

On memories and methods.

Eep Talstra

Last year I had the opportunity to present a paper when our research group, now called ETCBC, was celebrating its 40th birthday (October 31, 2017). In March I will participate in a symposium on the use of text databases in biblical scholarship, organized at the Fjellhaug International Academic College in Copenhagen. So, if you have been active for over 40 years in this discipline, such occasions cannot but provoke a special blend of remembering the romantic past of punched cards and of considering how to develop new methods in the era of internet. In this paper I will try to contribute to both. In the years past I have become aware of the strong interaction between the production of *instruments* and the continuing change of *methods* used for research in biblical texts. The development of computer based instruments and methods has moved from the imitation of instruments used in classical *philology* (concordances; text editions) to the design of new instruments for critical *linguistic* research (lexical and syntactic patterns). In our days our discipline moves on again: from linguistic analysis to *textual* interpretation (text structures and patterns of communication). More or less simultaneously one sees the change of techniques: from punched cards and main frames to CDRoms and query systems, and now to systems of open access cooperation. Within this history of changes the challenge has remained the same: how to combine methods and techniques?

1. Imitation: Instruments for philology

When we started to imitate classical concordances by computer, we first had to construct a database where to each word in a text (token) its corresponding lexeme (type) was added. That would allow one to sort a text according to the lexemes used there and to print a concordance of lexemes with all the tokens found in the *surface text*. This is in fact the algorithm implied in concordances of the type: Lisowsky. An advantage was the new option to produce specialized concordances, e.g. of one book or of some special chapters. And if you had in your database also marked clear syntactic units, such as clauses, the result would be a more or less classical concordance. For example, a concordance of the lexeme “father” in 1 Samuel 12, presenting the surface text forms (tokens) in the context of a clause. (I have replaced the transliterated text from the database by a translation and the Hebrew text.)

>B/ אָבִי [5] ‘father’

ISAM 12,06 and who brought your *ancestors* up out of the land of Egypt.

וְאֲשֶׁר הֶעֱלָה אֶת-אֲבוֹתֵיכֶם מֵאֶרֶץ מִצְרָיִם

ISAM 12,07 that he performed for you and for your *ancestors*

אֲשֶׁר-עָשָׂה אֲתֶכֶם וְאֶת אֲבוֹתֵיכֶם

ISAM 12,08 [who brought forth your *ancestors* out of Egypt

וַיּוֹצִיאֵנִי אֶת אֲבוֹתֵיכֶם מִמִּצְרָיִם

ISAM 12,08 then your *ancestors* cried to the LORD

וַיִּזְעֻקוּ אֲבוֹתֵיכֶם אֶל-יְהוָה

ISAM 12,15 then the hand of the LORD will be against you and your *ancestors*

וְהָיְתָה יַד יְהוָה בְּכֶם וּבְאֲבוֹתֵיכֶם

Traditionally this type of concordance was helpful for counting the words used in particular texts and to study them in terms of semantic or theological categories. Thus traditional concordances often have been used in interaction with a biblical theology on word level, or actually: on concept level, with a focus on theological statements, for example about ‘father’ as such. One can observe that in the discussion of the lemma “father” in THAT.¹ In summary: *“Father” can mean a, b, or c, etc. and the lemma “father” used in dtr. literature has a special theological function.* In this way a traditional concordance actually is an instrument for research in the history of religion collecting material for biblical theology.

As a user of this type of concordance you yourself have to observe that actually in 1 Samuel 12 we only have cases of “fathers” (plural) + sfx: “you” (plural). Therefore, researchers developing instruments for computer-assisted biblical research also have tried to imitate concordances of a different type, presenting surface forms (tokens) sorted according to additional features, such as singular, plural, pronominal suffixes, etc. This is the model Mandelkern. If you can add such morphological features to each word in a database, it will become much easier to observe, for example, that in 1 Samuel 12 we only find the form אבותיכם : fathers (plural) + pronominal suffix (second person plural). The series *“the Computer Bible”*, published in the seventies of the last century, gives a good impression of all the experiments that were needed before we were able to imitate both the Lisowsky-model and the Mandelkern-model and present concordances of tokens in a linguistically meaningful context.² Still, Mandelkern does not present us with ‘phrases’, so as a user even of this type of concordance you yourself should notice that in 1 Samuel 12 we find two cases of a complex phrase where the audience “you” and “your fathers” are addressed in combination. V.7: אבותיכם ואתם and v.15: ובאבותיכם, בכם (skipping here the text critical problem of את and את in verse 7).

2. Instruments for linguistic analysis: an individual text and the system of language

Here the real work began, for we had to leave the technique of imitating existing instruments and develop something new, i.e. the construction of a database with linguistic data beyond the level of lexemes. We needed a database identifying phrases, clauses, sentences (i.e. compound clauses) and text segments, such as narrative and direct speech sections,³ for otherwise it would not be possible to collect all other cases of compound phrases, as mentioned above, in the Hebrew Bible. Now such a text linguistic database exists, one can use the phrase “you and your

¹ E. Jenni, ‘אב ‘āb Vater’, in: Ernst Jenni, Claus Westermann (ed.), *Theologisches Handwörterbuch zum Alten Testament*, München/Zürich, 1971, Band I, kol, 1-17

² E. Talstra, ‘Exegesis and the Computer Science: Questions for the Text and Questions for the Computer’, *Bibliotheca Orientalis* 37 (1980) 120-129

³ Chr. Hardmeier, E. Talstra & B. Salzmann (ed.), *SESB: Stuttgart Electronic Study. Bible Instruction Manual. (Stuttgarter Elektronische Studienbibel. Handbuch)*, Stuttgart: Deutsche Bibelgesellschaft / Haarlem: Nederlands Bijbelgenootschap, 2004 [ISBN 3-438-06458-8] pp. 120 [2007: SESB, edition 2.0]

fathers” as the starting point for a query searching for similar cases. Can we find more cases of ‘fathers’ in complex phrases of the type: “prep. + suffix (2nd person plural) and (identical) prep. + fathers + suffix (2nd person plural)? It is done by composing queries as instructions to *search* for similar linguistic constructions in the database of analyzed texts.⁴

Here the discussion of methods of textual research comes up again. It is clear that both traditional and modern instruments of sorting and searching require from the user to carefully study the output data produced and the texts listed, in order to observe particular phenomena (as in Lisowsky or Mandelkern concordances), or to start thinking of ways of broadening or narrowing down the initial query. For example, can we also find cases of the same phrase type with 3rd person plural suffixes? Indeed such cases do exist, e.g. in Jeremiah 24:

Jer 24:10 And I will send sword, famine, and pestilence upon them, until they are utterly destroyed from the land that I gave **to them and their ancestors**.

In case the researcher precisely knows what s/he wants to search for, this query technique works very well and we need it. However, when doing linguistic or exegetical analysis one often would be pleased with a little more help with a presentation of relevant options and suggestions to decide what further steps to take. Therefore, can we write algorithms that themselves can sort the data and propose linguistic categories? To experiment with this I have expanded the concordance program mentioned above with a module by which the program itself registers and categorizes the various phrase constructions in which lexemes occur. This is a preliminary version since it looks for the individual lexemes composing a particular phrase, listing for each lexeme some relevant features: singular, plural, plural with pron. sfx., with proposition, etc. For example, for 1 Samuel 12 this concordance will list, of course, the same 5 cases of “father”, but now they will be presented in three groups:

>B/ אָבִי [5] ‘father’

[1] Noun + Sfx:2 Pl Mas

ISAM 12,08 then your ancestors cried to the LORD

וַיִּזְעֻקוּ אֲבוֹתֵיכֶם אֶל-יְהוָה

[2] Prep + Noun + Sfx:2 Pl Mas

ISAM 12,06 and who brought your ancestors up out of the land of Egypt.

וְאֲשֶׁר הֶעֱלָה אֶת-אֲבוֹתֵיכֶם מֵאֶרֶץ מִצְרַיִם

ISAM 12,08 who brought forth your ancestors out of Egypt.

וַיּוֹצִיאֵם אֶת אֲבוֹתֵיכֶם מִמִּצְרַיִם

[3] Prep + Sfx:2 Pl Mas +Cj +Prep +Noun +Sfx:2 Pl Mas

ISAM 12,07 that he performed for you and for your ancestors

אֲשֶׁר-עָשָׂה אֲתָכֶם וְאֶת אֲבוֹתֵיכֶם

ISAM 12,15 then the hand of the LORD will be against you and your ancestors.

וְהָיְתָה יַד יְהוָה בְּכֶם וּבְאֲבוֹתֵיכֶם

⁴ “ETCBC database of the Hebrew Bible with Query Saver” via: <http://shebanq.ancient-data.org>

From these categories a researcher can easily choose which one of these groups may be interesting for further searching in the book of Samuel or in the Hebrew Bible.⁵

3. From linguistic forms and functions to text level linguistic relations

In actual research we are entering again a next level of research, *i.e.*: the text, both as linguistic structure and as literary design. Until now the sorting and searching of linguistic data was based on *surface* forms (tokens) of particular lexemes (types) as they are actually present in the text. For further linguistic research (syntax) and for exegetical research we need to gain access to all linguistic phenomena that refer to anyone playing a role in the text, even when such references are ‘hidden’ in verbal forms or in pronominal references, For example, as readers we ‘know’ to whom the “your” in the expression “your fathers” refers to. But can one also construct algorithms to analyze the linguistic mechanism of such references and store the results in the database?

This is the task of *participant tracking* and it implies that we go beyond the study of linguistic *forms* or *functions* and enter the study of text level linguistic *relations*. Therefore, the existing database is now used to divide a text into domains (*i.e.* sections owned by a Speaker and an Audience). We first identify the participants active within a domain and after that we identify the **Speaker** and the **Audience** in the higher level domain, across the domain borders, with the **Sender** and the **Addressee** present in the embedded domain.⁶ See the example presented below. By this mechanism one can identify 1st and 2nd person forms with 3rd person references to the same participant in a text. Participants that occur in 3rd person both in higher level domains and in lower level domains (‘the Lord’ in the example below) can be identified directly. The result is an overview of participants active at text level, independent from the form they may take in a particular textual domain.

For example:

In 1 Samuel 12: 6-15 we find in Samuel’s speech mostly the suffix כֶּם : ‘you-plural’ to refer to the audience ‘people’. If, however, the program can find within the same domain a noun or a pronoun that matches with the suffix, that (pro)noun is taken as the label of the participant that is active within that domain. So in this domain the suffix כֶּם used in אֲבֹתֵיכֶם, could be identified with the pronoun אֲנִי in verse 14. After that the pronoun אֲנִי could be identified with the addressee “people” הָעָם in verse 6, by applying the rule that the אֲלֵ- complement used with a verb of speaking in a clause that introduces a direct speech section equals the “you” in that section. (compare the same procedure in verse 20)

In the text segment presented as an example below four of the participants, active in various domains, are marked by different colors: Samuel, the people, (your)ancestors, the LORD.

⁵ Some results I have presented in my speech at the 40th anniversary of ETCBC, October 31 2017: ‘Computers and Language, Literature and Living Tradition. “The fathers” as Data, Participants and Arguments in the Hebrew Bible.’

⁶ Discussed in: E. Talstra, ‘Approaching the mountain of Exodus 19: thou shalt explore syntax first’, *Hiphil Novum* 3/1 (2016) 2 – 24

1 Samuel 12: 6-15

Narrative: Narrator >> Reader

6	Samuel (Speaker) said to the people (Audience), direct speech: Sender= Samuel = "I" ; Addressee = people = "you" pl.
7 so that I may enter into judgment with you before the LORD , ... all the saving deeds he performed for you and for your ancestors .
8	When Jacob went into Egypt then your ancestors cried to the LORD
9	But they forgot the LORD their God;
10	Then they cried to the LORD , and said , direct speech: Sender = your ancestors = we ; Addressee = LORD / you 'We have sinned, because we have forsaken the LORD , but now rescue us out of the hand of our enemies,
11	And the LORD sent . . . and . . . Jephthah and Samuel and rescued you out of the hand of your enemies on every side;
12	When you saw King Nahash of the Ammonites came against you , you said to me , direct speech: Sender= people = "us" ; Addressee = Samuel = "I" 'No, but a king shall reign over us ,'
13	though the LORD your God was your king. See, here is the king
15	but if you will not heed the voice of the LORD , the hand of the LORD will be against you and your ancestors .

Once we know where the participants in a text are present in whatever linguistic marking, we are able to move our research from sorting surface text patterns into the level of linguistic roles and patterns of communication, e.g. listing a participant by all the linguistic markers used (3.1.), or listing the predication frames of all clauses where a participant has a particular role (3.2.).

3.1. Inventory of all *linguistic markers* of a participant in a text.

When we have successfully tracked all occurrences of ‘your fathers’ through the text, independent from the particular linguistic form used, we will be able to present a new kind of concordance that lists all cases of a participant, whether being referred to in 1st person, 2nd person, 3rd person, a proper name, or a pronominal suffix. As a consequence, instead of the 5 cases listed in the traditional concordance we now find 19 cases of an actual reference to “your fathers” in 1 Samuel 12 (accepting a verb and an explicit subject phrase as two occurrences). Since the participant tracking has found that the “your” in “your fathers” refers to Samuel’s actual audience: “the people”, the label produced for this Actor now is: “>B <M” אֲבֹתֵיכֶם i.e. “fathers” of the “people”.

Some selected lines, from verse 6-10, Samuel’s speech and from verse 10, the father’s speech: (output; Hebrew text added)

Actor: 14 [19 X] >B <M אֲבֹתֵיכֶם

Set 29 3pm=>B+>TM

ISAM 12,06 [W-<Cj>] [>CR <Re>] [H<LH <Pr>] [>T{91:29:14}>BTJKM<Ob>] [M->RY MYRJM<Co>]
וְאִשָּׁר הָעֵלָה אֶת אֲבֹתֵיכֶם מֵאֶרֶץ מִצְרָיִם

ISAM 12,08 [W-<Cj>] [JCBWM{125:29:14} <PO>] [B--MQWM H-ZH <Co>]

וַיֵּשְׁבוּ בְּמָקוֹם הַזֶּה

ISAM 12,09 [W-<Cj>] [JMKR <Pr>] [>TM{132:29:14} <Ob>] [B-JD SJSR> / FR YB> XYWR / ---

וַיִּמְכַר אֶתְכֶם בְּיַד סִיסְרָא שַׂר-צָבָא הַצּוֹר

ISAM 12,10 [W-<Cj>] [JZ<QW{146:29:14} <Pr>] [>L JHWH <Co>]

וַיִּזְעַקוּ אֶל יְהוָה

ISAM 12,10 [W-<Cj>] [J>MR{148:29:14} <Pr>]

וַיֹּאמֶר (וַיֹּאמְרוּ)

Set 36 1pc=

ISAM 12,10 [XV>NW{149:36:14} <Pr>]

חָטְאוּ

ISAM 12,10 [HYJLNW <PO>] [M-JD >JBJNW{160:36:14} <Co>]

הַצִּילֵנוּ מִיַּד אֲבוֹתֵינוּ

3.2. Predication frames: constituents and parsing.

Once we are able to identify all participants active in a text, we can also present the text’s clauses sorted by the *verbs* used, with the participant active as subject and the other clause constituents as the verb’s satellites. A program is collecting and sorting all clauses according to the verbs used and the subjects, either present as NP or implied by the verb. From there it can now construct the ‘*predication frames*’ used in the text.

Some examples of the verb **נציל** (hiph'il) in 1 Samuel 12 (the Hebrew text I have added to the output)

Verb + subject

NYL Hi =PRef 309: THW (NPSu=0) (NP-subject **תהו**, is absent, but is know from previous lines)
 ISAM 12,21 [W-<Cj>] [L> <Ng>] [JYJLW <Pr>]
 <Pr> **וְלֹא יִצִּילוּ** ISAM 12,21

Verb + subject + object + complement

NYL Hi =PRef 156: JHWH (NPSu=0) || Obj=sfx **נו** =>B <M || Comp= JD >JB +sfx **נו** =>B
 ISAM 12,10 [HYJLW <PO>] [M-JD >JBJNW <Co>]
 <Co> **מִיַּד אֹיְבֵינוּ** <Ob><Pr> **הַצִּילָנוּ** ISAM 12,10

Verb + subject + object + complement

NYL Hi =PRef 170: JHWH (NPSu=0) || Obj=171: <M || Comp=172: JD >JB <M
 ISAM 12,11 [W-<Cj>] [JYL <Pr>] [>TKM <Ob>] [M-JD >JBJKM / M-SBJB <sp>
 <Co> **מִיַּד אֹיְבֵיכֶם מִסָּבִיב** <Ob> **אֶתְכֶם** <Pr> **וַיִּצִּל** ISAM 12,11

In this presentation the emphasis is on the verbs and their constituents. That information is of importance for research in verbal valence and for translation proposals. For semantics it is important to know who the subject actually is, or what its semantic features are (human, divine, ...) even when the subject is not present as a NP, as it is the case with these examples.

Another option is to present the same predication frames from the perspective of text level communication: who is using a particular predication frame addressing whom?

For example:

domain: 7 [=direct speech domain in verse 10; "your fathers" address YHWH]
 =====

Speaker: 28: 3pm=>B+>TM Audience: 25: 3sm=JHWH

 <Co> **מִיַּד אֹיְבֵינוּ** <Ob><Pr> **הַצִּילָנוּ** ISAM 12,10

Predic	=PRef 156: NYL	
Subject	0	[= 2sm=]
Object	= Set 35: 1pc=	sfx
Complement	= Set 39:	JD >JB+1pc "WE" (our enemies = your fathers' enemies)
Adjunct	- -	
Time	- -	
Locative	- -	

domain: 6 [=direct speech domain in verse 10; Samuel addresses you (plural) = “the people”]

Speaker: 1: 3sm=CMW>L Audience: 24: 3pm=<M

<Co> מִיַּד אֵיבֵיכֶם מִסָּבִיב <Ob> אֶתְכֶם <Pr> וַיִּצֵל ISAM 12,11

Predic =PRef 170: NYL
 Subject 0 [= 3sm=JHWH]
 Object = Set 26: 2pm=>TM sfx
 Complement = Set 40: JD >JB+>TM (your enemies = the people’s enemies)
 Adjunct - -
 Time - -
 Locative - -

4. From language to *literature*.

Once we are able to sort linguistic patterns in terms of their text level function, we are getting closer to presenting material that will assist us in tracking processes of communication in a text. The use of the verb **נצַל** in verse 10 and 11 reveals that the transition from an analysis of linguistic *surface patterns* to an analysis of linguistic *relations* brings us in the domain of text as a literary composition. The text of Samuel’s speech appears inconsistent with linguistic patterns of communication. (10) “Your fathers” cried: **rescue us** out of the hand of **our** enemies! But Samuel concludes: (11) **The Lord rescued you** out of the hand of **your** enemies. (not: **them**)

1 Samuel 12

Narrative: Narrator >> Reader

6	Samuel said to the people,
	direct speech: Samuel = “I” >> people = “you” pl.
8	When Jacob went into Egypt
9	then your ancestors cried to the LORD
10	But they forgot the LORD their God;
	Then they cried to the LORD, and said,
	direct speech: your ancestors = we >> LORD = LORD / you
	‘We have sinned,
	because we have forsaken the LORD,
	but now rescue us out of the hand of our enemies,
11	And the LORD sent . . . and . . . Jephthah and Samuel and rescued you out of the hand of your enemies on every side;

Here we observe that occasionally literary patterns of communication overrule the linguistic system of text syntax. Samuel narrates (11) that the Lord reacted to the cry of the fathers by sending saviors, but then his narrative suddenly stops. He should have continued with: ‘and he saved **them** (your fathers) from **their** enemies’, but instead he re-addresses his audience: ‘and he saved **you** from **your** enemies’. It is a literary motivated shortcut which reveals that Samuel when narrating about the fathers actually is addressing his own audience. As a consequence the participant tracking programs conclude to two sets of enemies: “our enemies” = enemies of your fathers (10) and “your enemies” = enemies of his audience (11). This confirms in my view that the close connection of “you” and “your fathers” within one phrase (verse 7 and 15) is not a textual error, as LXX and several modern commentators suggest for verse 15, but is an intended idiom.⁷ As far as I can see scholars who propose to change the text of verse 15 did not see the linguistic inconsistency already present in verse 10 and 11.

This example may illustrate that computer-assisted text linguistic research can help us to reconsider our methods even on the level of textual and exegetical analysis. In classical commentaries research is dominated by questions of *history of religion*: who did speak to whom about what and when? Biblical scholarship needs to ask that kind of questions. However, with more linguistic data at our disposal we can also ask questions about how to understand texts in terms of the *process of communication*. One can do that by identifying textual participants in terms of the (embedded) domains they are active in together with the Speaker and Audience that ‘own’ the communication in those domains.

Below a few examples of texts where the participant ‘their/your fathers’ is presented in the context of one or more levels of communication by sets of speaker and audience.

Fathers as negative example and as warning to actual generation:

1 Sam. 12:8 When Jacob went into Egypt and the Egyptians oppressed them, then **your ancestors** cried to the LORD and the LORD sent Moses and Aaron, who brought forth **your ancestors** out of Egypt, and settled them in this place.

1 Sam. 12:9 But **they** forgot the LORD their God; and he sold **them** into the hand of Sisera, ...

Narrator >> Reader

[Samuel >> people

[your fathers cried to the Lord; they forgot the Lord

]

]

you should remember ...

In Jeremiah and Ezekiel we find more complex phrases explaining that the actual generation is not really different from the generation of the fathers. They have acted the same, they will be treated the same way. So Samuel’s argument: ‘please watch the dispute between God and your fathers’ appears to have failed.

⁷ discussed in H.J. Stoebe, *Das erste Buch Samuelis* (BKAT VIII.1], Gütersloh, 1973, 234.

Fathers as negative example and as examples of the history of God's dilemma's:

Ezekiel 20:18

prophet narrates: YHWH is narrating about himself

18 I said to **their children** in the wilderness: Do not follow the statutes of **your fathers**, nor observe their ordinances, nor defile yourselves with their idols.

Narrator (= "I" [Ezekiel]) >> Reader;

2 [Word YHWH >> me

5 ["I" >> elders of Israel

5 [YHWH >> elders of Israel

18 [YHWH >> children of Israel [sons]

[Do not follow the habits of **your fathers**

]

]

21 The sons refused. I, YHWH, planned to destroy them;
But I did not do it.

]

]

]

Fathers as part of all authorities that have failed:

Daniel 9:6 and 8

"we": prayer / confession about all generations and officials

Dan 9:6 We have not listened to your servants the **prophets**, who spoke in your name to our kings, our princes, and our **ancestors**, and to all the people of the land.

Dan 9:8 Open shame, O Lord, falls on us, our kings, our officials, and our **ancestors**, because we have sinned against you.

[Narrator: "I" [Daniel]) >> Reader

[Daniel ("I") >> YHWH

[we, kings, officials, **fathers** .. have sinned

]

]

]

One can see the literary or theological role of the fathers change, until God states that they are no longer decisive (Zecharia 1:1-6). To each of the examples I have added a short exegetical conclusion. Of course, here I have left the domain of linguistic analysis, but I hope the examples demonstrate some of the linguistic and literary methods made possible by computer-assisted tools of text level analysis.

This work on textual level still is very much experimental. We are considering how to make the preliminary results available and how to find ways to have colleagues from other institutes join us in doing these experiments in a shared electronic laboratory. That will be our next task.

March 17, 2018