Approaching the mountain of Exodus 19: thou shalt explore syntax first

Eep Talstra,

VU University Amsterdam, Faculty of Theology, De Boelelaan 1105, 1081 HV,

l.talstra1@hetnet.nl

Abstract: The objective of the research presented in this paper is to expand our existing database of Biblical Hebrew text with data and data structures to have an instrument that will allow a researcher access to various patterns of linguistic phenomena, such as patterns of clause relations, patterns of verbal valence and patterns of participant tracking in texts. It is the area where linguistic analysis and literary analysis meet and even have a considerable overlap. The question is: where can distributional linguistic analysis of patterns and system bring us? Identifying participants is in fact a part of a larger area of text-level research area of experimenting with pattern recognition and the assignment of linguistic functions to particular categories of data distribution. It appears, first, that pattern recognition as a technique to expand the database of classical Hebrew texts with more 'functional' features is effective. And, second, that this type of research is best approached from different angles at the same time: text grammar, valence research and participant tracking.

Keywords: text grammar, discourse analysis, pattern recognition, participant tracking.

1. Actual Research

The main goal of this paper is to describe the development of programs that contribute to the linguistic analysis of the actors or participants present in texts of the Hebrew Bible. The chapter chosen for this introduction is Exodus 19, in which are narrated the actions and dialogues between God, Moses and the people just before the proclamation of the Decalogue at mount Sinai. In this chapter one can observe a rather small set of actors as the main participants in the text: 'Moses', 'God', 'people of Israel' and 'the priests'. Also 'the mountain Sinai' could be mentioned: it is frequently merely the location of actions, but it can be a subject, too. Just like the people in verse 16 the mountain "trembles" as well (verse 18).

Composed of labels in English this listing of actors already reveals that a reader of a text usually makes a number of identifications of the participants in a more or less automatic way. The reader of the Hebrew text of Exodus 19 will observe that "God" is being referred to as יְהוָה, but also as הָאֱלֹהִים and in direct speech sections as "I", or as "you", or again as יְהוָה. The "people of Israel" are presented as בְּנֵי יִשְׂרָאֵל ("sons of Israel"), but also as יִשְׂרָאֵל ("Israel"), as בֵּית יַעֲקֹב ("house of Jacob") and as הָעָם ("the people"), and in direct speech sections also as אַתֶּם ("you") and as "we". The reader 'knows' that the "we" in "we will do" (verse 8) refers to the people, whereas "us" in "you have warned us" (verse 23) refers to Moses and the people together as a group. The question is: can one design algorithms to analyze linguistically such mechanisms of labelling and identification in order to store them as well-structured and easily accessible data? What is text grammar here and what is rhetorical or discourse analysis?[1]

---

[1] For the search for such data in Bible translation and in text pragmatic analysis see *e.g.*, L.J. de Regt, *Participants in Old Testament Texts and the Translator. Reference devices and their Rhetorical Impact*, SSN 39 (Assen: Van Gorcum, 1999); Nicolai Winther-Nielsen, "Towards the Peak of Mount Sinai: A Discourse-Pragmatic Analysis of Exodus 19." *SEE-J Hiphil 2* [http://www.see-j.net/hiphil] (2005); S. E. Runge, *A Discourse-*

From a linguistic point of view a search for such text grammatical mechanisms is relevant to help us understand more of text grammar and discourse analysis.[2] It involves, for example, the use of noun phrases and patterns of pronominalization. We find the clause וַיֹּאמֶר יְהוָה אֶל מֹשֶׁה ("YHWH said to Moses") in verse 9, 10, 21 and the clause וַיֹּאמֶר אֵלָיו יְהוָה ("YHWH said to him") with a different order of constituents and the pronominal suffix referring to Moses, in verse 24. Can one deduce a text grammatical rule that explains the reason for pronominalization in this clause type being applied here in verse 24 only?

Another example is the variation of patterns for introducing direct speech sections. Sometimes, as in the clauses just mentioned, both the speaker and the addressee are explicitly referred to, in other cases, *e.g.*, וַיַּעֲנוּ כָל-הָעָם יַחְדָּו וַיֹּאמְרוּ ("the people all answered as one and said", verse 8), there is gapping, *i.e.*, the absence of an explicit reference to the one who speaks or is being spoken to. For what linguistic reason should one assume Moses to be the addressee in verse 8? What is the discourse level information a reader is supposed to know or to build up during the process of reading?

Arguing from a literary point of view, however, an exegete of this chapter might wonder: if we as readers appear to know how to identify these participants in Exodus 19, why ask a computer to calculate these kinds of identifications for us? One answer to this question is that the identification of participants is not always self-evident, and any particular choice has an impact on translation and interpretation. So, it would be helpful to translators and exegetes to have direct access to comparable cases in the rest of the book of Exodus, or even in the entire Hebrew Bible. Secondly, one will always run into cases of a more or less unexpected usage of participants. For example, in Exodus 19:3 one finds בֵּית יַעֲקֹב ("house of Jacob") constructed parallel to בְּנֵי יִשְׂרָאֵל ("children of Israel"). How common is that in the book of Exodus or in the Hebrew Bible? Is it just a stylistic variation? Another case is Exodus 19:4 where Moses is commanded by YHWH to speak to Israel: "You yourselves have seen what I have done to Egypt". The usual linguistic pattern in such cases is that the "I" in direct speech sections refers to the one that is announced as its speaker. In this case that would be Moses. However, from reading the direct speech section it becomes clear that here the "I" expressed by Moses is YHWH. Is this a (socio)linguistic pattern which is used with other participants as well? If so, one might be able to propose a syntactic pattern for recognizing it. Or is it a literary pattern used only with YHWH and Moses or other prophets? Similarly, one might want to find more parallels to the unexpected shift from הָאֱלֹהִים ("God") in verse 2 to יְהוָה ("YHWH") in verse 3. That would be helpful when we ask questions such as: does this shift play any specific linguistic role in the composition of the chapter, *e.g.*, as a shift in view point? Or does this have a literary-historical background and not exclusively a linguistic one?

Therefore, when analyzing biblical texts one needs to find out whether the variety of denotations used for the same participant in a text is a matter of stylistic variation, of general syntactic patterns or of

---

Functional Description of Participant Reference in Biblical Hebrew, PhD dissertation University of Stellenbosch, 2006; F.H. Polak, "Participant Tracking, Positioning and the Pragmatics of Biblical Narrative" in *Biblical Hebrew Linguistics: Papers from the 16th WCJS, Jerusalem 2013*, ed. Adina Moshavi and Tania Notarius (Winona Lake: Eisenbrauns, (in press)); L.J. de Regt, "Participant Reference Devices and the Characterisation of Personages in 1 and 2 Samuel", in *The Books of Samuel. Stories – History – Reception History*, ed. W. Dietrich, BETL 284 (Leuven: Peeters, 2016), 257–270.

[2] E. Talstra, "Text linguistics: Biblical Hebrew" in *Encyclopedia of Hebrew Language and Linguistics III*, ed. G. Khan, EHLL (New York: Brill, 2013), 755–760.

syntactic patterns applied only in particular situations, or whether it could be an indicator of redactional elaboration of a text. This implies the question: do we observe regular linguistic patterns, or do we observe special literary phenomena as part of a particular textual design, or of the development of the text?

The research presented in this paper is not intended to produce *the* answers to such questions, rather its goal is to clarify how one can identify data and data structures that will allow the researcher to have access to various patterns of linguistic phenomena, as the ones mentioned above. Having access to such data implies the ability to compare, to test and to build a case for a linguistic or a literary hypothesis.[3]

1.1. The Database

Research in participants and participant tracking cannot start from scratch. The programs being designed for this research make use of a database which allows access to linguistic data at various levels: *words* (lexemes and their features), *phrases* (the lexemes they are built up from and phrase types such as: NP, PP, VP), *clauses* (the phrases they are built from, the parsing labels of the phrases (*e.g.*, subject, object, verbal predicate) and the type of clause, *e.g.,* nominal clause, verbal clause, attributive clause), *sentences* (the clauses they are built from, *e.g.*, verbal clause + attributive clause) and *textual domains* (the sentences they are built from and features, *e.g.*, narrative, direct speech, level of textual embedding).[4]

By way of introduction I present the text of Exodus 19:2-8 in a simplified version of phrases and clauses (Hebrew text and a translation added), with verse, text line, textual domain [D] and text type [N: narrative, Q: direct speech]:

וַיִּחַן שָׁם יִשְׂרָאֵל נֶגֶד הָהָר

```
Vs 2 l: 7 D1[1] N   [W-<Cj>] [JXN <Pr>] [CM <Co>] [JFR>L <Su>] [NGD H-HR <Lo>]
                    Israel camped there in front of the mountain
```

וּמֹשֶׁה עָלָה אֶל-הָאֱלֹהִים

```
Vs 3 l: 8 D1[1] N       [W-<Cj>] [MCH <Su>] [<LH <Pr>] [>L H->LHJM <Co>]
                        while Moses ascended towards Elohim.
```

וַיִּקְרָא אֵלָיו יְהוָה מִן-הָהָר

```
      l: 9 D1[1] N       [W-<Cj>][JQR> <Pr>][>LJW <Co>] [JHWH <Su>][MN H-HR <Lo>]
                         And YHWH called to him from the mountain,
```

לֵאמֹר

```
      l:10 D1[1] N       [L->MR <Pr>]
                          by saying:
```

כֹּה תֹאמַר לְבֵית יַעֲקֹב

```
      l:11 D2[2] NQ            [KH <Mo>] [T>MR <Pr>] [L-BJT J<QB <Co>]
                               "Thus you shall say to the house of Jacob,
```

וְתַגֵּיד לִבְנֵי יִשְׂרָאֵל

---

[3] About this research see: E. Talstra, "The Bible as Data and as Literature: The Example of Exod 16" in *A Pillar of Cloud to Guide. Text-critical, Redactional, and Linguistic Perspectives on the Old Testament in Honour of Marc Vervenne,* ed. H. Ausloos & B. Lemmelijn, BETL 269 (Leuven: Peeters, 2014), 549–568.

[4] The research in participant tracking is done from the basic data stored in the Unix system of our research center. A special version made for use by linguists, Bible translators and exegetes is the "ETCBC database of the Hebrew Bible with Query Saver": http://shebanq.ancient-data.org.

l:12 D2[2] NQ   [W-<Cj>] [TGJD <Pr>] [L-BNJ JFR>L <Co>]
and announce to the children of Israel:

אַתֶּם רְאִיתֶם

Vs 4 l:13 D3[3] NQQ   [>TM <Su>] [R>JTM <Pr>]
'You yourselves have seen

אֲשֶׁר עָשִׂיתִי לְמִצְרָיִם

l:14 D3[3] NQQ   [>CR <Cj>] [<FJTJ <Pr>] [L-MYRJM <Co>]
what I have done to Egypt

וָאֶשָּׂא אֶתְכֶם עַל-כַּנְפֵי נְשָׁרִים

l:15 D3[3] NQQ   W-<Cj>][>F> <Pr>][>TKM <Ob>][<L . . . <Co>]
and that I bore you on eagles' wings.

– – – –   . . .

אִם שָׁמוֹעַ תִּשְׁמְעוּ בְּקֹלִי

Vs 5 l:18 D3[3] NQQ   [>M <Cj>][CMW< <Mo>][TCM<W <Pr>][B-QLJ<Co>]
If you listen to my voice

וּשְׁמַרְתֶּם אֶת בְּרִיתִי

l:19 D3[3] NQQ   [W-<Cj>] [CMRTM <Pr>] [>T BRJTJ <Ob>]
and keep my covenant,

וִהְיִיתֶם לִי סְגֻלָּה מִכָּל הָעַמִּים

l:20 D3[3] NQQ   [W-<Cj>] [HJJTM <Pr>] [LJ <Co>] [SGLH <Co>]
then you will be for me a possession …. '

– – – –   . . .

אֵלֶּה הַדְּבָרִים

Vs 6 l:23 D2[4] NQ   [>LH <Su>] [H-DBRJM <PC>]
These are the words

אֲשֶׁר תְּדַבֵּר אֶל-בְּנֵי יִשְׂרָאֵל

l:24 D2[4] NQ   [>CR <Re>] [TDBR <Pr>] [>L BNJ JFR>L <Co>]
that you shall speak to the children of Israel."

וַיָּבֹא מֹשֶׁה

Vs 7 l:25 D1[5] N   [W-<Cj>] [JB> <Pr>] [MCH <Su>]
And Moses went

וַיִּקְרָא לְזִקְנֵי הָעָם

l:26 D1[5] N   [W-<Cj>] [JQR> <Pr>] [L-ZQNJ H-<M <Co>]
and he summoned the elders of the people

וַיָּשֶׂם לִפְנֵיהֶם אֵת כָּל הַדְּבָרִים הָאֵלֶּה

l:27 D1[5] N   [W-<Cj>][JFM<Pr>][L-PNJHM<Co>][>T KL H-DBRJM H->LH<Ob>]
and presented to them all of these words

אֲשֶׁר צִוָּהוּ יְהוָה

l:28 D1[5] N   [>CR <Re>] [YWHW <PO>] [JHWH <Su>]
Yнwн had commanded him.

וַיַּעֲנוּ כָל-הָעָם יַחְדָּו

Vs 8 l:29 D1[5] N   [W-<Cj>] [J<NW <Pr>] [KL H-<M <Su>] [JXDW <Mo>]
All of the people answered together

וַיֹּאמְרוּ

l:30 D1[5] N   [W-<Cj>] [J>MRW <Pr>]
and they said

כֹּל

```
l:31 D4[6] NQ          [KL <Ob>]
                       "All
```

אֲשֶׁר־דִּבֶּר <mark>יְהֹוָה</mark>

```
l:32 D4[6] NQ          [>CR <Re>] [DBR <Pr>] [JHWH <Su>]
                       that YHWH has said
```

נַעֲשֶׂה

```
l:33 D4[6] NQ          [N<FH <Pr>]
                       we will do."
```

To be able to work with these data it is important to know that the linguistic analysis applies two separate categories of labeling at each linguistic level of the text data: one for the linguistic elements in their distributional, *sequential* textual order and another one for the elements in their functional, *hierarchical* textual order. For example:

A text line is called a *clause atom*. Usually a line equals the functional category of a *clause*, except for cases of embedding (and special cases like left dislocation [*casus pendens*], ellipsis and such like). A clause is a construction with maximally one predication. In case of embedding some of the lines involved present only part of a clause as is demonstrated above in the section of lines (*clause atoms*) 31, 32, 33. Line (*clause atom*) 32 is an embedded, attributive clause. As a result, line 31 has no predication, only line 33 has. Thus, here we have in the database 3 *clause atoms* (lines), making 2 *clauses* (31 + 33 = verbal clause, and 32 = verbal clause; attributive to כֹּל "all") and 1 *sentence* (31 – 33).[5]

A similar distinction is made for *textual domains*. These are segments of text marked or characterized by one set of speaker and addressee (*e.g.*, Speaker: Narrator and Audience: Reader; or Speaker: Moses and Audience: Israel). Direct speech sections, at least in narratives, usually are embedded within a larger text. In terms of textual sequence the separate sections are called *domain atoms*; in terms of textual hierarchy they are regrouped and called *domains*. Thus, in our sample text *domain* 1, narrative [N], is composed of *domain atom* 1 (line 7 – 10) and *domain atom* 5 (line 25 – 30); *domain* 2, the first embedded direct speech [NQ] consists of *domain atom* 2 (line 11 - 12) and *domain atom* 4 (line 23-24). Domain 2 has a further embedded direct speech section [NQQ]: *domain atom* 3 (line 13 - 22) that equals *domain* 3.

In this way one can account both for the distributional, sequential order of the textual elements and for their functional, hierarchical analysis. This linguistic information allows programs to identify phrases that qualify as markers of participants and to propose connections between such phrases within a particular textual domain.

1.2. Method: Searching for Patterns

The linguistic database of the Hebrew Bible presented above has been built up by distributional linguistic analysis. This implies collecting and sorting lexical data, grammatical features and patterns of their distribution, in a production process that runs from the analysis of smaller elements (morphemes used to determine the features of lexemes) to the larger and more complex ones (phrases, clauses, sentences, textual domains). Especially in the case of a corpus of ancient languages where one cannot

---

[5] The same distinction applies also to the categories of phrases and sentences. They are not discussed in this paper.

rely on the information of 'native speakers' to evaluate proposals of linguistic analysis, the most effective way is the search for patterns and pattern recognition and to treat them consistently.

It is clear that when doing this type of computer-assisted analysis for building a database that is meant to be an instrument for research in text grammar and discourse analysis, one has entered a vast area of exploration. It is the area where linguistic analysis and literary analysis meet and even have a considerable overlap. The question is: where can distributional linguistic analysis of patterns and system bring us? Where do we need to add cultural information about communicative processes such as pragmatic or rhetorical effects, or semantic knowledge and information about persons and locations in the world beyond the text? The approach chosen here is one of trial and error. We have to find out how far the search for patterns can bring us. One has to bear in mind that this project is not about discourse analysis as such, but about methods and techniques to analyze linguistic data and to design data structures that should be sufficiently consistent to allow for further linguistic and literary research.

Therefore, before elaborating further on the process of participant tracking in particular, some remarks must be made on the techniques to be applied in this research. Tracking participants is in fact a part of a larger area of text-level research. It may be helpful to present first some examples of the search for patterns in related projects, *i.e.*, the search for patterns of clause relations and for patterns of verbal valence. Examples are taken from the text of Exodus 19: 3 – 7.[6]

1.2.1. Text Grammar: Patterns and Functions

Once we have completed *distributional* linguistic analysis to construct a hierarchical syntactic presentation of a text, the next matter of interest is whether this textual hierarchy will provide sufficient data to establish patterns of clause connections that allow us to make proposals for the *functional* labelling of verbal clauses in their context.

For example, in cases where a verbal form as such is not marked for modality, can one establish patterns of clause connections to be used as a syntactic argument for a modal interpretation of such verbal forms?

כֹּה **תֹאמַר** לְבֵית יַעֲקֹב

Vs 3 l:**11** D2[2] NQ   "Thus you *shall* say to the house of Jacob,

וְתַגֵּיד לִבְנֵי יִשְׂרָאֵל

l:**12** D2[2] NQ   and announce to the children of Israel:

[line 13 – 22 = *embedded direct speech [NQQ] = domain 3*]

אֵלֶּה הַדְּבָרִים

Vs 6 l:**23** D2[4] NQ   These are the words

אֲשֶׁר **תְּדַבֵּר** אֶל-בְּנֵי יִשְׂרָאֵל

l:**24** D2[4] NQ   that you *shall* speak to the children of Israel."

---

[6] E. Talstra, "Data, Knowledge and Tradition. Biblical Scholarship and the Humanities 2.0. Exodus 19 as a Laboratory Text" in *The Present State of OT Studies in the Low Countries; A Collection of Old Testament Studies Published on the Occasion of the Seventy-fifth Anniversary of the Oudtestamentisch Werkgezelschap*, ed. Klaas Spronk, OTS xx (Leiden: Brill, 2016), 228–247.

In the case of the 2nd person singular תֹאמַר in line 11 one can argue that the next line, a parallel constructed *weyiqtol* clause, provides us with the answer: line 11 and 12 are an instruction: "you shall …". But what about תְּדַבֵּר in line 24? In such cases patterns of clause level grammar and text level grammar can be found to cooperate. At clause level one can argue that patterns of the type "these are the words/commandments" followed by an attributive clause with a *yiqtol* in second person require a modal interpretation: "that you shall …". See, for example, Exodus 21:1 וְאֵלֶּה הַמִּשְׁפָּטִים אֲשֶׁר תָּשִׂים לִפְנֵיהֶם ("these are the ordinances that you shall present to them"). At text level one can find larger patterns that corroborate this. Line 11 and 12 appear to be continued by line 23 and 24. In the construction of the database, syntactic and lexical features (verbs of speaking and the 2nd person singular verbs) are sufficient markers to conclude that after the section with "I" and you-plural (line 13 – 22) the text from line 23 reconnects with line 11 and 12 (you singular), and thus belongs to *domain* 2. As a result, line 13 – 22 have been analyzed as an embedded direct speech section, *i.e.*, *domain* 3, and the modality of תְּדַבֵּר in line 24 as a continuation of the modality of line 11 and 12. See also כֹּה תֹאמַר in Exodus 20:22 where after an embedded section starting with you-plural the discourse is being continued by וְאֵלֶּה הַמִּשְׁפָּטִים אֲשֶׁר תָּשִׂים in 21:1 (mentioned above).

Another case is the order of particular clause types. In verse 2 and 3 we find three clauses in a row, each of which changes both the clause type and the explicit subject.

19:2 *wayyiqtol* [subj: people] – 19:3 *w-X-qatal* [subj. Moses] – *wayyiqtol* [subj YHWH]. Translations sometimes suggest that there are no changes in the text, *e.g.*, Revised Standard Version: "Israel encamped … And Moses went up … and the Lord called …". Only the capital "A" indicates the beginning of verse 3. The same is done with the last line of verse 17: *wayyiqtol* [subj: people] and the first line of verse 18: *w-X-qatal* [subj. the mount Sinai]. It implies that in such translations the verse division is dominating syntax. Interestingly, the New Revised Standard Version has in these cases: "Then Moses went up and Now Mount Sinai was wrapped up in smoke", which reveals that these *w-X-qatal* clauses with their introduction of new subjects have to be related to the preceding *wayyiqtol* clause. This means that we can propose a pattern of the type *wayyiqtol* [subj: X] – *w-X-qatal* [subj. Y] expressing simultaneous actions: X happened, while Y happened. This is to be tested in other textual hierarchies in the corpus.

Comparable is verse 5 (line 18-20), where an אִם *yiqtol* is continued by two *weqatal* clauses with the same subject: "if you obey … and keep … you will be ...". The final *weqatal* does not further elaborate the "if", but it expresses the result: "then …".

This also is a pattern to be tested in other textual hierarchies in researching the syntactic function of clauses and clause connections.[7]

By this type of pattern recognition one will be able to add new features to the database both at clause level syntax, *e.g.*, a *clause feature* such as "modality" and at text level syntax, *e.g.*, a *clause connection feature* such as "coordination" (line 11 and 12), and "result" (line 20, concluding line 18 and 19), or even a *paragraph feature* such as "simultaneous" (in the transition from line 7 to 8).

---

[7] This research has been done in the doctoral dissertation of G.J. Kalkman, *Verbal Forms in Biblical Hebrew Poetry: Poetic Freedom or Linguistic System?* (Amsterdam, 2015). To be consulted via the VU-repository: http://hdl.handle.net/1871/53295. Select: complete_dissertation.pdf.

1.2.2. *Patterns of Verbal Valence*

Another important area of experimenting with pattern recognition and the assignment of linguistic functions to particular categories of data distribution is the analysis of verbal valence.[8] The patterns to be identified, collected and interpreted in this research concern the usage of verbs in combination with their syntactically obligatory satellites: subject, object, indirect object and other complements. These patterns demonstrate that a verb has a particular meaning only in combination with particular sets of satellites with which it is connected in the texts. An example is the usage of the verb קָרָא in Exodus 19, 3, 7, 20, in combination with a prepositional phrase with אֶל or with לְ.

וַיִּקְרָא אֵלָיו יְהוָה מִן-הָהָר

Vs 3 l: 9 D1[1] N   And YHWH called to him from the mountain,

וַיִּקְרָא לְזִקְנֵי הָעָם

Vs 7 l:26 D1[5] N   and he called upon the elders of the people.

The combination of אֶל + קָרָא in verse 3 has been translated by: *called to him* both in the RSV and the NRSV.

The combination of לְ + קָרָא in verse 7 and 20 has been translated by: *he called + object* in the RSV. In NRSV the translation is: *he summoned + object*.

This rendering in NRSV demonstrates that when being aware of verbal valence a translator will be both more precise and more 'free' in proposing a translation.

Research into patterns of verbal valence identifies and systematically organizes such pattern into a new data set, to be added to the data base. This will assist translators and linguists to find a more consistent interface between the idiom of the source language and that of the target language. An example is the Hebrew idiom of the verb שִׂים + לִפְנֵי + pronominal suffix, referring to humans in verse 7. The literal "placed before them" can be rendered more idiomatically as:

וַיָּשֶׂם לִפְנֵיהֶם אֵת כָּל הַדְּבָרִים הָאֵלֶּה

Vs 7 l:27 D1[5] N   and (Moses) presented to them all of these words

RSV translates the expression in verse 7 more literally: "and set before them". NRSV also has: "and set before them". The comparable text of Exodus 21:1 וְאֵלֶּה הַמִּשְׁפָּטִים אֲשֶׁר תָּשִׂים לִפְנֵיהֶם, already mentioned above, speaks about regulations to be proclaimed to Israel. Here both the RSV and the NRSV also translate: "you shall set before them". Is that a valid rendering in English? Can one set instructions or words before someone, or is that merely an imitation of the Hebrew expression? One can set food before someone, but does one also set words or regulations before someone?[9] In such cases one rather would need other verbs, such as "present", "proclaim" of "prescribe".

---

[8] Much of this work is being done by Janet Dyk and a small team of colleagues. See: J.W. Dyk, "Traces of Valence Shift in Classical Hebrew" in *Discourse, Dialogue and Debate in the Bible. Essays in Honour of Frank H. Polak*, ed. Athalya Brenner-Idan (Phoenix Press: Sheffield, 2014), 48–65; J.W. Dyk et al., "Analysing Valence Patterns in Biblical Hebrew: Theoretical Questions and Analytical Frameworks", JNSL 40/1 (2014), 43–62.

[9] Examples from a lexicon on the internet (idioms.thefreedictionary.com): "I set the plate of sandwiches before the children and they were gone in a few minutes. Jane set the bowl of food before the cats."

In the Hebrew Bible one can find a number of texts where indeed food is set before someone. Genesis 24:33 שִׂים לֶחֶם וָמַיִם וייֹשׂם (וַיּוּשַׂם) לְפָנָיו לֶאֱכֹל NRSV: "then food was set before him". 2 Kings 6:22 לִפְנֵיהֶם NRSV: "set food and water before them".

In such cases the valence patterns in Hebrew and in English match. Interestingly, NRSV also translates with "set" in case the Hebrew text speaks about setting food in front of someone, but uses the verb נָתַן. Genesis 18:8 וַיִּקַּח … וַיִּתֵּן לִפְנֵיהֶם. NRSV: "he took … and set it before them". That is correct, but one wonders why the same strategy of translating in accordance with the difference of valence patterns in Hebrew and English has not been properly noted in the texts with the verb שִׂים mentioned above.

These examples may help demonstrate why the construction of a valence lexicon will be a useful contribution to the work of linguists and translators. To produce it, the database is used first to take stock of the patterns of verbs and their satellites. An inventory of patterns of clauses with a specific verb as its predicate, used with or without a subject, objects (phrases with or without a preposition), other complements, or adjuncts (non-obligatory elements referring to location, time, etc.) forms the basis of further analysis. Where relevant, the lexical characteristics of the objects referred to by the verbal satellites need to be noted. From the examples given it appears to be important to know whether subjects and objects have features such as: animate – non-animate, divine – human; human – non-human, or concrete – abstract.

After sorting the patterns and the features found, one can take the step from patterns to functions. Functions and translations proposed for a valence lexicon then have to be tested in concrete cases in the texts.

Here one can observe how the various lines of research have to interact. For example, if one wants to establish the lexical characteristics of a particular object while that object is present in the text only as a pronominal suffix, one needs participant tracking to find out whether the suffix refers to a person, an abstract entity, or something else. For example, in verse 4 the clause וָאָבִא אֶתְכֶם אֵלָי ("And I will bring you to me") only has markers of a first person singular and a second person plural. Once participant tracking has identified them as 'YHWH' and 'Israel', the (automated) analysis of the valence pattern used here can be completed.

For the moment one may conclude, first, that pattern recognition as a technique to expand the database with more 'functional' features is effective. And, second, that this type of research is best approached from different angles at the same time: text grammar, valence research and participant tracking.

2. Participant Tracking: Questions and Procedures

2.1. Questions

If one tries to identify 'who is who' in a text one rapidly is confronted with a great variation of linguistic phenomena in a literary composition. How does one identify the textual elements that qualify as possible referents to text-level participants? The variety of phenomena found make it imperative to establish a proper order, a particular hierarchy of the analyses to be made in order to distinguish between what is linguistic pattern and what is literary knowledge or cultural information.

To give some introduction into the area where the programs being designed have to operate, I present first a small inventory of observations of such phenomena in the text of Exodus 19 and the questions they provoke.

1. Clear cases of reference to participants are words or phrases with markers of person, number and gender, such as verbal forms, personal pronouns, pronominal suffixes. These cases, however, only cover a part of the material.

2. Cases of compound phrases or cases of a lack of congruency require additional analysis, since usually they contribute more than one element to the list of participant references. For example, in verse 7: לְזִקְנֵי הָעָם ("to the elders of the people") is in plural. It has a sub-phrase הָעָם ("the people") in singular, which is continued by כָל-הָעָם ("all of the people") in verse 8, where it, however, is constructed with a plural verb form: וַיַּעֲנוּ ("they answered").

3. Crossing over from narrative into direct speech domain (N > Q), one finds a transition of patterns of linguistic marking, *e.g.*, from וַיֹּאמְרוּ ("they said:") to נַעֲשֶׂה ("we will do") in verse 8. Are there syntactic patterns to identify the 3rd person plural in the N domain with the 1st person plural in the Q domain? What if the speaker in the narrative section is in singular, but speaks of "we" (*e.g.* verse 23)?

4. What kind of knowledge do we use when we find that, at least in Exodus 19, the phrases יִשְׂרָאֵל ("Israel"), בְּנֵי יִשְׂרָאֵל ("children of Israel") and בֵּית יַעֲקֹב ("the house of Jacob") all seem to refer to the same participant, and at what stage of the participant analysis do we apply that information?

5. How can one calculate and present the fact that all individual domains actually are text segments that 'belong to' particular participants that act as speaker and addressee? How could we compose an overview of the participants and their domains of communication as presented below? The example of domain 3 (mentioned in the introductory section of this paper), where Moses speaks by only quoting the words of God, illustrates the question of the interaction of linguistic patterns and literary design. So, an important objective is to calculate an overview of textual domains and their 'owners' as presented below.

```
L   7        Domain 1 Domain_atom 1 N:            Narrative
             Narrator - Reader;

L  11        Domain 2 Domain_atom 2 NQ:           direct speech (within N: Narrative)
             Narrator - Reader; YHWH - Moses;

L  13        Domain 3 Domain_atom 3 NQQ:          direct speech (within Q; within N: Narrative)
             Narrator - Reader; YHWH - Moses; Moses (=YHWH) - people

L  23        Domain 2 Domain_atom 4 NQ:           direct speech (within N: Narrative)
             Narrator - Reader; YHWH - Moses;

L  25        Domain 1 Domain_atom 5 N:            Narrative
             Narrator - Reader

L  31        Domain 4 Domain_atom 6 NQ:           direct speech (within N: Narrative)
             Narrator - Reader; people – Moses (?)
```

2.2. Procedures

The variety of textual phenomena mentioned above demonstrate that an attempt to construct programs for participant tracking is a process of both text-grammatical research (analyzing concrete texts) and system analysis (constructing a system of data types to store and test the analytical data found) to be performed at the same time. When analyzing (ancient) texts, we operate in the area where linguistic system and literary design overlap, where calculation and interpretation confront one another. Not much is known in advance about the logical design of text grammar, rather this work is a continuing process of trying out and testing one's ideas about what could work. How should one, then, proceed with participant analysis?

The analysis starts from a particular text with its syntactic hierarchy as present in the existing database (see section 1.1). In our case that is Exodus 19. To be precise: this textual hierarchy, having being produced before by distributional analysis of linguistic patterns, actually represents just an initial hypothesis about the delimitation and combination of clauses, sentences and domains. This no doubt will need refinement and correction as the participant analysis proceeds. In a way all analysis is a preliminary proposal to organize lower level data as consistently as possible into higher level categories. When one detects problems, either the existing analysis of lower level data needs correction, or the proposal for an analysis of higher level data needs revision. After a period of trial and error, a particular order of analytical procedures emerges.

First, I give a summary of the steps to be taken, which will be elaborated in the next sections (3.1-4).

1. Identify participant referent candidates [PRef] by selecting all lexemes, phrases or sub-phrases with *features* that mark person, number and/or gender.

2. Test by what mechanisms they refer back to one another, *e.g.*, matching by means of morphological patterns, syntactical patterns (such as subject – predicate), or lexical patterns (such as identical lexemes) (3.1.)?

3. Gather the various cases of back reference between PRef-candidates within each individual text domain and combine them into *sets* of participant referents [PSet]. Add a label, *e.g.*, 'Yʜᴡʜ' to occurrences of יְהוָה in the narrative domain and '>NKJ' to occurrences of the sender אָנֹכִי in domain 5, verse 9.

4 When crossing a domain border, identify the participant references and their sets that function as *Sender* and *Addressee* in a direct speech text segment (1ˢᵗ and 2ⁿᵈ person) and those that are used as *Speaker* and *Audience* (subject of the verb of speaking and the complement phrase connected to that verb) in the preceding and higher level text introducing the direct speech section (3.2.).

5. Try to connect the remaining PSets to PSets in the previous domains. This regards the sets that do not belong to the mechanism of speaker and addressee. Usually, they are PSets in 3ʳᵈ person with lexical identity beyond the domain borders. Some may have been found already with step 2 (see 3.1.), others will be found now. For example, יְהוָה in the direct speech section spoken by the people in verse 8 is now connected with the occurrence of יְהוָה in the narrative layer in verse 7.

6. Combine the connected PSets identified in the individual domains into text level *Actors* [PAct]. Here both the mechanisms of lexical identity and of *Speaker = Sender* and *Audience = Addressee* apply. As a result, one can now track these *Actors* through the text, independently of their appearance in 1ˢᵗ, 2ⁿᵈ or 3ʳᵈ person (3.3.).

7. Combine those PActs that have different labels, *e.g.*, 'JHWH' (ꜱʏʜꜱʜ) and '>LHJM' (אֱלֹהִים), but refer to the same referent in the text and assign to the combination a more abstract label, *e.g.*, 'God'. This step is taken in an interactive process.

8. Create clusters of PActs that appear to be interrelated in other ways, *e.g.*, by referring to a 'part of' another *Actor*, *e.g.*, 'mountain' הָהָר (verse 2) and 'top of the mountain' רֹאשׁ הָהָר (verse 20) (3.4.).

3. Participant Tracking in a Number of Steps

3.1. From Phrases to Participant referents (PRef) [steps 1-2]

In this first stage of the research two questions are asked:

1. How are PRefs to be identified? 2. How are they referring back to a previous PRef?

3.1.1. How are PRefs to be identified? [step 1]

Clear cases of PRefs are words or phrases with markers of person, number and gender, such as verbal forms, personal pronouns, pronominal suffixes. This also applies to elements with markers of number and gender only, *i.e.*, demonstrative pronouns, nouns and noun phrases. Such features may be derived from morphology, or, in the case of pronouns or proper names, from the lexicon. Proper names qualify by definition as PRefs.

Cases of compound phrases require additional analysis since one has to determine whether the various parts of a phrase (sub-phrases) could qualify as PRef, too. For example, the PP: לְזִקְנֵי הָעָם ("to the elders of the people") in verse 7 is in the plural. Since it has an internal genitive construction, there is a sub-phrase הָעָם ("the people") in the singular that also qualifies as a PRef. Also, the phrase כָל-הָעָם ("all of the people") in verse 8 is singular though it is constructed with וַיַּעֲנוּ ("they answered"), a plural verb. What should be done with the phrase אֶת כָּל הַדְּבָרִים הָאֵלֶּה ("all of these words") in verse 7? The head of the phrase, כָּל, is singular, but based on הַדְּבָרִים the phrase is plural; the latter is confirmed later when we see the back reference to אֵלֶּה הַדְּבָרִים ("these words") in verse 6. So, should we skip כָּל as head of a phrase in all cases? At the moment that is the policy adopted, except for cases where כָּל is used as an independent NP by itself.

3.1.2. How do PRefs refer back to a previous PRef? [step 2]

Test by what mechanisms PRefs refer back to one another. Do they match by morphological patterns (suffix matches suffix; suffix matches NP), by syntactic patterns (such as subject matches verbal predicate), or by lexical patterns (such as identical lexemes)? These categories of back reference, together with the PRef referred to, are stored with each PRef where they apply. This search turns out to be effective primarily within the same textual domain, because there are no disturbing shifts of person caused by a change of sender or addressee. Cases of *lexical* reference, however, will be found to easily transcend domain borders, but at this stage of the research such is accepted only in cases of 3rd person referents, for example: כָּל הַדְּבָרִים הָאֵלֶּה in verse 7, N-domain 1, matches with אֵלֶּה הַדְּבָרִים verse 6, Q-domain 2.

The results are given in a simplified schema below of line 25 – 33 (verse 7 – 8), with PRef 70 – 88, with references and categories of reference.

PRef	lex numbers	PersNumGen	Lexemes	Ref.;Ref.categ	(Hebrew text added)

```
domain:  1 = domain atom:  5
EXO 19,07  Line: 25
```

70	2- 2	PNG: 312	BW>[		וַיָּבֹא
71	3- 3	PNG:-112	MCH=/	70 S = P	מֹשֶׁה

```
EXO 19,07  Line: 26
```

72	2- 2	PNG: 312	QR>[	70 P = P	וַיִּקְרָא
73	3- 6	PNG:-132	L .ZQN/ .H .<M/		לְזִקְנֵי הָעָם
74	5- 6	PNG:-112	H .<M/		הָעָם

```
EXO 19,07  Line: 27
```

75	2- 2	PNG: 312	FJM[	72 P = P	וַיָּשֶׂם
76	3- 4	PNG: 000	L .PNH/+S		לִפְנֵיהֶם
77	4- 4	PNG: 332	sfx	73 sfx = NP	הֶם
78	7-10	PNG:-132	H .DBR/ .H .>LH		אֵת כָּל הַדְּבָרִים הָאֵלֶּה

```
EXO 19,07  Line: 28
```

79	2- 2	PNG: 312	YWH[		צִוָּהוּ
80	2- 2	PNG: 312	sfx	75 sfx = P	הוּ-
81	3- 3	PNG:-112	JHWH/	79 S = P	יְהוָה

```
EXO 19,08  Line: 29
```

| 82 | 2- 2 | PNG: 332 | <NH[| | וַיַּעֲנוּ |
| 83 | 4- 5 | PNG:-112 | H .<M/ | 82 S = P; 74= Lex | כָל-הָעָם |

```
EXO 19,08  Line: 30
```

| 84 | 2- 2 | PNG: 332 | >MR[| 82 P = P | וַיֹּאמְרוּ |

```
domain:  4 = domain atom:  6
EXO 19,08  Line: 31
```

| 85 | 1- 1 | PNG:-112 | KL/ | | כֹּל |

```
EXO 19,08  Line: 32
```

| 86 | 2- 2 | PNG: 312 | DBR[| | דִּבֶּר |
| 87 | 3- 3 | PNG:-112 | JHWH/ | 86 S = P; 81 = Lex | יְהוָה |

```
EXO 19,08  Line: 33
```

| 88 | 1- 1 | PNG: 130 | <FH[| | נַעֲשֶׂה |

3.2. From Participant Referents to Sets of Referents: PSets [steps 3-4]

3.2.1. Construct Sets of PRefs within each domain [step 3]

The various cases of back reference between PRef-candidates, established within each separate text domain, are gathered and combined into *sets* of participant referents [PSet]. In principle, only PRefs that refer back to a previous one are taken as part of a PSet. However, PRefs in 1st or 2nd person are always accepted as PSets, even if they have no back reference within the same domain. In most cases they will later be found to have a back reference beyond the domain boundary. This will be explained further below.

PSets are calculated now from the PRef connections constructed by step 2:

- PRefs 70, 71, 72, 75, 80 directly or indirectly refer to מֹשֶׁה (MCH).

- PRefs 73, 77 refer to זִקְנֵי הָעָם (ZQN <M)

- PRefs 74, 82, 83 to עָם (<M).

- In N-domain 1, PRefs 79, 81 refer to יְהוָה (JHWH).

- In Q-domain 4, PRefs 86 and 87 also refer to יְהוָה (JHWH). Identification of these PSets across the domain border will be made at the next stage.

- In Q-domain 4, PRefs 88 "we" will qualify as a PSet by itself. Connections of this PSet across the domain border will also be made at the next stage.

The sets of participant referents are numbered following the process of collecting the cases of back references when reading the text. As a result, the cases of 'MCH' (מֹשֶׁה) in the survey above in line 25 received the PSet number 5, since cases of 'MCH' already had been detected rather early in the N-domain of Exodus 19: from line 8 (verse 3). The cases of 'ZQN <M' (זִקְנֵי הָעָם), however, got the PSet number 17, since this only came up much later, in line 26. Sets, therefore, do not get a numbering at domain level, but at text level, in the linear order in which they have been detected in the text.

After having produced the various PSets, the program produces two labels for each PSet. *Label 1* contains the indications of person, number and gender valid for the PSet. The program tries to find out what indicators of person number and gender can be derived from all members of the set, *e.g.*, in the PSet 'MCH' the verb וַיָּבֹא (PRef 70: he went) has 3 sing masc., 'MCH' (מֹשֶׁה) (PRef 71) belongs to the same Set, so 3 sing. masc. is also valid for 'Moses'. In *Label 2* a name for the PSet is stored. In general, that label is not derived from the verb, but from the lexeme(s) of the first explicit subject phrase encountered: 'MCH' (מֹשֶׁה). As a result, the program can now add labels to each member of this PSet: `PSet 5:3sm=MCH`. In the same way, '3 sing masc.' and 'JHWH' are labels assigned to the cases of reference to יְהוָה in the narrative domain 1: `PSet 6:3sm=JHWH`. After that, the labels `PSet 20:3sm=JHWH` are assigned to occurrences of יְהוָה in domain 4. Later, the labels `PSet 24:1sm=>NKJ` are assigned to occurrences of the phrase אָנֹכִי in domain 5, verse 9. At the next level of analysis [step 5], we will find that `PSet 6` יְהוָה = `PSet 20` יְהוָה = `PSet 24` אָנֹכִי.

3.2.2. Connect Sets of PRefs across Domain Borders [step 4]
When in the process of textual analysis one crosses a domain border between N (narrative) and Q (direct speech), new patterns of linguistic connections appear to be active. In cases of 3rd person some PSets, present both in the N and the Q domain, can be connected by searching for lexical identity, such as יְהוָה in N-domain 1 (verse 7) and יְהוָה in Q-domain 4 (verse 8). To identify יְהוָה in N-domain 1 (verse 7 and 9) with the phrase אָנֹכִי in Q-domain 5 (verse 9) we need to apply other patterns of connections. First, we have to identify participant references and their PSets that are found to function as *Sender* or *Addressee* in a direct speech text segment (marked by 1st and 2nd person features) and, next, identify the PSets that are used as *Speaker* or *Audience* (subject of a verb of speaking and the ל-complement phrase connected to that verb) in the preceding and higher level text segment that introduces the direct speech section. The general pattern is:

```
Speaker [N] = Sender [Q] and Audience [N] = Addressee [Q]
```

For example:

> Verse 9: "YHWH [Sp] said to Moses [Aud]: 'I [Sen] am coming to you [Adr] …' "

> Verse 8: "they [Sp] said" = "we [Sen] will do."

Such combinations of PSets are registered and stored, *e.g.*, at the border between domain 4 and 1 (verse 8):

> Speaker `PSet19:3pm=<M` (domain 1) = Sender `PSet21:1pc` (domain 4)

At the border between domain 5 and 1 (verse 9):

> Speaker `PSet23:3sm=JHWH` (domain 1) = Sender `PSet24:1sm=>NKJ` (domain 5)
> Audience: `PSet5:3sm=MCH` (domain 1) = Addressee `PSet25:2sm=…` (domain 5)

As can be seen in the example of verse 8, frequently one encounters gaps in the texts. In verse 8 we have no explicit *audience* in the N-domain and no explicit *addressee* in the Q-domain. Further research is needed at the next stage: can we find information in the domains or in general patterns of communication to fill those gaps?

The results of step 3 and step 4 are listed in a simplified survey below:

```
domain:  1 = domain atom:  5

L:  25                                                          וַיָּבֹא מֹשֶׁה
  1N0   EXO 19,07 [W-<Cj>] [JB> <Pr>] [MCH <Su>]
 PRef: 70 JB>                  PSet:  5= 3sm=MCH    מֹשֶׁה
 PRef: 71 MCH                  PSet:  5= 3sm=MCH    מֹשֶׁה


L:  26                                                     וַיִּקְרָא לְזִקְנֵי הָעָם
  1N0   EXO 19,07 [W-<Cj>] [JQR> <Pr>] [L-ZQNJ H-<M <Co>]
 PRef: 72 JQR>                 PSet:  5= 3sm=MCH    מֹשֶׁה
 PRef: 73 L-ZQNJ H-<M          PSet: 17= 3pm=ZQN <M   זקן עם
 PRef: 74 H-<M                 PSet: 19= 3pm=<M       עָם
```

```
L:  27                                      וַיָּשֶׂם לִפְנֵיהֶם אֵת כָּל הַדְּבָרִים הָאֵלֶּה
  1N0   EXO 19,07 [W-<Cj] [JFM <Pr>] [L-PNJHM <Co>] [>T KL H-DBRJM H->LH <Ob>]
 PRef: 75 JFM                  PSet:  5= 3sm=MCH    מֹשֶׁה
 PRef: 76 L-PNJHM              no PSet
 PRef: 77 PNJHM:sfx            PSet: 17= 3pm=ZQN <M   זקן עם
 PRef: 78 H-DBRJM H->LH        PSet: 18= 0pm=DBR      (cf.vs.6; 3 pers. lexical identity)


L:  28                                                      אֲשֶׁר צִוָּהוּ יְהֹוָה
  1N0   EXO 19,07 [>CR <Re>] [YWHW <PO>] [JHWH <Su>]
 PRef: 79 YWHW                 PSet:  6= 3sm=JHWH   יְהֹוָה
 PRef: 80 YWHW:sfx             PSet:  5= 3sm=MCH    מֹשֶׁה
 PRef: 81 JHWH                 PSet:  6= 3sm=JHWH   יְהֹוָה
```

```
L:  29                                                    וַיַּעֲנוּ כָל-הָעָם יַחְדָּו
  1N0  EXO 19,08 [W-<Cj>] [J<NW <Pr>] [KL H-<M <Su>] [JXDW <Mo>]
 PRef: 82 J<NW            PSet: 19= 3pm=<M        עָם
 PRef: 83 H-<M            PSet: 19= 3pm=<M        עָם


L:  30                                                              וַיֹּאמְרוּ
  1N0  EXO 19,08 [W-<Cj>] [J>MRW <Pr>]
 PRef: 84 J>MRW           PSet: 19= 3pm=<M        עָם


domain:  4 = domain atom:  6
L:  31                                                                  כֹּל
  4Q1  EXO 19,08 [KL <Ob>]
 PRef: 85 KL                      no PSet


L:  32                                                     אֲשֶׁר-דִּבֶּר יְהֹוָה
  4Q1  EXO 19,08 [>CR <Re>] [DBR <Pr>] [JHWH <Su>]
 PRef: 86 DBR             PSet: 20= 3sm=JHWH    יְהֹוָה
 PRef: 87 JHWH            PSet: 20= 3sm=JHWH    יְהֹוָה


L:  33                                                                נַעֲשֶׂה
  4Q1  EXO 19,08 [N<FH <Pr>]
 PRef: 88 N<FH            PSet: 21= 1pc=          "we"
```

Sender PSet: 21= 1pc (domain 4) = Speaker PSet: 19= 3pm=<M (domain 1)

3.3. From Sets of Participant Referents to Actors: PActs [steps 5-6]

3.3.1. Connecting PSets not being Part of Patterns of Communication [step 5]
Once the patterns of Speaker = Sender and Audience = Addressee are applied to identify PSets from various domains, we try to connect the remaining PSets to PSets in previous domains. Usually they are PSets with lexical identity beyond domain borders, in 3rd person. Some may have been found already in step 2 (see 3.1.), others will be found now. For example, יְהֹוָה in the direct speech section spoken by the people in verse 8 is now connected with the occurrence of יְהֹוָה in the narrative layer in verse 7.

3.3.2. Identification of Actors [step 6]
Once we have found which PSets from various domains can be identified with each other, we can construct new sets of participant markers active at a higher linguistic level. The term used here is *Actors* (PAct). They represent syntactically calculated categories of participant references at text level. Here both mechanisms of identification beyond domain borders apply: lexical identity, and Speaker = Sender and Audience = Addressee.

In this process, we will come across a number of phenomena that require additional research, since it is difficult to determine whether we still will be able to find syntactic patterns, or whether we need to add semantic or literary knowledge. Such situations always reveal the iterativeness of this type of

research: after using syntactic patterns for building a database as an instrument for text-grammatical research, at a further stage we need the same database to discover more about how the various types of linguistic and literary knowledge interact in a textual composition.

- In a number of cases one will find changes in number. For example in verse 7 and 8: עָם as part of the phrase לְזִקְנֵי הָעָם is grammatically singular. However, in the next line it is constructed with a plural verb, while in the direct speech section that follows it is being referred to as "we", also a plural (see also the plural pronominal suffix referring to עָם in verse 10 and the construction in singular again in verse 16 and 23). As a result of the existing mechanism of labeling PSets and Pacts, the *actor* עָם is labeled as plural, whereas in the lower levels of the database the feature 'singular' will remain. In this way the data will allow for further study of such shifts in number.

- Another case needing further analysis is the difference between inclusive and non-inclusive "we" in verse 23 and verse 8. The difference can be found in an interactive process and then be stored, but can it also be calculated automatically?

One also finds cases of inconsistency, or at least, changes of patterns, occurring with the transition from narrative into direct speech, *e.g*.,

- Domains where, contrary to the usual patterns, the Sender in a direct speech domain is not the Speaker of the directly preceding domain, but is the Speaker introduced in a higher level domain: Moses is speaking, but he is only quoting Yʜᴡʜ (verse 4).

- Domains where the Sender equals the Speaker, however the Sender does not use the reference "I", but re-introduces the Speaker in 3ʳᵈ person (יְהוָה in verse 11 and 21).

For the moment, such phenomena can be stored in the data structures designed, but it remains to be seen how far one can go in detecting linguistic mechanisms to identify them.

Based on the data produced thus far, we can now expand the database in two ways. First, assign a separate PAct number to each participant referent in the text. Second, apply the PAct numbers also to the PSets that have been identified as Speaker, Audience, Sender or Addressee. In this way, we know who are the actors that are communicating in each individual domain. For example: is this a text domain of `Narrator` and `Reader`? Or is it a text domain of communication by יְהוָה and מֹשֶׁה?

The results of step 5 and step 6 are listed in the survey below:

```
domain:  1 = domain atom:  5
L:  25                                                                          וַיָּבֹא מֹשֶׁה
  1N0   EXO 19,07 [W-<Cj>] [JB> <Pr>] [MCH <Su>]
PRef: 70 JB>              PSet:  5= 3sm=MCH        PAct:  5= MCH                      מֹשֶׁה
PRef: 71 MCH             PSet:  5= 3sm=MCH        PAct:  5= MCH                      מֹשֶׁה


L:  26                                                                   וַיִּקְרָא לְזִקְנֵי הָעָם
  1N0   EXO 19,07 [W-<Cj>] [JQR> <Pr>] [L-ZQNJ H-<M <Co>]
PRef: 72 JQR>            PSet:  5= 3sm=MCH        PAct:  5= MCH                      מֹשֶׁה
PRef: 73 L-ZQNJ H-<M    PSet: 17= 3pm=ZQN <M     PAct: 11= ZQN <M             זקן עם
PRef: 74 H-<M           PSet: 19= 3pm=<M         PAct: 12= <M                        עָם
```

```
L:  27                                          וַיָּשֶׂם לִפְנֵיהֶם אֵת כָּל הַדְּבָרִים הָאֵלֶּה
 1N0  EXO 19,07 [W-<Cj] [JFM <Pr>] [L-PNJHM <Co>] [>T KL H-DBRJM H->LH <Ob>]
PRef: 75 JFM             PSet:  5= 3sm=MCH      PAct:  5= MCH         מֹשֶׁה
PRef: 76 L-PNJHM         no PSet
PRef: 77 PNJHM:sfx       PSet: 17= 3pm=ZQN <M   PAct: 11= ZQN <M      זקן עם
PRef: 78 H-DBRJM H->LH   PSet: 18= 0pm=DBR      PAct: 10= DBR         דבר


L:  28                                                     אֲשֶׁר צִוָּהוּ יְהוָה
 1N0  EXO 19,07 [>CR <Re>] [YWHW <PO>] [JHWH <Su>]
PRef: 79 YWHW            PSet:  6= 3sm=JHWH      PAct:  6= JHWH        יְהוָה
PRef: 80 YWHW:sfx        PSet:  5= 3sm=MCH       PAct:  5= MCH        מֹשֶׁה
PRef: 81 JHWH            PSet:  6= 3sm=JHWH      PAct:  6= JHWH        יְהוָה


L:  29                                                  וַיַּעֲנוּ כָל-הָעָם יַחְדָּו
 1N0  EXO 19,08 [W-<Cj>] [J<NW <Pr>] [KL H-<M <Su>] [JXDW <Mo>]
PRef: 82 J<NW            PSet: 19= 3pm=<M        PAct: 12= <M          עָם
PRef: 83 H-<M            PSet: 19= 3pm=<M        PAct: 12= <M          עָם


L:  30                                                              וַיֹּאמְרוּ
 1N0  EXO 19,08 [W-<Cj>] [J>MRW <Pr>]
PRef: 84 J>MRW           PSet: 19= 3pm=<M        PAct: 12= <M          עָם


domain:  4 = domain atom:  6
 --------------------------
Speaker :   84        = Set: 19 3pm= עָם      = Act: 12 עָם
Audience:    0        = Set:  5 3sm= מֹשֶׁה    = Act:  5 מֹשֶׁה
SenderPl:   88        = Set: 21 1pc= "we"     = Act: 12 עָם
Addressee:  unknown
L:  31                                                                   כֹּל
 4Q1  EXO 19,08 [KL <Ob>]
PRef: 85 KL                 no PSet

L:  32                                                        אֲשַׁר-דִּבֶּר יְהוָה
 4Q1  EXO 19,08 [>CR <Re>] [DBR <Pr>] [JHWH <Su>]
PRef: 86 DBR            PSet: 20= 3sm=JHWH      PAct:  6= JHWH         יְהוָה
PRef: 87 JHWH           PSet: 20= 3sm=JHWH      PAct:  6= JHWH         יְהוָה


L:  33                                                                  נַעֲשֶׂה

 4Q1  EXO 19,08 [N<FH <Pr>]
PRef: 88 N<FH          PSet: 21= 1pc="we"      PAct: 12= <M           עָם
```

As a result, one can now track these *actors* through the text, independent from their appearance in 1st, 2nd or 3rd person.

It also implies that one can produce a kind of concordance of *actors*, including the occurrences where they have not explicitly been marked in the surface text, such as in verbs without a NP as explicit

subject. For example, some lines with actor 5: `MCH` in `PSet 5` (domain 1) and `PSet 8` (domain 2).

[1] = Actor: 5 [55 X]   מֹשֶׁה

Set 5 3sm=MCH
Pref   PSet PAct ln Dom Ty Text
```
 21:  5: 5   8 d: 1 N  EXO 19,03                    וּמֹשֶׁה עָלָה אֶל-הָאֱלֹהִים
     [W-<Cj>] [MCH{21:5:5} <Su>] [<LH <Pr>] [>L H->LHJM <Co>]
 22:  5: 5   8 d: 1 N  EXO 19,03                    וּמֹשֶׁה עָלָה אֶל-הָאֱלֹהִים
     [W-<Cj>] [MCH <Su>] [<LH{22:5:5} <Pr>] [>L H->LHJM <Co>]
 25:  5: 5   9 d: 1 N  EXO 19,03                    וַיִּקְרָא אֵלָיו יְהוָה מִן-הָהָר
     [W-<Cj>] [JQR> <Pr>] [>LJW{25:5:5} <Co>] [JHWH <Su>] [MN H-HR <Lo>]
```

Set 8 2sm= . . . .
Pref   PSet PAct ln Dom Ty Text
```
 29:  8: 5  11 d: 2 Q  EXO 19,03                    כֹּה תֹאמַר לְבֵית יַעֲקֹב
     [KH <Mo>] [T>MR{29:8:5} <Pr>] [L-BJT J<QB <Co>]
 32:  8: 5  12 d: 2 Q  EXO 19,03                    וְתַגֵּיד לִבְנֵי יִשְׂרָאֵל
     [W-<Cj>] [TGJD{32:8:5} <Pr>] [L-BNJ JFR>L <Co>]
 67:  8: 5  24 d: 2 Q  EXO 19,06                    אֲשֶׁר תְּדַבֵּר אֶל-בְּנֵי יִשְׂרָאֵל
     [>CR <Re>] [TDBR{67:8:5} <Pr>] [>L BNJ JFR>L <Co>]
```

Set 5 3sm=MCH
Pref   PSet PAct ln Dom Ty Text
```
 70:  5: 5  25 d: 1 N  EXO 19,07                    וַיָּבֹא מֹשֶׁה
     [W-<Cj>] [JB{70:5:5} <Pr>] [MCH <Su>]
 71:  5: 5  25 d: 1 N  EXO 19,07                    וַיָּבֹא מֹשֶׁה
     [W-<Cj>] [JB> <Pr>] [MCH{71:5:5} <Su>]
 72:  5: 5  26 d: 1 N  EXO 19,07                    וַיִּקְרָא לְזִקְנֵי הָעָם
     [W-<Cj>] [JQR>{72:5:5} <Pr>] [L-ZQNJ H-<M <Co>]
 75:  5: 5  27 d: 1 N  EXO 19,07                    וַיָּשֶׂם לִפְנֵיהֶם אֵת כָּל הַדְּבָרִים הָאֵלֶּה
     [W-<Cj>] [JFM{75:5:5} <Pr>] [L-PNJHM <Co>] [>T KL H-DBRJM H->LH <Ob>]
 80:  5: 5  28 d: 1 N  EXO 19,07                    אֲשֶׁר צִוָּהוּ יְהוָה
     [>CR <Re>] [YWHW{80:5:5} <PO>] [JHWH <Su>]
 89:  5: 5  34 d: 1 N  EXO 19,08              וַיָּשֶׁב מֹשֶׁה אֶת דִּבְרֵי הָעָם אֶל  יְהוָה
     [W-<Cj>][JCB{89:5:5}<Pr>][MCH <Su>][>T DBRJ H-<M <Ob>][>L JHWH <Co>]
 90:  5: 5  34 d: 1 N  EXO 19,08              וַיָּשֶׁב מֹשֶׁה אֶת דִּבְרֵי הָעָם אֶל  יְהוָה
     [W-<Cj>][JCB <Pr>][MCH{90:5:5}<Su>][>T DBRJ H-<M <Ob>][>L JHWH <Co>]
 96:  5: 5  35 d: 1 N  EXO 19,09                    וַיֹּאמֶר יְהוָה אֶל-מֹשֶׁה
     [W-<Cj>] [J>MR <Pr>] [JHWH <Su>] [>L{96:5:5} MCH <Co>]
```

3.4. From Actors to a Network of Actors: Participants [steps 7-8]

3.4.1. Identification of Participants [step 7]
It is important to try a next step in the participants analysis, since the text level actors identified thus far still are patterns identified in the linguistic layer of a text. However, a reader of a text, using also semantic and cultural knowledge, will take a further step in decreasing the number of 'real' participants identified in a text. For example, when processing the information of a text, a reader will conclude that some actors with different PAct labels appear to refer to the same participant, *e.g.,* יְהוָה and אֱלֹהִים, although one should not too quickly identify them completely, since they also may express a change in viewpoint as the plot of a text develops. So we have to narrow down the number of actors into a smaller number of more abstract participants, even when we know that it will not always entail complete referential or semantic identity. The best approach here again seems to be experimentation. Let us assume that the referents of particular actors with different linguistic labels fully match, and then see whether that is a contribution to the text's consistency in terms of the interaction between actors, or not.

The procedure involves combining those PActs with different labels, *e.g.,* 'JHWH' and '>LHJM' (אֱלֹהִים), or 'BJT J<QB' (לְבֵית יַעֲקֹב) and 'BN JFR>L' (לִבְנֵי יִשְׂרָאֵל) that appear (or seem) to point to the same referent in the text. A more abstract label is then assigned to such combinations, *e.g.,* '*God*' or '*Israel*'. Finding and storing such identifications is done first in an interactive process. Once particular identifications have been established by the analysis of more texts, one may be able to construct a lexicon or a data set to be used by automatic procedures of identification.

Such 'final' products of combined PActs can be regarded as the text's *participants*.

3.4.2. Identification of clusters of Actors [step 8]
Finally, one can also create clusters of PActs that appear to be interrelated in different ways, *e.g.,* by exclusively referring to a 'part of' a main *Actor*. A cluster of PActs appears to be located around the Actor: 'mountain' הָהָר (verse 2). In Exodus 19 we find: 'mount Sinai' הַר סִינָי (verse 11), 'top of the mountain' רֹאשׁ הָהָר (verse 20), 'bottom of the mountain' תַּחְתִּית הָהָר (verse 17).

```
          SJNJ
      HR  SJNJ
      HR
  R>C HR
TXTJT HR
```

By finding a distinction between main actors and dependent actors, one is again able to narrow down the number of participants active in a text.

Even more than the work in steps 1 to 6, the research involved in steps 7 and 8 still is in its experimental stage. Here we are working both with linguistics and literary analysis at the same time. In the final paragraph I will give some indications of how the participant analysis may become a significant contribution to various domains of biblical research.

3.5. Summary of the analytical steps

Levels of analysis:

PHRASES	DOMAINS	TEXT	DISCOURSE

Data types produced:

PREFS	PSETS	PACTS	PARTICIPANTS

built from:

phrases in text	PRefs in domain	PSets in domains	PActs in text

identified by:

1. by PNG features 2. by: genitive connections or by coordination within phrases	1. Morphological or 2. lexical back reference within domains	1. lexical identity 2. crossing domain boundaries N-Q: Sp = Se; Aud = Adr	identical referents, 1. proposed by user; 2. or taken from a lexicon built from previous proposals

4. Remaining Tasks and Further Research

Following the steps described above, the programs being developed to identify participants in texts in Biblical Hebrew apply a strategy of bottom-up linguistic analysis. The first step is to collect all phrases and sub phrases with features that are linguistic markers of textual participants. The next steps using these data aim at establishing a hierarchy of relations between the elements of the first collection. It is important to bear in mind that these analytical procedures are used primarily to expand the syntactic information present in the database. Developing and applying them, however, makes clear that in developing these programs one is also experimenting with discourse linguistics.

4.1. Linguistics of Biblical Hebrew

Both the identification of relations within a textual domain (narrative or direct speech) and the identification of such relations while crossing the linguistic borders between such domains require interactive correction of machine made proposals, *e.g.*, cases where a text is complicated by the presence of a number of 3rd person pronominal suffixes.[10] The same is true for those cases mentioned above, where information not present in a particular text is needed to propose referential identity of particular actors, such as 'BJT J<QB' (לְבֵית יַעֲקֹב) and 'BN JFR>L' (לִבְנֵי יִשְׂרָאֵל). Therefore, a first task of further research is to study such cases of correction to see whether patterns could be found for such cases. It implies continued linguistic research in the domain of text grammar, semantics and discourse analysis.

An interesting case is found in those texts where the referential identity of particular actor labels is only disclosed gradually in the process of reading, *e.g.*, in Exodus 2 where the *name* "Moses" is only mentioned after he has been introduced first as "son", "child" and "boy". So, one has to find analytical procedures that are able to deal with shifts in the narrator's viewpoint and the elaboration of a plot.

---

[10] For example, Judges 3:5-6 discussed in E. Talstra, "Text linguistics: Biblical Hebrew", 755 (ftn. 2).

4.2. Bible Translations

Participant analysis is also a promising new tool for a critical assessment of translations. Modern translations of the Bible, trying to achieve the goal of being 'easy to read' exhibit a tendency not only to use a restricted set of words from a modern lexicon, but they also quite easily make literary or exegetically motivated changes in the plot of a text, sometimes overruling the analysis of participants.

An example is the text of Zechariah 1:1-6. In the opening section of the book, YHWH begins a dialogue with the generation in exile, summoning them to recall the words of the 'early prophets' to the previous generation and to see how long it took before that generation began to understand. Translators sometimes seem inclined to avoid the long direct speech section of words spoken by God, where (verse 6) he quotes the reaction by the previous generation: "Did not your fathers turn back and acknowledge …?" For instance, in the *Good News Bible*[11] a new paragraph is started there with, among others, a change of the address 'you'-plural into a narrative section speaking about "*they* repented …". This changes the entire pragmatics of the section, since God's exhortation to the present generation to reflect on what has been, is now aborted.

An even worse case is the change of participants in the modern translation into Dutch, the *Nieuwe Bijbelvertaling*.[12] Instead of having God continuing to be the speaker, the narrator takes over the last section: "Then the people returned and acknowledged: the Lord has done according to what he had announced." Thus "your fathers" has been changed into "the people", now meaning Zechariah's audience. The result is a classical pious text without any of the rhetorical tension of the Hebrew text: the generation of Zechariah heard the message and repented. That, however, would be the participant mentioned in verse 3, but the text has not a single linguistic sign indicating that 'they turned' in verse 6 should syntactically be connected back all the way up to the 'them' in verse 3. Nor is there any sign that God's direct speech has ended in verse 6a, so that the narrator would be the speaker of these final lines. On the contrary, in the text God's speech continues to the end of verse 6 and it is God who speaks about the fathers having finally returned. It makes God's speech a challenge to Zechariah's audience and a proper introduction to the book: having heard this, how will you respond?[13]

4.3. Literary Interpretation

Similar challenges are met when one applies the techniques of participant analysis to poetic sections which usually lack a narrative frame to help introduce and identify actors and participants. So initially, the actors found in such texts remain more or less abstract entities. For example, in Psalm 63, one finds a "you" (identified as 'God' and 'my God'), but also an "I", a "they", "the king" and "all (who swear)". Of course, in such cases participant analysis is not supposed to replace textual interpretation, for example, is the "I" to be identified with "the king"? Rather it provides us with an instrument that produces a linguistic frame for interpretation, since it can help characterize the actors by the language they use and by the position they take in the network of participants of the text. The "I" is most frequent, but is only marked by pronominal suffixes or by verbal morphemes. "God" as addressee is present in pronominal suffixes, once by the pronoun אַתָּה and once as the implied subject of a verb

---

[11] *Good News Bible (Today's English Version)* (American Bible Society, 1992).

[12] *Nieuwe Bijbelvertaling* (Haarlem: Nederlands Bijbelgenootschap, 2004): ' " [6]Toch hebben mijn woorden en de wetten die ik mijn dienaren de profeten had opgedragen te verkondigen, jullie voorouders getroffen."' Toen kwam het volk tot inkeer en erkende: 'De HEER van de hemelse machten heeft vanwege onze handel en wandel met ons gedaan wat hij zich had voorgenomen.' . The popular translation *Bijbel in Gewone Taal* (Haarlem: Nederlands Bijbelgenootschap, 2014), has elaborated the text further in this direction.

[13] E. Talstra, "Who is who in Zechariah 1:1-6? Text linguistics, participant tracking and our reading of Biblical texts" (forthcoming).

(verse 8). The "they" are only present by the pronoun הֵמָּה (verse 10). Once the actors in the text have been identified, we can also label the subjects of all verbs, even if in the text they have no explicit subject phrase, such as אֱלֹהִים = הָיִיתָ in verse 9 or יָבֹאוּ = הֵמָּה in verse 10. In addition to this, these actors can be further characterized by clusters (see § 3.4.2.), *e.g.* "I" and "my soul", "my flesh", or "God" = "you" and "your name", "your right hand". An interesting question for textual interpretation is: can one find similar clusters or a comparable network in other Psalms too?

Once the network of participants in the dialogue of Psalm 63:2-11 has been established, one can propose a relationship with the network of participants in the final section, verse 12, where we only find actors in third person: "the king", "all who swear" and also again אלהִים, who now no longer can be the addressee. So speaker and audience have changed here. Again, searching for a comparable transition of such networks of actors in other Psalms might be helpful for literary interpretation. Do other texts have a *participant* אֱלֹהִים where the text also reveals a transition from a second person *actor* אלהִים to a third person *actor* אלהִים? An example is Psalm 64:1,8.[14]

These experiments may demonstrate why the construction of participants as a new layer of data in the database is an important linguistic tool for textual analysis and for evaluating exegetical interpretation. Reading requires one to be aware of the interaction of participants, lexical expressions and rhetorical elements in the text. It is an interesting and stimulating experience to find out how far one could go in constructing the linguistic data needed, before further exegetical options of a particular text are explored.

---

[14] Christiaan Erwich and Eep Talstra, "The text as our teacher. Participant tracking in Psalm 64" in *Multiple Teachers and Disciples in Biblical Texts*, ed. Bert Koet and Archibald van Wieringen (forthcoming).