Aspects of *væ* ('and') as a discourse marker in Persian

Reza Kazemian[1] and Mohammad Amouzadeh[1,2]
[1] University of Isfahan | [2] Sun Yat-sen University

This study investigates the functions of *væ* ('and') as a discourse marker in Persian. More specifically, this study accounts for certain aspects of *væ* co-occurrences and their linearization order. Fraser's model (forthcoming) was mainly employed to classify the multiple functions conveyed by *væ*. A corpus-based approach was taken to provide an overview of *væ* co-occurrences with other discourse markers. The data were collected from both written and spoken corpora. Quantitative and qualitative analyses were conducted to examine the frequency and the functional differences in the use of *væ* in the data – namely, elaboration, inferential, contrast, and alternation. The results of the study indicate the mobile nature of *væ* in its co-occurrences with other DMs. The findings also show that some modifications to Fraser's (forthcoming) DM co-occurrence principles are required to handle certain cases of language-specific behavior of *væ* in Persian. The configuration suggested for *væ* uses and its multi-functionality will also shed some lights on cross-linguistic studies of its counterparts in other languages.

Keywords: *væ* discourse marker, DM co-occurrence, Pragmatic functions, Multifunctionality

1. Introduction

The importance of discourse markers (DMs) has attracted the attention of many scholars in linguistics with different perspectives (e.g., discourse connectivity and continuity, semantic change and grammaticalization, modal and pragmatic particles, etc.). This leads to diverse terms (see Dér 2010) and approaches for the investigation of DMs. Nonetheless, in most studies, they are mainly treated as communicative devices creating a connection between the preceding and following segments of discourse. The trend of investigation on DMs has not been confined within European major languages such as English (Aijmer 2002), German

Published online: 9 March 2022 *Pragmatics* 32:4 (2022), pp. 588–619.
ISSN 1018-2101 | E-ISSN 2406-4238

(Siebold 2021), Spanish (Pinto and Vigil 2020), French (Vanderbauwhede and Lamiroy 2020), but extended to other languages such as Arabic (Habib 2021) as well as Persian (Zoghdar-Moghdam and Dabirmoghdam 2002).

Persian DMs and their functions have been modeled through a number of diverse frameworks. However, earlier studies have primarily targeted written language and overlooked the fact that DMs can serve various functions in spoken language. Our survey also shows that certain important aspects of DMs in Persian remain understudied. One of these aspects relates to the pragmatic functions of *væ*. Despite extensive research on *and* in English (e.g., Sweetser 1990; Schiffrin 2006; Crible 2018), little research has been done in this area in Persian, except for the study conducted by Kassaei and Amouzadeh (2020). Yet, their study does not focus on *væ* and its multifunctionality (see Section 2.2). As a result, we have attempted to redress this gap by investigating the pragmatic functions of *væ* and its linearization order in two/multi-part co-occurrences. We hope to broaden our understanding of the complex phenomena of *væ* co-occurrences and multifunctionality. This may lay the groundwork for cross-linguistic investigations of 'and-constructions' in other languages (see Sweetser 1990).

2. Discourse markers and their co-occurrences

2.1 Discourse markers

Studies on DMs suffer from a lack of consensus at the level of definition. The question of how to define DMs has been a point of debate in research in this field (e.g., Fraser 1996; Schourup 1999; Schiffrin 2001; Crible 2017a; Heine et al. 2021). Accordingly, DMs can be viewed through the lens of two main criteria, viz. syntactic (integration and scope) and pragmatic (multifunctionality) (see Crible 2017a). Prototypically, they are syntactically optional elements and not an integrated part of the core syntax. They are relatively mobile and grammatically heterogeneous. DMs convey procedural rather than propositional meaning. Besides, what is more specific to DMs is their multifunctionality. This feature of DMs can be depicted in two forms: (1) they may have different functions on different occasions of use (see Section 4); or (2) they may simultaneously have different functions on a single occasion of use (see Section 5.1; cf. Dér 2010; Aijmer and Simon-Vandenbergen 2011).

The general thrust of a pragmatic approach is concerned with the meaning of utterances, particularly with how a DM in an utterance relates the message to that of a prior utterance. It presumes a separation between sentence (conceptual) and utterance (procedural) meaning. Thus, we start with a classification of pragmatic

meanings. To this end, we mainly use Fraser's (forthcoming) model to classify the DMs in our data. The reason for employing his model for this study is threefold. First, his model and method of classification provide a versatile analytical tool[1] that greatly helps to establish a firm footing for this study. Second, his classification is deeply rooted in a pragmatic approach, which holds a dominant position throughout this paper. Moreover, Kassaei and Amouzadeh (2020) have already found his model well-suited for the classification of Persian DMs. Note that we do not follow Fraser's model to the letter, so certain slight modifications will be made in order to meet our objectives. Specifically, we will employ his original model with two qualifications: (a) while Fraser's model is quite qualitative, the current study will be a combination of qualitative and quantitative analyses; and (b) since *væ* is strongly multifunctional and the notion of simultaneous multifunctionality of DMs is not considered in Fraser (forthcoming), this study tries to address the multifunctionality in question in terms of meaning potential.

Fraser (forthcoming) divides DMs into three major classes: retroactive (RDMs), linking (LDMs), and proactive (PDMs). The first class (RDMs) signal the speaker's perception of a prior utterance. For example, the bold parts in (1) below are treated as RDMs because they reflect the speaker's view of a former utterance. Yet, they designate different types of RDMs: 'oh', 'I see', 'well' and 'ok' are expressions, respectively, of surprise, recognition, and decision.

(1) A: I broke the window.
B: **Oh, I see. Well…Ok.** I guess you can pay for it. (Ibid.)

Fraser (forthcoming) takes the view that LDMs reflect the speaker's perspective of the relationship between earlier and forthcoming utterances. This class can further be divided into three main subclasses, each subclass being comprised of primary and secondary DMs (see Figure 1).

(2) A: I made Jake angry.
B: **And**, what did you say to him to make him mad? (Ibid.)

Finally, the third class of DMs are PDMs. As it is shown by Example (3), PDMs are the converse of RDMs as they signal the speaker's view of the following utterance rather than a preceding one. This class, similar to the other ones, comprises certain subclasses; namely, summarizing PDMs, attention-getting PDMs, commentary PDMs, illustrative PDM, and topic PDMs.

1. The versatility of this model is due mainly to investigating DM classifications and co-occurrences simultaneously, which is hardly seen in other studies. As the current study investigates both these issues, this model would be a firm foundation and fit this study neatly.

(3) A: There isn't any more food
B: **Anyway**, let's go home. (Ibid.)

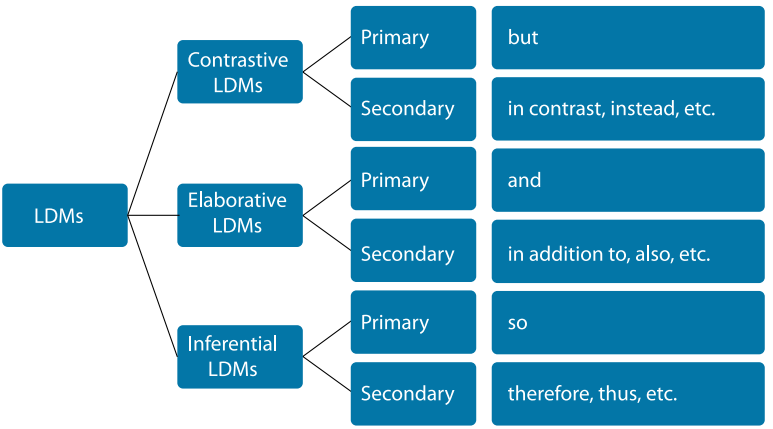

Figure 1. Fraser's Classification of Different types of LDMs

2.2 Discourse marker co-occurrences

DM co-occurrences are simply defined as contiguity of two or more DMs in an adjacency or non-adjacency order. This phenomenon has received increasing attention over the last decade since it is directly germane to the mobility and polyfunctionality of DMs. More recent studies on DM co-occurrences can be found in Oates (2000, 2001), Lohmann and Koops (2016), and Haselow (2019). Almost all of these studies examined the sequencing behavior and functional motivations underlying DM sequences from different perspectives. Certain motivations were proposed in earlier studies for such co-occurrences: (i) floor holding, (ii) functional specification, (iii) functional complementation (see Aijmer 2002; Haselow 2019).

The study of DM co-occurrences in Persian, however, seems to be limited, except for a few studies done by Ghaderi (2019), Kassaei and Amouzadeh (2020), and Ghaderi and Amouzadeh (2021). The more relevant study by Kassaei and Amouzadeh (2020) systematically investigated the combinations of Persian DMs. The authors (2020) examined all possible combinations of thirty Persian DMs (i.e., elaborative, contrastive, and inferential). They also argued that the order of DMs can, to a great extent, be predictable through certain frequent patterns. Moreover, they found that contrastive DMs are apt to combine with those from their own category, while elaborative and inferential ones are liable to take part in intra-category combinations. One of the main concerns in their work was the analysis of the combinatory behavior of *væ* DM. They claimed that *væ* appears

in an initial position in all its combinations with other DMs, and, consequently, conformed mainly to the findings of Oates (2000, 2001) and Fraser's (2009, forthcoming).

2.3 *væ* in Persian

Persian has three types of monosyndetic and bisyndetic coordinate conjunctions, namely conjunctive, adversative, and disjunctive (Stilo 2004, 271), and *væ* is treated as a monosyndetic coordinating conjunction. Although the coordinating conjunction *væ* and the connective clitic *-o* 'and'[2] are very close, and most people believe that they are two modes (spoken/written) of a single word, they vary etymologically. The former is derived from Arabic while the latter is a survival from middle Persian *u* (see Stilo 2004; Lambton 1953).

> Old Persian: utā > ud > uδ > Middle Persian: u > =o, 'and'
> (Kent 1953, 175; Horn 1893, 240)

This means that *væ* and *o* can conjoin any number of sentences, which are in a coordinating, causal or temporal relationship. In other words, they can be used to coordinate different elements (Mahootian and Gebhardt 1997):

- NP subjects/objects (two or more than two)
- VPs
- Attributive/predicate adjectives
- Attributive/predicate adverbs
- Adverbials (adverbs of manner and participial adverb constructions)

Almost all earlier studies on *væ* (e.g., Lazard 1992; Mahootian and Gebhardt 1997; Stilo 2004), have been confined to its syntactic and coordinative aspects, and heretofore none of them has been advertent to its discoursal use.

In this paper, we argue that *væ* cannot be simply and merely restricted to its coordinative nature. In most occurrences, it is semantically reduced and lacks propositional content. This gives rise to the emergence of pragmatic functions operating at higher levels, particularly at the discourse level. The functional properties of *væ* are by no means arbitrary, they are ascertained according to multiple features (e.g., prosodic, syntactic, pragmatic, etc.).[3] Therefore, the current study is the first of its kind that brings up *væ*, its multifunctionality and co-occurrences

---

2. The clitic-*o* is more common than the conjunction in informal speech (Mahootian and Gebhardt 1997).

3. These factors are also determinative of the order of DMs in sequences (see Crible and Degand 2021).

in terms of pragmatics, and it predominantly favors the pragmatic aspect of *væ*, which is rather overlooked in earlier studies.

3. Data and method

The data concerning *væ* occurrences and co-occurrences, differing in their modality and degree of formality, are based on both written and spoken corpora. TalkBank,[4] as a written corpus, is composed of over 474 million words, which makes it the largest Persian corpus.[5] It features diverse blog posts compiled from different Persian blog sites. In terms of style, its contents range from formal to informal, with various genres, such as politics, sports, economy, and culture. 1000 samples of *væ* tokens were identified randomly and extracted from the corpus for qualitative and quantitative analysis. The collected examples for this study were transliterated and translated into English. As a result of a widespread occurrence of DMs with various functions in conversations (see Bublitz, 2017; Crible and Cuenca 2017), and their pervasive co-occurrences, significantly in spoken data (Crible 2018), a spoken corpus was added to analyze *væ* with a finer granularity. The spoken data used in this study were compiled from approximately ten hours of conversation recorded in five sessions among eight native Persian participants: six adults and two children. It featured merely face-to-face conversations but included diverse interaction types (e.g., dialogue and group conversation). The conversations were recorded by using a cellphone in a similar setting (i.e., a park). To have access to natural and real-life data, the participants were kept unaware of the recording to the end of the data collecting task. Moreover, informed consent was obtained from all participants at the end of the task. After the process of anonymization, the illustrative examples required for this study were transcribed, based mainly on IPA and following the common conventions of conversational analysis as shown in Appendix A. Although the transcription method was verbatim,[6] only the parts containing *væ* occurrences and co-occurrences were transcribed.

---

4. This corpus can be accessed at https://www.sketchengine.co.uk.

5. There are other written Persian corpora, namely, the Bijankhan corpus and the corpus of the University of Leipzig, but we found the Persian TalkBank corpus more appropriate. First, the latter corpus is freely accessible to us through our institutional accounts. In order to show the different occurrences of *væ*, particularly when it co-occurs with other DMs, we needed to have a representative/balanced corpus. To this end, the TalkBank corpus was the right database to provide such diverse types of registers.

6. It characterizes all audible linguistic words and phrases as well as all non-linguistic vocalizations such as laughing and coughing.

Our bottom-up analyses of *væ* initially require identification of its actual uses in discourse. Thus, after the process of data collection, we identify all occurrences of *væ* DM entirely manually. Three positions at which it could occur are then distinguished. Note that *væ* can be used in these three positions, but not always with the same function. The next step is to provide a functional classification for different instances of *væ* based on the classification proposed by Fraser (forthcoming).

væ is one of the most frequent multifunctional Persian DMs (Kassaei and Amouzadeh 2020), for which Anvari (2001) provides 20 meanings. Based on Fraser's (forthcoming) model, to establish a functional classification for *væ*, we need to examine whether it can fulfill all LDM tasks. To achieve this goal empirically, 1000 *væ* tokens from the written corpus and 1000 tokens from the spoken data, were randomly selected and investigated. Functions carried by *væ* is not already apparent and, even in some cases, this task is highly complex and sometimes rather challenging. Therefore, in such circumstances, five native Persian speakers, being experts in linguistics, were asked to assist in determining the functions implicated by *væ* in order to improve the reliability and accuracy of the classification. The functions found in the data are presented in the next section.

4. Functions of væ as a DM

4.1 Elaborative function

The main function of *væ* can be described in terms of elaboration (see Kassaei and Amouzadeh 2020). It is considered to be a Persian Primary Elaborative Linking Discourse Marker (PELDM). As its name suggests, it signals that the following utterance provides an elaboration of a preceding one which can be viewed in different forms (e.g., addition, specification, temporal); in (4) the underlined utterance is regarded as an elaboration on the utterance preceding *væ*. This utterance is about the status of the U.S dollar currency in Iran and the elaboration provided by the next utterance falls within the same topic.

(4) *dær xosuse væz'yæt-e dolar(.)* ***væ*** *ælan væz'yæt hæm kæmi behtær*
in regard situation.EZ dollar(.) and now situation also little better
šode(.) bayæd beg-æm ke æz halæt-e hobab xaredʒ šode
get.PST.3SG must say.IMP-1SG that from form.EZ bubble out get.PST.3SG
About the status of the dollar (.) and now its status is better (.) I should say it is in a stable situation. (spoken data)

væ does not necessarily elaborate on a preceding utterance. In some cases, *væ* with an elaborative function operates beyond the sentence level and at the discourse

level. However, not all elaborative instances of *væ* can be seen as a single function since this includes some subfunctions, namely temporal, addition and specification. Yet, these subfunctions have been treated by some scholars as separate main functions of *and* (e.g., Halliday and Hasan 1976; Crible 2017). As Examples (5)–(9) show, all elaborative functions carried out by *væ* can be classified under one of these subfunctions as follows:

4.1.1 *Addition*

This is the prototypical function of *væ* that signals a simple addition to a previous utterance of information within the same topic. As (5) shows, the underlined utterance is believed to add some information to the preceding utterance. The utterance *dæst tænha hæstæm* ('I have no assistant') adds information regarding the status of *mæn kar daræm* ('I am busy').

(5) *væqti dærbareye æli porsid-æm(.) goft(.) mæn kar dar-æm*
when about-EZ Ali ask.PST-1SG (.) say.PST.3SG (.) I work have.PRS-1SG
væ *dæst tænha hæst-æm.*
and hand alone be.PRS-1SG
When I asked about Ali (.) he said (.): "I am busy **and** have no assistant".
(spoken data)

4.1.2 *Specification*

The specification function of *væ* resembles that of addition to some extent, particularly when it does not co-occur with specification DMs such as *bevize* ('especially') and *mæsælan* ('for example'). As a specificational marker, *væ* provides more details and examples in the segment following it. It applies when its following utterance describes the early situation in detail. Information stated in this utterance falls within the scope of the prior one. In (6) below, the specification reading of *væ* specifies one person (Javad) among many others. This specification reading utters in the process of pragmatic interpretation if there is an implied question (see Onea and Volodina 2011). In (6), we also observe that when the speaker introduces the proposition that "*many people came and went*" into the common ground, gives rise to an implied question, "*who exactly posed the problem?*". If such a question does not arise in the discourse, or if the speaker cannot assume that this question is present in the discourse, we may not see specificational *væ* in use.

(6) *xeili-ya umæd-æn o ræft-æn væ Jævad bud-e ke hæmiše*
many.PL come.PST-3PL and go.PST-3PL and Javad be.PST.3SG that always
moškel dorost kærd-e
problem correct do.PST-3SG
Many people came and went, and it was Javad who always caused the problem.
(spoken data)

4.1.3 *Temporal*

Elaborative sub-functions are quite similar to each other in such a way that making a distinction among them requires meticulous attention to available distinctive cues. Apropos of the temporal function, the most important cue to make the distinction is the specific stress carried by *væ* concerning a chronological order. Without this distinctive cue, one may not be able to discern the correct type of elaboration. As can be induced from (7), *væ* signals slight chronological stress on the order of the preceding and following segments. In this example, it is impossible to replace the utterance preceding *væ* with the one following it since chronological order would be halted in a real sense and, consequently, the sentence would be pragmatically inappropriate. In (7), it is quite clear that the action of using the mobile camera is a prerequisite for seeing digital content on the mobile screen. Cases like (7), in which changing the order of clauses would lead to a change in the interpretation of the sentence, have been called the asymmetric use of *væ* (see Sweetser 1990).

(7) *be ʤay-e an bærčæsb-ha, maikrosaft æz yek durbin-e*
to place-EZ that label-PL Microsoft from one camera-EZ
telephone-hamrah bæraye eskæn kærdæn-e barkod estefade mi-kon-æd
phone-accompany for scan do.INF-EZ barcode use IPF-do.PRS-3SG
væ *mohtæva-ye diʤital ru-ye telephone-hamrah næmayeš dade*
and content-EZ digital on-EZ phone-accompany show give.PST.PRF
mi-šæv-æd.
IPF-get.PRS-3SG
Microsoft uses the mobile camera instead of those labels in order to scan the barcode **and then** the digital content is shown on the mobile screen.
(written data)

4.2 Contrastive function

As mentioned above, it is difficult to draw a clear-cut distinction between addition and specification or addition and temporal. The same is true about contrast and concession. However, a distinction between contrast and concession is deemed

to be inevitable, and we make such a distinction whenever it is necessary, despite the fact that they are highly related to each other and that the distinction will be challenging. In the contrastive function, we merely compare two things in dissidence, while in a concession-based one, the observed situation is contrary to expectations. Furthermore, there is some counter-expectation in concessions, and the implied causal relation is canceled (see Webber et al. 2019).

In Example (8), *væ* functions as a contrastive marker since the segment following it is in contrast with the one preceding it; it signals a direct contrast between the preceding segment and the following one. In this example, the speaker believes everyone is making progress, however, her brother is doing worse than before.

(8) *hæme dar-æn pišræft mi-kon-ænd væ in dadaš-æm hær*
all have-3PL progress IPF-do.PRS-3PL and this brother-POSS.1SG each
ruz-eš bædtær æz diruz-e
day-POSS.3SG worse than yesterday-be
Everyone is progressing but my brother's situation is getting worse day by day.
(spoken data)

As noted above, the concession is conveyed when a causal relation based on one argument is canceled or denied. Correspondingly, Example (9) resembles a similar situation, which can be interpreted as follows: 'because I was sick, I was supposed not to come to your wedding party, and yet I came'. Here *væ* signals the segment following it as well as the fact that is the result of a canceled causal relation.

(9) *mæn mæriz bud-æm* ***væ*** *umæd-æm ærusi-tun.*
I sick be.PST-1SG and come.PST-1SG wedding-POSS.2PL
I was sick **yet** I came to your wedding party. (spoken data)

4.3 Inferential function

væ can also be used as an inferential marker. For example, in (10a) the utterance preceding *væ* is considered to be a reason for the one following it, while the following utterance is regarded as a result of the one before it. In Example (10a) people prefer *Ash* ('soup') and this preference is because of the cold weather:

(10a) *hæva særd šode (.)* ***væ*** *mærdom aš ra bištær tærʤih mi-dæn.*
weather cold get.PST.PRF(.) and people soup OBJM more Preference IPF-give.PRS.3PL
The weather has become cold (.) **and** people would rather eat soup.
(spoken data)

Akin to the other functions noted before, here the change in the order of the present segments occurring before and after *væ* is almost impossible. It stems from the fact that the segment following *væ* is regarded as a result of the segment preceding it; hence the result is not placed before the cause. By contrast, (10b) can be regarded as a typical example of the asymmetric use of *væ*. Although (10b) seems syntactically correct, it is semantically peculiar since the result has been placed before the cause, and the relationship between the two segments (following and preceding *væ*) is nonsensical.

(10b) ?*mærdom aš ra bištær tærʤih mi-dæn (.)* ***væ*** *hæva særd šod-e*
people soup OBJM more preference IPF-give.PRS.3PL(.) and weather cold get.PST-PRF
? People would rather eat soup (.) **and** the weather has become cold.

4.4 Alternative function

This function has been regarded as an elaborative function for a long time (cf. Fraser 2009; Kassaei and Amouzadeh 2020). It is here recognized as a single function, and not an elaborative sub-function. As the name of this function suggests, here *væ* is adopted to present alternatives. Unlike functions mentioned above, the utterances placed before and after *væ* can be used interchangeably with no alteration to the unmarked meaning. This may affect the level of emphasis or stress, as the first utterance is generally of primary importance. As is understood from (11a), the utterance following *væ* is an alternative to the previous one, both of them are underlined, and this does not present any elaboration, but only an alternation. In other words, *væ* in (11a) can be replaced by either *væ ya* ('and or') or *ya* ('or'). (11b) can display the palpable interchangeability of utterances in this function, which is quite impossible in previous ones.

(11) a. *ægær seda o sima film-ha-yi bær in æsas besaz-æd* ***væ***
If voice and face movie-PL-EZ on this base SUBJ.make-3SG and
filmnameh-ye bærxi film-ha ra be suræt-e ketab dæravær-æd
scenario-EZ some movie-PL OBJM to form-EZ book make.PRS-SG
mosælæman kudak-an esteqbal mi-kon-ænd
certainly child-PL welcome IPF-do.PRS-3PL

b. *ægær seda o sima filmname-ye bærxi film-ha ra be suræte ketab*
if voice and face scenario-EZ some movie-PL OBJM to form book
dæravær-æd ***væ*** *film-ha-yi bær in æsas besaz-æd mosælæman*
make.PRS-3SG and movie-PL-EZ on this base SUBJ.make-3SG certainly
kudak-an esteqbal mi-kon-ænd
child-PL welcome IPF-do.PRS-PL
If IRIB make movies on these bases **or** publish books out of certain scenarios, children will certainly like it. (written data)

4.5 Topic-shifting function

The functions of *væ* are not merely limited to LDM ones. In some cases, *væ* can also signal a topic PDM function. As Example (12) illustrates, while speaker (A) is talking about her life in Turkey, the second speaker shifts the topic from her life to her brother's. *væ* is used as a topic-shifter to change/start a new topic, where none of the other functions applies. Here, in contrast to the earlier functions denoting connectivity, *væ* makes a break with the previous utterance and introduces a new topic.

(12) A: *mæn ke tu torkie xeili æzyæt šod-æm*
I that in Turkey very irritate get.PST-1SG

B: ***væ*** *bæradær-et či kar mi-kon-e? (.)alman zendegi*
and brother-POSS.2SG what do IPF-do-3SG? (.)Germany life
mi-kon-e?
IPF-do.PRS-3SG?

A: I got really irritated in Turkey.

B: And what does your brother do? Is he living in Germany? (spoken data)

5. Results and discussion

væ shows a high degree of multifunctionality (see Table 1), and it can be used in different contexts to express diverse intentions. Each function carried out by *væ*

can be subsumed under one of these discourse functions: elaborative, contrastive, inferential, alternative, and topic-shifting.

Table 1. Frequency of *væ* LDM functions

væ functions	Written data (1,000 occurrences)	Verbal data (1,000 occurrences)
Elaborative	94.8%	95.8%
Contrastive*	3.4%	0.8%
Inferential	0.6%	1.3%
Alternative	1.2%	1.0%
Topic-shifting	0.0%	1.1%
Total	**100**	**100**

* As some studies (e.g., Crible 2017a; Fraser forthcoming) show, and as a consequence of the low frequency of *væ* serving a concessive function, this function has not been separated from the contrastive one.

Besides the aforementioned functions, two further sub-functions have been identified, which merit further consideration. The elaborative and contrastive functions are also divided into sub-functions, discussed in the previous section, and their frequencies are illustrated in Table 2. Unsurprisingly, addition is the most prevalent function among other ELDM subfunctions, and, correspondingly, among the other functions of *væ* in both written and verbal data. The prevalence of addition is genuinely consistent with Anvari's (2001) view and with the cross-linguistic studies of 'and-constructions' by Halliday and Hasan (1976) and Sweetser (1990). Addition must be the core meaning of *væ* for three reasons: (a) it is the most frequent function of *væ* (see Table 1); (b) it was registered as the first semantic entry for *væ* by Anvari (2001); (c) from a cross-linguistic perspective, the multiple interpretations of *væ* and its equivalents in other languages might be due to the iconic usage of a general concept of addition or connectivity (see Sweetser 1990).

Table 2. Frequency of *væ* ELDM sub-functions

væ ELDM sub-functions	Written data (948 occurrences)	Verbal data (958 occurrences)
Addition	84.4%	85.2%
Specification	3.2%	6.6%
Temporal	7.2%	4%
Total	**94.8%**	**95.8%**

5.1 Simultaneous multifunctionality of *væ*

In addition to the remarks above, there are different LDMs classified by their functions. The question then is whether one LDM can fulfill other functions. This leads us to the notion of multifunctionality, a perpetual feature of DMs; however, this important concept did not receive due attention by Fraser (forthcoming).

We view multifunctionality through the lens of meaning potential (see Norén and Linell 2007). Thus, we do not take a maximalist semantic approach to *væ*, signifying that *væ* does not characterize different stable lexical meanings performing different functions. However, according to the theory of meaning potential, *væ* is treated as a potentially polysemous word whose meanings are determined by the interface between meaning potential and contextual factors (e.g., co-text and situational conditions). In other words, the multiple meanings of *væ* are not considered to be its constant features.

It must be noted that *væ* as a multifunctional DM can have additional, specific, temporal, inferential, contrastive, alternative, and topic-shifting uses in different contexts. Its functions can be grouped into three domains of discourse represented in Figure 2. These domains were inspired mainly from Halliday and Hasan (1976), Redeker (1990), Sweetser (1990), and Gonzalez (2005); then they were revised and redefined by Crible (2017b, 107):

a. *ideational*: discourse relations between real-world events;
b. *rhetorical*: discourse relations between epistemic, speech-act events and metadiscursive functions;
c. *sequential*: structuration of discourse segments, both for local management of small units and macro-level organization.

If we regard 'addition' as the core meaning of *væ* (see the previous section), we can present its functional scopes covering from its core sequential domain down to less pragmatic uses in the ideational domain (see Figure 2).

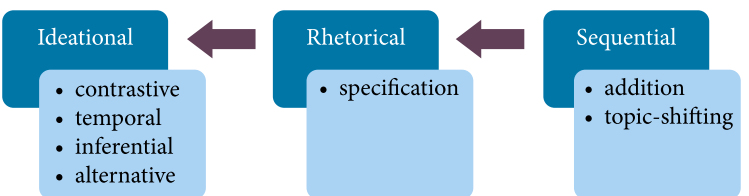

Figure 2. A Functional spectrum of *væ*

Now the question is whether *væ* can serve different functions simultaneously. A positive answer can be provided by the analysis of our data. The observed patterns of multifunctionality in our data are as follows:

i. *Additional-temporal function*

væ can designate the succession of events that helps discourse forward. It simultaneously adds more information to a preceding utterance. In other words, *væ* signals a sort of temporal relation between utterances, as well as adding more information within the same topic. In (13) below, not only does *væ*[7] show the sequential order of events (poisoning, killing, losing), but it also can signal addition to the current topic, which is the *leopard.*

(13) *do hæfte piš yek made pælæng dær hævali-ye rustay-i dær*
two week before one female leopard in around-EZ village-INDF in
hævali-ye Poldoxtær tævæsot-e šekarči-an mæsmum šod ***væ*** *be qætl*
around-EZ Poldoxtar by-EZ hunter-PL poisonous get.PST.3SG and to kill
res-id ***væ*** *tule-ha-ye in pælæng napædid šod*
reach.PST-3SG and cub-PL-POSS this leopard disappear get.PST.3SG
Two weeks ago, a female leopard was poisoned and killed by hunters around a village in Poldoxtær; then her cubs got lost. (written data)

ii. *Additional-contrastive function*

A close analysis of the instances of addition indicates that, in some cases, *væ* proves not to be restricted to the addition function and signals contrast concurrently. For example, in (14), while *væ* signals addition, it also reflects a sense of contrast with the previous utterance. Despite the fact that people are putting effort to solve the problem, the problem has not been eradicated yet.

(14) *in æfrad besyar tælaš mi-kon-ænd* ***væ*** *be sefr resid-æn-e in mozu'*
this people very attempt IPF-do.PRS-3PL and to zero reach-INF-EZ this issue
kar-e došvari æst.
activity-EZ difficult be.PRS.3SG
These people try a lot, but removing this issue is difficult. (written data)

iii. *Additional-inferential function*

An inferential function can also simultaneously co-occur with the core meaning of *væ*. That is to say, *væ* not only signals additional information to a preceding

7. One of the reviewers kindly pointed out that the second *væ* could also function as a consequential marker. This might be true if we assume that the killing of the leopard has led to her losing her cubs.

utterance or discourse, but also signals the causal role of such an utterance that could be inferred from the following one. For example, *væ* in (15) below shows that the utterance following *væ* adds more information to the current message as well as indicating the causal relation between the two utterances. Here, the fact that galaxies get far from each other is added information within the same topic. It is also the result of the previous utterance. Such binary functional behavior of *væ* is indicative of its simultaneous multifunctionality.

(15) *enerži-ye tarik be enbesat-e ʤæhan komæk mi-kon-æd* ***væ***
energy-EZ dark to expansion-EZ world help IPF-do.PRS-3SG and
kæhkešan-ha æz hæm dur mi-šæv-ænd
galaxy-PL from each other far IPF-get-3PL
Dark energy helps world expansion and galaxies get far from each other.
(written data)

iv. *Specific-contrastive function*

All instances of the simultaneous multifunctionality of *væ* are not always subject to constraints of addition. In some cases, we can see other functions signaled by *væ* concurrently. For example, in (16) below, *væ* in specificaiton reading takes a certain number out of a total amount (i.e., 30 chairs out of 130 chairs). As already mentioned, this kind of pragmatic interpretation is an answer given to an implied question underlying the discourse. Here, the implied question is '*how many seats did they expect to win?*'. By focusing on (16), we observe that *væ* simultaneously establishes a contrastive relation between the following and preceding utterances. This sense of contrast can be felt in a way that the political movement nominated one hundred and thirty candidates while their expectation was limited to only thirty seats.

(16) *in ʤæriyan sæd o si kandid-a ra mo'ærefi kærd*
this stream hundred and thirty candidate-PL OBJM nomination do.PST.3SG
væ *tænha be si dærsæd-e korsi-ha češm duxte bud*
and only to thirty percent-EZ seat-PL eye sew.PTCP be.PST.3SG
This (political) movement nominated one hundred and thirty candidates, but they only expected to win thirty seats. (written data)

The presence of multifunctionality in *væ* was delineated above. Interestingly enough, the two simultaneous functions of it come from different domains of discourse. As illustrated in Figure 3, at least two functional domains are required to be involved. In other words, it seems that functions from the same domain would not be simultaneously present.

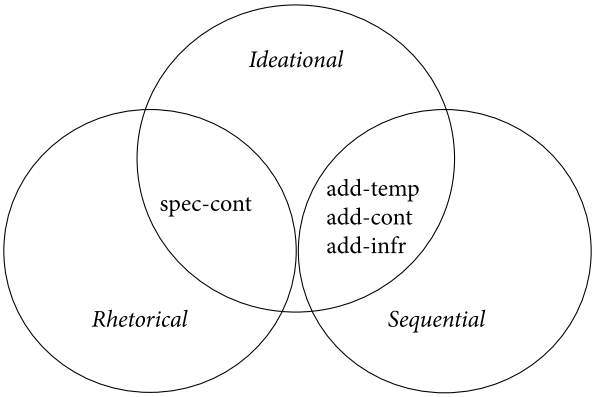

Figure 3. Domains of simultaneous functions

5.2 væ co-occurrences

As noted earlier, Fraser (forthcoming) provided a list of principles corresponding to DM co-occurrences. Here our aim is to revisit the ones pertinent to *væ* co-occurrences and to provide a detailed overview of its co-occurrences with other DMs. To the best of our knowledge and based on analyzed data, the general tendency towards *væ* co-occurrences can be listed as follows:

i. It can co-occur with all the secondary DMs from the same class. They can be additive, specific, or temporal DMs. For instance, *væ* in (17) co-occurs with an additive DM thereby the information, which is '*this winning could put our mind at ease*', is added to the results of Sepahan (an Iranian football club) win.

(17) A: *sepahan ba in bord 43 emtyaz-i šod o*
Sepahan with this winning 43 score-INF get.PST.3SG and
um-æd sædr-e ʤædvæl.
come.PRS-3SG top-EZ table
B: ***væ hæmčenin*** *xyal-e ma ra hæm rahæt kærd.*
and also mind-POSS we OBJM also easy do.PST.3SG
A: After winning this match, Sepahan came out at the top of the table with 43 scores.
B: **And also** it put our mind at ease. (spoken data)

However, *væ* may seem redundant when it co-occurs with other DMs that are nearly synonymous (e.g., *væ hæmčenin* 'and also'); this is almost identical to the mechanism proposed by Oates (2000), who pointed out that the impression of redundancy has not been made in this case. This stems from the fact that *væ* is bleached semantically in these co-occurrences and its potential meaning is directed by the DM, which follows or precedes it. The central

intriguing fact about these co-occurrences is that *væ* is bleached to the extent that it can be deleted without creating any flaw in discourse flow. Another point concerns the fact that deleting the second DM will not lead to any discourse gap if *væ* is used as a single DM either.

ii. As can be seen in Table 3, *væ* may co-occur with almost all secondary LDMs. These co-occurrences can function either following or preceding *væ*.

Table 3. *væ* DM co-occurrences with Persian SLDMs

væ	**Elaborative DMs**	*væ*	**Contrastive DMs**	*væ*	**Inferential DMs**	*væ*	**Alternative DMs**
✓	*hæmčenin* (also)	✓	*væli* (but)	✓	*bænabærin* (therefore)	✓	*ya* (or)
✓	*be viže* (especially)	✓	*æz suj-e digær* (on the other hand)	✓	*leza* (thus)	✓	*ya inke* (or that)
✓	*mæsælæn* (for example)	✓	*æz tæræf-e digær* (on the other hand)	✓	*dær nætiǯe* (therefore)		
✓	*mæxsusæn* (particularly)	✓	*dær moqabel* (by contrast)	✓	*be hæmin ellat* (because of this)		
✓	*xosusæn* (particularly)	✓	*ba in hal* (nevertheless)	✓	*be hæmin dælil* (because of this)		
✓	*bexosus* (particularly)	✓	*ba in voǯud* (nonetheless)	✓	*æz in ru* (thus)		
✓	*be ælave* (in addition)	✓	*ba voǯud-e in* (in spite of this)	✓	*æz in su* (because of this)		
✓	*be ezafe* (in addition)	✓	*bær æks* (on the contrary)	✓	*natiǯætæn* (therefore)		
✓	*æ'lave bær in* (moreover)			✓	*be hæmin xater* (because of this)		
				✓	*æz in xater* (because of this)		

iii. As noted above, *væ* can occur either before or after SELDM. However, Fraser (forthcoming), and Oates (2000), in English, assume that PLDMs are followed by SLDMs and that a reverse sequence is impossible. In terms of the linear order of markers, they argue that weak markers (e.g., *væ* 'and', *æmma* 'but', and *pæs* 'so') are followed by strong makers (*hæmčenin* 'also', *væli* 'but', and *bænabærin* 'thus') while in Persian this order is not perfectly genuine. In this connection, it can be argued that in Persian not only *væ* but also *æmma* ('but') and *pæs* ('so'), as PLDMs, can be either preceded or followed by SLDMs. This possibility can be observed in Examples (18a) and (18b) below, where *væ* and *hæmčenin* can be used interchangeably.[8]

(18) a. ***hæmčenin væ*** *be gozareš-e isna moʤtæba šærifi æz dæ'væt-e*
also and to report-EZ ISNA Mojtaba Sharifi from invitation-EZ
æli dayi væ sadeq dorudgær xæbær dad
Ali Dai and Sadeq Dorudgar news give.PST.3SG

b. ***væ hæmčenin*** *be gozareš-e isna moʤtæba šærifi æz dæ'væte*
and also to report-EZ ISNA Mojtaba Sharifi from invitation.EZ
æli dayi væ sadeq dorudgær xæbær dad.
Ali Dai and Sadeq Dorudgar news give.PST.3SG
And also, based on the report provided by ISNA, Mojtaba Sharifi announced that Ali Dayi and Sadeq Droudgar would be invited.
(written data)

iv. Fraser (forthcoming) argues that in sequences of PLDMS, particularly those involving *and,* only the sequence of *and so* is highly probable, but he controverts other alternative forms of co-occurrence. However, our examination of Persian data reveals that *væ* can also occur with all Persian PLDMs, which can be either in the form of following or preceding PLDMs. Examples (19a) and (19b) below are typical of the PELDM and PCLDM case of co-occurrences, where *væ* is following the PCLDM (19a) and is followed by that (19b). Note that the two DMs in (19a) do not form a single unit in terms of combination like in (19b). In other words, the co-occurrence of *æmma væ* ('but and') in (19a) does not constitute a combined discourse marker conveying a similar function.

8. This might have different reasons, but Persian's flexible word order and free structuring constituent (see Faghiri and Samvelian 2020, 8) seem a compelling reason for the mobility of DMs in sequences. Here we merely intend to show the possibility of change in the order of DMs, but it should be noted that it is certainly the case that, in some sequences, changes in position can lead to different pragmatic meanings.

(19) a. *dæqiqæn ne-mi-dun-am če zæman-i ra ruy-e dæstgah*
exactly NEG-IPF-know-1SG what time-INDF OBJM on-EZ system
tanzim kærd ***æmma væ*** *be mæn goft ke qæsd*
setting do.PST.3SG but and to me say.PST.3SG that intention
dašte ab be dzuš beres-e.
have.PST.3SG.PRF water to boiling reach.PRS-3SG

b. *dæqiqæn ne-mi-dun-am če zæman-i ra ruy-e dæstgah*
exactly NEG-IPF-know-1SG what time-INDF OBJM on-EZ system
tanzim kærd ***væ æmma*** *be mæn goft ke qæsd*
setting do.PST.3SG and but to me say.PST.3SG that intention
dašte ab be dzuš beres-e.
have.PST.3SG.PRF water to boiling reach.PRS-3SG
I don't know exactly what time he set the device, but he told me that he wanted to boil the water. (spoken data)

It is important to note that these two DMs can undergo incoordination[9] in some cases, when they function as a topic-shifting, turn-shifting, or turn-taking DM. As it can be gathered from (20), *væ æmma* ('and but') is employed to take a turn. This function can be seen in roundtable discussions by moderators. Here, two guests are discussing with each other, but the moderator halts the discussion to take the turn and deliver it to another speaker.

(20) A: [*in ke šoma mi-gid dʒænbe-ye fælsæfi-ye une*]
this that you IPF-say.PST.2PL aspect-EZ philosophical-EZ that
B: [*Næ (.) mæn æslæn kari be fælsæfe næ-dar-æm*]
no I absolutely work to Philosophy NEG-have.PRS-1SG
Moderator: ***væ æmma*** *begzar-id æz aqa-ye kælantæri be-porsæm*
and but let.IMP-3PL from sir.EZ Kalantari
A: [what you say is its philosophical perspective]
B: [NO (.) I absolutely don't care about the philosophical perspective]
Moderator: **but** let me ask Mr. Kalantari (spoken data)

This incoordination can be used in terms of topic shifting as well. As (21a) below shows, this co-occurrence acts as a device to change the topic. Speaker C has two

9. Here, the notion of incoordination for *væ æmma* traces back to what Kuteva et al. (2017) introduced as incoordinate sentences. In these sentences, the connectors 'and' and 'but' occur at the beginning of sentences while losing their original meaning and status as coordinative conjunctions. Kuteva et al. (2017) believe that when these conjunctions go through the incoordination process, they often transform into sentence particles with mirative values, namely, sudden discovery, surprise, unprepared mind, counter expectation, new information (for details on the category of mirativity, see Aikhenvald 2012, 437).

pieces of news, a good and a bad one. First she tells the bad one, which provokes hearers' (speaker A and B's) emotional reactions. To change the topic and lighten the hearers' somber mood, she benefits from the topic shifting nature of *væ æmma*.

(21a) A: *xeili narahæt šodæm bæra-š*
very angry get.PST.1SG for-OBJP
B: *are (.) bæd šod*
Yeah(.) bad get.PST.3SG
C: ***væ æmma*** *xæbære xub-æm*
and but news.EZ good-POSS.1SG
A: I felt really sad for her
B: yeah (.) it was really bad
C: but my good news (spoken data)

Moreover, an interesting fact about this co-occurrence with regard to the different functions it undergoes is that its linearization order can be flexible; that is, the function remains constant by the omission of *væ* or *æmma* ('but'). This possibility can be seen in Examples (21b) and (21c) below.

(21b) A: *xeili narahæt šodæm bæra-š*
very angry get.PST.1SG for-OBJP
B: *are (.) bæd šod*
Yeah(.) bad get.PST.3SG
C: ***væ*** *xæbære xub-æm*
and news.EZ good-POSS.1SG
A: I felt really sad for her
B: yeah (.) it was really bad
C: and my good news

(21c) A: *xeili narahæt šodæm bæra-š*
very angry get.PST.1SG for-OBJP
B: *are (.) bæd šod*
Yeah(.) bad get.PST.3SG
C: ***æmma*** *xæbære xub-æm*
but news.EZ good-POSS.1SG
A: I felt really sad about her
B: yeah (.) it was really bad
C: but my good news

Another mode of PLDM co-occurrence would be PELDM + PILDM. As was the case with the previous one, in this sequence, *væ* can be followed or preceded. However, the occurrence of *væ* in a preceding position is by far more frequent. In

Examples (22a) and (22b) below, although the original order belongs to the first one, the second one is possible as well.

(22a) A: *are (.) mæn dadaš-æm ke eqdam kærd bæra viza(.) mæn*
Yeah(.) I brother-POSS.1SG that action do.PST.3SG for visa(.) I
æm kærd-æm
also do.PST-1SG
B: *'e (.) **væ pæs** šoma æm ræftæni šod-i?*
Oh(.) and so you also going get.PST-2SG?

(22b) A: *are (.) mæn dadaš-æm ke eqdam kærd bæra viza(.) mæn*
Yeah (.) I brother-POSS.1SG that action do.PST.3SG for visa(.) I
æm kærd-æm.
also do.PST-1SG
B: *'e (.) **pæs væ** šoma hæm ræftæni šod-i?*
Oh(.) so and you also going get.PST-2SG?
A: yes (.) I also applied for a visa when my brother did so.
B: oh (.) **so** you are going as well? (spoken data)

An intriguing feature of PELDM + PILDM, not the reverse order, is that they can be nonadjacent and signal the same function, but they are mainly used in an informal context as illustrated in Example (22c) below.

(22c) A: *are (.) mæn dadaš-æm ke eqdam kærd bæra viza(.) mæn*
Yeah (.) I brother-POSS.1SG that action do.PST.3SG for visa(.) I
æm kærd-æm.
also do.PST-1SG
B: *'e (.) **væ** šoma hæm ræftæni šod-i?*
Oh(.) and you also going get.PST-2SG?

With regard to the last sequence of *væ* with PLDM, note that it co-occurs with PALDM *ya* ('or'). As Examples (23a) and (23b) show, the two above-mentioned features of PLDMs (linear order flexibility and optionality) can be seen in the case of the last sequence.

(23) a. *aya ne-mi-šæv-æd hæta dær ebteday-e hærekæt be an-ha*
QP[10] NEG-IPF-get.PRS-3SG even in beginning-EZ movement to that.PL
amuzeš dad ***ya væ*** *tæzækor-at-i dærmorede mævarede*
education give.PST.3SG or and warning-PL-INDF about case.PL.EZ
imeni be an-ha amuxt?
safety to that-PL teach.PST.3SG

b. *aya ne-mi-šæv-æd hæta dær ebteda-ye hærekæt be anh-a*
QP NEG-IPF-get.PRS-3SG even in beginning-EZ movement to that-PL
amuzeš dad ***væ ya*** *tæzækor-at-i dærmored-e mævared-e*
education give.PST.3SG and or warning-PL-INDF about-EZ case.PL-EZ
imeni be an-ha amuxt?
safety to that-PL teach.PST.3SG
Isn't it possible to train them or teach them some safety precautions, even at the beginning of their movement? (written data)

One should note that the results of corpus-based studies (see Kassaei and Amouzadeh 2020) on sequences of *væ* with other PLDMs indicate that conventional patterns, in which *vae* occurs earlier, are predominant. This might be due to the weak and ambiguous nature of *væ*; the weaker the DM for a given relation, the more it will be compensated by other DMs (Crible 2020).

DM co-occurrences do not always merely consist of two parts. There are instances of three-DM co-occurrences as well. As noted earlier, Fraser (forthcoming) assumed that PLDMs could be followed, but not preceded, by two SLDMs from the same class. However, *væ* in multi-part co-occurrences acts quite differently. It can be followed, preceded, or even be placed between DMs; such a possibility is evident below.[11] All instances below (a-f) seem theoretically possible in Persian and can be substituted with *væ hæmčenin be'ælave* ('and also in addition') in (24).

a. *væ hæmčenin be'ælave* (and also in addition)
b. *væ be'ælave hæmčenin* (and in addition also)
c. *hæmčenin væ be'ælave* (also and in addition)
d. *be'ælave væ hæmčenin* (in addition and also)
e. *hæmčenin be'ælave væ* (also in addition and)
f. *be'ælave hæmčenin væ* (in addition also and)

10. Question Particle

11. The point that should be raised here is that although all these three co-occurrences signal a single function, which is addition, there is a kind of variation in their prosodic structure when each combination is uttered.

(24) *bazikon-e tim-i hæmčon sepahan pæs æz do fæsl qæhremani-e*
player-EZ team-IND like Sepahan so from two season championship-EZ
motævali dær lig-e bærtær be læhaz-e arameš væ væz'yæt-e
continuous in league-EZ premier to aspect-EZ calmness and status-EZ
mætlubi ke dær Isfæhan væ ordugah-e tim-æš vodzud
good that in Isfahan and camp-EZ team-POSS.3SG existence
*dar-æd **væ hæmčenin be'ælave(.)** šærayet-e ideal in tim dær*
have.PRS-3SG and also in addition condition.PL-EZ ideal this team in
lig-e qæhreman-an-e asya ke mi-tævan-æd dær suræt-e edame
league.EZ champion-PL-EZ Asia that IPF-can-3SG in way-EZ continue
movæfæqyæt final væ qæhremani-e in dore ra be karname xod ezafe
success final and championship-EZ this term OBJM to resume self add
kon-æd.
do.PRS-3SG
A player of a team like Sepahan, which was the champion of the premier league for two seasons, because of favorable conditions in Isfahan and the camp, and also the ideal status of the team in Asian champion league, can have the honor of being in the final match and the championship of this course.
(written data)

With respect to multi-part co-occurrences, not only does *væ* co-occur with SELDMs, but it also might co-occur with heterogeneous SLDMs. This can be observed in different instances of (25a, b, c) below. Sequences of *væ* with other SLDMs bear a close resemblance to the linearization order that it follows in contiguity with SELDMs. To state it explicitly, *væ* might occupy all the three different positions in multi-part sequences with SLDMs.[12] Note that (25b) is the original example from our data and the two others show the mobility of *væ* when it co-occurs. Here, the speaker is worried as he and his colleagues quarreled with their boss, but on the other hand, he is happy that they could finally voice their concerns. The multi-part sequence of '*væli væ æz tæræfi*' was used to demonstrate this contrast.

12. Although all multi-part instances of *væ* co-occurrences might bear the same meaning, diverse pragmatic functions of *væ* make it necessary to note that we cannot afford to overlook the possibility of change in meaning in some cases, when it shifts from one position to the other one.

(25) a. ***væ væli æz tæræf-i*** *bæd hæm næ-šod(.)* *belæxereh*
and but from side-INDF bad also NEG-get.PST.3S(.) finally
hærfe-mun-a *zæd-im.*
word-POSS.1PL=OBJM hit.PST-1PL

b. ***væli væ æz tæræf-i*** *bæd hæm næ-šod(.)* *belæxereh*
but and from side.INDF bad also NEG-get.PST.3S (.) finally
hærfe-mun-a *zæd-im.*
word-POSS.1PL=OBJM hit.PST-1PL

c. ***væli æz tæræf-i væ*** *bæd hæm næšod(.)* *belæxereh*
but from side-INDF and bad also NEG-get.PST.3S (.) finally
hærfe-mun-a *zæd-im.*
word-POSS.1PL=OBJM hit.PST-1PL
But on the one hand, it wasn't bad (.) finally we could say what we intended to say. (spoken data)

As regards the motivations underlying *væ* co-occurrences with other DMs, some of the proposals made by earlier studies about DMs (see Section 2.2) appear to be fairly reasonable. The first underlying motivation, which was explained above, is known as functional specification. It suggests that in sequences of DMs, one DM, usually the stronger one, specifies the function of one with which it co-occurs (see Oates 2000; Haselow 2019). Sequences like *væ sepæs* ('and then'), can be explained as being motivated by restricting and specifying the vague meaning of *væ*. One could note that such motivation also confirms Gricean view of production and comprehension processes[13] (see Crible 2020). Alternatively stated, *væ* is of little informative value and does need to be enriched by other signals, which expedite the process of comprehension in discourse. The presence of compensating signals like *sepæs* ('then') might be mostly established when they co-occur with weak DMs, rather than with stronger ones. This potential factor is consistent with the Uniform Information Density Hypothesis (Levy and Jaeger 2007). According to this hypothesis, when an accurate interpretation can be plausibly deduced from one part of a sentence, the need for extra markers will be obviated.

In addition to functional specification, floor-holding can be viewed as another motivation underlying DM co-occurrences. From this point of view, *væ*

13. On the basis of the Gricean model, speakers/authors try to be as informative as required and produce informative utterances, such that the listeners or readers can understand the message (Grice 1975). Regarding his maxim of quantity, DM occurrences are subject to ambiguity and redundancy. To avoid redundancy, and as high-informative DMs can achieve an adequate level of interpretation, there is no need for another DM. In contrast, low-informative and ambiguous DMs such as *væ* are expected to be compensated by other DMs to be more transparent.

is indicative of one of the disfluencies occurring in online communication, as it demands instant cognitive processing and interlocutors' undivided attention (see Bortfeld et al. 2001; De Klerk 2005). In this contiguity, *væ* serves to bridge the gap during a cognitive planning pause. For example, in (26) below, *væ* within a lengthier span of production is used to increase the processing time of comprehension, and it is plainly void of semantic meaning. Here, the second speaker is thinking about the first speaker's ideas about new philosophical concepts, correspondingly *væ* is lengthened to fulfill this goal.

(26) A: *mæfahim-e novin-e fælsæfi æz dele dʒame'e birun*
concept.PL-EZ new-EZ philosophical from heart-EZ society outside
mi-yad
IPF-come.PST.3SG
B: ***væ:: pæs*** *to: ruykærd-e sonnæti ra qæbul*
and:: so you: approach-EZ traditional OBJM acceptance
næ-dar-i(.) *dorost mi-g-æm?*
NEG-have.PRS-3SG(.) right IPF-say.PRS-1SG
A: New philosophical concepts come out of society
B: So:: you: don't accept the classical approach(.) am I right? (spoken data)

Cognitive planning is not the only reason for floor holding motivations. In some cases, the speaker's cognitive orientation in discourse processing accompanies some strategies of online communication, such as turn-holding. As it is understood from (27), not only does speaker (A) use *væ* to buy time for his cognitive process, but he also attempts to keep the floor by repetition and sudden stress via a high pitch on the first *væ*. In (27), the second speaker wants to start vindicating himself immediately after hearing the first speaker's accusatory remarks, but the first speaker does not let him by repetitive use of *væ*.

(27) A: *kar-i kærdi ke næ-bayæd mi-kærd-i*
action-IND do.PST.2SG that NEG-must IPF-do.PST-2SG
B: *bebin*
look.IMP
A: **<u>væ</u>** ***væ: æmma*** *hala sæ'y kon dige del-e kæsi ra*
and and: but now try do.IMP.2SG anymore heart-POSS anybody OBJM
næ-škun-i
NEG-break.PRS-2SG
A: You did something that you shouldn't have done
B: Look!
A: but try to make no one sad anymore. (spoken data)

6. Conclusion

We have examined the *væ* functions and its co-occurrence. The results reveal that three types of functions for LDMs can be identified by *væ*. We have identified one more function for *væ*; namely, alternation, which has not been mentioned in Fraser (forthcoming). With respect to the multifunctionality of *væ*, we have shown that it can be simultaneously multifunctional, a fact that can be explained by recourse to the notion of meaning potential. The results have also shown that two simultaneous functions cannot originate from the same domain of discourse.

The second part of our analysis has dealt with DM co-occurrences. The empirical examination of *væ* co-occurrences has also revealed that *væ* co-occurrences do not perfectly conform in terms of DM combinations to the model proposed by Fraser (forthcoming). Moreover, *væ* co-occurrences can also cast doubt on proposals of the combined DMs made by scholars such as Oates (2000, 2001) and Fraser (forthcoming). They note that in DM co-occurrences, the first DM is typically a coordinator or a weak marker, while the second one is the more specific one or a strong marker. Their proposed degrees of integration are not completely consistent with the findings of the current study. That is to say, unlike, but complementary to adopted stances on the co-occurrence of 'and' and *væ* in earlier studies (Kassaei and Amouzadeh 2020), *væ* is not bound up with any position, and it can occur before and after other DMs. It is also true that every occurrence of *væ* does not appear with the same degree of frequency; normally, one variation is more or less common than the others. As our data show, the case of *væ* in terms of co-occurrences is not usual when it is compared with the established principles concerning DMs co-occurrences, which require further research.

A promising step towards finding what motivations and reasons lead to such co-occurrences and linearization order, in addition to what was mentioned in the earlier section, would be to examine factors such as functional, cognitive, and prosodic patterns associated with co-occurrences as well as the feature of multifunctionality. Another point that should be taken into account is the frequency of these co-occurrences that varies for different reasons and motivations. Despite the considerable progress made in studying *væ* and its co-occurrences, many important issues, particularly motivations behind the linearization order of DMs, remain unexplored. Last but not least, this article merely focused on the general and functional aspects of *væ* and its co-occurrences. Further research should be carried out on other Persian DMs individually, and in their occurrences in different sequences. The results of this study also suggest a number of new avenues for cross-linguistic and contrastive studies of DMs, which are direct equivalents of *væ* in other languages such as *und* (in German), *et* (in French), *ve* (in Turkish), etc.

Funding

Research funded by Alexander von Humboldt-Stiftung (Fo10416003) to Mohammad Amouzadeh.

Acknowledgments

We would like to express our deep gratitude to Tania Kuteva, Bernd Heine and Bruce Fraser for their invaluable and constructive suggestions on the earlier versions of our paper. We are also thankful to the Alexander von Humboldt Foundation for sponsoring this project. We would like to extend our thanks to the anonymous reviewers for their invaluable comments.

List of Abbreviations

EZ	ezafe marker
IMP	imperative
INDF	indefinite
INF	infinitive
IPF	imperfect tense
NEG	negative marker
OBJM	object marker
OP	object particle
PL	plural
POSS	possessive
PRF	perfect tense
PRS	present tense
PST	past tense
PTCP	participle
SG	singular
SUBJ	subjunctive

References

Aijmer, Karin. 2002. *English Discourse Particles: Evidence from a Corpus*. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.10

Aijmer, Karin, and Anne-Marie Simon-Vandenbergen. 2011. "Pragmatic Markers." In *Discursive Pragmatics* 8, ed. by Jan Zienkowski, Jan-Ola Östman, and Jef Verschueren, 223–247. Amsterdam: John Benjamins. https://doi.org/10.1075/hoph.8.13aij

Aikhenvald, Alexandra Y. 2012. "The Essence of Mirativity." *Linguistic Typology* 16 (3): 435–485. https://doi.org/10.1515/lity-2012-0017

Anvari, Hasan. 2001. *Færhænge Soxæn*. Tehran: Soxæn. [in Persian]

Bortfeld, Heather, Silvia D. Leon, Jonathan E. Bloom, Michael F. Schober, and Susan E. Brennan. 2001. "Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender." *Language and Speech* 44 (2): 123–147. https://doi.org/10.1177/00238309010440020101

Bublitz, Wolfram. 2017. "Oral Features in Fiction." In *Pragmatics of Fiction*, ed. by Miriam A. Locher, and Andreas H. Jucker, 235–263. De Gruyter Mouton. https://doi.org/10.1515/9783110431094-008

Crible, Ludivine. 2017a. "Discourse Markers and (Dis)fluency in English and French." *International Journal of Corpus Linguistics* 22 (2): 242–269. https://doi.org/10.1075/ijcl.22.2.04cri

Crible, Ludivine. 2017b. "Towards an Operational Category of Discourse Marker: A Definition and Its Model." In *Pragmatic Markers, Discourse Markers and Modal Particles*, ed. by Chiara Fedriani, and Andrea Sansó, 99–124. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.186.04cri

Crible, Ludivine. 2018. *Discourse Markers and (Dis)fluency*. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.286

Crible, Ludivine, and Maria-Josep Cuenca. 2017. "Discourse Markers in Speech Characteristics and Challenges for Corpus Annotation." *Dialogue and Discourse* 8 (2): 149–166. https://doi.org/10.5087/dad.2017.207

Crible, Ludivine. 2020. "Weak and Strong Discourse Markers in Speech, Chat, and Writing: Do Signals Compensate for Ambiguity in Explicit Relations?" *Discourse Processes* 57 (9): 793–807. https://doi.org/10.1080/0163853X.2020.1786778

Crible, Ludivine, and Liesbeth Degand. 2021. "Co-occurrence and Ordering of Discourse Markers in Sequences: A Multifactorial Study in Spoken French." *Journal of Pragmatics* 177: 18–28. https://doi.org/10.1016/j.pragma.2021.02.006

De Klerk, Vivian. 2005. "Procedural Meanings of 'Well' in a Corpus of Xhosa English." *Journal of Pragmatics* 37 (8): 1183–1205. https://doi.org/10.1016/j.pragma.2004.11.001

Dér, Csilla. 2010. "On the Status of Discourse Markers." *Acta Linguistica Hungarica* 57 (1): 3–28. https://doi.org/10.1556/ALing.57.2010.1.1

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. https://doi.org/10.5334/gjgl.1078

Fraser, Bruce. 1996. "Pragmatic Markers." *Pragmatics* 6 (2): 167–190. https://doi.org/10.1075/prag.6.2.03fra

Fraser, Bruce. 2009. "An Account of Discourse Markers." *International Review of Pragmatics* 1 (2): 293–320. https://doi.org/10.1163/187730909X12538045489818

Fraser, Bruce. forthcoming. "Canonical Sequences of Discourse Markers in English."

Ghaderi, Soleiman. 2019. Baresi Mo'tarezeh Are/Na dar Zabane Farsi [The Thetical Aspects of Are/Na (Yes/No), in Persian]. PhD Thesis, University of Isfahan.

Ghaderi, Soleiman, and Mohammad Amouzadeh. 2021. "Aspects of *Are* (Yes) in Persian Discourse: Its Functions, Positions, and Evolution." *Studia Linguistica* 75 (3): 623–658. https://doi.org/10.1111/stul.12173

González, Montserrat. 2005. "Pragmatic Markers and Discourse Coherence Relations in English and Catalan Oral Narrative." *Discourse Studies* 7 (1): 53–86. https://doi.org/10.1177/1461445605048767

Grice, Herbert P. 1975. "Logic and Conversation." In *Syntax and Semantics Vol. 3: Speech Acts*, ed. by Peter Cole, and Jerry L. Morgan, 41–58. Brill.

Habib, Rania. 2021. "The use of the Discourse Markers yaʕni and ʔinnu: 'I mean' in Syrian Arabic." *Journal of Pragmatics* 178: 245–257. https://doi.org/10.1016/j.pragma.2021.03.025

Halliday, Michael Alexander Kirkwood, and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Haselow, Alexander. 2019. "Discourse Marker Sequences: Insights into the Serial Order of Communicative Tasks in Real-time Turn Production." *Journal of Pragmatics* 146: 1–18. https://doi.org/10.1016/j.pragma.2019.04.003

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long. 2021. *The Rise of Discourse Markers*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108982856

Horn, Paul. 1893. *Grundriss der Neupersischen Etymologie*. Strassburg: Karl J. Trübner. https://doi.org/10.1515/9783111699394

Kassaei, Gholamreza, and Mohammad Amouzadeh. 2020. "The Combination of Discourse Markers in Persian." *International Review of Pragmatics* 12 (1): 135–163. https://doi.org/10.1163/18773109-01201102

Kent, Roland G. 1953. *Old Persian: Grammar. Texts. Lexicon*. New Haven: American Oriental Society.

Kuteva, Tania, Bernd Heine, Peter Austin, Seongha Rhee, Marine Vuilermet, and Domenico Niclot. 2017. "The 'Mirror' of Insubordination." *Linguistics Departmental Seminar Series, SOAS University of London*. Available at: https://www.youtube.com/watch?v=Jfq1KcoRens

Lambton, Ann KS. 1953. *Persian Grammar: Including Key*. Cambridge: Cambridge University Press.

Lazard, Gilbert. 1992. *A Grammar of Contemporary Persian*. Cosa Mesa: Mazda Publishers.

Levy, Roger, and T. Florian Jaeger. 2007. "Speakers Optimize Information Density through Syntactic Reduction." In *Advances in Neural Information Processing Systems* (NIPS), ed. by Bernhard Schölkopf, John Platt, and Thomas Hofmann, 849–856. MA: MIT Press.

Lohmann, Arne, and Christian Koops. 2016. "Aspects of Discourse Marker Sequencing – Emprical Challenges and Theoretical Implications." In *Outside the Clause: Forms and Functions of Extra-clausal Constituents*, ed. by Gunther Kaltenbock, Evelien Keizer, and Arne Lohmann, 417–446. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.178.14loh

Mahootian, Shahrzad, and Lewis Gebhardt. 1997. *Persian*. London: Routledge.

Norén, Kerstin, and Per Linell. 2007. "Meaning Potentials and the Interaction between Lexis and Contexts: An Empirical Substantiation." *Pragmatics* 17 (3): 387–416. https://doi.org/10.1075/prag.17.3.03nor

Oates, Sarah Louise. 2000. "Multiple Discourse Marker Occurrence: Creating Hierarchical for Natural Language." In *Procedding of the 3rd CLUK Colloquium*, 41–45. Brighton.

Oates, Sarah Louise. 2001. Multiple Discourse Occurrence: Creating Hierarchicies for Natural Languages Generation. MA dissertation. University of Brighton.

Onea, Edgar, and Anna Volodina. 2011. "Between Specification and Explanation: About a German Discourse Particle." *International Review of Pragmatics* 3 (1): 3–32. https://doi.org/10.1163/187731011X561036

Pinto, Derrin, and Donny Vigil. 2020. "Spanish Clicks in Discourse Marker Combinations." *Journal of Pragmatics* 159: 1–11. https://doi.org/10.1016/j.pragma.2020.01.009

Redeker, Gisela. 1990. "Ideational and Pragmatic Markers of Discourse Structure." *Journal of Pragmatics* 14 (3): 367–381. https://doi.org/10.1016/0378-2166(90)90095-U

Schiffrin, Deborah. 2001. "Discourse Markers: Language, Meaning, and Context." In *The Handbook of Discourse Analysis*, ed. by Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton, 54–74. Malden. Oxford: Blackwell.

Schiffrin, Deborah. 2006. "Discourse Marker Research and Theory: Revisiting and. In *Approaches to Discourse Particles*, ed. by Kerstin Fischer, 315–338. Oxford: Elsevier.

Schourup, Lawrence. 1999. "Discourse Markers." *Lingua*, 107 (3–4): 227–265. https://doi.org/10.1016/S0024-3841(96)90026-1

Siebold, Kathrin. 2021. "German *dann* – From Adverb to Discourse Marker." *Journal of Pragmatics* 175: 129–145. https://doi.org/10.1016/j.pragma.2021.01.010

Stilo, Donald. 2004. "Coordination in Three Western Iranian Languages: Vafsi, Persian and Gilaki." In *Coordinating Constructions*, ed. by Martin Haspelmath, 269–330. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.58.16sti

Sweetser, Eve. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620904

Vanderbauwhede, Gudrun, and Béatrice Lamiroy. 2020. "On Two French Discourse Markers and Their Dutch Equivalents: d'ailleurs and par ailleurs." *Journal of Pragmatics* 156: 168–175. https://doi.org/10.1016/j.pragma.2019.06.006

Zoghdar-Moghadam, Reza, and Mohammad Dabirmoghadam. 2002. "Contrastive Discourse Markers: The Case of "but" in English and "amma" in Persian." *Language Researches* 7 (12), 55–76. [in Persian]

Appendix A. Transcription Conventions

[ ]	overlap and simultaneous talk
=	latching
(.)	micro pause
(2.0)	measured pause
:,::	segmental lengthening according to duration
rea(hh)lly	laugh particles within talk
ABsolutely	strong, primary stress via loudness
<u>really</u>	stress via pitch or amplitude
.	falling intonation (terminal pitch)
,	continuing intonation
?	rising intonation
¿	a rise stronger than mid-level but weaker than high-terminal pitch

Publication history

Date received: 27 January 2021
Date accepted: 6 December 2021
Published online: 9 March 2022