word order variations. Below we present a summary of these studies and their main results.

Faghiri & Samvelian (2014) present a multifactorial study of word order preferences carried out on a sample of 905 NP-PP-V and PP-NP-V utterances extracted out of a journalistic corpus. They find a significant effect of heaviness, operationalized as the relative length between the NP and the PP in number of words,[30] corresponding to a long-before-short preference for bare and indefinite objects (Faghiri & Samvelian 2014: 226–227).[31] Importantly, in their sample, *rā*-marked objects occur in the NP-PP-V order in more than 95% regardless of the relative length (Faghiri & Samvelian 2014: 226).

In the vein of constrained production experimental studies on the heavy NP shift, Faghiri et al. (2014; 2018) carry out two experiments to test length-based ordering preferences observed in Faghiri & Samvelian's (2014) corpus study for indefinite and bare DOs respectively. They use a web-based version of the sentence production task used by Stallings et al. (1998) and Yamashita & Chang (2001) – that we are also using in our experimental studies (see Section 4.1). They study the relative order between the DO and the obligatory PP argument while manipulating the relative length between these two constituents by adding a relative clause to the PP and attributive (restrictive) modifiers to the DO.[32] Both experiments find a significant main effect of the relative length. In the case of indefinite DOs (Est. 1.75, SE = 0.299, $p < 0.001$), the rate of the NP-PP-V order increases from 55.7% to 80.3% with longer DOs (Faghiri et al. 2014: 232). For bare DOs (Est. 1.479, SE = 0.308, $p < 0.001$), on the other hand, the rate of the PP-NP-V order increases from 85.5% to 94.2% with longer PPs (Faghiri et al. 2018: 175).

Furthermore, Faghiri et al. (2018) present an additional experiment on the relative order between bare DOs and the PP argument. They focus on the behavior of simple bare DOs, ex. 'syrup', compared to modified bare DOs, ex. 'icy mint syrup', that are longer by two words, and report an important rise of shifted orders for the latter. The rate of the NP-PP-V order increases from 28.2% with bare single word DOs to 47.3% with bare modified DOs (Faghiri et al. 2018: 178).

Faghiri et al. (2018) also address the effect of weight on the relative order between the subject and the DO. They first note that the rate of the SOV order is overwhelmingly high, above 95%, in their corpus sample and that non-canonical (shifted) orders only occur with *rā*-marked DOs (see p. 171 for details). They carry out an experiment that allows them to test whether the rate of OSV increases with longer DOs, modified by a relative clause. Faghiri et al. (2018) use *rā*-marked human DOs to give the OSV order the highest chance (see p. 179). Their results show no weight-based effect rise in OSV orders (7.81% vs. 6.98%, respectively for long/complex vs. short/simple DOs).[33] It is important to note

[30] The estimated coefficient is 0.844 (SE = 0.261, $p < 0.01$), with an intercept of 1.593 (SE = 0.295, $p < 0.001$), when NP-PP-V is coded as success (Faghiri & Samvelian 2014: 228).

[31] Note that for simple bare (single word) nouns, not included in Faghiri & Samvelian's (2014) multifactorial analysis, Faghiri (2016) observes that the average length of the PP is significantly smaller in sentences that occur in the NP-PP-V order than in sentences in the reverse order (see Faghiri 2016: 151).

[32] Note that it very difficult to manipulate the length of indefinite and bare DOs in Persian by adding a relative clause, since the latter usually triggers *rā*-marking.

[33] The estimated coefficient for the main effect of length is not significantly different from zero ($p > 0.05$).

that their study has a high statistical power to detect a true small effect size (0.83 calculated with the power function in R for binary data with 877 observations and df = 1). Note also that the absence of a length-based effect is unlikely to be due to a ceiling effect, given that they do find a significant effect of animacy (Est. 0.4, SE = 0.172, $p < 0.05$): the rate of OSV rises from 4.2% with animate subjects to 10.64% with inanimate subjects (Faghiri et al. 2018: 180).

To resume, previous studies report length-based ordering variations corresponding to a long-before-short preference in the relative order between the DO and the PP argument, limited to non-*rā*-marked DOs. They show that 1) long PPs shift (leftward) more often than short PPs when the DO is bare or indefinite, 2) long bare DOs shift more often than short bare DOs. On the contrary, the relative length is reported to have zero effect on the linear position of *rā*-marked DOs (with respect to the PP argument) as well as on the relative order between the subject and the DO in transitive sentences.

In their attempt to provide an explanation for these observed length-based effects, Faghiri and colleagues favor a production-oriented account over parsing-oriented accounts. They argue that Yamashita & Chang's (2001) conceptual accessibility hypothesis provides an adequate account of the Persian data, in particular because it can provide a compelling explanation for the cases where no length-based effects are observed (Faghiri et al. 2018: 183), while these cases are problematic for dependency-length-minimizing models (Faghiri 2016: 268–269). They also rightly note that Hawkins's model falls short of accounting for the long-before-short preference in the preverbal domain in Persian (see e.g. Faghiri & Samvelian 2014: 228).

3.4 The focus of our study

Length-based ordering preferences depicted by available studies so far present some potential challenges for dependency-distance minimizing accounts.

1. The cases for which previous studies report zero length-based effects run counter to the predictions of DLM. According to the latter the long-before-short preference is expected to trigger ordering variations for all types of DOs.
2. Although MiD may account for the absence of a length-based effect in the case of *rā*-marked DOs (provided the latter are viewed as headed by *rā*), it makes wrong predictions with respect to the direction of length-based effects in the preverbal domain.
3. The relatively important rise in the rate of shifted orders observed for a two-word length difference, when the relative length is increased by adding two attributive modifiers to bare DOs, is intriguing.[34] Importantly, by comparing the size of length-based effects in the three experiments on the relative order between the NP and the PP argument, we observe that it is not possible to establish a correlation between the strength of the preference and the relative length.[35] This is potentially problematic for any dependency-distance minimizing account of length-based effects. Indeed, as we have seen in Section (3.2), the strength of a length-based preference is predicted to increase with the relative length between the two constituents regard-

[34] The effect size (Cohens w) of length, which we have calculated from the contingency table provided by Faghiri et al. (2018), is 0.214.

[35] A two- to four-word difference between the NP > PP and NP < PP conditions (by adding attributive modifiers to the NP and an RC to the PP) in Faghiri et al.'s (2014) experiment with indefinite DOs, triggers about 24% shifts. A three-word difference between the NP = PP and NP < PP conditions (by adding an RC to the PP) in Faghiri et al.'s (2018) experiment with bare DOs triggers less than 10% shifts. The effect size (Cohens w), calculated from the contingency tables provided by Faghiri et al., is respectively 0.255 and 0.139.

less of the direction of the difference. On the other hand, the strong effect observed when adding (restrictive) modifications to bare nouns may be viewed as triggered by the added information and the increase in their degree of specification rather than by length *per se*. Therefore, the question rises as to whether all length-based effects can be viewed as triggered by dependency-distance minimization. An alternative account would be to view the effect of adding attributive modifiers as an effect of conceptual enrichment, in line with the salient-first preference in the preverbal domain.

We have conducted two more production experiments to follow up on these observations. These experiments are carried out in order to 1) pin down the effect of phrasal length in terms of conceptual enrichment vs. dependency-distance by comparing simple/short vs. modified/long non-*rā*-DOs and 2) replicate the absence of length-based effect in the case of *rā*-marked DOs in order to assess the relevance of dependency-distance minimizing accounts.

4 Production experiments on word order variation

In this section, we present the results of two constrained sentence production experiments carried out to study the ordering preferences of Persian native speakers. The task used in these experiments is identical to the one used by Faghiri et al. (2014; 2018) in the previous experimental studies on Persian (see Section 3.3). We start by presenting this experimental protocol.

All our data are analyzed via statistical open access software R (R Core Team 2013), using the lm4 package for modeling (Bates et al. 2012). The response variable is of binary type, and logistic mixed effect modeling (hereafter GLMM) is used all along. In all models presented here, experimental variables are centered by using a sum-to-zero contrast coding, that is, the intercept corresponds to the grand (pooled) mean. We always report the results of the optimal model, that is, the maximal variance-covariance model supported by the data[36] and justified by model comparisons (Baayen et al. 2008). Also, we initially included the presentation order as a fixed effect in all models but only kept it when it yielded a significantly better fit. Note that this never had a meaningful impact on the results (p-values and the coefficients) of our target factors. For ease of reading, we only provide relevant result segments in the core text. The full summary of results for each model are provided in the appendices.

4.1 Procedure

Our experiments are implemented via web-based self-administrated questionnaires[37] conducted on the *Ibex Farm* (Drummond 2013). In these experiments, participants are asked to construct sentences with phrases that appear on the screen to complete a preamble.

This production paradigm is inspired by the *in lab* cued sentence recall production. This is a common experimental task in psycholinguistics for the study of sentence production and is used in experimental studies on word order preferences, especially on heavy constituent shift (e.g. Stallings et al. (1998); Yamashita & Chang (2001)). In these experiments, participants are asked to make a sentence with given constituents that appear on a computer screen and to produce it orally after a lapse of time during which, for distraction, they are presented with a basic arithmetic operation. This design is meant to encourage participants to produce their sentences from meaning (Yamashita & Chang 2001: 48).

[36] Models that did not converge or had perfect correlations between random effects were reduced.

[37] This means that these experiments have been designed specifically to be completed by a respondent (on their own computer) without intervention of the researchers.

In the design used in our experiments, participants are asked to make a sentence to complete a preamble with phrases that appear on the screen. Participants see simultaneously a preamble, an incomplete sentence represented by blank boxes and a vertical list of phrases (that appear in blue). A screenshot of a (filler) item is given in Figure 1.

Each experimental item contains a sentence in which three constituents are missing, represented by three blank boxes. To complete the sentence, participants are provided with four phrases (presented in a counterbalanced order). Participants are instructed to 1) read the preamble and the list of phrases, 2) complete the sentence with the most natural continuation that comes to their mind using three of the four given phrases, 3) fill in the blanks accordingly, and 4) click on "Continue" to go to the next sentence.

The list of options contains one element more than the number of blanks in order to prevent participants from guessing the purpose of the experiment and to push them to concentrate on the content of each sentence, in order to produce reasonably natural sentences. The relative order between these constituents (in the final sentence) is left to the participants and constitutes the response or the dependent variable.

4.2 Experiment 1

This experiment targets the effect of phrasal length by adding an attributive (restrictive) modifier to non-*rā*-marked (bare and indefinite) DOs. Faghiri et al. (2018) showed a particularly important effect of the phrasal length, when comparing the relative order of short bare DOs (one-word length ex. *gol* 'flower') with long bare DOs (three-word length, ex. *gol-e orkide-ye sefid* 'white orchid flower') with respect to inanimate PP arguments. The rate of the non-canonical NP-PP-V order increased from 28.2% to 47.3% for long bare DOs. The research question here is whether indefinite DOs show a comparable effect when lengthened by attributive (restrictive) modifiers.

Furthermore, in order to disentangle the effect of conceptual enrichment from the effect of increasing the dependency-distance, we manipulate the length by adding only a one-word attributive modifier. Recall that the strength of a dependency-length preference depends on the size of reduction in the length difference between two alternatives. A one-word length difference presents the least possible reduction in the dependency length. Note that the lowest rate of length-based shifts observed so far (in Persian data) is about 10% and was triggered by a three-word length difference (see Section 3.4). Consequently, a comparably strong effect in this configuration can hardly be viewed as a dependency-minimizing effect.

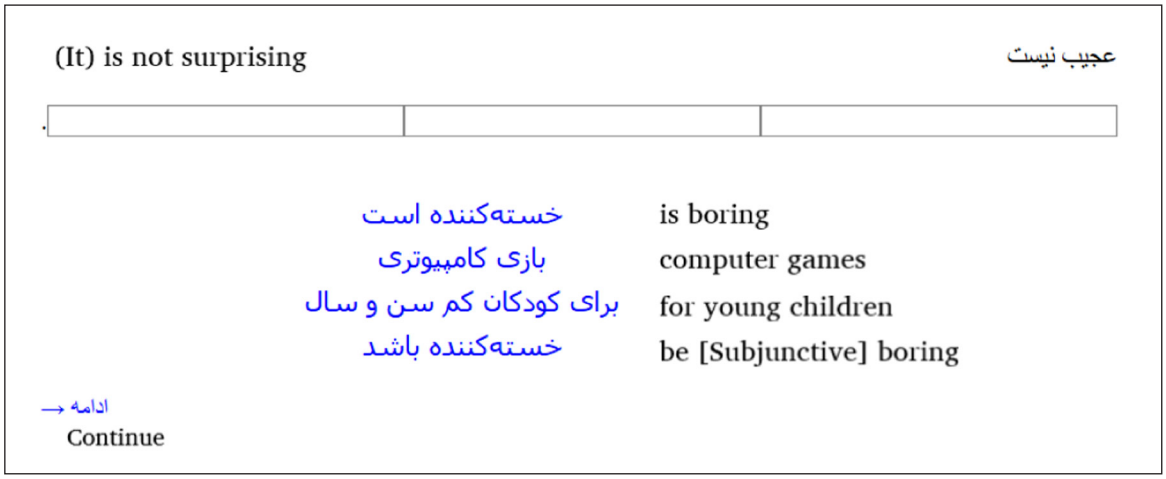


Figure 1: Screenshot of an experimental item (English equivalents are added).

Finally, in Faghiri et al.'s study, PPs were inanimate in all sentences and the results showed an overall higher rate of NP-PP-V order with respect to another experiment with bare DOs but including only animate PPs (see Faghiri et al. 2018: 177). Here, we include the animacy of the PP as a between-items (control) variable in order to neutralize its effect and also to see to what extent animate and inanimate PPs behave differently.

4.2.1 Method and materials

A set of 16 sentences was created for this experiment, following a 2 × 2 design with DO type (bare vs. indefinite) and DO length (simple vs. modified with a one-word attributive modifier) as within-item variables, where we prepared 4 versions of each sentence, as in (13). In half of sentences the PP was human (construed as a goal or a source argument and involving various prepositions, e.g. *be* 'to', *az* 'from' or *barāye* 'for') and in the other half inanimate (with the same proportion of different preposition types). DO type was treated as a between-subject variable: that is, one group of participants saw sentences with bare DOs and another group sentences with indefinite DOs.

Each experimental item was preceded by a preamble containing the subject and a (vertical) list of four constituents: a PP, two choices of formally identical NPs and a verb. The order of the list was counterbalanced (between PP over NP and NP over PP) across items. The dependent variable is the order between the three remaining constitutes (PP, NP and Verb) in sentences filled out by participants. However, expecting non-verb final orders to be scarce, the comparison will be limited to NP-PP-V vs. PP-NP-V.

(13) Ali hamiše …
Ali always …
[$_{NP1}$ (yek mošt) gerdu(=ye tāze)] [$_{PP}$ tu=ye sālād] [$_{V}$ mi-riz-ad]
a handful walnut=EZ fresh in=EZ salad IPFV-put.PRS-3SG
[$_{NP2}$ (yek meqdār) serke(=ye sib)]
a quantity vinegar=EZ apple
'Ali always puts (a handful of) (fresh) walnuts / (a small quantity of) (apple) vinegar in the salad.'

Summary of the experimental design:
V1: DO type: bare *gol* 'flower' vs. indefinite *yek daste gol* 'a bunch of flowers' (between-subjects variable)
V2: Length of the DO: simple *gol* vs. modified *gol-e orkide* 'orchid flower' (within-subject variable)
V3: Animacy of the PP: human vs. inanimate (between-items variable)
Dependent variable: relative order between the PP and the NP

These 16 experimental items were combined with 24 filler items. The final list of items, in which target items were spaced by at least one filler, was randomized for each participant individually. It contained an additional filler item appearing as the first item for all participants.

80 native speakers of Persian (39 women and 41 men; mean and median age: 33 and 31.5 years) volunteered to complete our web-based questionnaire (40 for each sub-experiment) – the exact number of participants was 97 but we discarded data from bi/multilingual speakers that did not declare Persian as their dominant language. Data from two participants in the indefinite sub-experiment was excluded from the final dataset because

they did not fill out sentences according to the instructions. There were also a few erroneous answers, which we marked as NA.

4.2.2 Results and discussion

Table 1 presents the frequencies of different ordering choices in the data, and Figures 2 and 3, the rate of NP-PP-V order (NPV in short), respectively by DO length and by animacy of the PP.

The overall rates for NP-PP-V are 12.5% for bare DOs, and 58.3% for indefinite DOs. These rates are in line with the rates reported for these DO types in previous experimental studies. We observe that the rate of NP-PP-V significantly increases with longer/modified NPs for both bare and indefinite types, respectively, from 8.0% to 17.0% ($\chi^2 = 11.454$, $p < 0.001$) and, from 51.0% to 65.6% ($\chi^2 = 12.59$, $p < 0.001$). Likewise, the rate of NP-PP-V is significantly higher for inanimate PPs with both bare and indefinite types, 7.7% vs. 17.2% ($\chi^2 = 12.693$, $p < 0.001$) and 54.1 vs. 62.5 ($\chi^2 = 3.976$, $p < 0.05$).

We analyzed all the data (a total of 1254 data points excluding miscellaneous orders) using an GLMM model including items and participants as random effects and DO type, DO length and PP animacy (sum-coded) as fixed effects. The NP-PP-V order is coded as success. We find a significant effect for the DO type (Est. = –1.94, SE = 0.22, $p < 0.001$) as well as for length (Est. = 0.55, SE = 0.11, $p < 0.001$), but no significant effect for animacy nor any significant interactions between the variables (ps > 0.10).

We then fitted two models separately for bare and indefinite DOs (a total of, respectively, 650 and 604 data points), to see whether there is a numerical difference between the estimated coefficients and in what direction. The estimated coefficient of length is greater for bare DOs (Est. = 0.64, SE = 0.17, $p < 0.001$) than for indefinite DOs (Est. = 0.38, SE = 0.11, $p < 0.01$).

Table 1: Distribution of word order for Experiment 1.

Bare Dos					
	NPV	**NVP**	**PNV**	**NA**	**Total**
DO Length					
Short	26	0	301	1	**328**
Long	55	1	268	4	**328**
PP Animcy					
Anim	25	0	300	3	**328**
Inan	56	1	269	2	**328**
Total	**81**	**1**	**569**	**5**	**656**
Indef Dos					
	NPV	**NVP**	**PNV**	**NA**	**Total**
DO Length					
Short	154	1	148	1	**304**
Long	198	1	104	1	**304**
PP Animcy					
Anim	164	0	139	1	**304**
Inan	188	2	113	1	**304**
Total	**352**	**2**	**252**	**2**	**608**

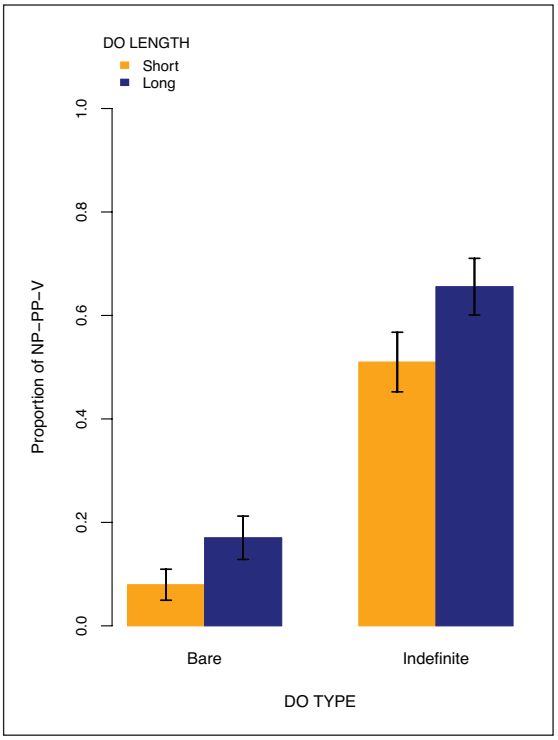


Figure 2: NP-PP-V rate by DO type and DO length in Experiment 1 (the error bars present 95% confidence intervals).

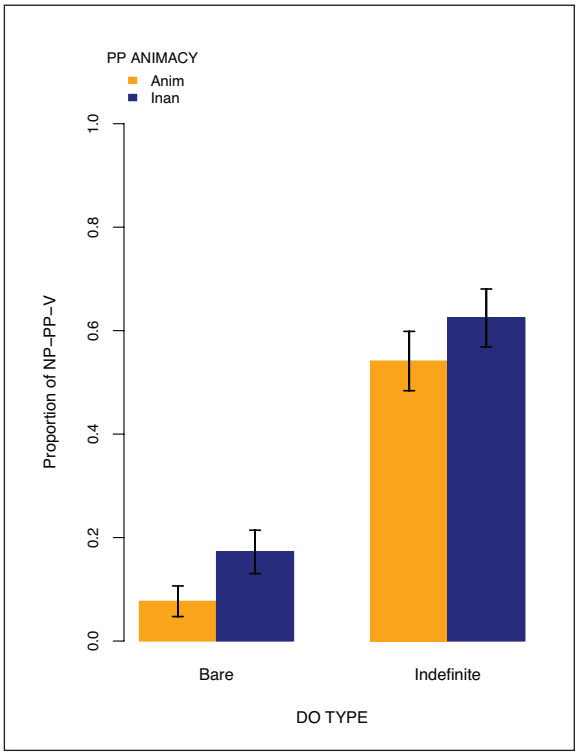


Figure 3: NP-PP-V rate by DO type and animacy of PP in Experiment 1.

In sum, our data show a robust effect of phrasal length for both non-*rā*-marked DO types. We observe that adding modifications to the DO increases NP-PP orders for both bare and indefinite types, and while there is no significant interaction between the effects of length and DO type, we can say that it is likely to have a larger effect size for bare DOs

than for indefinite ones. Indeed, if the effect of length is due to the increase in the degree of specification (of a non-specific object), it is safe to assume that with less specific objects (bare vs. indefinite) the contribution of length is likely to be more important.

With respect to the animacy of the PP, in line with the general animate-before-inanimate preference, we observe that overall in our data the rate of PP-NP-V orders increases with animate PPs. We are not going to comment further on the effect of animacy which is beyond the scope of this paper.

4.3 Experiment 2

In this experiment, we study the effect of the relative length (by adding a relative clause to the PP argument) for *rā*-marked vs. non-*rā*-marked DOs. In their experimental study, Faghiri et al. (2014) found an important effect of the relative length for indefinite (non-*rā*-marked) DOs, corresponding to a long-before-short preference (see Section 3.3). This preference is in contradiction with the prediction of MiD while it is in line with that of DLM (see Section 3.2). The latter predicts a long-before-short preference for *rā*-marked DOs as well. However, corpus studies found no length-based effects for these DOs suggesting that they may not be sensitive to length-based effects (Faghiri & Samvelian 2014: 230). Recall that MiD also predicts no length-based effects for *rā*-marked DOs in this configuration. The research question in this experiment is thus to test whether increasing the length of the PP argument favors more shifted orders for *rā*-marked NPs. Such result would provide further support for DLM while undermining Hawkins's MiD.

4.3.1 Method and materials

A set of 16 sentences was created for this experiment. Following a 2 × 2 design with DO type (*rā*-marked vs. indefinite) and PP length (simple vs. modified by a relative clause), we prepared 4 versions of each sentence, as in (14). In all sentences PPs were human, construed as goal or source arguments. They involved different prepositions: *be* 'to' (8 items), *az* 'from' (4 items) or *barāye* 'for' (4 items). Each experimental item was preceded by a preamble that contains the subject and a (vertical) list of four constituents: a PP, two choices of DOs (same type but lexically different) and a verb. The order between the PP and the two NPs was counter-balanced for all items. The dependent variable is the order between the three remaining constitutes (PP, NP and Verb) in sentences filled out by participants, coded as a binary variable – recall that as we are expecting non-verb final orders to be scarce, the comparison will be limited to NP-PP-V vs. PP-NP-V.

(14) Parvin …
Parvin …
[$_{NP1}$ yek arusak / arusak=rā] [$_{PP}$ barāye Zohre (ke dāšt gerye mi-kard)]
a doll / doll=RA for Zohreh that AUX cry IPFV-do.PST.3SG
[$_{NP2}$ yek ābnabāt / ābnabāt=rā]
a candy / candy=RA
[$_{V}$ xarid]
buy.PST.3SG
'Parvin bought the/a doll/candy for Zohreh (who was crying).'

Summary of the experimental design:
V1: DO type: *rā*-marked *arusak=rā* 'the doll' vs. indefinite *yek arusak* 'a doll'
V2: Length of the PP by adding a relative clause (of about four-word length)
Dependent variable: relative order between the PP and the NP

These 16 experimental items were combined with 24 filler items. The final list of items, in which target items were spaced by at least one filler, was randomized for each participant individually. It contained an additional filler item appearing as the first item for all participants.

34 native speakers of Persian (16 women and 18 men; mean and median age: 32 and 33 years) volunteered to complete our web-based questionnaire – the exact number of participants was 36 but we did not include bi/multilingual speakers who did not declare Persian as their dominant language. There were also a few erroneous or incomplete answers that we marked as NA.

4.3.2 Results and discussion

The overall rates of NP-PP-V order are 80.4% for *rā*-marked NPs, and 42.4% for indefinite NPs. The latter is lower than the rate reported previously for these NPs, but when we look at the distribution by the preposition involved (Figure 4), we observe a fairly important variation, suggesting that the default order is not the same depending on the preposition type. Importantly, for the preposition *be* 'to', the baseline rate tips for the NP-PP-V order in line with previous studies. Table 2 and Figure 5 present, respectively, the frequencies of different ordering choices in the data and the rate of NP-PP-V order by DO type and PP length. We observe that for both types the rate of NP-PP-V order is significantly higher with short PPs: 86.8% vs. 74.1% ($\chi^2 = 6.152$, $p < 0.05$) and, 52.6% vs. 31.9% ($\chi^2 = 11.068$, $p < 0.001$), respectively for *rā*-marked and indefinite NPs.

We analyzed all the data (a total of 541 data points excluding miscellaneous orders) using a GLMM model including preposition types, items and participants as random effects and DO type and PP length (both sum-coded) as fixed effects. The NP-PP-V order is coded as success. We find a significant effect for the DO type (Est. = 1.10, SE = 0.12, $p < 0.001$) as well as for PP length (Est. = –0.52, SE = 0.11, $p < 0.01$), but no significant interaction between the two variables ($p = 0.85$).

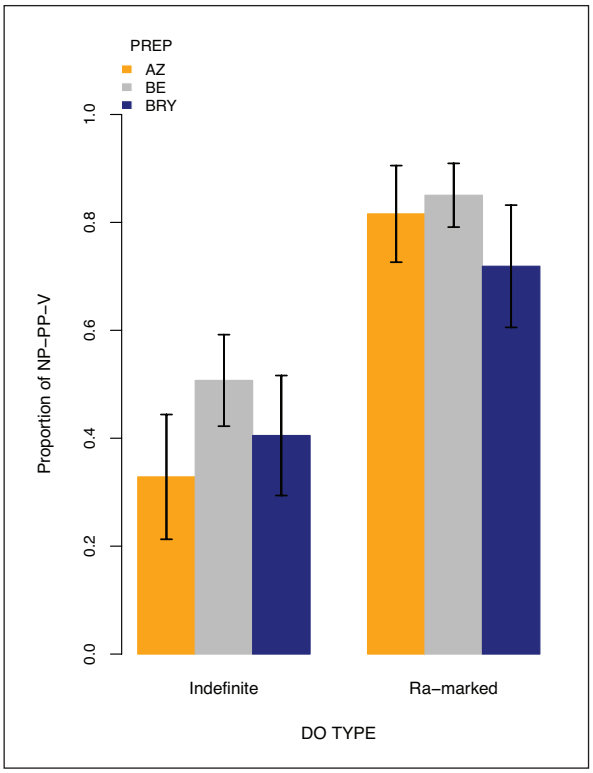


Figure 4: NP-PP-V rate by DO type and Preposition in Experiment 2.

Table 2: Distribution of word order for Experiment 2.

	NPV	NVP	PNV	NA	Total
Indefinite DOs					
Short PP	71	0	64	1	136
Long PP	43	1	92	0	136
Total	**114**	**1**	**156**	**1**	**272**
Rā-marked DOs					
Short PP	118	0	18	0	136
Long PP	100	0	35	1	136
Total	**218**	**0**	**53**	**1**	**272**

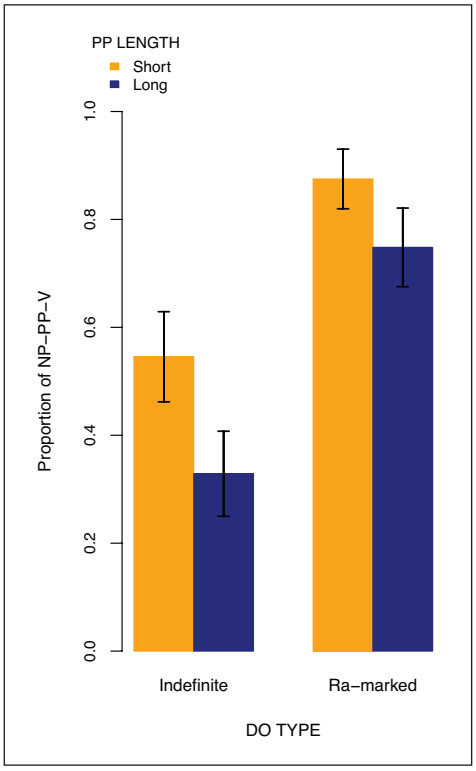


Figure 5: NP-PP-V rate by DO type and PP length in Experiment 2.

We also fitted the model separately for each DO type in order to check whether there is a numerical difference between the estimated coefficients of length between them. In addition, in order to have more homogeneous data for comparison, we also used a limited dataset including only items with the preposition *be*. In both cases, the estimated coefficient of PP length was slightly greater for indefinite DOs than for *rā*-marked DOs, respectively: –0.60 (SE = 0.16, $p < 0.001$) vs. –0.50, (SE = 0.18, $p < 0.01$), and –0.76 (SE = 0.23, $p < 0.01$) vs. –0.54 (SE = 0.276, $p < 0.05$).

In sum, our data show a robust length-based effect corresponding to the same long-before-short preference for both DO types. This effect is in accordance with the predictions of DLM, and importantly, it contradicts Faghiri & Samvelian's (2014) findings regarding the absence of a length effect on the position of *rā*-marked DOs. Nevertheless, while we find no significant interaction between the effects of relative length and DO

type, we observe that increasing the length of the PP is likely to have a larger effect in the case of indefinite DOs than *rā*-marked DOs. This is not surprising given that *rā*-marked DOs have a stronger bias toward the NP-PP-V order and are said to display less ordering variation than indefinite DOs (see e.g. Faghiri et al. 2018).

5 General discussion

Experiment 1 shows that there are strong length-based word order variations for which an account in terms of distance minimization is irrelevant. The effect of distance minimization on parsing is expected to be proportional to the relative length. In this experiment, the relative length between the PP and the NP varies by only one word between the two conditions. Hence, dependency-distance minimizing models do not predict a large effect, if any.

Meanwhile, an account in terms of the contribution of length to semantic enrichment/informativity is more satisfactory. One could safely assume that a restrictive modifier adds additional information to a non-specific NP, making its reference more specified/salient. In addition, we note that this effect conforms to observations made by Faghiri et al. (2014; 2018) and strengthens their claim that the position of the DO (with respect to the PP argument) depends on its degree of determination: the more determined and/or specific an NP, the more it is likely to precede the PP. If one considers that modification contributes to the degree of specification (and thus determination) of a referring expression, then these results are expected. Indeed, the comparison of the effect size between the DO type (bare vs. indefinite) and absence/presence of modification shows that, as claimed by Faghiri et al. (2018), the main predictor of the position of the DO is its degree of determination. Crucially, this result is compatible with a salient-first preference and supports Yamashita & Chang's (2001) conceptual accessibility hypothesis, which relates the long-before-short preference in OV languages to the semantic richness and informativity of longer constituents.

If we are on the right track, such manipulation of length should yield a fairly smaller effect in the case of (definite) *rā*-marked DOs, which are by definition specific referring expressions. The corpus data reported by Faghiri & Samvelian (2014) are compatible with this prediction, given that they do not show any effect of length on the order for *rā*-marked DOs (Faghiri & Samvelian 2014: 226). However, future experimental studies are required to test this hypothesis.[38]

In Experiment 2, the relative length between the PP and the NP is manipulated by adding a (four-word length) relative clause to the PP. The results clearly show that a long-before-short preference exists regardless of markedness and/or definiteness of the DO. Hawkins's EIC/MiD model falls short of accounting for the long-before-short preference altogether. The predictions of DLM, on the other hand, are met by the data: we find significant main effects for both relative length and DO type but no interactions between the two variables. This implies that the long-before-short preference is independent of the DO type.

It is worth noting that although distance minimization is relevant, we nevertheless observe that relative length has a much smaller effect size than DO type, which further supports the claim that the relative order between the PP argument and the DO is mainly determined by the DO type. Hence, it is not surprising that the rate of conformity to DLM, that is, the ratio

[38] Also, as mentioned by an anonymous reviewer, another rather straightforward way to test whether semantic richness per se plays a role on word order preferences is to manipulate the degree of informativity of nouns: e.g. "person" vs. "doctor" vs. "pediatrician".

of observed vs. expected shifts, is only 31% (13.6% and 48.1% respectively for *rā*-marked and indefinite DOs). In other words, overall DLM is not respected in our data.

At this point, it is interesting to compare Persian with other languages for which similar data are available. Below, we summarize data from studies on two SVO languages, French and English, and two SOV languages, Japanese and Basque. All these studies report data from comparable production experiments. The idea is to compare the rate of DLM-triggered shifts, that is, cases where there is a default/canonical word order and a heavy constituent shift (leftward or rightward) from its default position. We calculated the rate as the ratio of observed shifted cases, considering the expected order to be the non-shifted order. In Persian data, this rate, calculated for *rā*-marked DOs and corresponding to the rate of the PP-NP-V order with heavy PPs, is 26.3%.[39]

1. In a study on French focusing on indefinite DOs, the rate of conformity to DLM is about 70%, with about 53.7% of shifts for heavy NPs (Faghiri & Thuilier 2018: 10).
2. In a study on heavy NP shift in English by Stallings & MacDonald (2011: 182), the average rate of shifted heavy NPs is about 35.1% (for cases with comparable length difference between the two constituents) and can reach 40% when the length difference is the greatest (Stallings & MacDonald 2011: 185).
3. In the experimental study on Basque, the overall rate of shifts (to the left) with heavy NPs is above 60.1% (Ros et al. 2015: 1165).
4. In the experimental study on Japanese, the overall rate of shifts (to the left) with heavy NPs is 48.4% (Yamashita & Chang 2001: B54).

With respect to the relative order between the subject and the object in transitive sentences, based on the available data (see Section 3.3), we can safely conclude that the magnitude of the DLM effect is smaller in Persian transitive sentences compared to other SOV languages so far investigated. Recall that in their well-powered production experiment manipulating the length of the DO by adding a relative clause, Faghiri et al. (2018) did not find any effect of length on the rate of shifted orders. Yet, similar production studies (with comparable or even less statistical power) from Japanese and Basque find a clear long-before-short effect in the same configuration. The rate of shifted O-S orders with long DOs reached 27.5% (compared to 2.68% otherwise) in Japanese (Yamashita & Chang 2001: B50) and 19.7% (including both verb final and verb medial orders) contra 8.9% in Basque (Ros et al. 2015: 1165).

We can conclude from this comparison that the effect of DLM in Persian is less strong than in these languages. A similar observation has been made by Gildea & Temperley (2010: 304) "DLM is reflected much less strongly in German than in English".[40] The authors mention a number of specifics of German to explain this difference. Interestingly, there is one shared property between Persian and German that differentiates both of them from the other languages they are compared with, which is very likely to be relevant here: the fact that they both display mixed head direction while other languages are either consistently head-final or head-initial.

[39] It is important to bear in mind that in Basque (Ros et al. 2015: 1157) and Japanese (Yamashita & Chang 2001: B49–B50) the canonical order is assumed to be S-IO-DO-V. In French, like in English, the canonical order of postverbal complements is NP-PP.

[40] Note that their comparison is based on overall (uncontrolled) corpus-based calculation of the dependency length. Using data from treebanks for each language they calculate average dependency lengths for hypothetical optimized and random linearizations and compare them to the actual average dependency length.

Ros et al. (2015: 1168) also observe that dependency distance minimization is less effective in Basque (than Japanese and Korean),[41] however, they suggest that this difference is due to Basque's freer word order. They assume that when word order freedom increases, order becomes a less reliable parsing cue. As a result the impact of constituent length on word order might lessen. We think, on the contrary, that when word order is a strong parsing cue, shifted orders, even if they minimize the dependency length, are less efficient to process than non-shifted orders, and, consequently, dependency length is likely to have a lesser impact. Accordingly, it is expected that more grammaticalized/fixed word orders would reflect DLM less.[42] We get back to this issue in the conclusion.

6 Conclusion

In this paper, we have studied the effect of phrasal length on word order in Persian. Our data confirm a general long-before-short preference in line with other studies on SOV languages investigated so far (e.g. Basque, Korean and Japanese), contra the universal end-weight principle supported by availability-based models. This is important, because unlike previously studied OV languages, Persian is not consistently head-final and displays a mixed head direction.

We have provided solid experimental evidence for a dependency-distance-minimizing effect in Persian that previous corpus and experimental studies failed to detect (Faghiri & Samvelian 2014; Faghiri et al. 2014; 2018). Importantly, we have shown that Temperley's measure of dependency distance, DLM, based on Gibson's DLT (see Temperley 2007 and subsequent work) yields more accurate predictions than Hawkins's EIC/MiD (1994; 2004).

Furthermore, we have shown that there is also enough empirical evidence to maintain the conceptual accessibility hypothesis (Yamashita & Chang 2001), because we observe length-based effects that can hardly receive a dependency-length minimizing explanation, while they can be accounted for in terms of informativity and hence conform to the conceptual accessibility hypothesis. These findings imply that to explain word order preferences in Persian, and possibly in other languages, we need to take into account both a parsing-oriented account in terms of dependency distance minimization and a production-oriented account in terms of the conceptual accessibility hypothesis.

Finally, we have pointed out the fact that while dependency distance minimization is relevant for Persian, it is reflected less strongly in this language compared to other languages for which comparable data is available. In particular, in transitive sentences no dependency-distance-minimizing effect has been detected so far, which contrasts with what is reported for other studied SOV languages.

This is intriguing because Persian is considered an SOV flexible language, and, importantly, allows for different constituents to be placed in the postverbal domain. These findings may entail that the SOV order is more grammaticalized in Persian (than in these other languages) and thus constitutes a stronger parsing cue in this language.[43] In addition, differential object marking may also favor the reliance of parsing on word order, given the strong bias of *rā*-marked DOs towards a specific linear position (i.e. the NP-PP-V order), compared to non-*rā*-marked DOs, that display more variation.

[41] Their comparison with Korean is based on data from a similar production experiment by Dennison (2008) which only included transitive sentences.

[42] Note that this observation does not contradict the fact that word orders that comply with DLM are likely to become grammaticalized in a given language – for instance, one can argue that the postposition of subordinate clauses in Persian mentioned earlier (see Section 2.2 page 9) presents such a case.

[43] Importantly, for Basque as well, Ros et al. (2015: 1168) observe that distance minimization is less effective in transitive sentences where they argue that the SOV order seems to be grammaticalized.

Another difference between Persian and other investigated SOV languages is its "aberrant" properties with respect to word order typological universals, namely the fact that it displays also head-initial structures. Crucially, clausal verbal complements always occur post-verbally and a relative clause that modifies a preverbal constituent can be placed after the verb. Consequently, there is also enough evidence for a short-before-long preference in the postverbal domain that also shows a solid tendency for leftward ordering of heavy constituents. It could be the case that distance minimization is stronger in the postverbal domain that in the preverbal domain in Persian.

More experiments are required to test these assumptions and to pin down the respective contribution of different parsing cues. Also, crosslinguistic studies involving similar languages will certainly be promising in order to investigate the respective role of these parsing cue in relation to language specific typological properties.

Abbreviations

3 = third person, AUX = auxiliary, EZ = ezafe, IPFV = imperfective, PST = past, PRS = present, RSTR = restrictive, SG = singular

Additional Files

The additional files for this article can be found as follows:

- **Supplementary file 1.** List of experimental items. DOI: https://doi.org/10.5334/gjgl.1078.s1
- **Supplementary file 2.** Full summary of the GLMM models presented in Section 4. DOI: https://doi.org/10.5334/gjgl.1078.s2

Acknowledgements

We wish to express our gratitude to all participants who volunteered to take part in our experiments, as well as to friends who helped diffusing our call for participants in social networks. We are thankful to four anonymous reviewers whose comments and suggestions greatly improved the paper.

Funding Information

This work is partly supported by a public grant funded by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-LABX-0083).

Competing Interests

The authors have no competing interests to declare.

References

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483. DOI: https://doi.org/10.1023/A:1024109008573

Arnold, Jennifer E., Thomas Wasow, Anthony Losongco & Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76(1). 28–55. DOI: https://doi.org/10.2307/417392

Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4). 390–412. DOI: https://doi.org/10.1016/j.jml.2007.12.005

Bader, Markus & Jana Häussler. 2010. Word order in German: A corpus study. *Lingua* 120(3). 717–762. DOI: https://doi.org/10.1016/j.lingua.2009.05.007

Bates, Douglas, Martin Maechler & Ben Bolker. 2012. lme4: Linear mixed-effects models using s4 classes.

Behaghel, Otto. 1909. Beziehungen zwischen umfang und reihenfolge von satzgliedern. *Indogermanische Forschungen* 25(110–142).

Bever, Thomas G. 1970. The cognitive basis for linguistic structures. In John R. Hayes (ed.), *Cognition and language development*, 277–360. New York: Wiley & Sons, Inc.

Bock, J. Kathryn. 1982. Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review* 89(1). 147. DOI: https://doi.org/10.1037/0033-295X.89.1.1

Bock, J. Kathryn & David E. Irwin. 1980. Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior* 19(4). 467–484. DOI: https://doi.org/10.1016/S0022-5371(80)90321-7

Bock, J. Kathryn, Helga Loebell & Randal Morey. 1992. From conceptual roles to structural relations: bridging the syntactic cleft. *Psychological Review* 99(1). 150–171. DOI: https://doi.org/10.1037/0033-295X.99.1.150

Bock, J. Kathryn & Richard K. Warren. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21(1). 47–67. DOI: https://doi.org/10.1016/0010-0277(85)90023-X

Bock, J. Kathryn & Willem Levelt. 1994. Language production: Grammatical encoding. In Morton Ann Gernsbacher (ed.), *Handbook of psycholinguistics*, 945–984. New York: Academic Press.

Bossong, Georg. 1985. *Empirische universalienforschung*, vol. 14 (Ars linguistica). Tubingen: Gunten Narr.

Branigan, Holly P., Martin J. Pickering & Mikihiro Tanaka. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118(2). 172–189. DOI: https://doi.org/10.1016/j.lingua.2007.02.003

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

Chang, Franklin. 2009. Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language* 61(3). 374 – 397. DOI: https://doi.org/10.1016/j.jml.2009.07.006

Choi, Hye-Won. 2007. Length and order: A corpus study of Korean dative-accusative construction. *Discourse and Cognition* 14(3). 207–227. DOI: https://doi.org/10.15718/discog.2007.14.3.207

Dennison, Heeyeon Yoon. 2008. Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals. *Working Papers in Linguistics: University of Hawaii at Manoa* 39(2). 1–16.

De Smedt, Koenraad. 1994. Parallelism in incremental sentence generation. In Geert Adriaens & Udo Hahn (eds.), *Parallel natural language processing*, 421–447. Norwood, NJ: Ablex.

De Smedt, Koenraad & Gerard Kempen. 1987. Incremental sentence production, self-correction and coordination. In Gerard Kempen (ed.), *Natural language generation* (NATO ASI Series), 135. 365–376. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-009-3645-4_23

Dimmendaal, Gerrit Jan. 2011. *Historical linguistics and the comparative study of African languages*. Amsterdam: John Benjamins Publishing Company. DOI: https://doi.org/10.1075/z.161

Drummond, Alan. 2013. Ibex farm (free hosting for Internet-based experiments). Version 0.3. http://www.spellout.net/ibexfarm.

Faghiri, Pegah. 2016. *La variation de l'ordre des constituants dans le domaine préverbal en persan: approche empirique*. Paris: Université Sorbonne Nouvelle dissertation. https://tel.archives-ouvertes.fr/tel-01429416.

Faghiri, Pegah & Juliette Thuilier. 2018. Ordre des compléments postverbaux en francais: poids et accessibilité discursive. In *SHS web of conferences* (6e Congrès Mondial de Linguistique Française), 46. 14008. DOI: https://doi.org/10.1051/shsconf/20184614008

Faghiri, Pegah & Pollet Samvelian. 2014. Constituent ordering in Persian and the weight factor. In Christopher Pinon (ed.), *Empirical issues in syntax and semantics 10 (EISS 10)*, 215–232. CNRS. http://www.cssp.cnrs.fr/eiss10/.

Faghiri, Pegah & Pollet Samvelian. 2016. How much structure is needed: The case of the Persian VP. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Muller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 236–254. Stanford, CA: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2016/headlex2016-faghiri-samvelian.pdf.

Faghiri, Pegah, Pollet Samvelian & Barbara Hemforth. 2014. Accessibility and word order: The case of ditransitive constructions in Persian. In Stefan Muller (ed.), *Proceedings of HPSG 2014*, 217–237. Stanford, CA: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2014/fsh.pdf.

Faghiri, Pegah, Pollet Samvelian & Barbara Hemforth. 2018. Is there a canonical order in Persian ditransitive constructions? Corpus based and experimental studies. In Agnes Korn & Andrej Malchukov (eds.), *Ditransitive constructions in a cross-linguistic perspective*, 165–185. Wiesbaden: Reichert Verlag.

Ferreira, Fernanda. 1994. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language* 33(6). 715–736. DOI: https://doi.org/10.1006/jmla.1994.1034

Ferreira, Victor S. & Hiromi Yoshita. 2003. Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research* 32(6). 669–692. DOI: https://doi.org/10.1023/A:1026146332132

Frazier, Lyn & Janet D. Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6(4). 291–325. DOI: https://doi.org/10.1016/0010-0277(78)90002-1

Futrell, Richard, Kyle Mahowald & Edward Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences* 112(33). 10336–10341. DOI: https://doi.org/10.1073/pnas.1502134112

Garrett, Merrill F. 1980. Levels of processing in sentence production. In Brian Butterworth (ed.), *Language production. vol. 1: Speech and talk*, 177–220. London: Academic Press.

Gibson, Edward. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition* 68(1). 1–76. DOI: https://doi.org/10.1016/S0010-0277(98)00034-1

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita & Wayne O'Niel (eds.), *Image, language, brain: Papers from the first mind articulation project symposium*, 94–126. Cambridge, MA: MIT Press.

Gildea, Daniel & David Temperley. 2010. Do grammars minimize dependency length? *Cognitive Science* 34(2). 286–310. DOI: https://doi.org/10.1111/j.1551-6709.2009.01073.x

Gries, Stefan Th. 2003. Towards a corpus-based identification of prototypical instances of constructions. *Annual Review of Cognitive Linguistics* 1(1). 1–27. http://www.ingentaconnect.com/content/jbp/arcl/2003/00000001/00000001/art00001. DOI: https://doi.org/10.1075/arcl.1.02gri

Grieve, Robert & Roger J. Wales. 1973. Passives and topicalization. *British Journal of Psychology* 64(2). 173–182. DOI: https://doi.org/10.1111/j.2044-8295.1973.tb01340.x

Haitao Liu, Chunshan Xu & Junying Liang. 2017. Dependency distance: A new perspective on syntactic patterns in natural languages. *Physics of Life Reviews* 21. 171–193. DOI: https://doi.org/10.1016/j.plrev.2017.03.002

Hawkins, John A. 1990. A parsing theory of word order universals. *Linguistics Inquiry* 21(2). 223–261.

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

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199252695.001.0001

Hawkins, John A. 2007. Processing typology and why psychologists need to know about it. *New Ideas in Psychology* 25(2). 87–107. DOI: https://doi.org/10.1016/j.newideapsych.2007.02.003

Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199664993.001.0001

Karimi, Hossein & Fernanda Ferreira. 2016. Informativity renders a referent more accessible: Evidence from eyetracking. *Psychonomic Bulletin & Review* 23(2). 507–525. DOI: https://doi.org/10.3758/s13423-015-0917-1

Kempen, Gerard & Edward Hoenkamp. 1987. An incremental procedural grammar for sentence formulation. *Cognitive Science* 11(2). 201–258. DOI: https://doi.org/10.1207/s15516709cog1102_5

Kempen, Gerard & Karin Harbusch. 2003. Word order scrambling as a consequence of incremental sentence production. In Holden Hartl & Heike Tappe (eds.), *Mediating between concepts and grammar*, 141–64. Berlin: Mouton De Gruyter.

Kempen, Gerard & Karin Harbusch. 2004. A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In Thomas Pechmann & Christopher Habel (eds.), *Multidisciplinary approaches to language production*, 173–181. Berlin: Mouton de Gruyter.

Kimball, John. 1973. Seven principles of surface structure parsing in natural language. *Cognition* 2(1). 15–47. DOI: https://doi.org/10.1016/0010-0277(72)90028-5

Levelt, Wilem J. M. 1989. *Speaking: From intention to articulation* (ACL-MIT Press series in natural-language processing). Cambridge, MA: MIT Press. https://books.google.fr/books?id=LbVCdCE-NQAC.

Lohse, Barbara, John A. Hawkins & Thomas Wasow. 2004. Domain minimization in English verb-particle constructions. *Language* 80(2). 238–261. DOI: https://doi.org/10.1353/lan.2004.0089

Melnick, Robin Jeffrey. 2017. *Consistency in variation: On the provenance of End-Weight*. Stanford, CA: Stanford University dissertation.

Osgood, Charles Egerton & J. Kathryn Bock. 1977. Salience and sentencing: Some production principles. In Sheldon Rosenberg (ed.), *Sentence production: Developments in research and theory*, 89–140. Hillsdale: Lawrence Erlbaum.

Prat-Sala, Mercè & Holly P. Branigan. 2000. Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language* 42(2). 168–182. DOI: https://doi.org/10.1006/jmla.1999.2668

Quirk, Randolph, Sidney Greenbaum, Geoffrey Neil Leech, Jan Svartvik. 1972. *A grammar of contemporary English*. London: Longman.

R Core Team. 2013. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Rasekh-Mahand, Mohammad, Mojtaba Alizadeh-Sahraie & Raheleh Izadifar. 2016. A corpus-based analysis of relative clause extraposition in Persian. *Ampersand* 3. 21–31. DOI: https://doi.org/10.1016/j.amper.2016.02.001

Ros, Idoia, Mikel Santesteban, Kumiko Fukumura & Itziar Laka. 2015. Aiming at shorter dependencies: the role of agreement morphology. *Language, Cognition and Neuroscience* 30(9). 1156–1174. DOI: https://doi.org/10.1080/23273798.2014.994009

Samvelian, Pollet. 2018. Specific features of Persian syntax. The Ezāfe Construction, Differential Object Marking and Complex Predicates. In Anousheh Sedighi & Pouneh Shabani-Jadid (eds.), *The Oxford handbook of Persian linguistics*, 226–269. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198736745.013.9

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* 40(3). 177–187. DOI: https://doi.org/10.1006/jmla.1998.2586

Stallings, Lynne M., Maryellen C. MacDonald & Padraig G. O'Seaghdha. 1998. Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language* 39(3). 392–417. DOI: https://doi.org/10.1006/jmla.1998.2586

Tachihara, Karina & Adele E. Goldberg. 2020. Cognitive accessibility predicts word order of couples names in English and Japanese. *Cognitive Linguistics* 1(ahead-of-print). DOI: https://doi.org/10.1515/cog-2019-0031

Tanaka, Mikihiro N., Holly P. Branigan, Janet F. McLean & Martin J. Pickering. 2011. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language* 65(3). 318–330. DOI: https://doi.org/10.1016/j.jml.2011.04.009

Temperley, David. 2007. Minimization of dependency length in written English. *Cognition* 105(2). 300–333. DOI: https://doi.org/10.1016/j.cognition.2006.09.011

Temperley, David & Daniel Gildea. 2018. Minimizing syntactic dependency lengths: Typological/cognitive universal? *Annual Review of Linguistics* 4(1). 67–80. DOI: https://doi.org/10.1146/annurev-linguistics-011817-045617

Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9(1). 81–105. DOI: https://doi.org/10.1017/S0954394500001800

Wasow, Thomas. 2002. *Postverbal behavior*. Stanford: CSLI Publications.

Wasow, Thomas. 2013. The appeal of the PDC program. *Frontiers in Psychology* 4(236). DOI: https://doi.org/10.3389/fpsyg.2013.00236

Yamashita, Hiroko & Franklin Chang. 2001. "Long before short" preference in the production of a head-final language. *Cognition* 81(2). B45–B55. DOI: https://doi.org/10.1016/S0010-0277(01)00121-4

Yamashita, Hiroko & Franklin Chang. 2006. Sentence production in Japanese. In Mineharu Nakayama, Reiko Mazuka & Yasuhiro Shirai (eds.), *The handbook of East Asian psycholinguistics*, 291–297. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511758652.042

Yao, Yao. 2018. NP weight effects in word order variation in Mandarin Chinese. *Lingua Sinica* 4(1). 5. DOI: https://doi.org/10.1186/s40655-018-0037-8

Zoey Liu. 2019. Quantifying structural and lexical constraints in PP ordering typology. *Proceedings of the Society for Computation in Linguistics* 2(1). 306–309. DOI: https://doi.org/10.7275/ch62-ft92

How to cite this article: Faghiri, Pegah and Pollet Samvelian. 2020. Word order preferences and the effect of phrasal length in SOV languages: evidence from sentence production in Persian. *Glossa: a journal of general linguistics* 5(1): 86.1–33. DOI: https://doi.org/10.5334/gjgl.1078

Submitted: 31 August 2019 **Accepted:** 12 May 2020 **Published:** 31 August 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS