The Lester and Sally Entin Faculty of Humanities
School of Philosophy, Linguistics and Science Studies

Constructions "on the move": From independent-sentential to lexical constructions

Thesis submitted for the degree of
"DOCTOR OF PHILOSOPHY"

by

Israela Becker

Prepared under the guidance of
Prof. Mira Ariel and Prof. Rachel Giora

September 2023

In memory of
my father, William (Willy) Becker (1929-2001)
and
my grandmother, Ite-Reyze (Yeti) Becker (1895-1973)
(née Rosenblatt)
who did not survive the Holocaust

Ithaca

When you set out on the journey to Ithaca,
pray that the road be long,
full of adventures, full of knowledge.
The Laestrygonians and the Cyclopes,
the raging Poseidon do not fear:
you'll never find the likes of these on your way,
if lofty be your thoughts, if rare emotion
touches your spirit and your body.
The Laestrygonians and the Cyclopes,
the fierce Poseidon you'll not encounter,
unless you carry them along within your soul,
unless your soul raises them before you.

Pray that the road be long;
that there be many a summer morning,
when with what delight, what joy,
you'll enter into harbours yet unseen;
that you may stop at Phoenician emporia
and acquire all the fines wares,
mother-of-pearl and coral, amber and ebony,
and sensuous perfumes of every kind,
as many sensuous perfumes as you can;
that you may visit many an Egyptian city,
to learn and learn again from lettered men.

Ιθάκη

Σὰ βγεῖς στὸν πηγαιμὸ γιὰ τὴν Ἰθάκη,
νὰ εὔχεσαι νἆναι μακρὺς ὁ δρόμος,
γεμάτος περιπέτειες, γεμάτος γνώσεις.
Τοὺς Λαιστρυγόνας καὶ τοὺς Κύκλωπας,
τὸν θυμωμένο Ποσειδῶνα μὴ φοβᾶσαι,
τέτοια στὸν δρόμο σου ποτέ σου δὲν θὰ βρεῖς,
ἂν μέν' ἡ σκέψις σου ὑψηλή, ἂν ἐκλεκτὴ
συγκίνησις τὸ πνεῦμα καὶ τὸ σῶμα σου ἀγγίζει.
Τοὺς Λαιστρυγόνας καὶ τοὺς Κύκλωπας,
τὸν ἄγριο Ποσειδῶνα δὲν θὰ συναντήσεις,
ἂν δὲν τοὺς κουβανεῖς μὲς στὴν ψυχή σου,
ἂν ἡ ψυχή σου δὲν τοὺς στήνει ἐμπρός σου.

Νὰ εὔχεσαι νἆναι μακρὺς ὁ δρόμος.
Πολλὰ τὰ καλοκαιρινὰ πρωϊὰ νὰ εἶναι
ποῦ μὲ τί εὐχαρίστηση, μὲ τί χαρὰ
θὰ μπαίνεις σὲ λιμένας πρωτοειδωμένους.
νὰ σταματήσεις σ' ἐμπορεῖα Φοινικικά,
καὶ τὲς καλὲς πραγμάτειες ν' ἀποκτήσεις,
σεντέφια καὶ κοράλλια, κεχριμπάρια κ' ἔβενους,
καὶ ἡδονικὰ μυρωδικὰ κάθε λογῆς,
ὅσο μπορεῖς πιὸ ἄφθονα ἡδονικὰ μυρωδικά.
σὲ πόλεις Αἰγυπτιακὲς πολλὲς νὰ πᾷς,
νὰ μάθεις καὶ νὰ μάθεις ἀπ' τοὺς σπουδασμένους.

Always keep Ithaca in your mind.
To arrive there is your final destination.
But do not rush the voyage in the least.
Better it last for many years;
and once you're old, cast anchor on the isle,
rich with all you've gained along the way,
expecting not that Ithaca will give you wealth.

Ithaca gave you the wondrous voyage:
without her you'd never have set out.
But she has nothing to give you any more.

If then you find her poor, Ithaca has not deceived you.
As wise as you've become, with such experience, by now
You will have come to know what Ithacas really mean.

Πάντα στὸν νοῦ σου νἄχεις τὴν Ἰθάκη.
Τὸ φθάσιμον ἐκεῖ εἶν' ὁ προορισμός σου.
Ἀλλὰ μὴ βιάζεις τὸ ταξεῖδι διόλου.
Καλλίτερα χρόνια πολλὰ νὰ διαρκέσει.
καὶ γέρος πιὰ ν' ἀράξεις στὸ νησί,
πλούσιος μὲ ὅσα κέρδισες στὸν δρόμο,
μὴ προσδοκῶντας πλούτη νὰ σὲ δώσει ἡ Ἰθάκη.

Ἡ Ἰθάκη σ' ἔδωσε τ' ὡραῖο ταξεῖδι.
Χωρὶς αὐτὴν δὲν θἄβγαινες στὸν δρόμο.
Ἄλλα δὲν ἔχει νὰ σὲ δώσει πιά.

Κι ἂν πτωχικὴ τὴν βρεῖς, ἡ Ἰθάκη δὲν σὲ γέλασε.
Ἔτσι σοφὸς ποὺ ἔγινες, μὲ τόση πεῖρα,
ἤδη θὰ τὸ κατάλαβες ἡ Ἰθάκες τὶ σημαίνουν.

-- Konstantinos Petrou Kavafis, *The Collected Poems*,
(Translation: Evangelos Sachperosglou)

Acknowledgements

The subtitle of this dissertation should have been “the road *ultimately* taken”. And what a remarkable road it has been! As with every long (and challenging) road, one cannot possibly walk alone. I was fortunate to have several people walk with me along the way, people who had a significant effect on the final product.

First and foremost are my beloved advisors, Prof. Mira Ariel and Prof. Rachel Giora, to whom I am grateful beyond any measure. Mira and Rachel are the advisors a student could only dream of. The only thing that outranks their linguistic knowledge is their generosity. They have always known when they should allow me the time to withdraw into myself in order to come up with new ideas, process them, re-shape them, and finally write them down. But they have also known when they should interfere (in the gentlest manner) and offer their invaluable advice. And by doing so, they have made the way rewarding despite occasional bumps. They have made me a better linguist.

I am most grateful to Dr. Ruth Burstein. Ruti is not just a remarkably knowledgeable linguist, who has helped me a lot with issues which concern the fuzzy borders of Hebrew syntax, but she has also been a supportive friend all along the way.

I am most obliged to Prof. John W. DuBois (Jack) who made important suggestions that made me re-think terminology, and led me, eventually, to re-consider some of my ideas.

I am very thankful to Dr. Shir Givoni and Dr. Einat Kuzai with whom I had long productive discussions on various issues related to their fields of expertise, psycholinguistics and language change, respectively.

I am deeply indebted to Prof. Adele Goldberg and Dr. Nurit Melnik for reading my dissertation closely, for their insightful comments, and for their positive reviews which made me very happy.

And last but definitely not least is my dear and devoted brother, Yonatan, who encouraged me to follow my vocation and become a linguist. You are the best friend a person could wish for, and I am fortunate to have you as MY brother!

Throughout the doctoral years I benefitted from scholarships financed by Israel Science Foundation grant 540/19 awarded to Rachel Giora and Israel Science Foundation grant 1398/20 awarded to Mira Ariel, both matched by the School of Philosophy, Linguistics and Science Studies.

Abstract

For many years, the full sentence *xaval al hazman*, literally, 'it's a pity on the (wasted) time', has been a means to (indirectly) express a speaker's negative stance regarding some stance-object. By uttering *xaval al hazman*, the speaker intended to convey that the stance-object was not worthy of the addressee's (and speaker's) time (and attention), see (1). However, around 30 years ago, a new use of *xaval al hazman* appeared on the language scene. *Xaval al hazman* has changed its meaning completely, alongside its grammatical status. It is no longer only an independent sentence. It is also a full-fledged word which belongs to several word classes — an adjective (2a), an adverb (2b) and an intensifier (2c-d) — all conveying a highly intense evaluation. In Example (2a), *xaval al hazman* denotes 'amazing'; in Example (2b), it denotes 'amazingly'; and in Examples (2c-d) — 'extremely' and 'so much', respectively.

(1) One of the worst movies that I've seen lately!!!! […] In short, stay home, ***xaval al hazman***!!! [1]

(tinyurl.com/38syphbc)

(2) a. One of the best! A ***xaval al hazman*** movie! Strongly recommended!

(tinyurl.com/mpw32bjc)

b. A totally cool movie! […] Cillian Murphy plays ***xaval al hazman***! And Rachel McAdams […] is simply an excellent actress.

(tinyurl.com/ycyvcnvv)

c. A ***xaval al hazman*** funny movie, […] a surprising, light and very funny movie. Strongly recommended.

(tinyurl.com/5ymk7m9n)

[1] The examples throughout the dissertation are naturally occurring Hebrew examples, most of which were extracted from the web. For the sake of clarity, in Examples (1)-(4) only the English translation is presented. The link to the source of each example is provided in parentheses.

d. It's worth going to see this movie […] I laughed ***xaval al hazman***, and my 17-year old daughter enjoyed it too.

(tinyurl.com/2t3jkvp5)

Xaval al hazman attracted quite a lot of attention from Hebrew speakers presumably because it radically changed its polarity, and maybe also because it is a full sentence which turned into a word. This is a peripheral grammatical change which is rare in Hebrew, as well as in other languages.

This dissertation examines this peripheral phenomenon — lexicalization of a full sentence into a word — by using (quantitative) data from written corpora. **My main research question is: Under what conditions will a full sentence undergo lexicalization into a full-fledged word?**

The theoretical framework I adopted in order to provide a parsimonious—yet exhaustive—model that would account for this phenomenon is *Construction Grammar*. Construction Grammar considers every linguistic element, on every possible level — a morpheme, a word, a phrase, a clause and a multi-clausal sentence — a *construction* (e.g., Croft, 2001; Fillmore, Kay, & O'Connor, 1988; Goldberg, 1995; Langacker, 1987, 1991; Sag, 2012). Each construction constitutes a form-meaning pairing. The linguistic knowledge of the speakers is assumed to be a network of constructions (Goldberg, 2003). Each construction is linked to other constructions via *links* of various types, thus forming a multidimensional network (e.g., Diessel, 2023; Schmid, 2020). These links connect contructions that share formal and/or semantic features. Just as important and relevnat to my research question is another assumption underlying this theoretical framework. All the constructions are ordered on a continuum that includes both lexical and syntactic elements, that is, the lexicon and the syntax are not considered distinct entities (Goldberg, 2006: 220). This implies that every construction is free to move from the sentential—and therefore the more complex—end to the simple, idiosyncratic end of this continuum, namely, to undergo lexicalization.

To understand the nature of the change undergone by *xaval al hazman*, I examine its locus in the multidimensional network (that constitutes the linguistic knowledge of the speakers) from two complementary points of view. I start by examining the linguistic change from the point of view of the object of the linguistic change (here, *xaval al hazman*), and then I examine this change from the point of view of the linguistic context that enables this change. I do all this while underscoring the interactions between the various constructions — the sentence undergoing change and the linguistic context that enables this change. (These interactions are part of the links that form the multidimensional network.)

In order to present a theoretically solid analysis I show that although lexicalization of full sentences is rare, *xaval al hazman* is not one-of-a-kind. There are other sentences in Hebrew that have undergone a similar lexicalization process. In fact, there's a "family" of such sentences. The members of this family, which I dub the *Ultimate construction family* (due to the extreme message that they convey), are linked via *inheritance links* to a more abstract construction from which they inherit their

properties. Formally, it is a sentential construction headed by a predicate and not by a subject; functionally, this construction is of an evaluative nature (Kuzar, 2012). This family of sentences includes, apart from *xaval al hazman*, also *en dvarim ka'ele/u*, literally, 'there are no such things', *en milim*, literally, 'there are no words' and *ba livkot*, literally, 'it feels like crying', just to mention a few other members.

I argue that these sentences — the members of the Ultimate construction family — are not fundamentally (conceptually) different from VPs. Their constituents are highly relevant to each other. They are relevant to a point that they constitute a single 'interpretatively cohesive' unit, just like VPs. Hence, quite like VPs, they too tend to undergo semantic change and become semantically opaque. Moreover, by conveying an evaluation, they are necessarily relational, that is, they are 'semantically *in*complete' (despite their syntactic completeness). This is why they must be associated with an element in the preceding discourse to evaluate. In other words, they inherently trigger an inferred link to a preceding sentence where they find a relevant stance-object they can modify. In addition, being semantically opaque (e.g., 'amazing/ly' or 'extremely', rather than the literal 'it's a pity on the (wasted) time') and therefore mono-morphemic, as well as absolute newcomers to the lexicon, they make up excellent candidates to modify elements which belong to different word classes. In other words, they make up potential *flexible modifiers*.

I further present a model that highlights the critical role of the context in this lexicalization process. The fact that the phenomenon here examined involves a full sentence that becomes a word points to the need for a model of clause linkage. Such a model describes the transition from a paratactic sequence to a hypotactic sequence, specifically, how two independent sentences come to be linked in such a way that one of them is a main clause which scopes over the other, a subordinate clause. A relevant model has been put forward by Lehmann (1988). But Lehmann's model focused on the nominalization of subordinate clauses, which does not necessarily involve semantic change. Moreover, although these subordinate clauses start out as full sentences, the resulting words constitute just a fraction of these (full) sentences. The model that I propose is different. It provides an explanation for the sematic change undergone by the full sentences here examined — a semantic change which is critical to the lexicalization process — as well as for the change in the grammatical status of these full sentences in their entirety.

The context that enables the semantic change of the members of the Ultimate construction family is, obviously, every context of a highly emotive nature. But the only contexts that meet this criterion and also enable the change in the grammatical status of the members of the Ultimate construction family are exclamative sentences. Not every type of exclamative sentences, but rather Anaphoric degree-adverb (i.e., *such* and *so*) exclamative sentences, see (3). These exclamative sentences invoke a Correlative endpoint resultant-state clause which compensates for the natural loss of the emotive nature of the degree-adverbs, see (4). This is the only strategy to do so, since *such* and *so* do not usually reduplicate (as does *very*, for example) in order to compensate for this loss.

(3) This is **such** a funny movie.

(https://tinyurl.com/59eu6uyy)

(4) This is **such** a funny movie, (up to a point) that ***xaval al hazman***.

The members of the Ultimate construction family occupy the slot of the Correlative endpoint resultant-state clause, because despite the semantic change, their sentential status did not change right away. This kind of link between a potential slot and the members of the Ultimate construction family is a *filler-slot link*. This is how a pragmatic motivation can explain the incorporation of the members of the Ultimate construction family into the preceding Anaphoric degree-adverb exclamative sentence.

In addition, positing these exclamative sentences as the enabling context for the change in grammatical status allows for a straightforward explanation for the status of the members of the Ultimate construction family as flexible modifiers. The exclamative sentences are cast in the form of a sentential construction headed by a subject and not by a predicate. Some exclamative sentences are nominal sentences where the focus of the sentence is a noun or an adjective. Others are verbal sentences where the focus is a verb. The invoked Correlative endpoint resultant-state clause modifies these foci, regardless of their word class. The grammatical role of the members of the Ultimate construction family is then determined by the modified element, whether a noun, an adjective or a verb. After the members of the Ultimate construction family are incorporated into the preceding sentence, the emotively bleached anaphoric degree-adverb (e.g., *such* in 3 above) may become redundant, which is why it is omittable. Once the relativizer *that* (see 4 above) is also omitted, the members of the Ultimate construction family are reanalyzed as direct modifiers of the focus of the exclamative sentence, as full-fledged words.

My analysis shows that the various members of the Ultimate construction family do not undergo the lexicalization process to the same extent. The model can account for this by reference to the hypothesized *horizontal links* with competing sentences on the local network of the members of the Ultimate construction family. I argue that it is the competition between the alternatives that limits the tendency of some members of the Ultimate construction family to incorporate into the Anaphoric degree-adverb exclamative sentences and undergo the change in grammatical status.

I also show that the specific context that triggers the change in grammatical status — the degree-adverb exclamative sentence — must dissolve (i.e., become optional rather than obligatory) in order for the members of the Ultimate construction family to undergo further constructional changes (such as derivation and inflection, when relevant). If this context does not dissolve, then no further constructional change will take place.

Apart from my attempt to answer my main research question — Under what conditions will a full sentence undergo lexicalization into a full-fledged word? — I had to tackle the lack of a large, spoken, diachronic and accessible corpus at the very beginning of my research. Such a corpus is vital for a research which deals with linguistic change. To handle this issue, at least with regard to the semantic change of

the members of the Ultimate construction family, I had to devise alternative methods to substantiate semantic change. **My secondary research question is therefore: How can one substantiate the presence of semantic change in the absence of a diachronic corpus?**

I show that metalinguistic activity of speakers, whether explicit or implicit, allows to substantiate the presence of semantic change. The methods proposed build on the cognitive aspect and the sociopragmatic aspect (Schmid, 2016 [2011]) of metalinguistic activity of speakers who sense semantic change.

The cognitive aspect builds on the working of three psycholinguistic theories in discourse: The Graded Salience Hypothesis (Giora, 1997, 2003), The Low-Salience Marking Hypothesis (Givoni, 2020; Givoni, Giora, & Bergerbest, 2013) and the Optimal Innovation Hypothesis (Giora et al., 2004). I show that speakers who are sensitive to semantic change mark it explicitly. The contents of this kind of marking testify to the direction of the semantic change, as well as to the relative salience of the various meanings. I also show that this relative salience can be supported by examining wordplay that speakers produce spontaneously.

The sociopragmatic aspect builds on the assumption that certain conservative speakers intentionally avoid the use of innovative meanings which they consider a threat to their social identity. Specifically, I examine the willingness of speakers to adopt neologisms. I compare the lexical choices of speakers from the Jewish ultra-orthodox community, known to be lexically conservative, to the lexical choices of speakers from the general population of Hebrew speakers. I show that the differential use of innovative meanings between the two communities attests to semantic change.

In sum, the goals of the two parts of this dissertation are different from one another, and so are the research questions. But they complement each other. The picture that emerges from the two taken together — the model that provides a natural account for the lexicalization process undergone by a full sentence, as well as speakers' sensitivity to the semantic *rather* than to the grammatical aspect of change — supports the claim of Construction Grammar that there is no division between syntax and lexicon. Indeed, the various constructions differ in terms of complexity, but they form part of the same level of representation.

Table of contents

Preface

Parts of this dissertation were published as articles or accepted for publication in peer reviewed journals:

Chapter 3 was adapted into an article and appeared as

> Becker, Israela. 2023. It's all about the sentential construction: Lexicalization of *complete* mono-clausal sentences into words – Evidence from Hebrew. *Studies in Language*, **47**(2), 463-504. https://doi.org/10.1075/sl.21006.bec

Chapter 4 was adapted into an article by the name "Scaffolding the sentential *Ultimate* construction into a word: A link-based analysis" authored by Becker, Israela and Ariel, Mira. It has been accepted for publication in *Constructions and Frames*.

Chapter 6 was adapted into an article and appeared as

> Becker, Israela. 2024. Let my speakers talk: Metalinguistic activity can indicate semantic change. *Corpus Linguistics and Linguistic Theory*, **20**(2), 289-319. https://doi.org/10.1515/cllt-2023-0022

Chapter 1: Introduction

In nova fert animus mutatas dicere formas corpora;
-- Ovid, *Metamorphoses*, book I, lines 1-2, 8 AD/1998

(Of bodies changed to other forms I tell)

1.1 The phenomenon

Around thirty years ago (in the mid-nineties of the twentieth century) *xaval al hazman* splashed into our language scene, the language scene of Hebrew speakers. The good old full sentence *xaval al hazman* 'it's a waste of time' (*xaval* 'it's a pity', *al* 'on', *hazman* 'the time'), as in Example (1.1), gave way to new *xaval al hazman*s, full-fledged grammatical words: an extreme positive adjective, equivalent to 'amazing' (1.2a); an extreme positive manner adverb, equivalent to 'amazingly' (1.2b); and an intensifier of adjectives and verbs, equivalent to 'extremely' or 'so much', (1.2c and 1.2d, respectively). Examples (1.1) and (1.2) differ by the semantics of *xaval al hazman* as well as by its grammatical statuses.

(1.1) I can't get it how anyone could recommend this movie. A depressing, slow and endlessly long movie. ***Xaval al hazman*** ['it a waste of time' – IB].[2]
(tinyurl.com/4wtctnc)

(1.2) a. This is a ***xaval al hazman*** ['amazing' – IB] movie. If you haven't watched it yet, you should do it right away!!!
(tinyurl.com/3rhkfr4h)

[2] The examples throughout the dissertation are naturally occurring Hebrew examples, most of which were extracted from the web, following the by-now well-established Web-as-Corpus approach (see Hundt, Nesselhauf, & Biewer, 2007). For the sake of clarity, in Examples (1.1) and (1.2) only the English translation is presented. Most of the remaining examples, however, contain both the Hebrew original alongside the English gloss and paraphrase. Different levels of glossing are used based on relevance to the subject matter. The link to the source of each example is provided in parentheses.

b. Listen, Charlize hasn't won an Academy Award for her part for nothing. She plays ***xaval al hazman*** ['amazingly' – IB] there.

(tinyurl.com/hdhr5w4f)

c. The movie is based – so I take it – on a real story but ***xaval al hazman*** ['extremely' – IB] terrifying.

(tinyurl.com/5cw29b3p)

d. It's worth watching this movie […]. I laughed ***xaval al hazman*** ['so much' – IB], and my 17-year-old daughter enjoyed it too.[3]

(tinyurl.com/2t3jkvp5)

This particular event of linguistic change attracted unusual attention by Hebrew speakers.[4] After all, it is not every day that a full sentence undergoes lexicalization — semantic change (here, a radical change of polarity) followed by change in grammatical status — to become a content word, whether in Hebrew or cross-linguistically.[5] This is quite a marginal phenomenon. But a marginal phenomenon can be a golden opportunity to take a second look at long-established linguistic theories through an exceptional prism, such that would allow to (re-)test those theories, refine them, or even extend them, if needed. And this is exactly what I do in this dissertation.

1.2 The issues raised in this dissertation and the perspectives taken to look at them (and solve them)

In this dissertation,

(i) I distill the preconditions that a full sentence should meet in order to become a (potential) candidate to undergo lexicalization (other than being of high frequency). In doing so, I examine the motivation for lexicalization — both the semantic change and the change in grammatical status — from the point of view of the full sentence undergoing lexicalization;

(ii) I further study the interaction between the full sentence undergoing lexicalization and the surrounding context which partakes in the lexicalization process. In doing so, I examine the change in grammatical status of the sentence undergoing lexicalization from the point of view of the surrounding context.

[3] Note that Example (1.1) — the *original* meaning of *xaval al hazman* — was produced in 2015, whereas Examples (1.2a-d) — the *new* meanings of *xaval al hazman* — were produced earlier, 2003, 2004, 2007 and 2010 respectively. This implies that the original meaning is still alive and kicking, although it is not the exclusive meaning anymore, and hardly the salient one, as will be shown in Chapter 6.

[4] See Chapter 6 for how this unusual attention on the part of non-linguist speakers attests to semantic change.

[5] Lexicalization is "a process by which new linguistic entities, be it simple or complex words or just new senses, become conventionalized on the level of the lexicon" (Blank, 2001: 1603).

(iii) Finally, I try to overcome the lack of a diachronic corpus (as was the case at the early stages of my research), required to substantiate a claim about semantic change, by introducing new methods for detecting semantic change based entirely on a synchronic corpus.

In terms of Weinreich, Labov and Herzog's (1968) influential paper "Empirical foundations for a theory of language change", what I do in (i) is provide an answer to the *constraint* problem which "[…] inquire[s] into the set of possible changes and possible conditions for changes which can take place in a structure of a given type" (p. 101); in (ii) I provide an answer to the *transition* problem which "[…] ask[s] about intervening stages which can be observed […] between any two forms of a language […]" (p. 101); in (iii) I present new methods for detecting semantic change, which rely on the notion of evaluation derived from the *evaluation* problem where the "[…] changes [can] be evaluated in terms of their effects upon linguistic structure, upon communicative efficiency (as related, e.g., to functional load), and on the wide range of nonrepresentational factors involved in speaking" (p. 101). The latter is, basically, "the level of social awareness of a linguistic change" (Labov, 2017: 263).

Following Schimd's (2016 [2011]) approach to describing the process of the establishment of complex lexemes, in (i) and (ii), I take a *structural* perspective which views "the internal structure of the word itself with regard to changes in its form, meaning and dependence on the immediate linguistic context" (p. 71). In (iii) I adopt both a *cognitive* perspective and a *sociopragmatic* perspective of semantic change. The cognitive perspective views "the word in the minds of the speakers with regard to its entrenchment in the individual mental lexicons of the speakers and the conceptual status it has achieved there" (p. 71). The sociopragmatic perspective views "the word in the speech community with regard to the extent of its spread and diffusion, i.e. the degree of use and familiarity among the members of the speech community" (p. 71). Figure 1.1 aligns each problem addressed in this dissertation with the respective perspective(s) taken to look at (in order to eventually solve the problem).

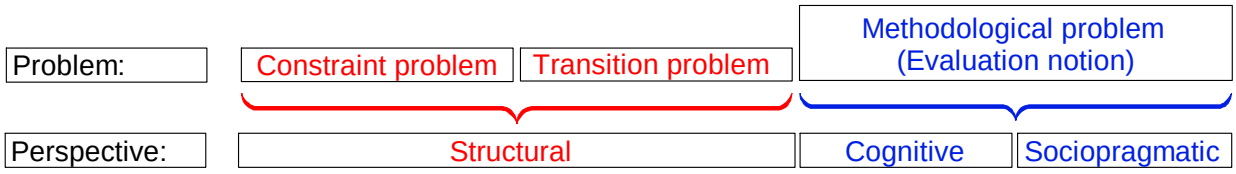

Figure 1.1: The problems addressed in this dissertation and the respective perspectives taken to look at them

This phenomenon of a full sentence becoming a word (or words) will be accounted for under the framework of Construction Grammar. I will show (as the dissertation unfolds) that it fits it like a glove. I therefore start by describing briefly what Construction Grammar is, and show that although originally a synchronic theory, it

perfectly suits the task of accounting for linguistic changes (the one here studied included).

1.3 Construction Grammar and linguistic change

1.3.1 Construction Grammar in brief

Construction Grammar theories (e.g., Croft, 2001; Fillmore, Kay, & O'Connor, 1988; Goldberg, 1995; Langacker, 1987, 1991; Sag, 2012) consider language an inventory of constructions called *Construct-i-con* (Goldberg, 2003). The constructions are coupled pairs of form and meaning/function, on every conceivable linguistic level — from morphemes, through words, phrases, clauses to multi-clausal sentences. The form and meaning/function pairings show varying degrees of schematicity/abstractness. A schematic/abstract (and therefore relatively productive) construction is, for example, the Ditransitive construction which points to an intentional transfer, as in *Mary gave John the book.* Here the specific lexical items, although constrained by the function of the construction, are relatively free. Non-schematic, idiosyncratic constructions are idioms such as *kick the bucket* as well as single words. Here the choice of lexical items is heavily restricted. Since meaning is mapped directly onto form, Construction Grammar theories are necessarily *non-derivational*.

The Construct-i-con itself is not just an inventory, but a structured inventory of interconnected nodes which constitute a hierarchical, taxonomic *network* (Croft, 2001; Goldberg, 1995, 2006; Langacker, 2008).[6] Each node represents a construction (of any degree of schematicity/abstractness) and constitutes a complex of semantic, pragmatic, syntactic, morphological and phonological features. This complex representation indicates that Construction Grammar theories are necessarily *non-modular*.

The links between nodes represent similarity relations between the constructions that the nodes represent, that is, the properties they have in common. The literature on constructional links focused mainly on *vertical*, inheritance links (e.g., Boas, 2013; Goldberg, 1995: Ch. 3; Hilpert, 2014: Ch. 3; Traugott & Trousdale, 2013: Ch. 2). An inheritance link between two constructions indicates that the dominated—necessarily more concrete—construction inherits some of its features from, at least, one dominating—necessarily more abstract and therefore more general—construction. Inheritance links are largely tied to form and differ from one another on the meaning relation between the dominating and the dominated nodes (see Hilpert, 2014: 60-65 for a summary). For example, a *polysemy link* connects the abstract Ditransitive construction which has the basic sense of 'X causes Y to receive Z', exemplified by, e.g., *Mary gave John the book*, and an extended sense 'X enables Y to receive Z', exemplified by, e.g., *The doctor allowed me a full meal* (Hilpert, 2014: 60).

The inheritance model has been expanded in the last decade or so, when non-vertical (non-inheritance) links were given proper consideration (e.g., Diessel, 2015, 2023; Perek, 2012; Sommerer, 2020; Van de Velde, 2014). *Horizontal* links "connect

[6] But see Bybee (2010).

constructions which show the same degree of abstractness and which are related to each other because they share similar formal and/or semantic features" (Sommerer, 2020: 92). For example, the abstract Ditransitive construction mentioned above (exemplified by *Mary gave John the book*) is horizontally linked to its paraphrase, the Transfer-caused-motion construction (exemplified by *Mary gave the book to John*), because they both share the basic meaning of 'X causes Y to receive Z' ('transfer of possession'; Perek, 2012). Other non-vertical links are *syntagmatic relations* which "connect symbolic units, for example words or phrases, that are frequently used together in sequential order" (Diessel, 2023: 16). For example, in the Ditransitive construction, the verb entails two arguments, an agent and a patient, typically encoded by NPs. Additioanl non-vertical links are *filler-slot relations* which "specify associations between the slots of constructional schemas and particular lexical or phrasal fillers" (Diessel, 2023: 16). For example, the verb slot in the Ditransitive construction usually hosts verbs, such as *give*, which convey some sense of transfer. Taken together, the various kinds of links (vertical as well as non-vertical) constitute a multidimensional network (Diessel, 2020; Schmid, 2020; Smirnova & Sommerer, 2020).

Many contructional grammarians maintain that "most information that is supposed to be stored in that network is stored in the nodes" (Hilpert 2018: 31). It should be noted, however, that some, such as Diessel (2020) grant greater weight to the links than to the nodes. Schmid (2020) even goes a step further to reject the notion of nodes and suggest that all the linguistic information is kept in the links ("an associative network", as he calls it).

1.3.2 Construction Grammar and diachrony

Now, Construction Grammars were designed for studying the synchronic aspects of language, but in the last three decades, constructional theories have also been applied to the study of linguistic diachrony. Barðdal and Gildea (2015) provide a detailed overview of Construction Grammar principles relevant to diachrony, the first of which, (presumably the most essential one) is the coupling of form and meaning/function which is no stranger to the phenomenon of linguistic change, for linguistic change of a linguistic unit involves either a change of form (phonological, morphological or syntactic), or a change of meaning (semantic or pragmatic), or both. But more than that, the assumption underlying Construction Grammar theories that all constructions lie along a continuum of schematicity/abstractness and there is no impenetrable wall separating sentences from single words (as endorsed by proponents of Syntax-and-Lexicon theories; Chomsky, 1965, 1970; Jackendoff, 1977), enables construction grammarians to provide (i) an account of the phenomenon of full sentences inserted into slots reserved for full-fledged lexemes (as Finkbeiner & Meibauer, 2016; Shirtz & Goldberg, submitted for publication, do), as well as (ii) an account of the lexicalization process that these full sentences may undergo. This is so, because each construction is free to "move" from the more schematic end of the continuum to the less schematic—more idiosyncratic—end (i.e., undergo lexicalization), or the other way around (i.e., undergo grammaticalization). And since under Construction Grammars, speakers'

"knowledge of a construction is the sum total of [their] experience with that construction" (Hilpert, 2014: 2), each meaning shift along each form shift, in the course of lexicalization (or grammaticalization), can be pragmatically rationalized.

The acknowledgement, then, that (any) linguistic change can be conceived as a change that a construction undergoes is rather straightforward, as agreed on by construction grammarian in the field of Diachronic Construction Grammar, e.g., Bergs and Diewald (2008), Bybee (2010), Coussé, Andersson, and Olofsson (2018), Fried (2009), Hilpert (2013), Sommerer, Gildea, Barðdal, and Smirnova (2015), Sommerer and Smirnova (2020) and Traugott and Trousdale (2013), to mention just a few.

1.3.3 What is 'constructional change'?

A definition of 'constructional change' was offered by Hilpert (2013):

> [C]onstructional change selectively seizes a conventionalized form-meaning pair of a language, altering it in terms of its form, its function, any aspect of its frequency, its distribution in the linguistic community, or any combination of these. (p. 16)

And in light of the undisputed dominance of the systematic study of grammaticalization processes (Brinton & Traugott, 2005: 2; Traugott & Trousdale, 2013: 32) in comparison to other processes of linguistic change, Hilpert (2013) also specified explicitly that

> [c]onstructional change is more encompassing than the changes that characterize grammaticalization. Specifically, it includes processes of lexicalization, processes of syntactic change that do not instantiate grammaticalization, processes within derivational morphology, and processes of frequency change that are unrelated to grammaticalization. (p. 8-9)

Hilpert further indicated that constructional change is not language change. Language change, which is a global change across many constructions, he argued, requires high-level generalizations. Such generalizations are highly abstract and as such have no concrete actualization. But actualization, i.e., actual utterances produced by actual speakers, is a must for a change to take place. Therefore, a constructional change is a local—rather than a global—language change. As such, it affects generalizations, but on a relatively low level of abstraction. (Note, however, that Barðdal and Gildea, 2015, as opposed to Hilpert, regard large scale language changes as constructional changes too.)

1.3.4 The 'construction' in 'constructional change'

It is noteworthy that a covert premise behind Hilpert's definition(s), as well as the studies listed at the end of Section 1.3.2, is that the 'construction' in a 'constructional change' is the linguistic unit which is the subject of change. But other linguists analyzing diachronic change (e.g., Bybee, 2003; Diewald, 2006; Traugott, 2003) maintain that the 'construction' in a 'constructional change' is not the linguistic unit

undergoing change, but rather the specific context in which the linguistic unit undergoing change is embedded. The 'construction' (i.e., specific context), in this sense, is just the trigger for linguistic change of a linguistic unit to take place, and it is actually intact.

The phenomenon analyzed in this dissertation complies with both views of 'constructional change'. The full sentences which are the subject of change are considered constructions, as is the context which enables this change (as will become clear in the next chapters). After all, there are "constructions all over" (as the name of the volume edited by Schönefeld, 2006 suggests).

1.3.5 'Constructional change' or maybe 'constructionalization'?

Traugott and Trousdale (2013) distinguished between 'constructional change' and 'constructionalization' (a term introduced initially by Noël, 2007). 'Constructional change', they argued, "is a change affecting one internal dimension of a construction. It does not involve the creation of a new node" (p. 26). 'Constructionalization', on the other hand, is "the creation of a $form_{new}$-$meaning_{new}$ pairing, in other words, the development of a new sign" (p. 22). Accordingly, constructional changes then "precede enable or 'feed' constructionalization" (p. 27). Traugott and Trousdale further argued that constructional changes preceding constructionalization are 'pre-constructionalization constructional changes', but constructional changes can also follow constructionalization. They dubbed the latter 'post-constructionalization constructional changes'.

The distinction between 'constructional change' and 'constructionalization' has triggered criticism on both the theoretical and empirical levels (e.g., Börjars, Vincent, & Walkden, 2015; Flach, 2020; Hilpert, 2018; Smirnova & Sommerer, 2020). Smirnova and Sommerer (2020) argue against this distinction focusing on three issues:

(i) The basic tenet of Construction Grammar is the form-meaning pairing which is what defines a construction. If only one of the two undergoes change ('constructional change' in terms of Traugott and Trousdale), then a new form-meaning pairing comes into being. A new—and necessarily different—form-meaning pairing implies a new—and necessarily different—construction, in fact, a new node in the network. As such, it *must* be the outcome of constructionalization. But then this change is taken (by Traugott and Trousdale) as 'constructional change'.

(ii) If constructionalization involves the creation of a new node, then it *must* be abrupt (no node → a node). The notion of abruptness is in conflict with the notion of gradualness inherent to linguistic change.

(iii) Constructionalization must "be accompanied by changes in degree of schematicity, productivity, and compositionality" (Traugott & Trousdale, 2013: 22). However, not all three parameters always apply to every construction on every level of the constructional hierarchy.

Börjars et al. (2015), Hilpert (2018) and Flach (2020) also criticized the (forced) distinction between 'constructional change' and 'constructionalization'. They pointed out that 'constructionalization' is a relative term that depends on the starting point that the linguist chose for the evolution of a given item. Therefore, "[t]he difference between constructionalization and constructional change is thus not a matter of different linguistic processes, but rather a distinction that lies in the eye of the beholder" (Hilpert, 2018: 29). Flach (2020) suggested that "constructionalization is useful if it refers to its point reading, while its process reading is subsumed under 'constructional emergence' " (p. 46). I side with Börjars et al., Hilpert, Flach, and Smirnova and Sommerer, and adhere to the 'constructional changes' terminology. This choice is coherent with the emphasis I put in this dissertation on in the links between the constructions, rather than on the nodes.

1.4 The constructional models adopted

In order to decide which constructional models are best suited to handle the linguistic change studied in this dissertation, one should consider the fact that not only the (immediate) context of change is made of sentences, but also the linguistic units undergoing change are (full) sentences. The adopted models should then be able to handle linguistic change at the sentence—rather than the word or phrase—level. One should also bear in mind that the models should take into account the discourse function of the changing linguistic unit (here, a full sentence) in relation to the embedding context, for linguistic change always results from the interaction between the two (e.g., Ariel, 2008: Ch. 5; Diewald, 2006; Fried, 2009).

As for the linguistic unit undergoing change, the two aspects — its discourse function along with its sentential form — are addressed by the constructional model introduced by Kuzar (2012). It is based on the distinction between *categorical* and *thetic* propositions (Chafe, 1974; Kuno, 1972; Kuroda, 1972; Lambrecht, 1994: Ch. 4; 2000; Sasse, 1987; and see Firbas, 1974 for a review of the Functional Sentence Perspective). The essentials of this model — those which are relevant to the present study — are described and exemplified in Chapter 3.

As for the embedding (sentential) context, the two aspects are addressed by Michaelis' (2001) and Michaelis and Lambrecht's (1996) comprehensive analysis of the Exclamative sentence construction, described and exemplified in Chapter 4.

Kuzar's model rests on Croft's (2001) Radical Construction Grammar. Michaelis' (2001) and Michaelis and Lambrecht's (1996) analyses are associated with Goldberg's (1995, 2006, among many others of her publications) Cognitive Construction Grammar. Both, however, assign equal importance to the semantic and formal aspects of constructions. As such, they faithfully serve my account of the relevant linguistic

change, which, quite like any linguistic change, assigns equal importance to semantics and form.[7] Moreover, Both adopt a thoroughly usage-based approach, and so do I.

1.5 A few more words about terminology

Before I set out to resolve the issues raised in Section 1.2, a few words about terminology are in order.

1.5.1 'Word' or 'lexeme'?

Throughout this dissertation, I use the term 'word' rather than 'lexeme'. A 'lexeme' includes "different forms of the same word" (Huddleston & Pullum, 2002: 27) which are of importance within a morphological context. This is not the case in the present study, where sentences are examined against intra-sentential lexical elements, and therefore 'word' is a better term.

But the term 'word' itself is not unproblematic. The numerous attempts to provide a typologically valid definition of a grammatical word seem to have failed (see Dixon & Aikhenvald, 2002; Haspelmath, 2011 for critical reviews). Many criteria were suggested, but none were found to be both necessary and sufficient, even within the same language. I will, nevertheless, adopt Dixon and Aikhenvald's suggested criteria for a grammatical word, which are not universal (as Dixon and Aikhenvald themselves explicitly note), yet seem to serve the purpose of the present study (and also seem to apply to Hebrew):

> A grammatical word consists of a number of grammatical elements which:
> (a) always occur together, rather than scattered through the clause (the criterion of cohesiveness);[8]
> (b) occur in a fixed order;
> (c) have a conventionalized coherence and meaning. [9] (p. 19)

By "conventionalized coherence and meaning" (criterion (c)), Dixon and Aikhenvald refer to the general propensity of words for non-compositionality.[10] This criterion has

[7] Langacker's (1987, 1991) Cognitive Grammar lacks syntax which is vital to my analysis. Fillmore et al.'s (1988) "Berkeley" Construction Grammar and Sag's (2012) Sign-based Construction Grammar are far too formal (for my purposes).

[8] See also, e.g., Booij's (2009) notion of 'no manipulation' and Ramat's (2016) notion of 'cohesion'.

[9] See also, e.g., Ramat's (2016) notion of 'opaqueness', essential for a word to be highly 'wordy', that is prototypical.

[10] Non-compositionality is a complex term, as problematized by Svensson (2008) with regard to fixed expressions. The dichotomy compositional/non-compositional, she argues, can be associated with any of the four following dichotomies or their combinations: Motivation/non-motivation, transparency/opacity, analyzability/unanalyzability and literal/figurative meaning. In this dissertation, non-compositionality is taken to be equivalent to opacity (and accordingly – compositionality to transparency), since it applies to both words and to *all* the full sentences undergoing lexicalization studied here, whereas the other dichotomies may apply to some—but not all—sentences.

been discussed by earlier authors (Harris, 2000; Kanerva, 1987; Zwicky & Pullum, 1983) as applying to languages typologically unrelated to one another.[11] The most non-compositional units, Bybee (1985) suggested, are lexical units (i.e., words) where several semantic elements are fused together to produce *mono-morphemic* units which bear no predictable meaning.

On the face of it, this criterion (non-compositionality) seems somewhat inappropriate for Hebrew, a Semitic language of rich and transparent morphology (both concatenative and non-concatenative; see Berman-Aronson, 1978: Ch. 3; Ravid, 1990). Apparently, only a small number of derivational devices — a vocalic tier and/or affixes — have a clear semantic profile (e.g., Nir, 1993: 26, 46; Rabin, 1985; Ravid, 1999; Shatil, 2006), and when they combine with a base or a root (i.e., a consonantal skeleton), the meaning of the resulting product is often not predictable.

1.5.2 'Sentence' or 'clause'?

According Haspelmath (2019), "[a] clause is a combination of a predicate (full verb or nonverbal predicate) and its arguments, plus modifiers" whereas "[a] sentence is a maximal clause, i.e. *a clause that is not part of another clause*" (emphasis mine). In this dissertation, I use the term 'sentence' quite often for two reasons: (i) the newly evolved words here studied have originated from independent matrix sentences, "maximal clauses" in Haspelmath's terms; (ii) the specific contextual scaffolding here argued for is *multi*-clausal, and therefore necessarily sentential. In fact, multi-clausality is a crucial factor in the lexicalization process here studied, as will be clarified in the next chapters.

Obviously, the term 'clause' is not irrelevant. In Chapters 3 and 4, it alternates with 'sentence' depending on the context. In Chapter 6, however, I use the more general term 'syntagma' rather than 'sentence' (or 'clause'). This is because the methods I propose for detecting semantic change in the absence of a diachronic corpus (to which Chapter 6 is dedicated) are not limited to the sentences (or clauses) that have undergone semantic change. They can be equally applied to phrases (and even words).

1.5.3 'Lexicalization' or maybe 'desentientialization'?

I have opted for the use of 'constructional change' *in lieu* of 'constructionalization' above. Note, however, that the specific case of constructional change here examined is lexicalization. I therefore use the term 'lexicalization' predominantly. I use the term 'constructional change' occasionally, only when its serves a specific goal better than 'lexicalization'.

But then, if we take another look at the title of this dissertation, we realize that 'lexicalization' could have been potentially replaced with 'desentientialization' a term originally introduced by Lehmann (1988), for the two — lexicalization and desentientialization — constitute two sides of the same coin. And still, I prefer 'lexicalization'. This is because, based on the vast literature recording very many cases of 'desentientialization' across languages, desentientialization is associated with the

[11] This intuition is referred to by Haspelmath (2011: 4) as a "(quite possibly correct) feeling".

specific loss of explicitly marked sentential features. This is not the case here, as will become clear in the following chapters. Moreover, 'desententialization' implies that sententiality is something that needs to be disposed of in order for a sentence to become a word. But I must disagree. The persistent sententiality of the sentences here studied is crucial for them to eventually become words. Hence, I opt for 'lexicalization' rather than 'desententialization'.

Now, just to make things clear, in the next section I briefly review a phenomenon which seems to be similar to the phenomenon studied in this dissertation but, in fact, is not.

1.6 Non-lexicalized full sentences embedded in slots reserved for single words

A seemingly related phenomenon is exemplified in (1.3), where a full sentence, a clause, or a phrase fills a slot reserved for a full-fledged word (representing many other examples in the literature).

(1.3) A ***God-is-dead*** approach

This phenomenon, known as "phrasal compounding" where the non-head constituent is an element which modifies the head, is attested in several languages – For Afrikaans see Botha (1981); for Japanese see Shibatani & Kageyama (1988); for Mandarin Chinese see Wiese (1996); for German see Meibauer (2007) and Trips and Kornfilt (2015); for English see Trips (2012), Trips and Kornfilt (2015), and Shirtz and Goldberg (submitted for publication); for Turkish and (possibly) Sakha see Trips and Kornfilt (2015); for Hebrew see a brief mention in Shirtz and Goldberg (submitted for publication); many additional references are listed in Bruening 2018: 12.

Phrasal compounds are *conscious ad-hoc* creations (Meibauer, 2007), where a quote or a quote-like sentence/clause/phrase is mobilized for special pragmatic effects such as conveying a witty, sarcastic message (Meibauer, 2007; Shirtz & Goldberg, submitted for publication). Only a miniscule minority, such as *"I have a dream" speech* (repeatedly mentioned in the relevant literature) are lexicalized. In contrast, the focus of this dissertation are lexicalized full sentences which have undergone a gradual, *unintended* lexicalization process. This shift from a full sentence into a syntactic word serves a different goal, which motivates its different nature.

1.7 Outline of this dissertation

Since the analysis I propose is a quantitative analysis, I start by describing, in Chapter 2, the sources of data used. Then, in Chapters 3, 4 and 5, I present my account of the phenomenon presented in Section 1.1 above. In Chapter 6, I propose alternative solutions to the challenging problem I confronted at the early stages of my research — lack of available diachronic data. I conclude with Chapter 7.

Chapter 2: Data

"Data! Data! Data!" he cried impatiently. "I can't make bricks without clay."
-- Arthur Conan Doyle, *The Adventure of the Copper Beeches*, 1892

The study of linguistic change requires available diachronic data, preferably spoken. However, diachronic corpora of spoken Hebrew are few, of small size, cover a limited period of time, and are of limited access. Having no choice, I resorted to written corpora. None of them meet all the requirements from an ideal corpus (spoken, large, diachronic, and accessible), but they complement each other in such a way that allowed me to propose a well-reasoned model of linguistic change, statistically supported by data.

Given that linguistic change is a phenomenon associated with speech, the data I have secured are the closest thing to spoken data. The web-based corpora I used represent semi-spoken language (see, e.g., Danet, 2001, and for Hebrew see Vaisman & Gonen, 2011: Ch. 1 & 2). In the journalistic corpora I used, which is of a rather formal nature (Rubinstein, 2019), I made special efforts to spot examples which seem to be reconstructions of real-life speech events (Culperer & Kytö, 2010), many of which are marked by quotation marks.

In what follows, I briefly describe each of the five corpora I used, which together cover a period of around 170 years, from the revival of Modern Hebrew up until the present day.

2.1 Historical Jewish Press corpus

Historical Jewish Press[12] is a diachronic corpus of Jewish journals published around the world from the mid-nineteenth century up to the mid-eighties of the twentieth century. Most of them, but not all, were published in Hebrew. Rubinstein (2019) estimated that the total number of tokens of the Hebrew section of this corpus is ~1.3 ×109. As a journalistic corpus it is edited/standardized by professional editors. This corpus is not tagged for part-of-speech or morphologically annotated.

[12] https://www.nli.org.il/he/discover/newspapers/jpress/about#whatis

2.2 Yedioth Ahronoth corpus

Since Historical Jewish Press does not include *Yedioth Ahronoth* 'latest news', which has been the most popular daily in Palestine and later on in Israel for many years,[13] I used the digitized archives of the latter.[14] The archives comprise of issues from 1935 up until now, which implies that it is a dynamic corpus. The corpus size is not available. As a journalistic corpus — just like Historical Jewish Press — it is edited/standardized by professional editors. But unlike Historical Jewish Press, Yedioth Ahronoth corpus contains the contents of a single daily only, and therefore the number of writers is limited. This corpus too is not tagged for part-of-speech or morphologically annotated.

The period covered by Yedioth Ahronoth partially overlaps the latest period covered by Historical Jewish Press (i.e., the mid-eighties of the twentieth century), but also fills the gap between the mid-eighties of the twentieth century and the beginning of the twenty-first century. The beginning of the twenty-first century marks the beginning of the Web-2.0 era where the web started to be fed by ordinary people's contributions (Blank & Reisdorf, 2012; O'Reilly & Battelle, 2009), here represented by *IsraBlog*, *Seret* 'movie' and *HeTenTen* corpora described below.

2.3 IsraBlog corpus

IsraBlog is a web-based diachronic corpus containing the contents of a web-site by the same name (www.israblog.co.il) which represents semi-spoken *non*-edited language, and which I scraped on July 2017.[15] It comprises of ~740×10^3 blogs, implying the same number of distinct bloggers (assuming one blog per person), each tagged for gender and age. The number of writers is considerably larger than the number of writers in Yedioth Ahronoth corpus. The earliest blogpost dates back to 2001. The corpus contains ~15×10^6 posts, altogether ~168×10^6 tokens (and ~1.5×10^6 types). This corpus, however, is not tagged for part-of-speech or morphologically annotated.

2.4 Seret corpus

Seret (the Hebrew word for 'movie') is also a web-based diachronic corpus containing the contents of a web-site by the same name (www.seret.co.il),[16] which I also scraped on July 2017. It comprises of movie reviews written both by professional critics and movie audience (alongside other kinds of information about movies and TV series). The earliest movie review found in Seret dates back to 1999. All movie reviews (~4.5×10^3) contain ~1.5×10^6 tokens (and ~1.1×10^5 types), written by ~8,000 distinct authors, each of which is tagged for gender and age. This corpus, just like the four previous ones, is not tagged for part-of-speech or morphologically annotated.

[13] https://www.jewishvirtuallibrary.org/the-israeli-press

[14] http://192.115.83.120/Olive/APA/Test/#panel=search

[15] Access to the raw data is available upon request.

[16] Access to the raw data is available upon request.

2.5 HeTenTen corpus

HeTenTen is a web-based corpus tagged for part-of-speech and morphologically annotated (unlike the four previous corpora).[17] It is presently the largest web-corpus of Modern Hebrew available, scraped around 2014. It contains ~1.0×10^9 tokens arranged within ~1.2×10^6 web documents. HeTenTen corpus, however, is neither tagged for the production date of each instance (and therefore cannot be considered diachronic), nor for speakers' age.

The time span covered by the various corpora alongside their characteristics are summarized in Figure 2.1.

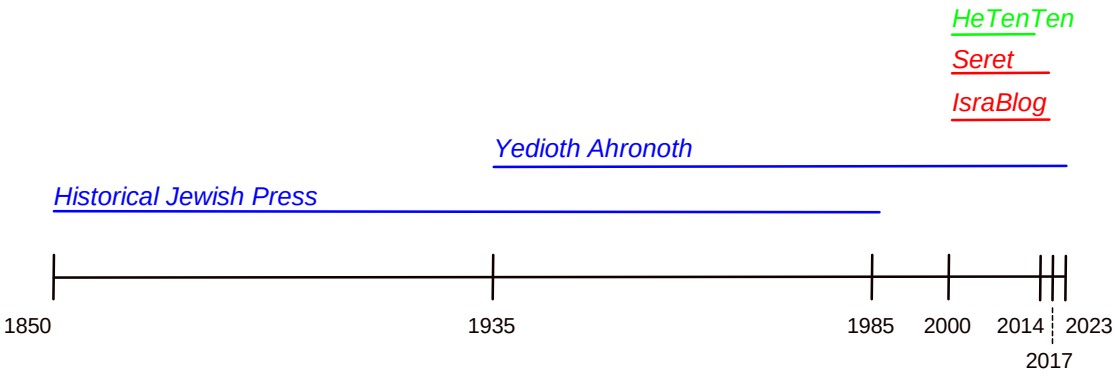

Figure 2.1: The time span covered by the corpora used in this dissertation: blue = formal language, tagged for date of production; red = semi-spoken language, tagged for date of production; green = semi-spoken language, not tagged for date of production (but tagged for part-of-speech and morphologically annotated)

2.6 Corpus selection

In the analyses presented in this dissertation, I gave precedence to IsraBlog, which is both semi-spoken (*non*-edited language), diverse in terms of number of writers, and tagged for date of production. When IsraBlog failed to produce data, for reasons which have to do with the time span it covers, I switched to the relatively formal (and edited) Yedioth Ahronoth and Historical Jewish Press, both tagged for date of production. When data tagged for part-of-speech was required, but date of production was not as important, I switched to the (partially) semi-spoken HeTenTen. The semi-spoken Seret was used just as a means to replicate results.

I considered *all* the query results from IsraBlog, Yedioth Ahronoth and HeTenTen (and Seret), ranging from a handful of results to several hundred (and sometimes even several thousand).

[17] http://www.sketchengine.co.uk

2.7 Rosenthal's comprehensive dictionary of Israeli slang

The phenomenon studied in this dissertation is a case of neologism which started out as slang. In case it needs to be examined against a corpus of other slangy neologisms, the contents of Rosenthal's (2005/2018) comprehensive dictionary of Israeli slang were used, for this dictionary is the most comprehensive and updated "corpus" of its kind to date (~10,000 entries). This "corpus" is not tagged for part-of-speech or morphologically annotated, and neither is it tagged for date of production.

2.8 Some additional experimental data

The lion's share of the data I used comes from corpora. However, the results of a small-N design experiment (reported in Section 3.3.2.2), as well as the results of two exploratory one-item tests (reported in Section 5.2.1.6), provide some complementary data.

Small-N design experiments usually focus on ten or fewer participants whose behavior (or outcome) is measured repeatedly and compared over time. No suggestion as to the participants' naivety is made (e.g., Smith & Little, 2018). In fact, in the experiment described in Section 3.3.2.2, participants were chosen due to their linguistic expertise. The same holds for the participants of the exploratory one-item tests described in Section 5.2.1.6.

Having presented the sources of data used in this dissertation, I am ready to tackle the three issues raised in Section 1.2 in turn.

Chapter 3: The *preconditions* required of a full sentence to become a word – An answer to the CONSTRAINT problem

Sometime, before they reached the crossroad, Sophie and Max met. We could have talked about how people meet and about what makes them wrap around each other. But if one stops and dwells on every single point, one cannot get anywhere. What's clear is that after they had met, they became one administrative entity, and after some time they reached the crossroad. A gate that they could not have seen, that everything would change after they go through, has opened before them.

-- Ilana Rudashevski, *Taska, 2022* [18]

Weinreich, Labov and Herzog's (1968) CONSTRAINT problem "[…] inquire[s] into the set of possible changes and possible conditions for changes which can take place in *a structure of a given type*" (p. 101; emphasis mine). In this chapter I attempt to solve the CONSTRAINT problem with respect to the linguistic change *xaval al hazman*, originally, 'it's a waste of time', has undergone. I also attempt to formulate a generalization about the preconditions which must be met for an independent full sentence to turn into an intra-sentential element, a word.

3.1 "Two are better than one […] and a threefold cord is not quickly broken"[19]

An attempt to account for the (rare) phenomenon presented in Section 1.1 — an independent full sentence, *xaval al hazman*, becoming a full-fledged word — will

[18] מתישהו, לפני שהגיעו אל פרשת הדרכים, סופי ומקס נפגשו. אפשר לדבר על איך אנשים נפגשים ועל מה שגורם להם להיכרך אחד בשני. אבל אם עוצרים ומתעכבים ככה בכל נקודה, לא מגיעים לשום מקום. מה שברור זה שאחרי שנפגשו, הם הפכו לישות מנהלית אחת, ואחרי איזה זמן הגיעו אל פרשת הדרכים. שער שלא יכלו לראות אותו, שהכול ישתנה אחרי שיעברו בו, הופיע מולם. (אילנה רוּדַשֶׁבְסקי, טַסקָה, 2022: 61)

[19] Ecclesiastes 4: 9, 12.

certainly benefit from the existence of similar cases. A set of cases, similar to one another in some respects, while at the same time different from one another in other respects, is essential to formulating robust generalizations (not just in linguistics, obviously). Luckily, *xaval al hazman* is not as unique in Hebrew as it seems. There's a family of similar Hebrew sentences which have undergone—or are on the brink of undergoing—the same linguistic changes *xaval al hazman* has undergone (to different degree, as I will show in Chapters 4 and 5), producing similar quadruplets as those exemplified in (1.2) above — positive adjectives, positive manner adverbs, and (neutral) intensifiers — all amplifying modifiers. These full sentences, which have attracted far less attention than *xaval al hazman*, are listed in Table 3.1 alongside their original compositional meaning. After undergoing the semantic change, they all point to an extreme state of affairs, an endpoint, accompanied by a strong speaker's stance.

#	Sentence	Literal, original meaning
1	*xaval al hazman*	'it's a waste of time'
2	*xaval al hamilim*	'it's a waste of words'
3	*ba livkot*/*lamut*	'it feels like crying/dying'
4	*efšar lehištage'a*/*lamut*[20]	'it's possible to go crazy/die'
5	*en ma lehagid*/*ledaber*	'there's nothing to say/speak'
6	*en milim*[21]	'there are no words'
7	*en dvarim ka'ele/u*	'there are no such things'

Table 3.1: Full Hebrew sentences which turned—or are on the brink of turning—into words, alongside their original compositional meaning

Once changed, all the sentences listed in Table 3.1 manifest lexicalization. Lexicalization is the process by which complex syntagmas change their meaning along with a change in grammatical status, forming new unanalyzable mono-morphemic content words. Cases of lexicalization are widely attested in many languages. Notably, however, their sources are predominantly phrases – NPs, PPs, AdjPs and VPs (as discussed and exemplified by, e.g., Bauer, 1983: Ch. 3; Blank, 2001; Brinton & Traugott, 2005: 48-49; Lipka, 1992: Ch. 3). See, for instance, early MidEng. *to morrow*]$_{PP}$ 'on the morning' > ModEng. *tomorrow*]$_{ADV}$ 'the day after the present day' (OED, 2019: s.v. tomorrow). Examples of full sentences serving as raw material from which new content words evolve are rarely provided in the literature (but see Section 3.6).

טוֹבִים הַשְּׁנַיִם מִן־הָאֶחָד אֲשֶׁר יֵשׁ־לָהֶם שָׂכָר טוֹב בַּעֲמָלָם: כִּי אִם־יִפֹּלוּ הָאֶחָד יָקִים אֶת־חֲבֵרוֹ וְאִילוֹ הֶאָחָד שֶׁיִּפּוֹל וְאֵין שֵׁנִי לַהֲקִימוֹ: גַּם אִם־יִשְׁכְּבוּ שְׁנַיִם וְחַם לָהֶם וּלְאֶחָד אֵיךְ יֵחָם: וְאִם־יִתְקְפוֹ הָאֶחָד הַשְּׁנַיִם יַעַמְדוּ נֶגְדּוֹ **וְהַחוּט הַמְשֻׁלָּשׁ לֹא בִמְהֵרָה יִנָּתֵק**: (קהלת ד', 12-9).

[20] Possibly from Yiddish: *מע קען משוגע ווערן [me(n) ken mešuge ver(ə)n]* 'one can go crazy'.

[21] Possibly from Russian: *Нет слов [najt slof]* 'no words'.

The prevalence of phrases over sentences as a source for new content words suggests that phrases make up better candidates than sentences for undergoing lexicalization. One explanation for this difference could be that phrases are, by and large, shorter than sentences. They are therefore usually more frequent (as well as more prosodically compact) and thus more amenable to lexicalization. But this raises the following interrelated questions:

> Why don't sentences as frequent and as prosodically compact as phrases, such as the Hebrew *ani lo xošev* 'I don't think so' (pronounced /anloxošev/), undergo lexicalization? What, if any, are the preconditions for an independent full sentence to undergo lexicalization? Why are there so few full sentences that undergo lexicalization, even sentences as short as phrasal syntagmas?

If I treat all the sentences in Table 3.1 together, although their semantic change is a result of different inferential steps and mechanisms (as described in Appendix A), I may provide an answer to these questions, using, of course, the appropriate constructional model. In the next section I describe exactly such a model.

3.2 The constructional model adopted

The fact that the syntagmas undergoing change (here, lexicalization) are sentences, determined my choice of the constructional model to follow. As indicated already in Chapter 1, the model pursued must be able to handle linguistic change on the sentence—rather than the word or phrase—level. In addition, the model must be able to take into account the discourse function of the changing linguistic unit(s) when embedded in context. After all, linguistic change always takes place in context (e.g., Ariel, 2008: Ch. 5; Diewald, 2006; Fried, 2009).

The constructional model that I adopt here is based on the distinction between two types of propositions (Chafe, 1974; Kuno, 1972; Kuroda, 1972; Lambrecht, 1994: Ch. 4; 2000; Sasse, 1987; and see Firbas, 1974 for a review of the Functional Sentence Perspective of the Prague School of Linguistics):

- *A categorical* proposition[22] — an informationally bi-partite structure — consisting of a focal portion (also known as *rheme* or *comment*) and a non-focal portion (also known as *theme* or *topic*).
- *A thetic* proposition[23] — an informationally mono-partite structure — where the entire proposition functions as a unit which constitutes the focal information.

[22] A categorical proposition has also been referred to as 'thematic sentence' (Kuno, 1972) and 'predicate-focus construction' (Lambrecht, 1994, 2000).

[23] A thetic proposition has also been referred to as 'neutral description' (Kuno, 1972), 'all-new sentence' (Allerton & Cruttenden, 1979) and 'sentence-focus construction' (Lambrecht, 1994).

Each of these proposition types is associated with a specific sentential construction presented in detail in the next section.

3.2.1 The constructional (structural and functional) properties of categorical and thetic propositions in Hebrew

All linguists (listed in the previous section, and many others) who distinguish categorical and thetic propositions associate the discourse function of each with its distinct structural properties. This is also the case in Hebrew, where the two proposition types are distinguished by the linear order of their components. A Subject-Initial sentential construction encodes a categorical proposition, whereas a Predicate-Initial sentential construction, encodes a thetic proposition (Kuzar, 1989, 1992a, 2002, 2006a, 2006b, 2012; Melnik, 2002, 2006).

I argue that the lexicalization process studied here is a function of the type of proposition, that is, the linear order of the elements the proposition is composed of and their nature, as well as its discourse function. I therefore present the detailed formulae of the sentential constructions encoding categorical and thetic propositions (see Figures 3.1-3.2 and 3.3 below, respectively), as proposed by Kuzar (2012: 55, 59, 94, 104) and termed *S(entence)-patterns*. Each S-pattern consists of several consecutive slots. Each slot is a multivariate slot which can accommodate different parts of speech, irrespective of their lexical category (see also Izre'el, 2012). Some slots are obligatory whereas others are optional.

Figure 3.1 presents the *Verbal Subject-Initial S-pattern* (henceforth, Verbal S1 S-pattern) associated with the categorical proposition in Hebrew. The top row represents the categorial affiliation of the words that can fill each slot; the second row specifies their sentential functional roles; and the remaining rows display the application of this construction to Examples (3.1) and (3.2) below. The obligatory slots in this specific sentential construction are those of the verb and the preverbal NP, marked here as 'subject'. The other slots are optional and therefore shadowed.

NP	[$_{vp}$	V	NP/PP	NP/PP	]
Subject		Predicate	Object$_1$	Object$_2$	
hi		*katva*			
hi		*katva*	*mixtav*	*li-vna*	

Figure 3.1: The Verbal Subject-Initial S-pattern associated with the categorical proposition in Hebrew

(3.1) *hi* *katva.*
she wrote[24]

'She wrote.'

(tinyurl.com/2beudvx7)

(3.2) *hi* *katva* *mixtav* *li-vna.*
she wrote a.letter to-her.son

'She wrote a letter to her son.'

(tinyurl.com/2zhy2czj)

Kuzar (2012: 56) suggested that in a narrative context, "[t]he Verbal S1 S-pattern hosts all actions and events constituting the storyline of the narrative", or "[t]he foregrounded actions of a storyline" (p. 191).

Figure 3.2 presents the *Copular Subject-Initial S-pattern* (henceforth, Copular S1 S-pattern) also associated with the categorical proposition in Hebrew. The top row represents the categorial affiliation of the lexical units that can fill each slot, the second row specifies their sentential functional roles, and the remaining rows display the application of this construction to Examples (3.3)-(3.5) below. The obligatory slots in this specific sentential construction are those of the pre-copular NP, marked here as 'basis subject', and the 'assigned term'. The other slots are optional and therefore shadowed.

NP	COP	NP/PP	P/PNP[25]	NP/AP
Basis Subject	Assigner	Affectee	Relation	Assigned Term
exad me-hem	*hu*			*ben*
hu		*bišvili*	*kmo*	*ben*
hu				*nifla*

Figure 3.2: The Copular Subject-Initial S-pattern associated with the categorical proposition in Hebrew

[24] Throughout the dissertation, different levels of glossing are used based on relevance to the subject matter.

[25] PNP stands for a composite preposition such as *al pney* 'on the face of', *be'emca'ut* 'by means of', *al saf* 'on the verge of' (Kuzar, 2012: 59).

(3.3) *exad* *me-hem* *hu* *ben.*
one of-them is a.boy

'One of the them is a boy.'

(tinyurl.com/4vwbat2m)

(3.4) *hu* *bišvili* *kmo* *ben.*
he for.me like a.son

'He is like a son to me'.'

(tinyurl.com/y8m84n5k)

(3.5) *hu* *nifla.*
he wonderful

'He is wonderful.'

(tinyurl.com/23k5xxvb)

Kuzar (2012) suggested that in a narrative context,

> [t]he copular sentence expresses a state in which an Assigned Term is assigned to a Basis Subject as being its equivalent, its substitute, or in a relation with it. […] [T]he Copular S-pattern, like its English counterpart, does not have a directly designated narrative function, but based on its constructional function, namely the expression of states, it is often used to supply the background and the circumstances pertaining to the storyline. (p. 59-60)

Figure 3.3 presents the *Predicate-Initial S-patterns* (henceforth, P1 S-pattern) associated with the thetic proposition in Hebrew: (a) is the S-pattern relevant to Example (3.6) below, a possessive proposition based on an existential proposition which incorporated a dative possessor (Hebrew is a *non-habere* language; and see Section 3.2.5 below); and (b) is the S-pattern relevant to Example (3.7) below which is an evaluative proposition. The top row in each of (a) and (b) represents the categorial affiliation of the words that can fill each slot; the middle row specifies their sentential functional roles; and the third row displays the application of these constructions to Examples (3.6) and (3.7), respectively. The obligatory slots in these specific S-patterns are those of the predicate and post-predicate NP or nominalization, marked here as 'existent' or 'evaluee', respectively. The remaining slot is optional and therefore shadowed.

Crucially, in the P1 S-pattern no syntactic role is allocated for a subject, since the components that Kuzar names 'existent' and 'evaluee' do not maintain the properties of 'real' subjects. I will consider this point in more detail as the chapter unfolds.

a

[PredP	Unq//V/A/P	NP/PP	] NP/Nominalization
	Existential Predicate	Object/Possessor	Existent/Possessed
	yeš	*li*	*mixtav*

b

[PredP	Unq/V/N/A/P	NP/PP	] NP/Nominalization
	Evaluative Predicate	Object/Affectee	Evaluee
	kef	*li*	*li-kro* *et hamilim* *ha-nehedarot* *ha-ele*

Figure 3.3: The Predicate-Initial S-patterns associated with the thetic proposition in Hebrew: (a) the existential/possessive Predicate-Initial S-pattern; (b) the evaluative Predicate-Initial S-pattern[26]

(3.6) *yeš* *li* *mixtav*.
there.is to.me a.letter

'I have a letter.'

(tinyurl.com/5n8b47yv)

(3.7) *kef* *li* *likro* *et* *ha-milim* *ha-nehedarot*
it's.a.pleasure to.me to.read ACC the-words the-wonderful

he-ele.
the-these

'It's a pleasure for me to read these wonderful words.'

(tinyurl.com/yddzr3s4)

Kuzar (2012: 68) suggested that "[t]hrough the use of the existential predicate, the Existential S-pattern announces the existence of the existent." He further suggested that "[t]hrough the use of the evaluative predicate, the Evaluative S-pattern evaluates the evaluee" (p. 80).

[26] 'Unq' stands for a morphologically unique predicate.

The sentences studied here (see Table 3.1) are *all* instantiations of thetic propositions, all encoded in the P1 S-pattern.

In the next sections, 3.2.2-3.2.5, I present each of the predicates of the sentences studied here in greater detail — *xaval* 'it's too bad/a pity', *efšar* 'it's possible', *ba* 'it feels like', and *en* 'there is/are no'. The actual sentences studied here are exemplified in Section 3.2.6, and will be analyzed in detail in Section 3.3 and onwards.

3.2.2 *xaval* 'it's too bad/a pity'

Xaval is a morphologically unique predicate of an attitudinal nature, meaning 'it's too bad/a pity'. It is endemic to the evaluative thetic proposition (see Figure 3.3b). *Xaval* 'it's too bad/a pity' can be followed by an optional dative participant, either a pronoun or a lexical constituent (e.g., *li* 'to me', *lakahal* 'to the audience', respectively). The optional dative participant is followed by a nominalization, either an infinitival phrase as in Example (3.8) or a 'that'-clause as in Example (3.9).

(3.8)

xaval	*(li)*	*lehafsid*	*otax*.
it's.too.bad	(to.me)	to.lose	you

'It's too bad (for me) to lose you.'

(tinyurl.com/2p89d4u4)

(3.9)

xaval	*(li)*	*še-hifsadeti*	*gam*	*otax*.
it's.too.bad	(to.me)	that-I.lost	too	you

'It's too bad (for me) that I lost you too.'

(tinyurl.com/yckr45pd)

Xaval 'it's too bad/a pity' has a variant in which the non-dative constituent following *xaval* is a PP with the preposition *al* 'on', as in Example (3.10).

(3.10)

xaval	*(li)*	*al*	***ha-hefsed***	***b-a-xaci^***	***ha-gmar.***[27]
it's.too.bad	(to.me)	on	**the-loss**	**in-the-half**	**the-final**

'**The loss in the semi-final** is too bad (for me).'

(tinyurl.com/5ebbyy6r)

In Example (3.10), the NP (marked in bold) within the PP is the evaluee (in Kuzar's terms, see Figure 3.3b). The entire sentence implies that the loss in the semi-final is an unfortunate and a distressing event.

This specific variant of the *xaval* 'it's too bad/a pity' sentence in Example (3.10) has a sub-construction in which the NP within the PP is a 'waste'-related commodity —

[27] The notation ^ stands for a construct state.

time, money, or effort. In this specific sub-construction, 'waste' is strongly implied. It is therefore often omitted, while leaving the remaining noun, e.g., 'time', 'money' or 'effort', as the sole NP inside the PP. This is a case of metonymy, where the NP represents an entire event of wasting a certain commodity (Kuzar, 2012: 107). In such cases, *xaval al NP* means 'it's too bad the waste of an NP', as in Example (3.11).

(3.11)	*xaval*	*(li)*	*al*	*ha-zman/kesef/ma'amac.*
	it's.too.bad	(to.me)	on	the-time/money/effort

'it's a waste of time/money/effort (for me).'

3.2.3 *ba* 'it feels like'

Ba is the present tense, 3rd person, singular, masculine form of the verbal root √*b.a* 'come'. When inserted into the P1 S-pattern, *ba* is coerced by the construction and interpreted as 'it feels like', which expresses volition. [28] It is then attitudinal.

Ba 'it feels like' too can be followed by an optional dative participant.[29] The element following the dative participant is either an infinitival phrase, see Example (3.12) or a 'that'-clause, see Example (3.13). *Ba* followed by an NP is also possible, see Example (3.14). Just as in the *xaval* 'waste'-related variant (see Example 3.11 above), here too metonymy is involved, where the NP represents an entire event associated with the NP. In the case of Example (3.14), it is the act of eating.

(3.12)	*ba*	*(li)*	*le'exol*	*mašehu*	*ta'im*	*u-mefanek.*
	it.feels.like	(to.me)	to.eat	something	delicious	and-indulgent

'I feel like eating something delicious and indulgent.'
(tinyurl.com/ynrr7jau)

(3.13)	*ba*	*(li)*	*še-noxal*	*yaxad*	*soufflé.*
	it.feels.like	(to.me)	that-we.will.eat	together	a.soufflé

'I would like us to eat a soufflé together.'
(tinyurl.com/3y54zwna)

[28] Note that when embedded in a Verbal S1 S-pattern (encoding a categorical proposition), this verbal root, √*b.a*, bears the meaning of 'come', as in *I come home at 8pm* (tinyurl.com/ytsf8jc4). The case of √*b.a* is then a case of polysemy coerced by different S-patterns, as in other cases mentioned by Kuzar (2012: Ch. 3). For example, Hebrew *car* is interpreted as 'narrow' when embedded in the Copular S1 S-pattern, but as 'distressful/sorry' when embedded in the Evaluative P1 S-pattern.

[29] *Ba* 'it feels like' is more commonly followed by a dative participant, but *ba* without a dative participant does occur.

(3.14) *ba* *(li)* *soufflé^* *šokolad.*
it.feels.like (to.me) a.soufflé chocolate

'I feel like (having) a chocolate soufflé.'

(tinyurl.com/59fh2hk2)

3.2.4 *efšar* 'it's possible'

Efšar, like *xaval* 'it's too bad/a pity', is a morphologically unique predicate endemic to the evaluative thetic proposition. It is of an epistemic nature and interpreted as 'it's possible'. Unlike *xaval* 'it's too bad/a pity' and *ba* 'it feels like', it is never followed by a dative participant. But similar to *xaval* 'it's too bad/a pity' and *ba* 'it feels like', the nominalization that follows *efšar* 'it's possible', is either an infinitival phrase, see Example (3.15), or a 'that'-clause, see Example (3.16). Note that the comparable *xaval* 'it's too bad/a pity' sentences are those exemplified in (3.8) and (3.9) above, not the 'waste'-related variant in Example (3.11). The comparable *ba* 'it feels like' sentences are (3.12) and (3.13).

(3.15) *efšar* *lištot* *tmisat^* *cianid* *ve-lehiša'er* *b-a-xa'im.*
it's.possible to.drink solution Cyanide and-to.stay in-the-life

'It's possible to drink a Cyanide solution and stay alive'

(tinyurl.com/y95ywnpn)

(3.16) *tarimi* *telephon.* *efšar* *še-nište* *te.*
lift phone it's.possible that-we.will.drink tea

'Give me a call. We might have tea (together).'

(tinyurl.com/mzb3msuz)

Efšar 'it's possible' followed by an NP is also possible (see Example 3.17 below). This is also a case of metonymy, where the NP represents an entire event, as in the *xaval* 'it's too bad/a pity' 'waste'-related variant in Example (3.11) and the *ba* 'it feels like' in Example (3.14) above. Here the typical event would be the speaker's offer to provide the addressee with a certain item, such as *te* 'tea' in Example (3.17).

(3.17) *efšar* *te* *xam.*
it's.possible tea hot

'It's possible (to have) hot tea.'

(tinyurl.com/zvdph5t2)

3.2.5 *en* 'there's/are no'

En too is a morphologically unique predicate. It is interpreted as 'there's/are no', and endemic to the existential—rather than the evaluative—thetic proposition (see Example 3.18 below). It is the counterpart of the morphologically unique predicate *yeš* 'there is/are', also endemic to the existential thetic proposition. *En* 'there's/are no' can be followed by an optional dative participant, thus expressing negative possession rather than existence (see Example 3.19). Last comes the obligatory existent (in Kuzar's terms, see Figure 3.3a) which could be either an NP or a nominalization. In Examples (3.18) and (3.19) the existent is an NP (*giša* 'access').

(3.18) *en* *giša* *l-a-internet.*
there's.no access to-the-internet

'There's no access to the internet.'

(tinyurl.com/2p82zu2b)

(3.19) *en* *li* *giša* *l-a-internet.*
there's.no to.me access to-the-internet

'I have no access to the internet.'

(tinyurl.com/2p8kpv8b)

When the existent is a nominalization, *en* 'there's/are no' takes only an infinitival phrase (and not a 'that'-clause). This variant, however, is irrelevant to my study and therefore will not be presented.

3.2.6 The actual sentences studied here

The actual sentences studied here are summarized in (3.20)-(3.23) along with their aligned glosses.

(3.20) ***xaval***$_{PREDCATE}$ $\varnothing_{AFFECTEE}$ ***[al ha-zman / ha-milim]***$_{EVALUEE}$
it's.too.bad on (the waste of) the-time / the-words

'It's a waste of time/words.'

(3.21) ***ba***$_{PREDCATE}$ $\varnothing_{AFFECTEE}$ ***[livkot / lamut]***$_{EVALUEE}$
it.feels.like to.cry / to.die

'It feels like crying/dying.'

(3.22) ***efšar***$_{PREDCATE}$ ***[lehištage'a / lamut]***$_{EVALUEE}$
it's.possible to.go.crazy / to.die

'It's possible to go crazy/die.'

(3.23) ***en***$_{PREDCATE}$ ∅$_{AFFECTEE}$ ***[milim / dvarim ka'ele / ma lehagid / ma ledaber]***$_{EXISTENT}$
there's/are.no words / things such.as.these / what to.say / what to.speak

'There is/are no words/no such things/nothing to say/nothing to speak.'

Note that all (3.20)-(3.23) lack a dative participant, a fact that will prove essential for the lexicalization process, as I will suggest later in this chapter.

Now, in terms of Goldberg's (1995) classical hierarchical network model, the Construct-i-con (see Section 1.3.1), and as sketched in Figure 3.4 below, the actual sentences studied here (3.20)-(3.23) are instantiations (see 'instance links'[30] in Figure 3.4 below) of the more abstract constructions presented in Sections 3.2.2-3.2.5 above. The latter inherit their properties from the even more abstract Evaluative P1 S-pattern (Figure 3.3b) or the Existential P1 S-pattern (Figure 3.3a) via 'polysemy links'.[31] The Evaluative P1 S-pattern and the Existential P1 S-pattern, in turn, inherit their properties from the even more abstract P1 S-pattern also via 'polysemy links'. (Admittedly, this sketch is partial. But in the context of the proposed constructional account, there is no need for a more detailed description.)

Indeed, the meaning of existential propositions seems to differ from that of evaluative propositions, as reflected in Figure 3.4. Kuzar (2012: 110-111), however, considers the two of them members of the same category, a *composite* category. Some members of the composite category, such as the *en* 'there's/are no' sentences studied here (see (3.23) above), combine an existential with an evaluative meaning, which motivates my treatment of all the sentences studied here as evaluative.

Since all the sentences studied here are instantiations of the same sentential construction, and after undergoing semantic change, they also share the very same function (all point to an extreme state of affairs, an endpoint, accompanied by a strong speaker's stance; see Section 1.1), I refer to them from this point on as the *Ultimate construction family*.[32]

[30] "[An] 'instance link' […] shows that a construction is a special case of another construction in the sense that it is a more fully specified version of the other construction" (Boas, 2013: 184).

[31] " [A] 'polysemy link' […] represents relations between subtypes of constructions that exhibit the same syntactic specifications but differ in their semantics" (Boas, 2013: 184).

[32] I thank John W. Du Bois for suggesting this neat label.

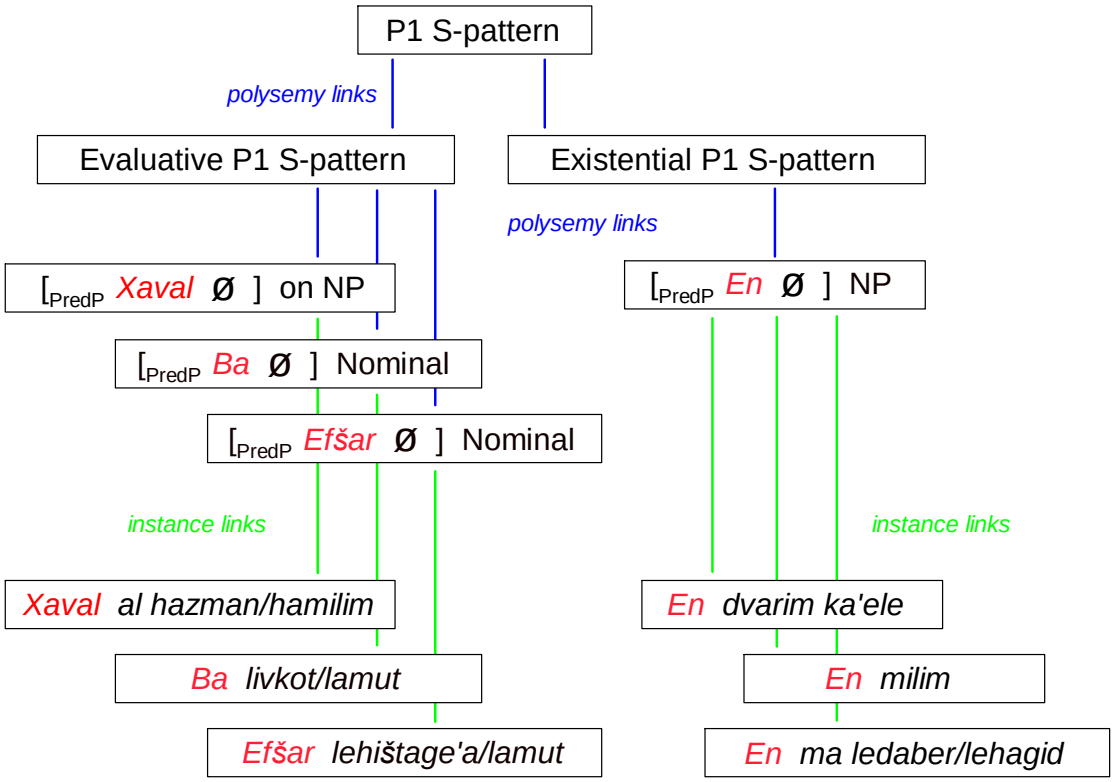

Figure 3.4: A segment of the Construct-i-con relevant to the P1 S-pattern in the context of the present study

In the next sections, I propose the preconditions that the members of the Ultimate construction family must meet in order to become words. I start with the precondition which drives the semantic change. Specifically, I argue for the resemblance of the *dativeless* thetic (3.20)-(3.23) to VPs. VPs are mono-partite by nature, and as such constitute 'interpretatively cohesive' units which are prone to undergoing semantic change to become semantically opaque. If the dativeless (3.20)-(3.23) are also mono-partite, just like VPs, and therefore 'interpretatively cohesive' (as I will show), there's no reason why they would not undergo semantic change as well.

3.3 'Interpretative cohesion' enabled by theticity drives the semantic change

3.3.1 Phrases are interpretatively cohesive units and therefore tend towards semantic opacity

Pioneering studies by Fodor and Bever (1965) and Johnson (1965) demonstrated that syntactic phrases of all kinds (i.e., NPs, PPs, AdjPs, and VPs), have a psychological reality: Speakers are sensitive to phrase boundaries; phrases form processing units made of smaller units, i.e., words; the words in the phrase are more relevant to each other than to other words outside the phrase.

The evidence adduced to support this argument focused on the *a*symmetry between Verb-Object combinations and Subject-Verb combinations, in terms of semantic

opacity. This interpretative *a*symmetry, Marantz (1984) suggested, is reflected in the tendency of Verb-Object combinations to have a "slightly or highly unusual semantics from what one would expect from the canonical uses of the verb" (p. 27). This interpretative *a*symmetry, Marantz (1984: 27ff.) further suggested, can even account for the tendency of Verb-Object combinations to idiomatize (viz., become semantically opaque) more easily than Subject-Verb combinations. Similar arguments were made by Tomlin (1986: Ch. 4) and O'Grady (1998). The former added evidence supporting this claim from a variety of typologically-diverse languages.[33]

These data indicate that the VP is an *interpretatively cohesive* unit, which naturally motivates its higher tendency to become opaque. In the next section I show that this very interpretative cohesion is also manifested in thetic propositions encoded in P1 S-pattern lacking a dative participant. (In Section 3.4 I explain why the semantic change may benefit from the lack of the dative participant.)

3.3.2 Dativeless thetic propositions are interpretatively cohesive units and therefore lend themselves to semantic opacity

In the following sections, I explain the nature of the semantic bonding between the predicates *xaval* 'it's too bad/a pity', *efšar* 'it's possible', *ba* 'it feels like', and *en* 'there's/are no', and the respective non-dative elements they combine with. Each of these predicates, I argue, forms an interpretatively cohesive unit with its respective adjacent non-dative element, which, much like VPs, can turn semantically opaque. In Section 3.3.2.1 I address *en* 'there's/are no' (3.23) and *xaval* 'it's too bad/a pity' (3.20), and in Section 3.3.2.2 I address *ba* 'it feels like' (3.21), and *efšar* 'it's possible' (3.22).

3.3.2.1 The predicate and the NP in dativeless thetic propositions form an interpretatively cohesive unit

My claim that the elements of (3.23) — $en_{PREDICATE}$ $[milim/dvarim\ ka'ele\ /\ ma\ lehagid/ledaber]_{EVALUEE}$ 'There are no words/such things' / 'There is nothing to say/to speak' — constitute an interpretatively cohesive unit, is based on Lambrecht (2000). Lambrecht examined constructions encoding thetic propositions versus constructions encoding categorical propositions across many languages, and found that thetic propositions consistently cancel the morpho-syntactic (and prosodic) properties associated with the post-predicative NP. Although this NP is traditionally considered the "NP subject", it is quite unlike the subject of categorical propositions in that it is seldom agentive, and/or topical. Lambrecht further showed that de-topicalization is not only manifested through cancelling of (topical) subject-marking features, but (often) also by marking the non-topical "NP subject" of the thetic propositions with formal features associated with direct objects of categorical propositions.

In Hebrew too, the "NP subject" of thetic propositions does not quite pattern with prototypical subjects. Kuzar (2002) and Melnik (2006) showed that the "NP subject" manifests many direct object properties: It assumes a post-verbal position in a language

[33] Semantic opacity is not VP-unique. Titone and Connine (1994) listed syntagmas other than VPs which are semantically opaque, all of which are phrases — NPs, PPs, AdjPs.

(i.e., Hebrew), in which the unmarked subject position is pre-verbal. There's often no morphological agreement between this "NP subject" and the preceding predicate. Finally, this "NP subject", when definite, is often case-marked as a direct object. It's not surprising, then, that Kuzar (2012) chose to label this "NP subject" an 'existent' (see Figure 3.3a), for it is not a "true" subject.

Is this also the case with the specific "NP subjects" in (3.23) — *milim/dvarim ka'ele / ma lehagid/ledaber* 'words/such things / what to say/speak'? Indeed, data from HeTenTen corpus (Kilgarriff et al., 2014; see Section 2.5) provide an affirmative answer.

I extracted all sentences in the form of both (3.27a) and (3.27b) from HeTenTen corpus. (3.27) are modified versions of (3.26). In (3.27), the indefinite "NP subjects" of (3.26) — *milim/dvarim* — are marked as definite with the accusative marker *et*. (3.27a) is a dativeless sentence, whereas in (3.27b) the dative participant is present. Note that I considered (3.27b), because according to the current literature, the status of these "NP subjects" — *milim/dvarim ka'ele* — as direct objects is not expected to be affected by the presence of a dative participant.

(3.26) a. ***en***$_{PREDICATE}$ ∅$_{AFFECTEE}$ ***[milim/dvarim ka'ele]***$_{EXISTENT}$
there's/are.no words/things such.as.these

'There are no words/such things.'

b. ***en***$_{PREDICATE}$ *li*$_{AFFECTEE}$ ***[milim/dvarim ka'ele]***$_{EXISTENT}$
there's/are.no to.me words/things such.as.these

'I have no words/such things.'

(3.27) a. ***en***$_{PREDICATE}$ ∅$_{AFFECTEE}$ ***et*** ***[ha-milim/ha-dvarim]***$_{EXISTENT}$
there's/are.no ACC the-words/the-things

'There aren't the words/the things.'

b. ***en***$_{PREDICATE}$ *li*$_{AFFECTEE}$ ***et*** ***[ha-milim/ha-dvarim]***$_{EXISTENT}$
there's/are.no to.me ACC the-words/the-things

'I don't have the words/the things.'

The query produced 14 instances of accusative-marked *hamilim* 'the words' as in (3.27a) and 22 instances of accusative-marked *hadvarim* 'the things' in (3.27a). It further produced 26 instances of accusative-marked *hamilim* 'the words' in (3.27b) and 10 instances of accusative-marked *hadvarim* 'the things' in (3.27b). These 72 instances (14+22+26+10=72) indicate that Hebrew speakers take the specific *definite* "NP subjects" in (3.27) — *hamilim/hadvarim* 'the words/the things' — not quite as subjects

but rather as direct objects (of categorical propositions), and by implication also their indefinite counterparts in (3.26), when embedded in this P1 S-pattern (Figure 3.3a).

This procedure does not apply to *ma lehagid/ledaber* 'what to say/speak', because it cannot bear the definite article, and should therefore be examined differently. The only reference I know of which considers *ma lehagid/ledaber* 'what to say/speak' when embedded in a P1 S-pattern is Rosén (1977: 216). Rosén saw *ma lehagid/ledaber* 'what to say/speak' as an infinitival group with an interrogative element (here, *ma* 'what') not placed at the beginning of the sentence. But he did not specify the functional role of *ma lehagid/ledaber* 'what to say/speak'. Epstein (1971) who analyzed similar sequences embedded, however, in the (biblical) S1 S-pattern where the predicate requires an NP as a complement, suggested that *ma lehagid/ledaber* 'what to say/speak' may have the functional role of an object. Moreover, if we attempt to replace *ma* 'what' in *en ma lehagid/ledaber* with the lexical noun *davar* 'thing', there seems to be no change of meaning of *en ma lehagid/ledaber* (Ruth Burstein, p.c., 20 June 2022). *Davar lehagid/ledaber* 'thing to say/speak', as any other "NP subject" of thetic propositions, can be taken as a direct object, and by implication *ma lehagid/ledaber* 'what to say/speak' as well.

In light of these results, I suggest that the direct-object-like NPs in (3.23) (*milim/dvarim ka'ele / ma lehagid/ledaber* 'words/such things / what to say/speak') form a VP-like unit together with the predicate *en* 'there's/are no'. Much like "true" VPs, which make up interpretatively cohesive units, these VP-like units, I suggest, are prone to semantic change, to idiomatization.

Unlike the predicate *en* 'there's/are no', the predicate *xaval* 'it's too bad/a pity' is not directly followed by an NP, but by a PP, specifically, an NP preceded by the preposition *al* 'on' (3.20). Such a case of a thetic proposition was discussed only by Kuzar (2002), who (correctly) argued, that the fact that the argument is a PP following a predicate, indicates that it cannot be the subject, but rather the object of the predicate. It is even more special than the case in which the element directly following the predicate is just an NP (as in the *en* 'there's/are no' sentences, (3.23)). This combination of a predicate and a PP is definitely a VP. Although not a Verb-Direct Object combination, it is still a more cohesive unit than a Subject-Verb combination. As such, I suggest, it is prone to semantic change, to idiomatization.

Recall also that this particular thetic *xaval* 'it's too bad/a pity' proposition is of the 'waste'-related kind (3.11), where 'waste' is strongly implied and therefore omitted, while leaving the wasted commodity as the sole NP inside the PP. The absence of an explicit 'waste' contributes to the semantic opacity of this specific thetic proposition.

In the next section I show that the remaining predicates and the respective infinitives that follow them, in (3.21) and (3.22), also form interpretatively cohesive units.

3.3.2.2 The predicate and the infinitive in dativeless thetic propositions form an interpretatively cohesive unit

The syntactic status of the infinitive in thetic propositions has triggered a lively debate among linguists who study Modern Hebrew. Stern (1983) argued that the infinitive is

the subject of the predicate in thetic propositions, because it scores positively on two tests: (i) It can be paraphrased by an NP. This NP together with an adequate predicate (often a predicative adjective) can form an alternative categorical proposition with an equivalent meaning. In this alternative categorical proposition, the NP serves as the subject; (ii) similar to an NP, the infinitive can also serve as an answer to the question 'what X?', where 'X' stands for the predicate in the thetic proposition.

Rubinstein (1968: 167-178) too considered the infinitive as the subject of thetic propositions. He argued, however, that in the presence of an expletive (optional, in Hebrew), the expletive becomes the subject, relegating the infinitive to an appositive. Berman (1980), on the other hand, rejected the analysis of the infinitive as a subject of thetic propositions, because it fails to meet the criteria of a subject: It lacks referentiality often associated with "true" or lexical subjects; it fails to trigger agreement with the predicate; and its basic position is post-predicative, not sentence-initial.[34] Berman classified such thetic propositions as subjectless altogether. Mor and Pat-El (2016) and Pat-El (2018) suggested that the introduction of the infinitive by a preposition (*le-* 'to') could explain why it is regarded as the complement of the predicate rather than its subject. Unlike the others, Kuzar (2002) preferred to leave the question open. He argued that since an infinitive bears no morphological signs whatsoever, its status cannot be determined with absolute confidence. This is likely the reason why Kuzar (2012) chose to label this infinitive an 'evaluee' (see Figure 3.3b above).

The unresolved syntactic status of infinitives in thetic propositions prompted me to follow another line of analysis — that of Divjak and Janda (2008, and Janda and Divjak 2015). Divjak and Janda examined a specific form of impersonal constructions in Russian, those that contain a finite verb, a dative participant, and an infinitive (much like the sentences studied here). They showed that this sequence of elements is shared by two constructions of an identical form *yet* bearing different semantics. They analyzed the role of the infinitive with respect to the finite verb, and showed that in the first construction type, the finite verb introduces a syntactic slot for a nominative. The infinitive easily fits into this syntactic slot, thus serving as an alternative to an NP. This infinitive, being inanimate, is indeed a non-typical subject, but it is still a subject. In the second construction type, however, the finite verb is a morphologically defective verb that does not introduce any nominative slot. This defective finite verb integrates with the infinitive, which carries most of the semantic load of this construction, to produce a 'complex event', similar to modal verbs "that merely modify the infinite verb" (Divjak & Janda, 2008: 169).

In order to determine whether a given finite verb belongs to the first or to the second construction type, Divjak and Janda ran a small-N design experiment. They asked native Russian speakers to judge the acceptability of schematic constructions they were presented with (termed 'pro-form constructions'). If the constructions were judged acceptable, then the speakers were asked to produce an example of their own with

[34] Recall Kuzar (2002) and Melnik (2006) above, who used the very same arguments to show that the NP following the predicate in thetic propositions is an object rather than an "NP subject".

lexical elements. For example, a triggering schematic construction could have been *it's nice what/to do*, where the pro-noun *what* stands for a noun, and the pro-predicate *to do* stands for an infinitive. If both pro-form constructions, with *what* and with *to do*, are judged acceptable, then the infinitive is an argument of the finite verb, specifically, its subject. However, if only the *to do* pro-form construction is judged acceptable, then the infinitive is not an argument of the finite verb. The infinitive is, in fact, the semantic kernel of the construction, whereas the finite verb is just modal-like. The results of the experiment supported Divjak and Janda's hypothesis that a single sequence of a finite verb, a dative participant, and an infinitive, is shared by two constructions of an identical form yet bearing different semantics.

Given the Russian facts and the well-known effect of Russian on Modern Hebrew,[35] it is not implausible to assume that, similarly to Russian, in Hebrew too, a single sequence comprising of a finite verb, a dative participant, and an infinitive, is shared by two constructions of an identical form but bearing different semantics. Such an assumption can account for the longstanding debate over the syntactic status of the infinitive in thetic proposition (i.e., whether it is a subject, or not). Divjak and Janda's testing methodology can naturally help us decide this matter. Specifically, their methodology can help determine the degree of semantic bonding between each predicate and the respective infinitive in (3.21) and (3.22).

To this end, I replicated Divjak and Janda's experiment (see Appendix B for details). Six native Hebrew speakers were asked to decide whether *efšar* 'it's possible' and *ba* 'it feels like' introduce a nominative slot or not. They were requested to produce sentences according to certain pro-form constructions they were presented with, and also to judge the acceptability of made-up sentences.

Results show that *efšar* 'it's possible' can be followed only by an infinitive, rather than by an NP. In other words, a nominative is blocked here. But this is not at all surprising, since *efšar* 'it's possible" is a full-fledged modal. *Ba* 'it feels like' behaves similarly to *efšar* 'it's possible', but it does not (yet) have an indisputable status as a modal. Still, both produce a 'complex event' with the respective infinitives, where the infinitives are "the center of gravity" (Divjak, 2010: 39), and *efšar* 'it's possible' or *ba* 'it feels like' are just modifiers[36] (see Croft's 2001: 216-220 process of clause collapsing). By definition, such a 'complex event' is a kind of a VP. Much like "true" VPs, which make up interpretatively cohesive units, these 'complex events', VP-like units, I suggest, are prone to semantic change, to idiomatization.

[35] Russian is indeed typologically remote from Hebrew, but it had a considerable influence on the evolution of Modern Hebrew (Dubnov, 2005b; Izre'el, 2002; Kuzar, 2001: 120-134; Wexler, 1990). In particular, it influenced the use and productivity of thetic propositions in Hebrew (Dubnov, 2005a, 2005b).

[36] Interestingly, in Russian, as in Hebrew, the counterpart of *ba* 'it feels like' is a modal-like defective verb of the volitional kind which forms a 'complex event' with the infinitive that follows (Divjak & Janda, 2008: 170).

I have so far shown that the sentences analyzed here, all thetic propositions, are either of the Verb-Object like kind (Section 3.3.2.1) or of the Verb-Infinitive—the 'complex event'—kind (Section 3.3.2.2). Importantly, in both cases, the non-dative elements are not the subjects of the thetic proposition. They are elements that team up with the predicates to produce an interpretatively cohesive unit which tends to become idiomatic.[37] At this stage, after semantic change has taken place, the newly evolved entities are *idiomatic sentences*, but not yet full-fledged words.

Now we may ask: Why, during this stage of the lexicalization of the evaluative thetic propositions, i.e., idiomatization, is the dative participant (which fills an optional slot in the P1 S-pattern) left out? And, could it be the case that the semantic change benefits from its absence? The answer is provided in the next section.

3.4 The semantic change may benefit from the absence of the dative participant

According to Kuzar (2012), already quoted above, the function of the Evaluative P1 S-pattern is as follows:

> Through the use of an evaluative predicate, the evaluative S-pattern evaluates the evaluee. **The evaluation is made relevant to an affectee**, if present. (p. 80; emphasis mine)

This definition suggests that the affectee, here the dative participant, may be somewhat of a topic, for according to Lambrecht (1994),

> [a] referent is interpreted as the topic of a proposition if **in a given discourse** the proposition is construed as being **about** this referent, i.e. as expressing information which is **relevant to** and which increases the addressee's **knowledge of** this referent. (p. 127; original emphasis)

The suggested topicality of the dative participant in the Evaluative P1 S-pattern is not surprising because the dative participant in the Hebrew P1 S-pattern (of all kinds) is mainly the speaker (see Dattner's 2019 quantitative analysis), and therefore necessarily human. Human participants, whether dative-marked or not, tend to be construed as topical (e.g., Brown, 1983; Comrie, 1989: 197-198; Dahl & Fraurud, 1996; Givón, 1976: 152; Haspelmath, 2001).[38] This has been supported by Mor and Pat-El (2016) with regard to the specific dative participant in the Evaluative P1 S-pattern in Hebrew. Kuzar (2012: Ch. 8) even termed the dative participant in the P1 S-pattern a

[37] Noteworthy in this context is Vilkuna's (1989: 156) concept of 'semantic bonding' — the interdependency between the verb and the following NP — in Finnish existential thetic propositions. The outcome of such 'semantic bonding', she suggested, is the 'idiomatic flavor' (p. 166) of existential thetic propositions. This intuition is shared by Marantz (1984: 27) re VPs (see Section 3.3.1).

[38] Human participants are topical because they observe the principle that "grammar reflects both human perception of the world and human anthropocentric interests" (Wierzbicka, 1988: 250).

secondary topic, after Lambrecht (1994: 147), as opposed to the *primary topic* in the S1 S-pattern. Kuzar went on to suggest that the secondary topic

> [i]s a referent whose topicality is not established as part of the prototypical instantiation of the primary topic in a syntactic construction, namely subjecthood, but rather as a participant, whose information status is calculated by the hearer in real time. (p. 195)

In light of the topical status of the dative participant in the P1 S-pattern, I suggest that the dative participant may interfere with the semantic change of the members of the Ultimate construction family. Put differently, the semantic change may benefit from the absence of the dative participant. From the point of semantic change,

(i) the dative participant may be less relevant to the predicate and the non-dative element than the two are to each other. It would therefore not (so easily) lend itself to an interpretatively cohesive unit;
(ii) it may also attenuate the high intensity of the newly evolved idiomatic sentences (all members of the Ultimate construction family).

I tackle each of these two issues in turn.

I start by justifying my argument in (i) above. According to Lambrecht (2000), the dative participant does not rule out the status of a sentence as a thetic proposition, as focal. Consider Examples (3.28) from Italian (=Lambrecht's (45)).

(3.28) a. *Mi si è rotta la macchina*
to.me itself is broken the car
'My CAR broke down'

b. L' ha lasciata il marito
her has left the husband
'HER HUSBAND left her'

Lambrecht noted that

> "[t]he presence of the **topical dative** (*mi*) or accusative (*l'*) pronouns in (45) [here (3.28) – IB] does not preclude SF status [i.e., theticity – IB] of the two sentences. This freedom of co-occurrence is due to the fact that the object constituent has no prosodic nor syntactic focus properties, hence **does not compete with the inverted subject** [i.e., the non-dative element - IB] **for focus status**." (p. 648; emphasis mine)

Hence, the dative element is not focal. Lambrecht and Polinsky (1997) even defined the focus domain of the thetic proposition as "the sentence **minus** any topical non-subject arguments" (p. 2; emphasis mine), here, the dative participant. It therefore seems that

the presence of the dative participant does not rule out the theticity of Example (3.28), while at the same time, it is not considered part of the focus.

If the dative participant is indeed not part of the focus, then the obvious question to ask would be whether the conceptual bonding between the dative participant and the predicate, or the non-dative element, or both, is as tight as the conceptual bonding between the predicate and the non-dative element (which made them team up to form an interpretatively cohesive unit; see Section 3.3). In order to check this, one needs to examine the extent to which propositions in the form of a P1 S-pattern *with* and *without* a dative participant map onto the S1 S-pattern, which is normally associated with a categorical proposition. The subject of the S1 S-pattern, the (nominative-marked) primary topic, is a potential counterpart of the dative participant in the P1 S-pattern (for a similar view see Melnik, 2014). If the relation between the dative participant and remaining elements in the P1 S-pattern is similar to the relation between the topic of the S1 S-pattern and the comment, then the prediction is as follows: Speakers are expected to paraphrase propositions with a dative participant encoded in the P1 S-pattern by using the S1 S-pattern more often than their dativeless counterparts. However, if the relation between the dative participant and the remaining elements is as tight as the relation between the predicate and the non-dative element, then speakers are expected to paraphrase propositions in the form of a P1 S-pattern by using the S1 S-pattern to the same degree, regardless of the presence or absence of the dative participant.

In order to decide this matter, I compiled a list of all 168 propositions encoded in the P1 S-pattern, all colloquial neologisms from Rosenthal's (2005/2018) Comprehensive Dictionary of Israeli Slang (~10,000 entries; see Section 2.7). I divided them into two groups — those with a dative participant and those without a dative participant. I then checked how each item was paraphrased — whether by an S1 S-pattern, or by the P1 S-pattern, or simply by a phrase. For instance, Example (3.29a) is a proposition cast in the form of a P1 S-pattern appearing *with* a dative participant. It is paraphrased by a proposition (3.29a') cast in the form of a Verbal S1 S-pattern. Example (3.29b), on the other hand, is a proposition cast in the form of a P1 S-pattern *with* a dative participant which is paraphrased by a different proposition cast in the form of a P1 S-pattern (3.29b'). Example (3.30a) is a proposition cast in the form of a P1 S-pattern *without* a dative participant which is paraphrased by a proposition cast in the form of a Verbal S1 S-pattern (3.30a'). Example (3.30b) is a proposition cast in the form of a P1 S-pattern *without* a dative participant which is paraphrased by a proposition also cast in the form of a P1 S-pattern (3.30b').

(3.29) a. *en*$_{PREDICATE}$ *lo*$_{AFFECTEE}$ *elohim*$_{EXISTENT}$
there's.no to.him God

a'. paraphrase (Verbal S1 S-pattern):
hu *po'el* *lelo* *ma'acorim* *musari'im*
he operates without restraints moral

'He operates beyond the pale of any acceptable human conduct.'

b. *nišbar*$_{PREDICATE}$ *lo*$_{AFFECTEE}$
breaks to.him

b'. paraphrase (Evaluative P1 S-pattern):
kaca *nafšo*
has.enough.of his.soul

'He has enough of (something).'

(3.30) a. *tov*$_{PREDICATE}$ ∅$_{AFFECTEE}$ *lada'at*
good to.know

a'. paraphrase (Verbal S1 S-pattern)
lamadeti *mašehu* *mo'il*
I.learned something useful

'It's good to know.'

b. *efšar*$_{PREDICATE}$ ∅$_{AFFECTEE}$ *lišmo'a* *sika* *nofelet*
it's.possible to.hear pin falling

b'. paraphrase (Existential P1 S-pattern):
sarera *dmama* *muxletet*
prevailed silence utter

'There was utter silence.'

A Chi-square test of homogeneity shows that significantly more propositions cast in the form of a P1 S-pattern *with* a dative participant (see Example 3.29) (64/103=62%) than propositions cast in the form of a P1 S-pattern *without* a dative participant (see Example 3.30) (13/67=19%) were paraphrased by propositions cast in the form of the S1 S-pattern, χ^2 (1, N = 170) = 29.9, $p = 4.5\times10^{-8}$ (φ = 0.42, a large effect size). These findings suggest that the dative participant in thetic propositions in the form of a P1 S-pattern is perceived by speakers as close to the nominative-marked topic in the S1 S-pattern. Therefore, it stands to reason that the relation between this topical dative and

the remaining elements in the P1 S-pattern is similar to the relation between the nominative-marked topic in the S1 S-pattern and the respective comment. The latter rarely fuse together to idiomatize (see Section 3.3.1). This may also be the case with the dative participant and remaining elements in the P1 S-pattern.[39]

I move on to justifying my argument in (ii) above. As evidenced from Examples (1.2) and more examples (C1-C10) in Appendix C, all the members of the Ultimate construction family, idiomatic sentences by now, are modifiers conveying a highly intense evaluation. Let's take *en milim* 'there are no words' and *ba livkot* 'it feels like crying' as representative examples, and see how they gained intensifying meaning (also briefly mentioned in Appendix A) and how the presence of the dative participant could have suppressed or, at least, attenuated it.

En milim 'there are no words' is an existential proposition interpreted as evaluative (see Section 3.2.5). *En milim* indicates that there are no words (whatsoever) to express the speaker's amazement/shock regarding a stance-object, for the said stance-object is of such extreme quality that no words are suitable for the task. The high intensity of *en milim* 'there are no words' is a result of an intended exaggeration, for it is obviously unlikely that suitable words are nowhere to be found, such that would be capable of evaluating a stance-object, however amazing/shocking it may be. If a dative participant is introduced into this evaluation, *en li milim* 'I have no words', then the evaluation, though of high intensity, represents the opinion of the stance-taker alone. In other words, it is only the stance-taker that lacks suitable words to evaluate the amazing/shocking stance-object. So, in fact, such words could possibly exist. The intense evaluation no longer gains its strength from having no exception.

Ba livkot 'it feels like crying' became an intensifier via a cognitive-affective model of negativity bias, where the semantics of negative emotion are metonymically mapped into intensifiers (e.g., Jing-Schmidt, 2007). As in the case of *en milim* 'there are no words', in the absence of a dative participant, *ba livkot* 'it feels like crying' represents a shared negative emotion. This emotion is obviously attenuated if attributed exclusively to a single person, the stance-taker, here the dative participant.

In sum, I suggest that the semantic change here described may benefit from the absence of the human dative participant. This absence guarantees that the thetic proposition introduces only a minimal sentence, where all elements semantically bond to produce an interpretatively cohesive unit. Such a unit is a potential candidate to undergo semantic change, idiomatization, to eventually form an idiomatic sentence.[40]

---

[39] In this context, it's worth mentioning (3.21) the *ba* 'it feels like' sentence. In this sentence, where the predicate and the infinitive combine to produce a 'complex event' (see Section 3.3.2.2 above), the dative participant is taken as an 'agentive experiencer' (Divjak & Janda, 2008: 163). It functions as the agent of the event expressed by the infinitive and simultaneously as the experiencer of the (defective) finite verb (i.e., *ba*). Agentivity is a subject property, and subjects, in turn, often function as topics.

[40] The stage of being an idiomatic sentence (rather than a word) is a well-defined stage as attested by evidence provided in Chapter 4.

Now it's time to ask what affects/triggers the change in grammatical status of these idiomatic sentences. In the next section I suggest a partial answer. A comprehensive answer will be provided in Chapter 4.

3.5 'Semantic Incompleteness' enabled by theticity drives the grammatical change

Categorical and thetic propositions are each associated with a distinct discourse function, as already mentioned in Section 3.2.1 above. Indeed, propositions in the form of the Verbal S1 S-pattern provide the foreground actions of the storyline, whereas propositions in the form of the Copular S1 S-pattern provide the background information. Both, however, are associated with topics already active in the storyline. The function of evaluative thetic propositions in the form of a P1 S-pattern is to evaluate the situation encoded in the evaluee, which is the stance-object (Kuzar, 2012: 191-192). Consider (3.20)-(3.22), repeated here for convenience. In (3.20), the evaluee '(waste of) words/time' is a distressing event; in (3.21), the infinitival evaluee 'to cry/die' is an instinctual event; and in (3.22), the infinitival evaluee 'to go crazy/to die' is a possible event.

(3.20) ***xaval***$_{PREDCATE}$ $\varnothing_{AFFECTE}$ ***[al ha-zman / ha-milim]***$_{EVALUEE}$
it's.too.bad on (the waste of) the-time / the-words

'It's a waste of time/words.'

(3.21) ***ba***$_{PREDCATE}$ $\varnothing_{AFFECTE}$ ***[livkot / lamut]***$_{EVALUEE}$
it.feels.like to.cry / to.die

'It feels like crying/dying.'

(3.22) ***efšar***$_{PREDCATE}$ ***[lehištage'a / lamut]***$_{EVALUEE}$
it's.possible to.go.crazy / to.die

'It's possible to go crazy/die.'

But if the elements of each of (3.20)-(3.22) make up a single interpretative unit which becomes an idiomatic sentence, as I showed in Section 3.3, then the evaluee is no longer the stance-object. (3.20)-(3.22) are still evaluative and each of them functions as a modifying unit conveying a property. As any unit that conveys a property, they must be *relational*. "Relations are conceptually dependent, i.e. one cannot conceptualize interconnections without conceptualizing the entities that they interconnect" (Langacker, 1987: 215). So if a concept is relational, it "inherently requires reference to another concept" (Croft, 2001: 87). In other words, the newly evolved evaluating/modifying idiomatic sentences are *semantically incomplete* and in need of a head to modify. Such a head can be found in the very preceding sentence, should it

contain an explicit modifiable element which the newly evolved modifier may attach to.

The analysis of (3.23), also repeated below for convenience, is much simpler. By virtue of being existentials functioning as evaluating modifiers, the modifiable element of (3.23) must be external. As a unit which comprises of an anaphoric element, *en dvarim* ***ka'ele/u*** 'there are no **such** things' (specifically) even points explicitly to a previously mentioned external modifiable element, a stance-object (see the examples in Appendix C.10).

(3.23)	***en***$_{PREDCATE}$	$\varnothing_{AFFECTE}$	***[milim / dvarim ka'ele / ma lehagid / ma ledaber]***$_{EXISTENT}$
	there's/are.no		words / things such.as.these / what to.say / what to.speak

'There is/are no words/no such things/nothing to say/nothing to speak.'

The modifying idiomatic sentences evolved out of the members of the Ultimate construction family are 'flexible modifiers', that is, they are syntactically and/or semantically flexible (McNabb, 2012; Salazar-García, 2010; see Chapter 4 for more details). Their flexibility is not actually surprising, since "it has been shown that the categorial specificity of linguistic units increases—resulting in a decrease in flexibility—when they become structurally more complex" (Van Lier & Rijkhoff, 2013: 23). As semantically opaque idiomatic sentences, the members of the Ultimate construction family are mono-morphemic and therefore not structurally complex. Consequently, they can make up good flexible modifiers. As flexible modifiers, they can modify nouns, verbs, or adverbs. And since Hebrew modifiers follow the modified element,[41] the newly evolved modifiers function as adjectives — when following a noun, as adverbs — when following a verb, or as intensifiers — when following an adjective or a verb. Admittedly, this is a simplistic description of motivation for the change in grammatical status that the newly evolved idiomatic sentences undergo. I here consider only the motivation on the part of the idiomatic sentences. Obviously, the role of the context in the change in grammatical status is as crucial. The motivation on the part of the context which affects/triggers the change in grammatical status is described in detail in Chapter 4.

In sum, I suggest that the semantic change here described is a consequence of elements which form an interpretatively cohesive unit. This, in turn, constitutes a good candidate for idiomatization, eventually (potentially) leading to the creation of an idiomatic sentence. The newly evolved idiomatic sentence functions as a modifier (expressing a highly intense evaluation). As a modifier, it is semantically incomplete,

[41] Hebrew adjectives and adverbs always follow the noun or verb they modify, respectively. An intensifier can either precede or follow the adjective.

necessarily in need of a head in prior discourse to modify. This state of affairs motivates the change in grammatical status.

3.6 Some related phenomena

Thus far, I have focused on thetic propositions in Hebrew which have turned into intra-sentential elements, words. Evidence from other languages is reminiscent of this phenomenon, and may provide support for my claim, that it is interpretative cohesion on the one hand and semantic incompleteness on the other that enable the lexicalization of thetic propositions. Three cases are considered in the following sections.

3.6.1 Sentence adverbs in European languages

The lexicalization of sentence adverbs in European languages was described in great detail by Ramat and Ricca (1998). Relevant to my study are not the derivational formations (e.g., Eng. *–ly*, It. *–mente*), but rather idiosyncratic formations of sentence adverbs, and in particular those formed by fusion (univerbation, in their terminology) of the elements of the sentences. Modal epistemic adverbs display the greatest number of idiosyncratic formations, those denoting 'perhaps' (e.g., Eng. *(it) may be*, Fr. *(il) peut-être*, Nor. *(Det) kan ske* 'it can/may happen'), 'probably' (e.g., Lith. *turbūt* 'it must be') and 'apparently' (e.g., Ger. *scheints* 'seems it'). Many of these are the outcome of fusion of main clauses or parentheticals. The latter are indeed full sentences in their own right. Based on the list presented by Ramat and Ricca, it seems that many of these formations meet the criteria required for lexicalization (of full sentences): Many (but not all!) originate from thetic propositions cast in the form of a P1 S-pattern, thus making an interpretatively cohesive unit. As adverbs, they are semantically incomplete by nature. No wonder, then, that these adverbs may fit as parentheticals into another sentence in the most natural way.

It should be noted that the epistemic adverbs denoting 'perhaps' and 'probably' have not been considered full sentences in the literature (except for Ramat and Ricca), because, in order to count as such, they require an expletive subject, which was often left out during the fusion of the other elements. Instead, they have been considered VPs (Ježek & Ramat, 2009: 400). The status of those adverbs — whether sentences or VPs — underscores the structural and functional resemblance between thetic propositions (P1 sentences) and VPs (see Section 3.3) and provides further support for my analysis.

3.6.2 Nouns in European languages

Cases where nouns evolved from imperatives are attested in European languages, e.g., Eng. *forget-me-not*, Ger. *Tunichtgut* 'a person who is good-for-nothing', Fr. *tromp-l'oeil* 'painting-related visual deception intended to create an illusion of certain spatial qualities' and It. *battiloro* 'gold-beater, lit. beat the gold' (Blank, 2001: 1602; Gaeta & Ricca, 2009: fn. 13; Lehmann, 2020: fn. 27).

For the specific cases where the "the verb form is ambiguous between imperative, third person singular and the bare verb stem", as is the case with the abovementioned examples, Lehmann (2020) suggested that "[…] a structural position in which just a verb stem is needed is occupied by the imperative, regardless of its meaning" (p. 20).

Gaeta (2015: 120) preferred the term ‘verbal themes’ (i.e., a root+thematic vowel=stem) over ‘imperative’, implying that these sequences, whether they have undergone lexicalization or conversion (here, nominalization), are but VPs, not (full) sentences.

These nouns, quite like the adverbs in the previous section, evolved from syntagmas which are Janus-faced. They can be considered full sentences, and at the same time VPs. The latter are definitely interpretatively cohesive units, and as such can undergo semantic change. VPs are already sentence-internal, so semantic incompleteness (which triggers/affects the change in grammatical status from an independent sentence to a word) is irrelevant in the current case.

3.6.3 Nouns in Native American languages

Cases of nouns evolved from full sentences are attested in Native American languages (Mithun, 2006, 2014, 2020), e.g., Mohawk ‘he argues’=‘lawyer’, ‘one wipes with it’=‘towel’. In those languages, the verb is obligatorily marked for its arguments, using referential pronominals. This is why such verbal forms can also function as full sentences. When such entities become nouns, the external argument is either the unmarked third person singular or simply generic, and the aspect is most often habitual, thus alluding to nominalization of VPs.

Once again, these nouns evolved from syntagmas which are Janus-faced. They can be considered full sentences, and at the same time VPs. The latter are definitely interpretatively cohesive units, and as such can undergo semantic change. As sentence-internal, their semantic incompleteness is of no relevance to the undergoing nominalization.

3.7 Summary and conclusions

The goal of this chapter was to understand why full sentences rarely become intra-sentential elements (i.e., words), namely, undergo lexicalization. By analyzing a set of full sentences in Hebrew that have become, or are on the verge of becoming words — in fact, the only sentences in Hebrew to have done so — I argued that the lexicalization process is construction-dependent.

I showed that it is the type of proposition and its constructional properties — both form and function — that account for the linguistic change or lack thereof. I contrasted thetic with categorical propositions and showed that only the former can evolve into words:

(a) In order for a proposition to undergo semantic change, specifically turning semantically opaque, just like words, it must be mono-partite. Only thetic —not categorical—propositions constitute mono-partite units. Thetic propositions introduce only a comment, while categorical propositions introduce both a topic and a comment. The elements within a comment (similarly to the elements within a topic) are relevant to each other more than to elements outside of it. This mutual relevance implies that these elements are semantically bondable. As such, they make up an interpretatively

cohesive unit, which can further become semantically opaque to produce an idiomatic sentence. Obviously, this precondition does not entail semantic change of every single thetic proposition. It just marks them as potential candidates to undergo semantic change, once the relevant cognitive-functional mechanism comes into play.

(b) In order for the newly evolved idiomatic sentence to undergo change in grammatical status, from an extra-sentential to an intra-sentential element, it must relate to an external concept from the prior discourse. For that, it must be relational/semantically incomplete. Only idiomatic sentences which evolved out of evaluative thetic propositions having undergone semantic change to become modifiers are semantically incomplete. As modifiers — relational elements by definition — they are in search of a modifiable element in prior discourse to attach to. The same cannot be said for categorical propositions.

I also suggested that the semantic change may benefit from the absence of the dative participant, a secondary topic. In the absence of a dative participant, there is no potential topic that disrupts the potential semantic bonding between the remaining elements (the predicate and the non-dative element).

Given this analysis, the phenomenon of full sentences in Hebrew, which have become—or are on the verge of becoming—words (i.e., undergoing lexicalization), is not as puzzling as it seemed to be at the beginning of this dissertation. This process/phenomenon depends on the likelihood of all the sentential elements to make up an interpretatively cohesive unit, where the various elements may fuse to make an idiomatic sentence. In that sense, a fused thetic proposition cast in the form of an P1 S-pattern is not too different from a VP. Once construed as a VP, it is not unreasonable for a thetic proposition to undergo lexicalization. The contrast between categorical and thetic propositions is not at all different from the contrast between Subject-Verb combinations and Verb-Object combinations with respect to fusion and further idiomatization (see Section 3.3.1). After all, S1 S-pattern categorical propositions are Subject-Verb combinations, and P1 S-pattern dativeless thetic propositions have been shown here to bear resemblance to Verb-Object combinations (i.e., VPs).

Again, I must underscore that in this chapter I considered only the motivation to undergo lexicalization on the part of the members of the Ultimate construction family. But as in every case of linguistic change, it takes two to tango. The motivation on the part of the context must also be considered. This motivation is dealt with in the next chapter.

Chapter 4: The *contextual conditions* required of a full sentence to become a word – An answer to the TRANSITION problem

Scaffolding

Masons, when they start upon a building,
Are careful to test out the scaffolding;
Make sure that planks won't slip at busy points,
Secure all ladders, tighten bolted joints.
And yet all this comes down when the job's done
Showing off walls of sure and solid stone.
So if, my dear, there sometimes seem to be
Old bridges breaking between you and me
Never fear. We may let the scaffolds fall
Confident that we have built our wall.

-- Seamus Justin Heaney, in *Death of a Naturalist*, 1966: 37

Weinreich, Labov and Herzog's (1968) TRANSITION problem "[…] ask[s] about the intervening stages which can be observed […] between any two forms of a language […]" (p. 101). The present chapter and the next chapter are dedicated to solving the transition problem of the members of the Ultimate construction family. In the present chapter I present an account of the change in the grammatical status undergone by *xaval al hazman*, as well as the other members of the Ultimate construction family, to have become intra-sentential elements (eventually, words). I emphasize the critical constructional role of the context (pun intended) as scaffolding the change in grammatical status, as well as its critical role in the inception of 'flexible modifiers'. In the next chapter (Chapter 5), I show that the constructional scaffolding context must come down once "wordification" is complete in order to allow further developments which indicate the depth of lexicalization.

4.1 "[…] and a time to build up"[42]

In the previous chapter I offered a motivated account for the typologically rare case of full sentences turning into words. I outlined the preconditions that must be met by full sentences in order to qualify as candidates for lexicalization — interpretative cohesion and semantic incompleteness — arguing that it is the specific construction (the form/function association) of the sentences in Table 3.1 that makes them good candidates for lexicalization. I examined this lexicalization process from the point of view of the linguistic units which are the subject of change, i.e., the members of the Ultimate construction family. However, one cannot overlook the role of the context in the process of linguistic change (Diewald, 2006; Evans & Wilkins, 2000; Heine, 2002, *inter alia*). In other words, the preconditions I proposed in the previous chapter only constitute necessary but not sufficient conditions for this lexicalization process.

In this chapter I provide an account of the linguistic change here studied from the point of view of the context embracing the newly evolved idiomatic sentences. I focus on the phase of change in grammatical status, where the (already) semantically opaque idiomatic sentences are integrated into the preceding sentence as a *bona fide* syntactic constituent (an adjective, an adverb or an intensifier; see Examples (1.2a-d)), shifting from what Haspelmath (2022) terms the *inventorium*, i.e., the set of "expression[s] with idiosyncratic, not fully predictable properties", to the *lexemicon*, i.e., "the set of all lexemes of a language, i.e. the members of the major lexical categories noun, verb and adjective."[43] As such, the evolution here described is an instance of complexity building (e.g., Biber & Gray, 2016; Du Bois, 2003; Givón, 2009), specifically of *compactization*, where the same amount of information originally spanning over two separate utterances — one introducing the modifiable head, and the other introducing a modifier expressing an extreme stance about the head — is squeezed into a single sentence.

Naturally, this account too is a constructional account, emphasizing the scaffolding role of the context in supporting the change in grammatical status leading to lexicalization. It also accounts for the—just as important—fact that the newly evolved words are, in fact, flexible modifiers. I ask the following questions:

Qi: What exactly is the context that mediates/d the change in grammatical status from an extra-sentential modifying idiomatic sentence to a modifier — an adjective, an adverb or an intensifier — of a single distinct constituent within the boundaries of a sentence?

Qii: What is it that stimulates/d the flexibility of the newly evolved modifiers?

---

[42] Ecclesiastes 3: 3

עֵת לִפְרוֹץ **וְעֵת לִבְנוֹת:** (קהלת ג', 3)

[43] In the present study, however, only the newly evolved adjectives belong to a major lexical category. The newly evolved adverbs and intensifiers belong to a minor one.

Qiii: Which (amplifying) intra-sentential element — an adjective, an adverb or an intensifier — is/was the first to evolve? And is there any clear line of trans-categorization?

In order to provide a theoretically sound and a statistically grounded model of this linguistic change, I will analyze not just *xaval al hazman* but three additional members of the Ultimate construction family listed in Table 3.1: *en dvarim ka'ele/u*, originally, 'there are no such things', *en milim*, originally, 'there are no words', and *ba livkot*, originally, 'it feels like crying'.

I start by describing the model that (I believe) will provide answers to all three research questions above.

4.2 The model proposed to account for the change in grammatical status

The contrast between the grammatical statuses of the starting point and the endpoint of the lexicalization process here studied — *xaval al hazman* as an independent sentence in Example (1.1) as opposed to full-fledged words in Examples (1.2a-d) — is a contrast between parataxis and hypotaxis. It is therefore suggestive of a (lexicalization) process which involves clause linkage.

In his seminal paper about the typology of clause linkage as a way to build complex sentences, Lehmann (1988) put forward a six parameter model designed to characterize every possible complex sentence. The parameters are described in terms of continua which "extend from a pole of maximal elaboration to a pole of maximal compression (or condensation) of lexical and grammatical information" (p. 216) and are correlated with one another (some more than others):

1. the hierarchical downgrading of the subordinate clause,
2. the main clause syntactic level of the subordinate clause,
3. the desententialization of the subordinate clause,
4. the grammaticalization of the main verb,
5. the interlacing of the two clauses, and
6. the explicitness of the linking.

Two of these parameters suggest that Lehmann's model could successfully account for the phenomenon here studied, and provide answers to the three research questions posed above. The first relevant parameter is the second one on the list — the syntactic level of the subordinate clause in the main clause. This parameter specifies the constituent of the embedding main clause that the subordinate clause integrates with. This constituent could be the entire embedding main clause or any part of it. At the extreme, maximally compressed end of this particular continuum, stands the word. The second relevant parameter is the third one on the list — desententialization of the subordinate clause. Desententialization is the reduction process by which a subordinate clause loses the properties of a clause. The components of the clause which are dropped are those which allow reference to a specific state of affairs, such as illocutionary force,

mood, tense, aspect, and participants. At the extreme, maximally compressed end of this particular continuum, stand nouns, but it could also be adverbs. And I believe that adjectives can also be included, simply because they are functionally related to adverbs.

If clause linkage is indeed the key for the linguistic change here studied, then the members of the Ultimate construction family can serve the role of the subordinate clause, because they are the units undergoing change. But what kind of sentence serves the role of the embedding main clause, the contextual sentence which provides the scaffolding for the change in grammatical status to take place?

The way to target this (kind of) sentence is predicated on the notion that "[t]he act of combining the clauses and signaling this combination linguistically is grounded in rhetorical production strategies" (Hopper & Traugott, 2003 [1993]: 177). If the members of the Ultimate construction family point to an extreme state of affairs, an endpoint, accompanied by a strong speaker's stance, then they can serve, for example, as a means to overcome the common tendency of other intensifiers to lose their emotive force over time (see, e.g., Hopper & Traugott, 2003 [1993]: 122; Klein, 1998: 26; Méndez-Naya, 2003). This suggests that the contextual, embedding main sentence should be of the kind that conveys a strong speaker's stance too. In fact, the contextual sentences should be of the kind that enables compensation over the loss of emotive force of the intensifiers that it hosts not by reduplicating the intensifier, but only by incorporating a further reinforcing (subordinate) clause. After all, the members of the Ultimate construction family are semantically opaque idiomatic sentences (by now), but still clauses. The (Hebrew) *Anaphoric degree-adverb exclamatives*,[44] exemplified in (4.1), a family of constructions in itself (Michaelis, 2001; Michaelis & Lambrecht, 1996), make up such potential contextual candidates. This is because they are often accompanied by an elaborating continuation, a *Correlative endpoint resultant-state clause*, as in Example (4.2) (Glinert's 1989: 218-219 unnumbered example; see Henkin 1994 for similar observations), introduced by *še* 'that'. This Correlative endpoint resultant-state clause was originally headed by *ad še* 'up.until that', which now seems to be "missing" the appropriate adverbial conjunctional head (*ad* 'up.until'), as explicitly suggested by Glinert (1989: 218-219) and Kuzar (1992b: 78), and implicitly by Henkin (1994: 135). The Correlative endpoint resultant-state clause elaborates on the evaluation involved in the preceding clause, testifying to the high intensity of the proposition conveyed (Mor, 1992).

(4.1) a. *hu* ***kol-kax*** muxšar!
he **so** talented

'He is **so** talented!'

[44] They are termed "anaphoric" because the intensifier has evolved from an originally deictic term *kol-kax* 'so' and *kaze/kazot/ka'ele/ka'elu* 'such$_{SG.M/SG.F/PL.M/PL.F}$'.

b. *hu* ***kol-kax*** caxak!
he **so** laughed

'He laughed **so much**!'

c. *hu* ***kaze*** muxšar!
he **such** talented

'He is **that** talented!'

d. *hu* ***kaze*** baxur!
he **such** a.young.man

'He is **such** a young man!'

(4.2) *kol-kax* *kar* ***še-kaše*** ***lišon***
so cold **that-it's.hard** **to.sleep**

'It's so cold **that it's hard to sleep**'

Note that unlike the Anaphoric degree-adverb exclamatives, see Example (4.2), other sentences (exclamatives such as *What a boy!* and non-exclamatives such as *He's a really wonderful boy*, alike) do not invoke a resultant-state clause. If they do, then it's possibly a consequence of analogy (see corroborating evidence in Section 4.3.3 below).

In terms of network links, the members of the Ultimate construction family and Correlative endpoint resultant-state clause have *filler-slot relations*, such that "specify associations between the slots of constructional schemas and particular lexical or phrasal fillers" (Diessel, 2023: 16), as illustrated in Figure 4.1. Note that the filler here is not "lexical or phrasal", but rather clausal (as described in Diessel, 2023: 50-53).

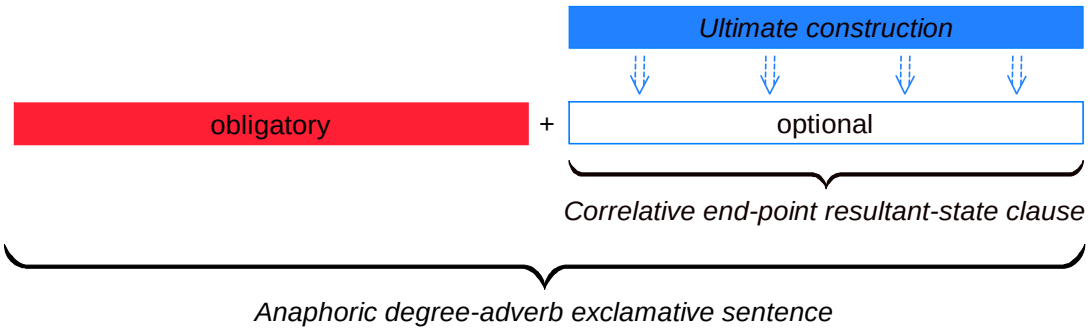

Figure 4.1: The filler-slot relations between the constructions involved in the lexicalization process described in this dissertation

Another—just as important—reason for proposing the Anaphoric degree-adverb exclamatives as the contextual scaffolding for the phenomenon here studied is the fact that focus of the variants of this construction (see Examples 4.1 above) could be a noun, a verb or an adjective. This focus is, in fact, the constituent eventually intensified by the Correlative endpoint resultant-state clause, which may explain the categorial flexibility of the newly evolved amplifiers — an adjective, an adverb or an intensifier, respectively — embedded in the slot of the Correlative endpoint resultant-state clause.

Following these considerations, I propose a four-stage evolutionary model of lexicalization for the members of the Ultimate construction family turning into intra-sentential flexible modifiers. I focus on the scaffolding role of the Anaphoric degree-adverb exclamative construction, and use *xaval al hazman* as a representative example:

Stage I:	*Xaval al hazman* is an independent idiomatic sentence preceding *or* following contexts expressing a speaker's strong stance on some state of affairs.
Stage II:	*Xaval al hazman* is incorporated into the preceding sentence which is an Anaphoric degree-adverb exclamative. The integrated *xaval al hazman* functions as a Correlative endpoint resultant-state clause, introduced by *še* 'that'.
Stage III :	Same as II, except that the preceding sentence is no longer necessarily an exclamative.
Stage IV:	Clause incorporation no longer requires *še* 'that'.

Stage I. Contexts (cited in English for convenience) preceded or followed by an independent idiomatic sentence *xaval al hazman* convey a strong speaker's stance on some state of affairs. This state of affairs typically involves a noteworthy fact, and can be expressed by an exclamative construction (see Ia and Ib below), an intensified non-exclamative construction (see 'extremely' in Ic below), or simply by means of lexical choice (see 'impressive' in Id below). This fine categorization is of no significance at this stage, but will prove to be significant in the next one.

(Ia)	She is **so** impressive! ***xaval al hazman*.**
(Ib)	**What** an impressive woman! ***xaval al hazman*.**
(Ic)	She is **extremely** impressive. ***xaval al hazman*.**
(Id)	She is **impressive**. ***xaval al hazman*.**

Stage II. As noted above, clause linkage is a syntactic means for incorporating one clause into another. As already mentioned, the only sentence type which invokes an additional clause (here a Correlative endpoint resultant-state clause) as a means of reinforcement of an intensifier losing its emotive force, is the Anaphoric degree-adverb

exclamative (Ia above). In other words, the Anaphoric degree-adverb exclamative embedding the Correlative endpoint resultant-state clause (by means of *še* 'that') may serve as scaffolding for the introduction of *xaval al hazman* (and other members of the Ultimate construction family) into the main sentence (see IIa below).

The other variants (Ib-Id above) may also incorporate the members of the Ultimate construction family via a Correlative endpoint resultant-state clause into the sentence (IIb and IIc below), but this is quite uncommon (see corroborating evidence in Section 4.3.3 below).

(IIa)	She is **so** impressive ***še-xaval al hazman*!**
(IIb)	**What** an impressive woman ***še-xaval al hazman*.**
(IIc)	She is **extremely** impressive ***še-xaval al hazman*.**

Note that my claim that stage II is an instance of clause linkage has implications for stage I. The independent idiomatic sentence *following* the contexts expressing a speaker's strong stance on some state of affairs would be preferable over a *preceding* one, in terms of the transition from stage I to stage II. The former, paratactic sequence (stage I) overlaps with the equivalent hypotactic complex sentence where the idiomatic sentence serves the role of a subordinate clause (stage II). In both cases — parataxis of stage I and hypotaxis of stage II — the idiomatic sentences share the backwards orientation towards the preceding clause or discourse,[45] thus "facilitating" the transition from stage I to stage II.

Stage III. The semantically "fresh" amplifying *xaval al hazman* can offer an alternative to the anaphoric degree-adverb (of the exclamative) which has probably lost its strong emotive force (see, e.g., Hopper & Traugott, 2003 [1993]: 122; Klein, 1998: 26; Méndez-Naya, 2003). The omission of the anaphoric degree-adverb marks the beginning of the falling apart of the contextual scaffolding (see III).

(III)	She is ∅ impressive ***še-xaval al hazman*.**

Stage IV. The final sign that the contextual scaffolding is no longer needed is the loss of *še* 'that'. This loss possibly results from the disappearance of the anaphoric degree-adverb (or the Anaphoric degree-adverb exclamative sentence construction) which provided the original slot for a Correlative endpoint resultant-state clause.[46] This is how the remaining *xaval al hazman* comes directly to modify the focus of the sentence it has been incorporated into (see IV), and it is reanalyzed as a modifier.

[45] In Hebrew, the Correlative endpoint resultant-state clause — a result clause — always follows the main clause, as is the case in many other languages (Hetterle, 2015: 124).

[46] The loss of *še* 'that' is further motivated by the fact that *še* 'that' loss in Modern Hebrew — whether a relativizer or a complementizer — is not uncommon (see Neuman, 2017 who argues that this loss is a manifestation of the structural residues of a substrate language, Moroccan Arabic; see also brief mentions by Zadka, 1991 and Blau, 1999: footnote 6).

(IV)	She is ∅ impressive ∅-***xaval al hazman*.**

In sum, I claim that it is the Anaphoric degree-adverb exclamative (IIa) that introduces/d the members of the Ultimate construction family into the preceding sentence, by means of a Correlative endpoint resultant-state clause. This paved the way for modification/amplification of some focal element within the embedding main clause. The diversity of focal elements dictates/d the word class of each newly evolved word, which accounts for the flexible nature of *xaval al hazman* as a modifier.

Now, in order to support my model, I need to show that,

(i) at stage I, the members of the Ultimate construction family follow the contexts expressing a speaker's strong stance (on some state of affairs) significantly more than precede them. This will explain a natural transition from stage I to stage II;
(ii) whereas non-exclamatives that host intensifiers can simply reduplicate the intensifier in order to overcome the loss of emotive force, this is not the case for Anaphoric degree-adverb exclamatives;
(iii) Anaphoric degree-adverb exclamatives (IIa) are more frequent than any other alternative (of stage II);
(iv) Anaphoric degree-adverb exclamatives (IIa) showed up chronologically earlier than sentences lacking the anaphoric degree-adverb (stage III), exclamatives and non-exclamatives alike.

One last thing before I turn to the actual analysis. Recall that in Section 4.1 above, I declared a combined analysis of *xaval al hazman* and three additional members of the Ultimate construction family, in order to present a firmly established model. To this end, I wish to substantiate the absolute functional equivalence of the additional members of the Ultimate construction family — *en dvarim ka'ele/u*, originally, 'there are no such things', *en milim*, originally, 'there are no words', and *ba livkot*, originally, 'it feels like crying' — with *xaval al hazman*. Examples (4.3a-c), where each of the three co-occurs alongside *xaval al hazman* (presumably as a means to underscore the speaker's strong stance), indeed show that each of the additional members of the Ultimate construction family is functionally equivalent to *xaval al hazman*. All four therefore deserve to be included in a combined analysis as suggested above.

(4.3) a. *šama'anu* *be-mikre* *mišehu* *mesaper* *be-hitlahavut*
we.heard in-coincidence someone recounting in-enthusiasm

še-hu *maca* *dira* ***še-xaval al hazman***
that-he found apartment **that-it's a waste of time → amazing**

še-en dvarim ka'ele *be-merkaz^* *ha-ir.*
that-there are no such things → amazing in-center the-city

'We overheard, unintentionally, someone recounting in great enthusiasm that he had found an **amazing**, **amazing** apartment in the city center.'
(tinyurl.com/2s4c2pwj)

b. *ha-xatuna* *hayta* *be'emet* *madhima*
the-wedding was truly amazing

en milim, ***xaval al hazman.***
there are no words → amazing **it's a waste of time → amazing**

'the wedding was truly amazing, **amazing**, **amazing**.'
(tinyurl.com/yypbd7mh)

c. *yafot* *yafot* *kol-kax* *yafot*
beautiful beautiful so beautiful

še-ba livkot…
that-it feels like crying → extremely

***še-xaval al hazman*!!!**[47]
that-it's a waste of time → extremely

'beautiful, beautiful, **so very** beautiful…**very much**!!!'
(tinyurl.com/2v7x5wh4)

I now turn to presenting the data that provide support for the proposed model. I address issues (i)-(iv) above one at a time.

4.3 Data supporting the model proposed to account for the change in grammatical status

4.3.1 The positions of the newly evolved idiomatic sentences with respect to the relevant context(s) (i)

First, I need to establish the preferred position of each independent idiomatic sentence with respect to the contexts that it evaluates, those expressing a speaker's strong stance on some state of affairs. As suggested above, a postposed position would facilitate a transition from the paratactic sequences of stage I to the hypotactic one(s) of stage II. And indeed, as is evident from IsraBlog data — semi-spoken, non-edited/standardized and diverse (in terms of number of writers) — presented in Figure 4.2, there are significantly more cases of *xaval al hazman* following (n = 365) than preceding (n =

[47] Interestingly, *ba livkot* and *xaval al hazman* are separated by three dots underscoring the fact that they are functionally equivalent and cannot occupy the same syntactic slot.

68) the contexts that they evaluate (binomial test, $p = 1.64\times10^{-50}$).[48] This is also the case with *en milim* — there are significantly more cases of *en milim* following ($n = 236$) than preceding ($n = 48$) the contexts that they evaluate (binomial test, $p = 2.76\times10^{-31}$). And the same holds for *en dvarim ka'ele/u* — there are significantly more cases of *en dvarim ka'ele/u* following ($n = 228$) than preceding ($n = 14$) the contexts that they evaluate (binomial test, $p = 3.27\times10^{-53}$). *Ba livkot* is not considered here, since it started out as a stage II item (for reasons that will be clarified at the end of Section 4.3.5.2).

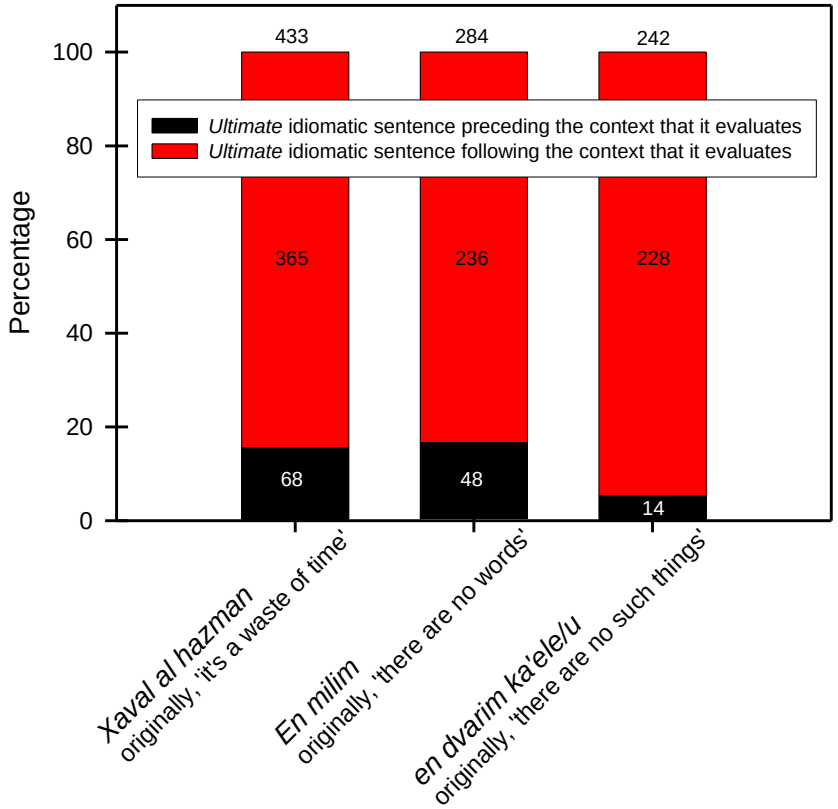

Figure 4.2: A snapshot (as of July 2017) of the position of extra-sentential idiomatic sentences (i.e., items of stage I) *xaval al hazman*, *en milim* and *en dvarim ka'ele/u* relative to the context that they evaluate — following (marked red) or preceding it (marked black). Data extracted from IsraBlog corpus.

From these results I conclude that a transition from a paratactic sequence (stage I) to a hypotactic sequence (stage II), where the members of the Ultimate construction

[48] Since IsraBlog is not a spoken corpus, but rather an informal written one, a formal sign for *xaval al hazman* as an independent sentence was at least one of the following conditions: *xaval al hazman* (i) preceding—rather than following—the sentence that contains the stance-object; or (ii) following the sentence that contains the stance-object, but distinctly marked by any of the following punctuation marks: a full stop, a comma, a colon, a semi-colon, three dots, a hyphen and sometimes even several spaces, rather than a single space, between *xaval al hazman* and the sentence to which it pertains.

family are incorporated into the preceding sentence as a subordinate clause, is plausible. In the next section I address issue (ii).

4.3.2 Strategies to compensate for the loss of emotive force of intensifiers (ii)

I further need to show that the Correlative endpoint resultant-state clause is (almost) the only strategy for reinforcing the Anaphoric degree-adverbs exclamatives, in order to overcome the common tendency of intensifiers (here, Anaphoric degree-adverbs) to lose their emotive force over time. As already mentioned, non-exclamatives hosting other intensifiers can simply reduplicate the intensifier in order to overcome this loss. To this end, I extracted from (part-of-speech tagged) HeTenTen corpus all instances of the anaphoric degree-adverbs in Examples (4.1a-d), as well as some other non-anaphoric degree-adverb intensifiers when *followed* by an adjective: both when a single instance of each intensifier precedes the adjective, and when a reduplicated intensifier precedes the adjective. The findings in Table 4.1 attest to the likelihood of intensifiers of the non-anaphoric degree-adverb kind to reduplicate in order to compensate (most probably) for the loss of (their) emotive force over time. This is not the case with anaphoric degree-adverbs, which suggests why they make up plausible candidates to use a Correlative endpoint resultant-state clause as an alternative strategy to compensate for this loss.

Type of intensifier	**Intensifier**	**# of *single* intensifier followed by adjective**	**# of *double* intensifier followed by adjective**	$\boldsymbol{double/single}$**(%)**
Anaphoric degree-adverb	*ka'ele/u* 'such$_{PL}$'	18,652	0	0
	kaze 'such$_{SG.M}$'	16,863	2	0.01
	kazot 'such$_{SG.F}$'	3,189	2	0.06
	kol-kax 'so'	1,010	2	0.20
Non-anaphoric degree-adverb	*nora* 'very'	12,736	442	3.50
	meod 'very'	266,781	10,730	4.00
	mamaš 'really'	77,124	3,295	4.30

Table 4.1: Single and reduplicated intensifiers. Data extracted from HeTenTen corpus.

From these results, I conclude that the Correlative endpoint resultant-state clause would be more associated with Anaphoric degree-adverbs exclamatives than with non-exclamatives. This state of affairs predicts that stage II Anaphoric degree-adverb exclamatives hosting the members of the Ultimate construction family would be more frequent than any other alternative of stage II, and therefore make up the critical context

for the change in grammatical status of the members of the Ultimate construction family. This issue is further addressed in the next section (iii).

4.3.3 The frequency of stage II Anaphoric degree-adverb exclamatives with respect to any alternative of stage II (iii)

In order to show the dominance of stage II Anaphoric degree-adverb exclamatives over any other stage II alternatives, I classified the data from IsraBlog corpus for *xaval al hazman*, *en dvarim ka'ele/u*, *en milim* and *ba livkot* into the nine distinct categories I-IV of the proposed model listed in Section 4.2: four stage I contexts (Ia-Id), three stage II contexts (IIa-IIc) and two more contexts, one representing stage III and the other stage IV (see p. 49-51). It will be instructive to examine the various stage II (alternative) contexts against the background of stage I contexts, all presented in Figure 4.3.[49]

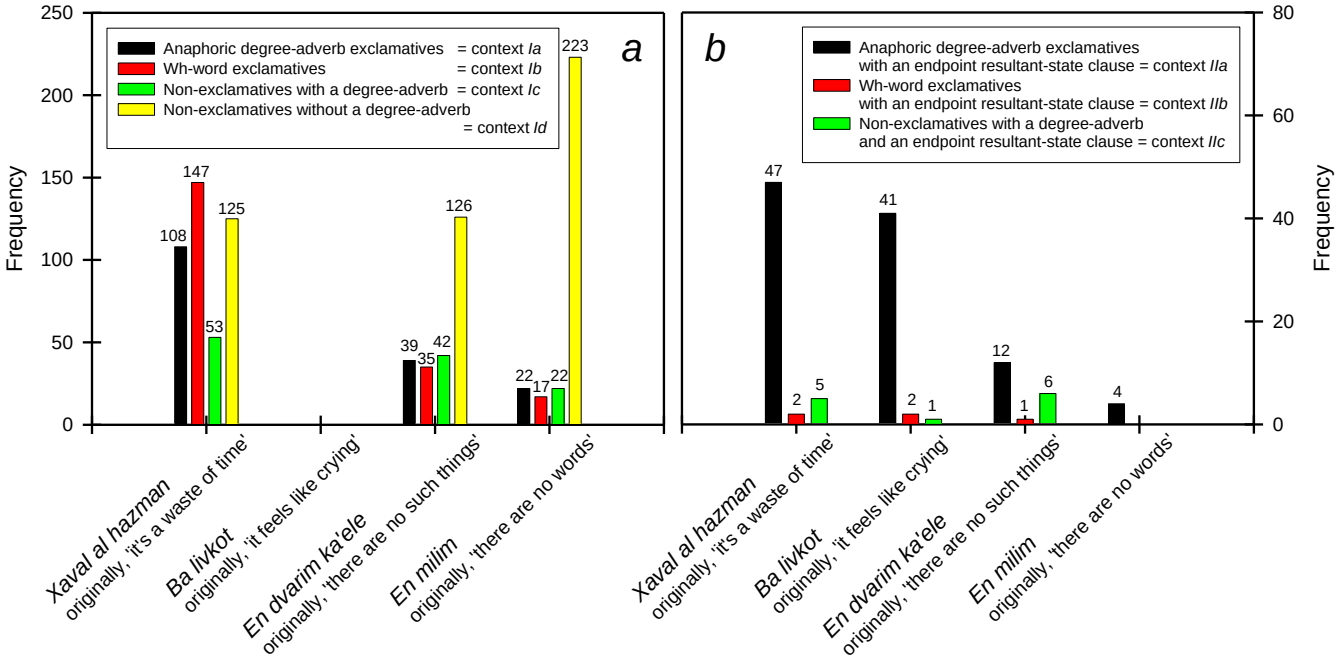

Figure 4.3: A snapshot (as of July 2017) of the frequencies of the various contexts of the Ultimate construction family: (a) when the family members are still independent idiomatic sentences (stage I); (b) when the family members are incorporated into the preceding sentence by means of a Correlative endpoint resultant-state clause (stage II). Data extracted from IsraBlog corpus.

It is reasonable to assume that all four stage I contexts of the model prior to the change in grammatical status, Ia-Id above (see p. 49), affect/trigger/ed the *semantic* change (exemplified, e.g., by the meaning contrast between Example (1.1) and Examples (1.2a-d) above). Each of these contexts — whether exclamative or not — is

[49] Note that due to the Layering principle (Hopper, 1991) items of distinct stages on the evolution path — here stage I items alongside stage II items — co-exist.

a context where a stance-taker conveys a very strong stance with regard to some state of affairs. As a consequence, the frequency of the Anaphoric degree-adverb exclamatives in stage I is not expected to be significantly higher than any of the other contexts.[50] Let's focus on Figure 4.3a and examine the frequency of the Anaphoric degree-adverb exclamatives with respect to the other contexts:

a. In the case of *xaval al hazman*: A χ^2 goodness-of-fit test was performed to examine the difference between the four different contexts and a discrete uniform distribution. Results show that the difference is statistically significant (χ^2 (3, N = 433) = 44.66, $p = 1.1\times10^{-9}$), that is, the frequencies do not follow a discrete uniform distribution. Post-hoc pairwise two-sided binomial tests were conducted to compare the Anaphoric degree-adverb exclamatives (black) and the Wh-word exclamatives (red) groups, the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives with a degree-adverb (green) groups, and the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives without a degree-adverb (yellow) groups, using a Bonferroni-adjusted alpha level of 0.016 (0.05/3). The difference between the black group (42.4% out of the group sum, 95% CI [0.36, 0.48]) and the red group is statistically marginal (n_{black} = 108, n_{red} = 147, p = 0.052), but note that the black group is less frequent than the red one. The difference between the black group (46% out of the group sum, 95% CI [0.40, .53]) and the yellow group is not statistically significant (n_{black} = 108, n_{yellow} = 125, p = 0.88). The difference between the black group (67% out of the group sum, 95% CI [.88, 1]) and the green group is statistically significant (n_{black} = 108, n_{green} = 53, $p = 5.25\times10^{-4}$), but note that the black group is less frequent than the green one, and not the other way around. In no case is the Anaphoric degree-adverb exclamative (black group) significantly more frequent than any of the other three groups (contexts).
b. In the case of *en dvarim ka'ele/u*: A χ^2 goodness-of-fit test was performed to examine the difference between the four different contexts and a discrete uniform distribution. Results show that the difference is statistically significant (χ^2 (3, N = 242) = 132.29, $p < 2.2\times10^{-16}$), that is, the frequencies do not follow a discrete uniform distribution. Post-hoc pairwise two-sided binomial tests were conducted to compare the Anaphoric degree-adverb exclamatives (black) and the Wh-word exclamatives (red) groups, the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives with a degree-adverb (green) groups, and the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives without a degree-adverb (yellow) groups, using a Bonferroni-adjusted alpha level of 0.016 (0.05/3). The difference between the black group (53% out of the group

[50] If any of the four contexts is significantly more dominant than the other, this is expected to be the non-exclamative without any degree-adverb due to its syntactic unmarkedness.

sum, 95% CI [.41, .64]) and the red group is not statistically significant (n_{black} = 39, n_{red} = 35, p = 0.73). The difference between the black group (48% out of the group sum, 95% CI [.37, .60]) and the green group is also not statistically significant (n_{black} = 39, n_{green} = 42, p = 0.82). The difference between the black group (21% out of the group sum, 95% CI [.15, 0.28]) and the yellow group is statistically significant (n_{black} = 39, n_{yellow} = 146, p = 2.85×10^{-15}), but note that the black group is significantly less frequent than the yellow one, and not the other way around. In no case is the Anaphoric degree-adverb exclamative (black group) significantly more frequent than any of the other three groups (contexts).

c. In the case of *en milim*: A χ^2 goodness-of-fit test was performed to examine the difference between the four different contexts and a discrete uniform distribution. Results show that the difference is a statistically significant (χ^2 (3, N = 284) = 434.11, $p < 2.2\times10^{-16}$), that is, the frequencies do not follow a discrete uniform distribution. Post-hoc pairwise two-sided binomial tests were conducted to compare the Anaphoric degree-adverb exclamatives (black) and the Wh-word exclamatives (red) groups, the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives with a degree-adverb (green) groups, and the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives without a degree-adverb (yellow) groups, using a Bonferroni-adjusted alpha level of 0.016 (0.05/3). The difference between the black group (56% out of the group sum, 95% CI [.40, .72]) and the red group is not statistically significant (n_{black} = 22, n_{red} = 17, p = 0.52). The difference between the black group (50% out of the group sum, 95% CI [0.35, 0.65]) and the green group is also not statistically significant (n_{black} = 22, n_{green} = 22, p = 1). The difference between the black group (8.9% out of the group sum, 95% CI [0.06, .13]) and the yellow group is statistically significant (n_{black} = 22, n_{yellow} = 223, p = 4.81×10^{-43}). However, in this case the black group is significantly less frequent than the yellow one, and not the other way around. In no case is the Anaphoric degree-adverb exclamative (black group) significantly more frequent than any of the other three groups (contexts).

d. The case of *ba livkot* is not relevant, since it started out only as a stage II item (see details below).

In sum, no member of the Ultimate construction family here examined — *xaval al hazman*, *en dvarim ka'ele/u*, *en milim* — shows an Anaphoric degree-adverb exclamative (stage I context) which is significantly more frequent than any of the other three contexts. This is in line with my assumption that any type of context which conveys a very strong stance can affect/trigger the semantic change of the members of the Ultimate construction family.

However, if we focus on the frequencies of stage II contexts presented in Figure 4.3b (the stage that — as I propose — affects/triggers/mediates/d the change in grammatical

status), we immediately realize that the Anaphoric degree-adverb exclamative, marked black, quite consistently constitutes the most frequent context here:

a. In the case of *xaval al hazman*: A χ^2 goodness-of-fit test was performed to examine the difference between the three different contexts and a discrete uniform distribution. Results show that the difference is statistically significant (χ^2 (2, N = 54) = 70.34, $p = 5.34\times10^{-16}$), that is, the frequencies do not follow a discrete uniform distribution. Post-hoc pairwise two-sided binomial tests were conducted to compare the Anaphoric degree-adverb exclamatives (black) and the Wh-word exclamatives (red) groups, and the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives with a degree-adverb (green) groups, using a Bonferroni-adjusted alpha level of 0.025 (0.05/2). The difference between the black group (96% out of the group sum, 95% CI [.86, 0.99]) and the red group is statistically significant (n_{black} = 47, n_{red} = 2, $p = 4.36\times10^{-12}$). The difference between the black group (90% out of the group sum, 95% CI [0.79, 0.97]) and the green group is also statistically significant (n_{black} = 47, n_{green} = 5, $p = 1.28\times10^{-9}$). In both cases, the Anaphoric degree-adverb exclamative context (black group) makes up the dominant context.

b. In the case of *ba livkot*: A χ^2 goodness-of-fit test was performed to examine the difference between the three different contexts and a discrete uniform distribution. Results show that the difference is statistically significant (χ^2 (2, N = 44) = 70.955, $p = 3.91\times10^{-16}$), that is, the frequencies do not follow a discrete uniform distribution. Post-hoc pairwise two-sided binomial tests were conducted to compare the Anaphoric degree-adverb exclamatives (black) and the Wh-word exclamatives (red) groups, and the Anaphoric degree-adverb exclamatives (black) and the non-exclamatives with a degree-adverb (green) groups, using a Bonferroni-adjusted alpha level of 0.025 (0.05/2). The difference between the black group (95% out of the group sum, 95% CI [.84, .99]) and the red group is statistically significant (n_{black} = 41, n_{red} = 2, $p = 2.15\times10^{-10}$). The difference between the black group (90% out of the group sum, 98% CI [.87, 1.0]) and the green group is also statistically significant (n_{black} = 41, n_{green} = 1, $p = 1.96\times10^{-11}$). In both cases, the Anaphoric degree-adverb exclamative context (black group) makes up the dominant context.

c. In the case of *en dvarim ka'ele/u*: A χ^2 goodness-of-fit test was performed to examine the difference between the three different contexts and a discrete uniform distribution. Results show that the difference is statistically significant (n_{black} = 12, n_{red} = 1, n_{green} = 6, p = 0.005), that is, the frequencies do not follow a discrete uniform distribution. Post-hoc pairwise two-sided binomial tests were conducted to compare the Anaphoric degree-adverb exclamatives (black) and the Wh-word exclamatives (red) groups, and the Anaphoric degree-adverb exclamatives (black) and the non-

exclamatives with a degree-adverb (green) groups, using a Bonferroni-adjusted alpha level of 0.025 (0.05/2). The difference between the black group (92% out of the group sum, 95% CI [.64, 1.0]) and the red group is statistically significant (n_{black} = 12, n_{red} = 1, p = 0.003). The difference between the black group (90% out of the group sum, 98% CI [0.41, 0.87]) and the green group is not statistically significant (n_{black} = 12, n_{green} = 6, p = 0.238).[51]

d. In the case of *en milim*, there are too few tokens to run a statistical analysis (N = 4). Nevertheless, all of them are of the Anaphoric degree-adverb exclamative type.

From these results we can conclude that stage II Anaphoric degree-adverb exclamatives (IIa) are significantly more frequent than any alternative of stage II, regardless of (their) lack of dominance as stage I contexts, almost without an exception. These results provide further support for the issue addressed in Section 4.3.2 above — the Correlative endpoint resultant-state clause is indeed associated with Anaphoric degree-adverb exclamatives more than with any other kind of exclamatives or non-exclamatives. This is so, presumably because it is the sole means to compensate for the loss of emotive force of anaphoric degree-adverbs. It therefore stands to reason that the Anaphoric degree-adverb exclamatives constitute the context that affects/triggers/ mediates/d the change in grammatical status of the members of the Ultimate construction family by incorporating them into the slot of the Correlative endpoint resultant-state clause. In the next section I address issue (iv).

4.3.4 The timeline of Anaphoric degree-adverb exclamatives (IIa) and their counterparts lacking the anaphoric degree-adverb (III) (iv)

I now need to show that stage II Anaphoric degree-adverb exclamatives (IIa) showed up chronologically earlier than sentences lacking the anaphoric degree-adverb (III). To this end, I had to switch to Yedioth Ahronoth corpus, where each token of *xaval al hazman* is tagged for date of production. Although IsraBlog corpus is also tagged for date of production, it was not adequate for this specific task, because the semantic/functional (and possibly also the grammatical) change of *xaval al hazman* occurred during the nineties of the twentieth century, as evidenced from Yedioth Ahronoth data presented in Figure 4.4 below, whereas IsraBlog corpus was only launched towards the end of 2001. Note that in Figure 4.4 the 'old meaning' of *xaval al hazman* is the negative (almost) compositional meaning exemplified in (1.1). The 'new meaning' is the modifying, mostly positive one conveying a highly intense evaluation,

[51] I here switched to a multinomial test instead of a χ^2 goodness-of-fit test, because the latter requires that the expected value would be larger than 5. But here, under the assumption of a discrete uniform distribution, the expected value is ~5 ($N * p_i = (12 + 1 + 6) * \frac{1}{3}$).

as exemplified in (1.2). (See also Appendix A for Ariel's, 2017 suggestion as to the inferential steps leading to this semantic change.)

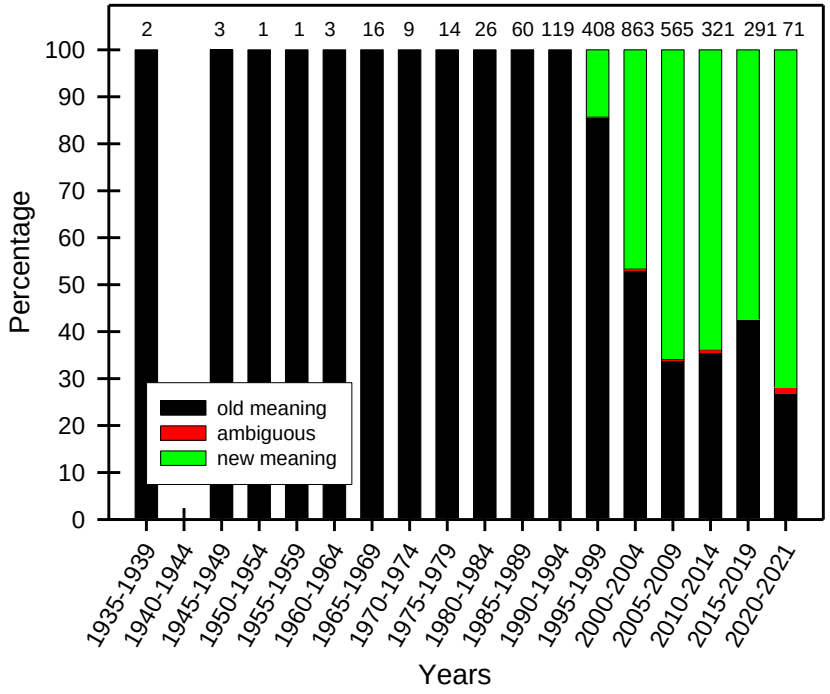

Figure 4.4: The distribution of the old versus the newly evolved meaning(s) of *xaval al hazman* as a function of time. The number of counts for every five-year interval is marked on top of each bar.[52] Note that items of different grammatical statuses, as well as items which have not yet changed their grammatical status (the latter preceded by *še* 'that'), are considered together. Data extracted from Yedioth Ahronoth corpus.

The earliest example of the newly evolved *xaval al hazman* in Yedioth Ahronoth corpus — Example (4.4) — is indeed an instance of an independent idiomatic sentence (stage I). It is a highly positive review of an LP that was released by a well-known singer.

(4.4)	***xaval al hazman.***	*'xam*	*al*	*ha-yare'ax'*	*hu*
	it's a of waste time → outstanding	hot	on	the-moon	is

*yecirat^*	*mofet.*
artwork	exemplary

[52] I assume that the change in total counts is possibly a result of change in the overall number of tokens. Since the overall number of tokens is not available, this assumption is based on the changed number of articles printed in Yedioth Ahronoth over the years.

‘**Outstanding**. ‘Hot on the moon’ [the name of an LP – IB] is a masterpiece.’
(December 1995, Yedioth Ahronoth)

The next example — Example (4.5) — is an instance of an intra-sentential ***še-xaval al hazman*** ’**that** it’s a waste of time’ which appeared in an excerpt from a highly positive restaurant review, three years later than the previous example (Example 4.4). *Xaval al hazman* here functions as a noun modifier. Note, however, that *xaval al hazman* is here already incorporated into a preceding sentence which is not an Anaphoric degree-adverb exclamative, context IIa, but rather a stage III context.

(4.5) *tuna* *na’a* *be-taxmic* ***še-xaval al hazman***.
tuna raw in-marinade **that-it’s a waste of time → amazing**

‘[…], **amazing** raw tuna in marinade, […]’[53]
(June 1998, Yedioth Ahronoth)

In the next example — Example (4.6) — *xaval al hazman* is incorporated into the preceding sentence without the scaffolding *še* ‘that’, functioning as a full-fledged adjective (stage IV). This example postdates the previous one (Example 4.5) only by a month.

(4.6) *šlomo* *arci* *kotev* *širim* ***xaval al hazman***.
Shlomo Artsi writes songs **it’s a waste of time → amazing**

‘[…], Shlomo Artzi writes **amazing** songs, […]’[54]
(July 1998, Yedioth Ahronoth)

No trace of any Anaphoric degree-word exclamative (context IIa) is to be found in this corpus up until more than two years later, in an introduction to a recipe, see Example (4.7).

(4.7) *lehalan* *matkon* *kaley-kalut* *ve-**kol-kax*** *ta’im*
the.following recipe easy-peasy and-**so** delicious

še-xaval al hazman.
that-it’s a waste of time → amazingly

[53] Note that alternatively, *xaval al hazman* can be interpreted here as an adjective modifying the NP *taxmic* ‘marinade’.

[54] Note that alternatively, *xaval al hazman* can be interpreted here as a manner adverb modifying the VP *kotev širim* ‘writes songs’.

‘Here is the easiest recipe and **so amazingly** delicious.’

(December 2000, Yedioth Ahronoth)

The data in Examples (4.4)-(4.7) seem to challenge the model I proposed in Section 4.2, because a stage II context postdate both stage III and stage IV items. But it may also indicate that Yedioth Ahronoth corpus is not entirely adequate for the task. After all, exclamatives, which, as I claimed, affect/trigger/mediate/d the change in grammatical status, are instances of spoken informal speech, and Yedioth Ahronoth, a corpus of journalistic writing, represents a (more) formal (and edited) register (Rubinstein, 2019). I therefore turned again to the informal web-based IsraBlog corpus. As explained above, I could not here examine data relevant to *xaval al hazman*, since its semantic/functional and grammatical changes occurred during the nineties of the twentieth century, as attested by data from Yedioth Ahronoth corpus (see Figure 4.4), whereas IsraBlog corpus was only launched in late 2001. I therefore examined another, later member of the Ultimate construction family, the functionally similar *en dvarim ka’ele/u* (see Example 4.3a above). The semantic/functional change of *en dvarim ka’ele/u* occurred sometimes during the first decade of the twenty-first century as evidenced from Yedioth Ahronoth data presented in Figure 4.5 below. This period is well represented in the informal IsraBlog corpus. Note that the “old meaning” of *en dvarim ka’ele/u* is the literal meaning referring to an object which does not exist, either locally or globally, namely ‘there are no such things’. The “new meaning” is the modifying, mostly positive one conveying a highly intense evaluation, as exemplified in Appendix C.10.

I therefore extracted all instances of *en dvarim ka’ele/u* in IsraBlog corpus, and classified them into the nine distinct categories I-IV of the proposed model listed in Section 4.2 (see p. 49-51). No instance of context IIc items (non-exclamatives with a degree word) was found, and there were only three instances of stage IV *en dvarim ka’ele/u*. I here consider together (1) all cases of *en dvarim ka’ele/u* as an independent idiomatic sentence preceding or following contexts expressing a speaker’s strong stance (stage I); (2) the Anaphoric degree-adverb exclamatives with a Correlative endpoint resultant-state clause (context IIa); and (3) non-exclamatives with a Correlative endpoint resultant-state clause (stage III).

I tagged each example according to both year and week of its production rather than just the year, in order to make sure I did not miss any rapid change. The distribution of the variable ‘Date of production’ is crucial for choosing the appropriate statistical method. So a Shapiro-Wilk test was performed and showed that indeed the distribution of stage II and stage III items is normal (stage II: $W(12) = 0.906$, $p = 0.19$; stage III: $W(49) = 0.958$, $p = 0.079$). However, the distribution of stage I items is not normal ($W(242) = 0.898$, $p = 9.37\times10^{-12}$). (In addition, no outliers were found for any group.) In such a case, the appropriate test is a Kruskal-Wallis test (Kruskal-Wallis H test), which shows that there is a statistically significant difference between the groups ($H(2, N = 303) = 23.8$, $p = 6.9\times10^{-6}$) with a mean rank of 140.0 for the date of production of stage I items, 181.2 for the date of production of stage II items, and 204.8 for the date

of production of stage III items. A moderate effect size was detected, $\varepsilon^2(H) = 0.073$, 95% CI [0.03, 0.13]. Additionally, post-hoc pairwise Mann-Whitney tests were conducted to compare all group pairs, using a Bonferroni-adjusted α level of 0.0166 (0.05/3). These tests show that,

a. The median of stage I items (*Mdn* = 2009.33) is significantly lower from the median of stage III items (*Mdn* = 2012.08), $U(n_{stageI} = 242, n_{stageIII} = 49) = 3386$, $p = 6.6\times10^{-6}$, 95% CI [-2.5, -1.0];
b. the median of stage I items (*Mdn* = 2009.33) is not significantly different from stage II items (*Mdn* = 2010.33), $U(n_{stageI} = 242, n_{stageII} = 12) = 1058$, $p = 0.339$, 95% CI [-2.42, 0.167];
c. the median of stage II items (*Mdn* = 2010.33) is not significantly different from stage III items (*Mdn* = 2012.08), $U(n_{stageII} = 12, n_{stageIII} = 49) = 250$, $p = 1.0$, 95% CI [-2.58, 1.0].

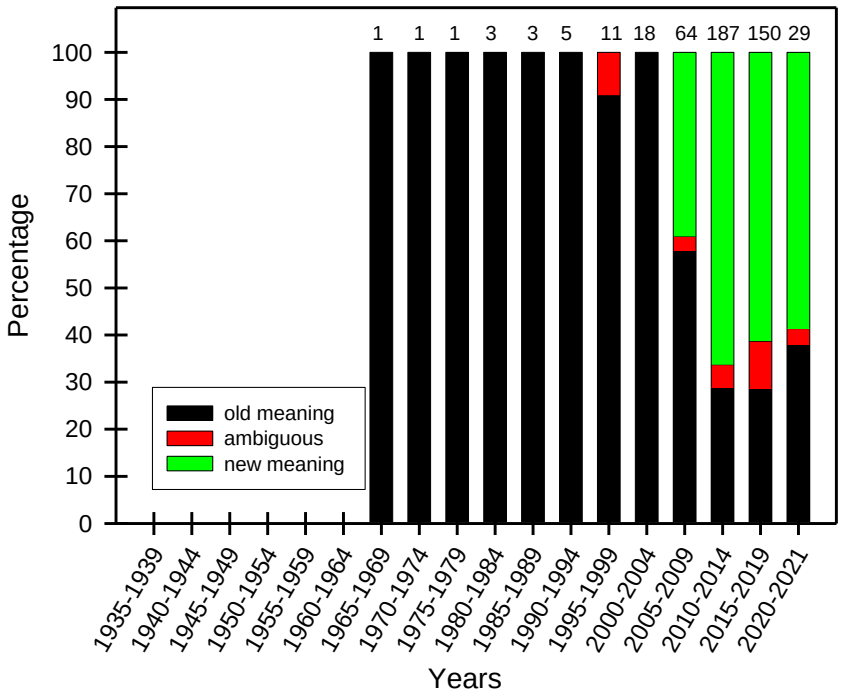

Figure 4.5: The distribution of the old versus the newly evolved meaning(s) of *en dvarim ka'ele/u* 'there are no such things' as a function of time. The number of counts for every five-year interval is marked on top of each bar.[55] Note that items of different grammatical statuses, as well as items which have not yet changed their grammatical status (the latter preceded by *še* 'that'), are considered together. Data extracted from Yedioth Ahronoth corpus.

[55] I assume that the change in total counts is possibly a result of change in the overall number of tokens. Since the overall number of tokens is not available, this assumption is based on the changed number of articles printed in Yedioth Ahronoth over the years.

These results are depicted in Figure 4.6 below, and may point to the timeline of the evolution of *en dvarim ka'ele/u*, and by implication, possibly of *xaval al hazman* as well. In particular, these data suggest that Anaphoric degree-word exclamatives with a Correlative endpoint resultant-state clause such as *She is **so** impressive **še-xaval al hazman*** (IIa), showed up chronologically earlier than non-exclamatives with a Correlative end point resultant-state clause as in *She is ∅ impressive **še-xaval al hazman*** (III).[56] (An alternative analysis of these data, represented by an association plot, produced identical results. It is presented in Appendix D.1.)

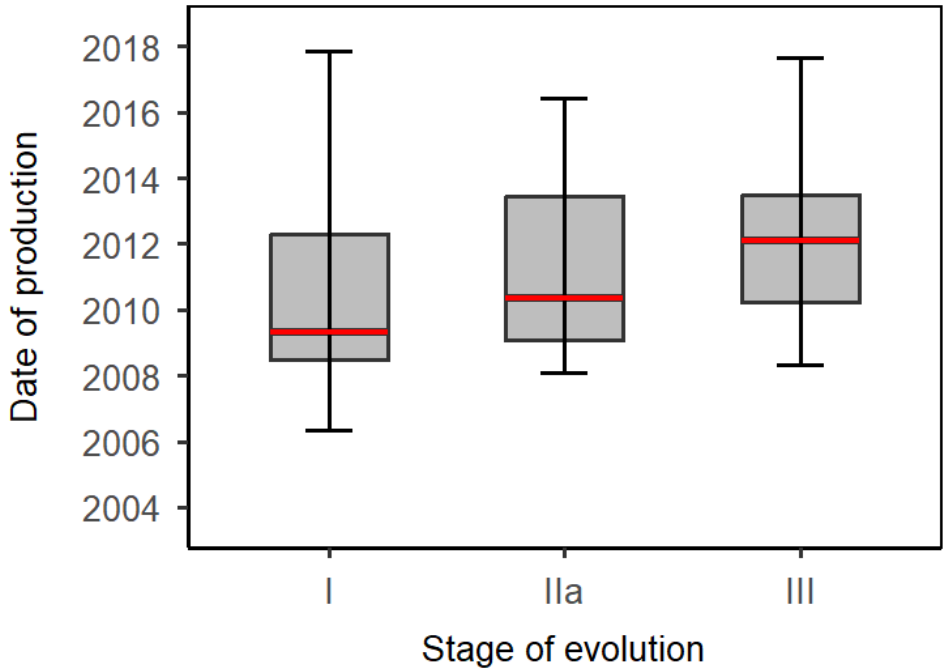

Figure 4.6: Box plot representing the timeline of evolution of *en dvarim ka'ele/u*: from an independent idiomatic sentence (I), through an idiomatic sentence integrated into the preceding Anaphoric degree-adverb exclamative by means of a Correlative endpoint resultant-state clause (IIa), and then an idiomatic sentence integrated into the preceding, non-exclamative sentence with no degree-adverb, again by means of a Correlative endpoint resultant-state clause (III). The red horizontal lines indicate medians. Data extracted from IsraBlog corpus.

With the data presented in Sections 4.3.1–4.3.4, I can now provide an answer to the first question posed in Section 4.1 above:

[56] All four stage I groups were collapsed due to theoretical considerations (see the proposed model in Section 4.2). Note that this procedure is line with the results of a Kruskal-Wallis test which shows that there is no statistically significant difference between any two of stage I contexts ($H(3, N = 242) = 0.71$, $p = 0.87$).

Qi: What exactly is the context that mediates/d the change in grammatical status from an extra-sentential modifying idiomatic sentence to a modifier — an adjective, an adverb or an intensifier — of a single distinct constituent within the boundaries of a sentence?

Ai: The recurrent Anaphoric degree-adverb exclamative is what affects/ triggers/mediates/d the change in grammatical status of the members of the Ultimate construction family from extra-sentential to intra-sentential elements. This Anaphoric degree-adverb exclamative often invokes a Correlative endpoint resultant-state clause, designed to further reinforce the fading extreme propositional content of the exclamative. In fact, the Correlative endpoint resultant-state clause is virtually the only reinforcement strategy available here. Note that *functionally/ semantically*, the members of the Ultimate construction family have turned into (opaque) idiomatic sentences already when commenting on an adjacent, prior sentence, when they still constituted an independent utterance. But syntax often lags behind semantics. Thus, *syntactically*, the members of the Ultimate construction family maintained their sentential status. As such, they fit perfectly the slot of the Correlative endpoint resultant-state clause (invoked by the Anaphoric degree-adverb exclamative), which requires a clause depicting an endpoint.

4.3.5 A further look at the claim about a strong association between the members of the Ultimate construction family and the Anaphoric degree-adverb exclamatives

Before concluding this section, I would like to re-examine the association between the members of the Ultimate construction family and the Anaphoric degree-adverb exclamatives. This association has been claimed to be strong, and consequently responsible for the change in grammatical status of the members of the Ultimate construction family. If indeed the members of the Ultimate construction family and the Anaphoric degree-adverbs exclamatives are so strongly associated, then the transition from stage I to stage II would be equally plausible for all members of the Ultimate construction family. But if we take a closer, more critical look at the rightmost end of Figure 4.3b, we realize that *en milim* is hardly associated with the Anaphoric degree-adverbs exclamatives, four instances in total. This is a disturbing finding, which may argue against my proposed model in Section 4.2.

This finding is even more peculiar in light of the diachrony of *en milim*. Consider Example (4.8), the earliest instance of *en milim* as an idiomatic sentence spotted in Yedioth Ahronoth corpus. In this example, which dates back to 1973, *en milim* is uttered by the then Israeli prime-minister Golda Meir at a welcome home reception for Israeli soldiers who had spent four years in captivity. The speaker is obviously very emotional as evidenced from the tears in her eyes.

(4.8) *pašut,* ***en milim,*** *hi* *omeret* *u-dma'ot*
simply **there are no words** she says and-tears

nikvot *be-eneha.*
well.up in-her.eyes

'"**There are** simply **no words"**, she says and tears well up in her eyes.'
(Yedioth Ahronoth)

Clearly, this use of *en milim* emerged a little more than 20 years earlier than *xaval al hazman* and 35 years earlier than *en dvarim ka'ele/u*. Still the data in Figure 4.3b indicate that *en milim* has hardly been incorporated into the Anaphoric degree-adverb exclamatives.

A possible reason would be the position of stage I independent *en milim* relative to the context that it evaluates. Perhaps the independent *en milim* has a preference for preceding rather than following the context that it evaluates (for some reason). If this is the case, then the transition from stage I to stage II — from a paratactic sequence to a hypotactic one — would not be as natural as suggested in Section 4.2. This is, however, not the case as evidenced from Figure 4.2 above. The vast majority of the independent *en milim* follows the context that it evaluates.

The patterning of *en milim* is even more puzzling when compared with *xaval al hazman*. *Xaval al hazman* and *en milim* show similar rates of stage I items following—rather than preceding—the context that they evaluate, 84.3% ($^{365}/_{433}$) and 83.1% ($^{236}/_{284}$), respectively, as shown in Figure 4.2. A Chi-square test of homogeneity shows that there is no significant difference between them, χ^2 (1, N = 717) = 0.104, p =.75, 95% CI [-0.070, 0.46].[57] Why is it, then, that *en milim* which exhibits similar initial conditions to those of *xaval al hazman*, doesn't proceed from stage I to

[57] In this context it is worth noting that *en dvarim ka'ele/u* shows the highest rate of stage I items following—rather than preceding—the object that they evaluate, 94.2% ($^{228}/_{242}$), as compared to its counterparts, 84.3% ($^{365}/_{433}$) and 83.1% ($^{236}/_{284}$) for *xaval al hazman* and *en milim*, respectively. A Chi-square test of homogeneity shows there is a significant association between the kind of Ultimate construction used and its inclination to follow the object that it evaluates , χ^2 (2, N = 959) = 19.8, p = 5.1×10^{-5} (φ = 0.64, a large effect size). *En dvarim ka'ele/u* is responsible for this significant difference. Post-hoc pairwise two-sided Chi-square tests of homogeneity were conducted to compare *en dvarim ka'ele/u* and each of *xaval al hazman* and *en milim* using a Bonferroni-adjusted alpha level of 0.025 (0.05/2). The difference between *en dvarim ka'ele/u* and *xaval al hazman* is statistically significant (χ^2 (1, N = 675) = 17.0 , p = 3.8×10^{-5}, 95% CI [0.061, 0.171]). The difference between *en dvarim ka'ele/u* and *en milim* is also statistically significant χ^2 (1, N = 526) = 15.9 , p = 6.8×10^{-5}, 95% CI [0.057, 0.150]). But this state of affairs is not at all surprising, since *ka'ele/u* 'such as these' is anaphoric *in se*, and as such it is expected to follow predominantly its antecedent. The few cases of cataphoric *en dvarim ka'ele/u* are headlines which serve another pragmatic purpose, curiosity arousal (Kronrod & Engel, 2001).

stage II (i.e., incorporated into the preceding sentence, the Anaphoric degree-adverb exclamative), whereas *xaval al hazman* does?

I suggest that the evolution of *en milim* remained stagnant as a result of *constructional competition* (e.g., Berg, 2014; De Smet, D'hoedt, Fonteyn, & Van Goethem, 2018; Sommerer, 2020; Van Goethem, Vanderbauwhede, & De Smet, 2018). *En milim* embedded in the Anaphoric degree-adverb exclamative is competed by another construction on its local *constructional network*, as I explain in the next sections.

4.3.5.1 Constructional network and constructional competition

As briefly stated in Section 1.3.1, any construction in the Construct-i-con is linked to other constructions. Chapter 3 builds on the *vertical* (inheritance) links between the members of the Ultimate construction family and more abstract dominating constructions on the hierarchy, the Evaluative P1 S-pattern and the P1 S-pattern (see Figure 3.4).

Constructions are also linked to each other by *horizontal* links. Horizontal links can link constructions that share both their form and meaning as a consequence of being licensed by a common dominating construction(s) (and are therefore related to one another through inheritance), just like the various members of the Ultimate construction family which are all licensed by the Evaluative P1 S-pattern. But horizontal links can also link constructions that share either their form or their meaning, as a consequence of being licensed by more abstract dominating constructions, this time distinct from one another (e.g., Cappelle, 2006; Perek, 2012; Sommerer, 2020; Ungerer, 2021, 2024; Van de Velde, 2014; Zehentner & Traugott, 2020). In this case, the related constructions are considered "neighbors" (Diessel, 2023: 71-74).[58] Constructions having this kind of relations have been termed *degenerate* constructions (Van de Velde, 2014) or *allostructions* (Cappelle, 2006).[59] Allostructions have been shown to have psychological reality. Speakers are aware that constructions considered as allostructions "have the same 'descriptive' meaning, i.e. that they can be used to describe the same set of situations" (Perek, 2012: 606). Following Goldberg (1995: 91), I refer to this specific kind of horizontal link a 'semantic synonymy' link.[60]

The semantic synonymy between allostructions may bring about constructional competition (e.g., Berg, 2014; De Smet et al., 2018; Sommerer, 2020; Van Goethem et

[58] Cappelle (2006: 25) suggested a middle way between these two diametrically opposite categories — "two (or more) different formal versions of one and the same underspecified pattern". So did Perek (2012) who dubbed this higher level underspecified alternation-based pattern 'constructeme'. The underspecified patterns suggested by Cappelle and Perek seem to generalize over the constructions that they examine rather neatly. But this is not always the case.

[59] While Van de Velde (2014: 173) asserts that "degeneracy mostly consists of many-to-many relationships between form and meaning", he does not preclude the many-to-one relationship between form and meaning.

[60] Note that there seem to be horizontal links between constructions which are not (near-) synonymous, and cannot be considered allostructions (Ungerer, 2021).

al., 2018). This competition may end up by the omnipresence of one allostructions at the expense of the other(s). This is exactly what happens in the case of an Anaphoric degree-adverb exclamative specifically hosting *en milim* and an alternative sentence-level construction, an *allosentence* (a term introduced by Daneš, 1966 and later on used by Lambrecht, 1994). The competing allosentence has the upper hand, reducing the frequency of the Anaphoric degree-adverb exclamative hosting *en milim*, thus slowing down the change in the grammatical status undergone by *en milim* (from a grammatical sentence to a word). This state of affairs is described in the next section.

4.3.5.2 The allosentence competing with the Anaphoric degree-adverb exclamative hosting en milim

When I examined the data which pertains to *en milim* in IsraBlog corpus, a potential allosentence of the Anaphoric degree-adverb exclamatives hosting *en milim* has stood out. The said allosentence has a long history, even from before it could have been considered an allosentence of any other sentence. Example (4.9) — a lamentation over a close friend from 1865 — is representative.

(4.9)	*en*	*milim*	*be-fi*	*leta'er*	*ecvoni*	*al*
	there's/are.no	words	in-my.mouth	to.describe	my.agony	on

*mot^*	*yedid^*	*ne'uray.*
death	friend	my.youth

'I have no words to describe my agony over the death of my boyhood friend.'

(Historical Jewish Press)

In this example *en milim* is followed by two consecutive (and interchangeable) slots, the one slot hosting the word *be-fi* 'in my mouth' (among other alternatives such as *ba-pe* 'in the mouth', *be-finu* 'in our mouth', *be-fiv* 'in his mouth', *be-šum safa* 'in no language'); the other slot hosting an infinitival clause such as *leta'er* 'to describe' or an equivalent 'that'-clause, e.g., *še-yeta'aru* 'that will describe' (among other options such as *levate* 'to express', *lehabi'a* 'to convey', *lehasbir* 'to explain', *lehagdir* 'to define', *lecayen* 'to note', *lehagid* and *lomar* 'to say'). When these verbs team up with *en milim*, they convey the speaker's evaluation, specifically the heightened emotional state of mind of the speaker with respect to some stance-object. The said stance-object is the NP in the scope of the infinitival *leta'er* 'to describe' (and similar verbs), *ecvoni* 'my agony' in Example (4.9). In fact, sentences like Example (4.9) are presumably the source of the bare *en milim*.

Now, Example (4.9) and similar examples can be deemed exclamatives. Within the "many-to-one mapping of form to function" makeup of the Exclamative sentence (super) construction (Michaelis, 2001: 1041), which is a heterogeneous category of sentences within and across languages (Michaelis, 2001) — see Figure 4.7 below —

these exclamatives roughly fall under the category which Michaelis (2001) dubbed "complementation structures involving factive epistemic matrix verbs". Michaelis (2001) and Michaelis and Lambrecht (1996) mentioned two potential matrix predicates — *I can't believe* and *I'm amazed at*. These matrix predicates induce an exclamative reading in Hebrew too. I suggest one more matrix predicate, not mentioned in the literature, that does the very same thing — *en milim (befi) leta'er* 'there are no words (in my mouth) to describe'. The matrix predicate conveys the speaker's surprise at some non-canonical situation involving some referent in the scope of the matrix verb, implying that the presupposed scalar property associated with that referent exceeds her expectations. She is therefore speechless.

I suggest that as an exclamative, any matrix clause in the form of *en milim (befi) leta'er* scoping over a stance-object is horizontally linked to an Anaphoric degree-adverb exclamative hosting *en milim* via semantic synonymy links, marked turquoise in Figure 4.7. The two are indeed formally distinct, since the former inherits its formal properties from the P1 S-pattern whereas the latter from the S1 S-pattern (see Figure 3.3, and 3.1 and 3.2, respectively). But they are semantically and discoursally equivalent. They may therefore be considered allosentences and may well serve as constructional competitors.

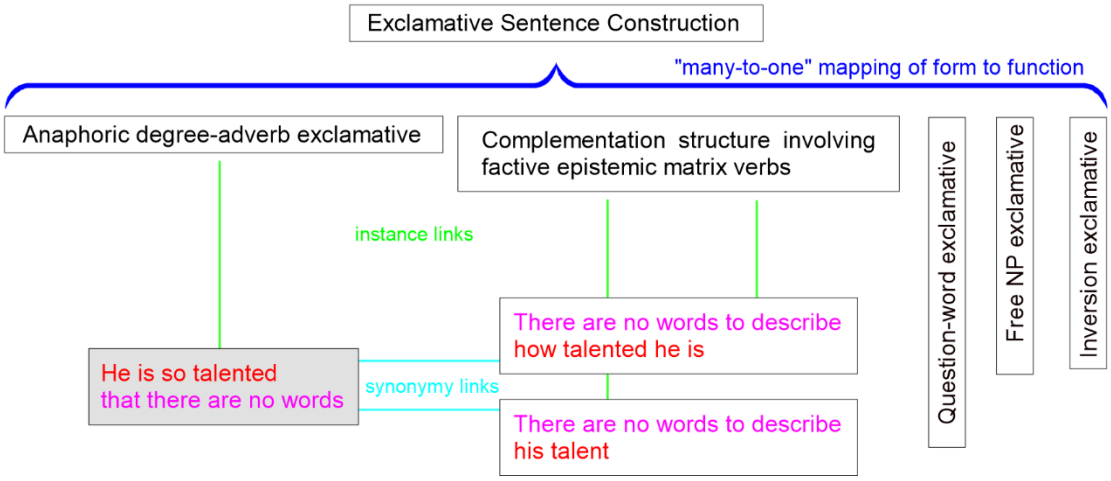

Figure 4.7: The local network of the Anaphoric degree-adverb exclamative hosting *en milim* (shadowed)

If we examine the frequencies of the bare *en milim* and *en milim (befi) leta'er* in Yedioth Ahronoth corpus, see Figure 4.8, we realize that, by and large, the frequency of *en milim* changes over the years, but so does the frequency of *en milim (befi) leta'er* (presumably as a result of the change in the overall number of tokens in Yedioth Ahronoth). In no case is *en milim* (red bars) favored by speakers more than *en milim (befi) leta'er* (black bars). They are either equally favored, or *en milim (befi) leta'er* is significantly more favored than its counterpart. In Appendix E, I present a similar

analysis where the *en milim (befi) leta'er* set subsumes several other verbs which are interchangeable with *leta'er* 'to describe' — *levate* 'to express', *lehabi'a* 'to convey', *lehasbir* 'to explain', *lehagdir* 'to define', *lehagid* and *lomar* 'to say'. In Figure E1, no single time interval shows a significantly higher frequency of *en milim* as compared to *en milim (befi) leta'er*, just like in Figure 4.8.[61]

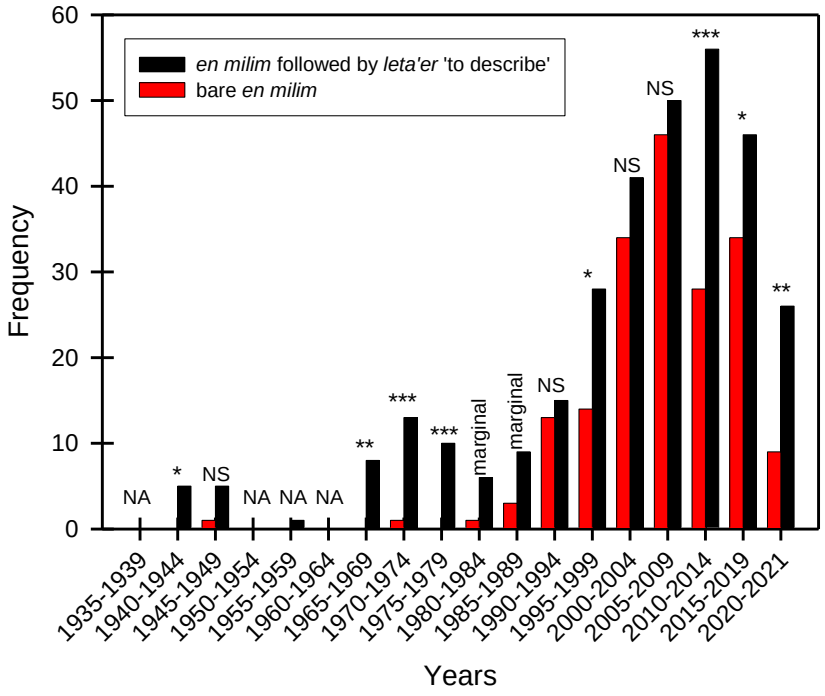

Figure 4.8: The frequencies of bare *en milim* and *en milim (befi) leta'er* over time. The level of significance for each five-year interval (which is the outcome of a binomial test) is marked above each pair of bars. '*' = $p < 0.05$, '**' = $p < 0.01$; '***' = $p < 0.001$; 'NA' = Not Applicable; 'NS' = Not Significant. Data extracted from Yedioth Ahronoth corpus.

These results attest to the status of the *en milim (befi) leta'er* scoping over a stance-object as a potential competitor to the Anaphoric degree-adverb exclamative hosting *en milim*. Not only are they both semantically and discoursally equivalent (by virtue of being exclamatives), but they are also compact in the sense that the stance-object and the element that expresses the speaker's stance are both forced into a single sentence. Recall that a possible motivation for moving on from stage I to stage II is complexity building via compactization, as mentioned in Section 4.1. As a potential competitor, *en*

[61] Note that in this analysis I consider *en milim* any case of *en milim* — whether an independent idiomatic sentence (i.e, stage I) or *en milim* hosted by Anaphoric degree-adverb exclamative. If only the latter was considered, then there would only be nine relevant examples overall, only three of which host the bare *en milim*.

milim (befi) leta'er scoping over a stance-object construction may have blocked the coming into being of the Anaphoric degree-adverb exclamative hosting *en milim* (stage II) which is the stepping stone to the next stage of this evolution process. In other words, a critical mass of Anaphoric degree-adverb exclamatives hosting *en milim* items, required to drive this evolution process towards stage III (and onwards), failed to build up.

This kind of blocking is reminiscent of the phenomenon of lexical statistical preemption of new forms and patterns by extant ones (Goldberg, 2006: Ch. 5) which — quite like the case presented here — is a function of frequency (e.g., Bybee, 2006) and possibly of social conventions (e.g., Traugott & Trousdale, 2013: 206).

There could also be blocking on the part of *xaval al hazman* and *en dvarim ka'ele/u* hosted by the Anaphoric degree-adverb exclamative. It's not impossible that *xaval al hazman* and *en dvarim ka'ele/u* appropriated the slot of the Correlative endpoint resultant-state clause, making it difficult for other members of the Ultimate construction family, such an *en milim*, to get in.

My *pro forma* attempts to look for a conspicuous counterpart of the new *xaval al hazman*, which together with *xaval al hazman* hosted by the Anaphoric degree-adverb exclamative may have formed a pair of competing allosentences, were unsuccessful. This is, of course, not in the least surprising in light of the meteoric evolution of *xaval al hazman*.

As for *en dvarim ka'ele/u* — because of its inherent anaphoric nature, it tends to follow its stance-object rather than precede it significantly more than *xaval al hazman* and *en milim*, (see Figure 4.2 and the analysis in fn. 57). This, of course, can account for the lack of any competitor for *en dvarim ka'ele/u* where *en dvarim ka'ele/u* scopes over a stance-object that follows it.

As for *ba livkot* — even if it had an allosentence, I would have predicted that the Anaphoric degree-adverb exclamative hosting *ba livkot* would have had the upper hand. This is because *ba livkot* appeared on the language scene as a stage II item, an element tightly linked with the Anaphoric degree-adverb exclamative, mentioned for the first time in a popular song from the late-sixties of the twentieth century. [62] In that song, a womanizer confesses that he is emotionally moved by the presence of attractive women so much that he feels like bursting into tears: *kol-kax yafot še-ba livkot* 'so beautiful up to a point that it feels like crying'. The iconic status of this entire syntagma *kol-kax yafot še-ba livkot*, may have blocked the rise of any competitor.

It seems that so far, the model I have proposed establishes firmly the context associated with the change in grammatical status of the members of the Ultimate construction family, by taking into account several relations and interactions in the Construct-i-con. In the next section I attempt to provide answers to the second and third questions posed in Section 4.1 of this chapter.

[62] The lyrics of *yafot, yafot*: http://tinyurl.com/3898mhpf

4.3.6 The flexibility of the newly evolved modifiers and the timeline of emergence of adjectives, adverbs and intensifiers

The remaining questions posed in Section 4.1 concern the flexibility of stage IV items of the model that I proposed:

Qii: What is it that stimulates/d the flexibility of the newly evolved modifiers?

Qiii: Which (amplifying) intra-sentential element — an adjective, an adverb or an intensifier — is/was the first to evolve? And is there any clear line of trans-categorization?

I have argued above that the Anaphoric degree-adverb exclamatives provide the mediating context for the change in grammatical status of the idiomatic sentences (the members of the Ultimate construction family which are still independent sentences) to words. Specifically, it is the recurrent accompanying Correlative endpoint resultant-state clause, which is interpreted as reinforcing the anaphoric degree-adverb (and the entire exclamative). The Anaphoric degree-adverb exclamatives are instantiations of the S1 S-patterns (see Figures 3.1 and 3.2). As such, they have different foci. In the case of the Copular S1 S-pattern, the foci — dubbed the 'Assigned Term' by Kuzar (2012) — are an adjective in Examples (4.1a) and (4.13c) and a noun in Example (4.1d) (repeated here for convenience). In the case of the Verbal S1 S-pattern, the focus is a verb in Example (4.1b).

(4.1) a. *hu* ***kol-kax*** muxšar!
he **so** talented

'He is so talented!'

b. *hu* ***kol-kax*** caxak!
he **so** laughed

'He laughed so much!'

c. *hu* ***kaze*** muxšar!
he **such** talented

'He is that talented!'

d. *hu* ***kaze*** baxur!
he **such** a.young.man

'He is such a young man!'

All of these foci may be assigned positions on a relevant scale (for the case of nouns, see Fillmore et al., 1988). This is why all of these foci can be amplified, regardless of their word class, by means of the Correlative endpoint resultant-state clause.

The members of the Ultimate construction family, on their part, are potential 'flexible modifiers' (McNabb, 2012; Salazar-García, 2010). A flexible modifier is a modifier of flexible semantics which can modify gradable and non-gradable properties alike, as well as individuals, situations or propositions, and therefore exhibits a wide syntactic distribution. This has been shown to apply to Hebrew *mamaš* 'really' (Bardenstein & Ariel, 2022; McNabb, 2012) and *legamrey* 'completely' (Bardenstein & Ariel, 2022; Shaviv, 2018), English *–ass* (Irwin, 2015), and Italian suffix *–issimo* (Beltrama & Bochnak, 2015). Beltrama and Bochnak, for example, who studied the wide categorial distribution of the suffix *–issimo* in Italian, argued that in the case of *–issimo*, this distribution is a result of its expressive layer, that is, the speaker's "heightened emotional status about the content of the proposition" (p. 847). The Italian *–issimo* is a case of *pragmatic*—not *grammatical*—intensification:

> Whereas grammatical intensification targets specific degree scales lexicalized in gradable expressions, pragmatic intensification is more general and targets a **contextual** variable implicated in the interpretation of some expression, whether or not that expression is grammatically gradable (i.e., introduces a degree variable). (p. 876; emphasis mine)

The members of the Ultimate construction family are potential flexible modifiers because they are (i) structurally simple, i.e., mono-morphemic, thus bearing no specific predictable meaning (as noted in Section 3.5), and (ii) lexicon-external, thus not subject to specific grammatical constraints dictated by the modified head. As such they can modify foci of different word classes — an adjective as in Examples (4.1a) and (4.1c), a verb as in (4.1b), and a noun as in (4.1d).

We here have another layer of *filler-slot* relations (see Figure 4.1) beyond the purely syntactic one and the purely semantic one. One that has to do with the potential scalarity of the modifiable foci and the scalarity of the members of the Ultimate construction family.

With these data I can now provide an answer to the second question posed in Section 4.1 of this chapter:

Qii: What is it that stimulates/d the flexibility of the newly evolved modifiers?

Aii: The newly evolved idiomatic sentences are potential flexible modifiers, which can modify any word class — nouns, verbs and adjectives alike — because as mono-morphemic they bear no specific predictable meaning. Moreover, originally they were lexicon-external, thus not subject to rigid grammatical conventions which dictate the specialization of modifiers according to the syntactic category of the modified element. The foci of

> the Anaphoric degree-adverb exclamatives, on their part, are of various word classes — nouns, verbs and adjectives. This incorporation of the newly evolved idiomatic sentences into the Anaphoric degree-adverb exclamatives by means of the Correlative endpoint resultant state clause, lead to fruition of the (potential) flexibility of the members of the Ultimate construction family.

This state of affairs could have been seen as a case of *coercion-by-override* (Audring & Booij, 2016; Booij & Audring, 2018; Michaelis, 2004) where "[m]orphological schemas may change the semantic class of the base word" (Booij & Audring, 2018: 209). But note that coercion-by-override involves overt morphological marking or conversion of an existing word. In contrast, the members of the Ultimate construction family undergoing lexicalization are brand new newcomers to the lexicon. None of them have ever been a full-fledged word before being incorporated by the Anaphoric degree-adverb exclamatives. What we have here is not a change of word class, but rather the *inception* of new words.

Now, since there's no reason to assume that speakers convey a strong stance about a stance-object based on the categorial membership of the focus, no focus — a noun, a verb or an adjective — is expected to be preferred over another (see Examples 4.1a-d above). As a result, the respective newly evolved modifiers/amplifiers — an adjective, an adverb and an intensifier — must have come into being simultaneously. In order to test this claim, I focused again on the data of *en dvarim ka'ele/u*, specifically on sentences where ***še**-en dvarim ka'ele/u* '**that** there are no such things' is already integrated into sentences which are no longer exclamatives (stage III). At this stage, ***še**-en dvarim ka'ele/u* is just one step away from being reanalyzed as a full-fledged word (stage IV).

I classified the 49 instances into three groups: modifiers of nouns (n = 16), modifiers of adjectives (n = 23) and modifiers of verbs (n = 10). (No manner adverbs were found.) I tagged each example according to the week of its production rather than just the year, in order to make sure I did not miss any rapid change. Again, the distribution of the variable 'Date of production' is crucial for choosing the appropriate statistical method. So a Shapiro-Wilk test was performed. It did not show evidence of non-normality for any group (modifiers of nouns: $W(16) = 0.93$, $p = 0.254$; modifiers of adjectives: $W(23) = 0.95$, $p = 0.361$; modifiers of verbs: $W(10) = 0.90$, $p = 0.222$). A Levene's test was also performed to check the state of homoscedasticity. Levene's test statistic (based on medians) is not significant ($F(2,46) = 0.017$, $p = 0.983$), so homoscedasticity can be assumed. In addition, no outliers were found for any group. A one-way ANOVA was then performed to examine the difference in temporal means (i.e., date of production) between groups. Results indicate that the differences between the groups are not statistically significant ($F(2, 46) = 0.012$, $p = 0.99$), as depicted in Figure 4.9. (An alternative analysis of these data, represented by an association plot, produced identical results. It is presented in Appendix D.2.)

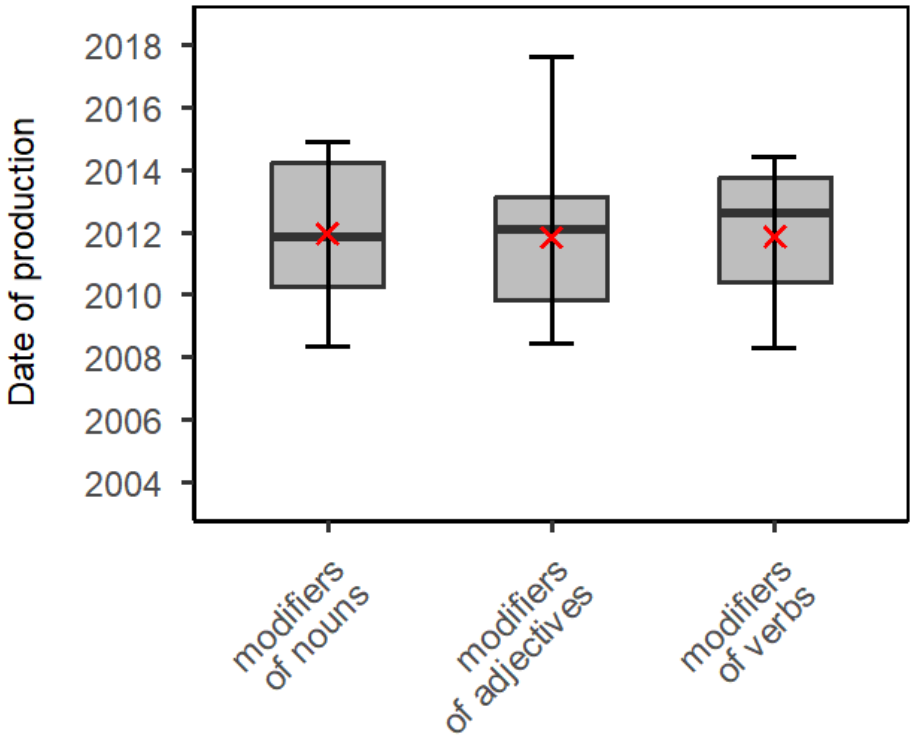

Figure 4.9: Box plot representing the simultaneous inception of the three categories of stage III *en dvarim ka'ele/u*. The red Xs represent the mean of each box plot. Data extracted from IsraBlog corpus.

With these data I can provide an answer to the third research question posed in Section 4.1 of this chapter:

Qiii: Which (amplifying) intra-sentential element — an adjective, an adverb or an intensifier — is/was the first to evolve? And is there any clear line of trans-categorization?

Aiii: The newly evolved idiomatic sentences are flexible modifiers, which seem not to be biased towards any specific modifiable head of the Anaphoric degree-adverb exclamatives (i.e., its focus) based on its categorial membership. So they came into being simultaneously. Trans-categorization, then, is rejected.[63]

[63] On the face of it, the suggestion regarding the simultaneous coming into being (of words from several word classes) and the supporting results, seem to beg the typologically related question whether Hebrew, a Semitic language, is capable of including flexible items, not just as anomalous exceptions. But in light of Salazar-García (2010) who maintained that 'flexibility' is not a property of a language as a whole, but rather a property of lexical units, this question seems superfluous. Interestingly, Salazar-García noted that flexibility is a dominant strategy in the field of the main degree words in Romance languages, showing that "the syntactic slot occupied by the head affects the functional characterization of the quantificational modifier" (p. 212). "In other words, the part-of-speech category of flexible lexical items

To sum up, Section 4.3 provided a parsimonious model that accounts for both the change in grammatical status of the members of the Ultimate construction family, as well as their semantic and syntactic flexibility. This model draws on the notion of clause linkage, specifically clause linkage of the Anaphoric degree-adverb exclamatives and the members of the Ultimate construction family (via the Correlative endpoint resultant-state clause). A divergence from the model (i.e., the case of *en milim*) has also been accounted for.

4.4 What's next?

So far I have shown how the contextual constructional scaffolding (i.e., the Anaphoric degree-adverb exclamatives) is responsible for incorporating the extra-sentential idiomatic sentences — the members of the Ultimate construction family — into the preceding sentence. Naturally, the scaffolding falls out of use once the word status of the new intra-sentential elements is established. When scaffolding is no longer required, the new word can expand into additional syntactic slots, other than those exemplified in (1.2). After all, "linguistic forms which have reached a well-defined grammatical phase don't necessarily live happily ever after in their grammatical niches" (Ariel, 2008: 250). In the next chapter I argue (and provide support for my argument) that further developments on the lexicalization path and their rates are conditioned by the collapse of the special contextual scaffolding (here, Anaphoric degree-adverb exclamatives).

depends on the role they play in each context" (p. 210), rather than the type of language. And this may also account for the readiness of Hebrew to entertain flexible modifiers.

Chapter 5: Further developments on the lexicalization path of the newly evolved words – More about the TRANSITION problem

Die Proletarier haben nichts in ihr zu verlieren als ihre Ketten.
Sie haben eine Welt zu gewinnen.
-- Karl Marx and Friedrich Engels, in *Manifest der Kommunistischen Partei*, 1848

(*The proletarians have nothing to lose but their chains.*
They have a world to win.)

5.1 What may happen after the collapse of the scaffolding context?

As I have shown in the previous chapter, complete "wordification" of the members of the Ultimate construction family is accompanied by the collapse of the constructional scaffolding context (stage IV in the proposed model). The loss of constructional scaffolding, I here argue, enables further constructional changes of the newly evolved word/s. By "further constructional changes" I mean that the members of the Ultimate construction family may display, for example, morphological behaviors typical of full-fledged words of the relevant categories. For example, an adjectival token may add on an adjective suffix and/or inflect for grammatical gender, as full-fledged adjectives do in Hebrew. I also argue that the rate at which the new words occur without any scaffolding support (stage IV) is a predictor of the rate of further constructional changes — syntactic and morphological developments — and therefore of degree and depth of lexicalization.

Let's take a look at Figure 5.1 below which is a snapshot of the frequencies of the various members of the Ultimate construction family as a function of stage of evolution. *Xaval al hazman* (bottom panel) shows the highest frequency of advanced, totally scaffoldless (stage IV) uses (marked yellow). If indeed the frequency of scaffoldless stage IV items is a predictor of further constructional changes, then *xaval al hazman* is expected to display the largest selection of further constructional changes. Next comes *en dvarim ka'ele/u* (second panel from the bottom) which shows fewer signs of scaffoldless intra-sentential uses (stage IV), and is therefore expected to lag behind *xaval al hazman* on the number of further constructional changes. *Ba livkot* (second

panel from top) shows some tokens with only partial scaffolding (stage III *še* 'that', marked green), but no signs of giving it up (stage IV). *En milim* (top panel) has hardly reached stage II, let alone stage III, as described at length in Section 4.3.5. These two — *ba livkot* and *en milim* — which lack stage IV items, for opposite reasons, are not expected to show any further constructional changes whatsoever. The next sections show that, indeed, the frequency of scaffoldless stage IV tokens is a predictor of syntactic expansion and morphological adaptation — constructional changes — which attest to degree of lexicalization. I begin with *xaval al hazman*.[64]

5.2 Further developments of *xaval al hazman*, originally, 'it's a waste of time'

As predicted, due to the high frequency of scaffoldless stage IV items, *xaval al hazman* shows a wide gamut of additional behaviors typical of full-fledged words. I start by describing the fully spelled (and most probably fully pronounced) *xaval al hazman* (Section 5.2.1), followed by its acronymic variant, *XVL"Z*, pronounced /xavláz/ (Section 5.2.2).

5.2.1 Fully spelled (and most probably fully pronounced) *xaval al hazman*

5.2.1.1 Xaval al hazman is a 'central' adjective

Given the relative frequency of intra-sentential elements without any scaffolding (stage IV) as compared with partial scaffolding (stage III) (marked yellow and green, respectively, in Figure 5.1), *xaval al hazman* seems to be a well-established adjective. If so, I can test it against the four characteristic features of adjectives proposed by Quirk, Greenbaum, Leech, and Svartvik (1985: 7.1-7.3), listed in (5.1).

(5.1) a. Adjectives can freely occur in attributive position.
b. Adjectives can freely occur in predicative position.
c. Adjectives can be premodified by the intensifier *very*.[65]
d. Adjectives can take comparative and superlative forms.

[64] Note that my goal here is not to analyze the actualization path of any of the members of the Ultimate construction family along the lines of De Smet (2012). "Actualization is traditionally seen as the process following syntactic reanalysis whereby an item's new syntactic status manifests itself in new syntactic behavior" (p. 601). Rather, I wish to examine a more specific aspect of actualization, namely the rate of constructional changes of the newly evolved word as a function of the presence—or lack—of the constructional scaffolding, which (also) indicates the degree of lexicalization.

[65] In Hebrew, this modification can be either pre-adjectival or post-adjectival.

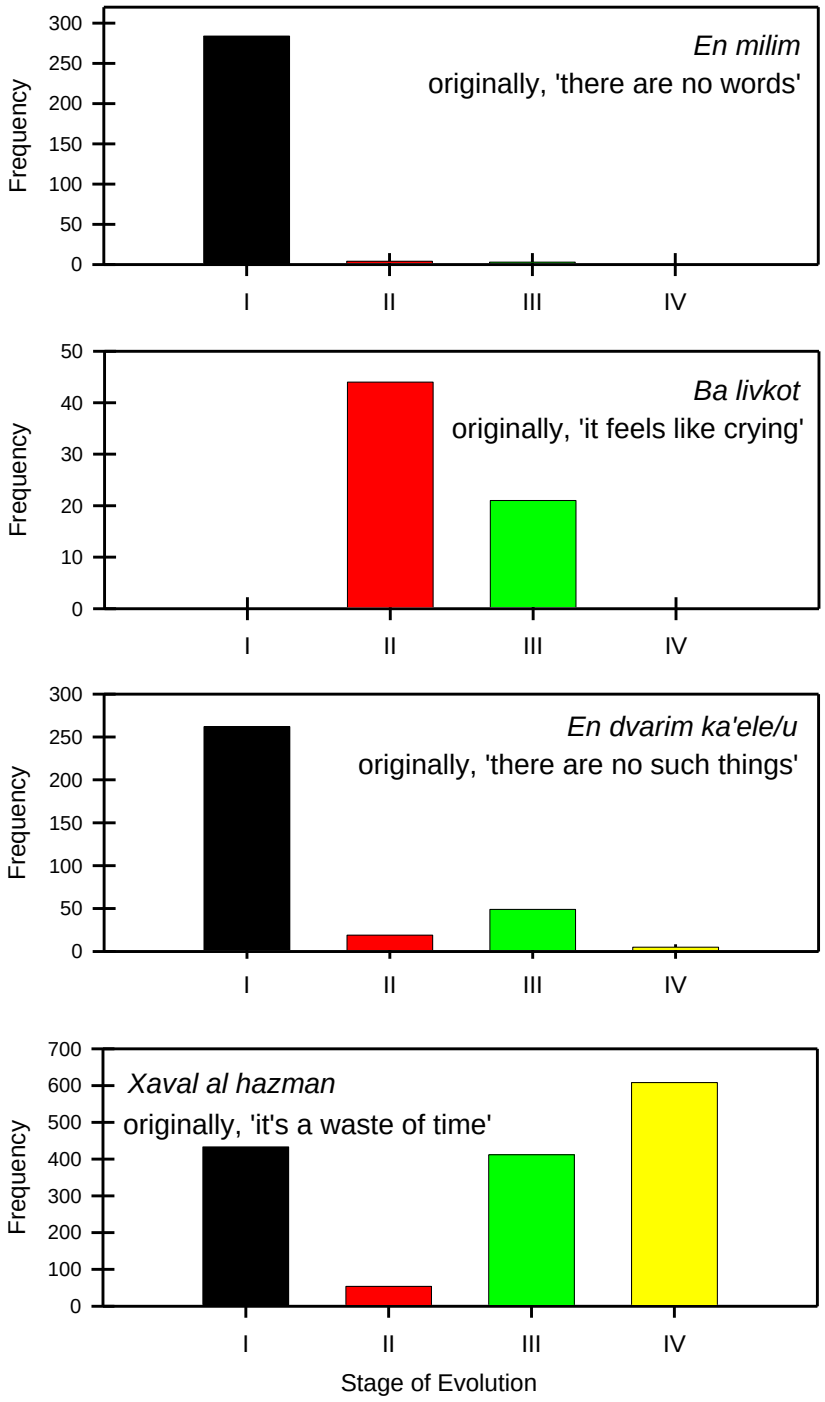

Figure 5.1: A snapshot (as of July 2017) of the frequencies of the various members of the Ultimate construction family as a function of stage of evolution.[66] Data extracted from IsraBlog corpus.

[66] Note that due to the Layering principle (Hopper, 1991) items of distinct stages on the evolution path — here stage I, II, III and IV — co-exist.

Quirk et al. argued that the first two features (5.1a-b) are more fundamental than the other two (5.1c-d), for they can distinguish between an adjective and an adverb. Importantly, adjectives that satisfy both (5.1a) and (5.1b) are considered 'central', whereas adjectives that satisfy only one of (5.1a) or (5.1b) are considered 'peripheral'. The data from IsraBlog corpus show 144 attributive adjectival tokens and 97 predicative adjectival token of *xaval al hazman*[67] (the ratio is a matter of genre; see Englebretson, 1997), thus attesting to *xaval al hazman* behaving just like a central adjective. Examples (5.2a-b) are representative examples of attributive (5.2a) and predicative (5.2b) uses.

(5.2) a.

Rami	*hu*	*tabax*	*anak.*	*mexin*	***oxel***
Rami	is	a.cook	gigantic	prepares	**food**

xaval al hazman.
it's a waste of time → amazing

'Rami is an outstanding cook. He prepares **amazing food**.'
(tinyurl.com/5n9x8ff6)

b.

ha-melcarim	*omrim*	*še-**ha-mana***	***ha-zot***	***hi***
the-waiters	say	that-**the-dish**	**the-this**	**is**

xaval al hazman.
it's a waste of time → amazing

'The waiters say that **this dish is amazing**.'
(IsraBlog)

Adjectival *xaval al hazman* also satisfies criterion (5.1c), but a slight qualification is in order. In the only relevant example I spotted, (5.3), *xaval al hazman* is preceded by *mamaš* 'really'. Although *mamaš* is a lexicalized intensifier (Bardenstein & Ariel, 2023), here it can be also interpreted as a counter-loosener (Bardenstein & Ariel, 2022) which is an intermediate phase between 'truth' interpretation and genuine 'intensification'. No instances of *xaval al hazman* amplified by the more lexicalized intensifiers *meod*, *nora* or *be-yoter*, all denoting 'very', were detected.

[67] My classification follows the discourse definition of attributivity/predicativity (Englebretson, 1997; Ferris, 1993: 39; Thompson, 1990). An attributively used adjective introduces a new discourse referent, while a predicatively used adjective modifies an already established discourse referent.

(5.3) *ani* *xayav* *lehagid* ***sefer*** ***mamaš***
I must to.say **a.book** **really**

xaval al hazman.
that-it's a waste of time → outstanding

'I must say: **A really outstanding book**…'

(tinyurl.com/byjdj465)

Examples of adjectival *xaval al hazman* embedded in comparative and superlative constructions were also found, as in Examples (5.4a) and (5.4b), respectively.

(5.4) a. *haxi* *xašuv* *bišul* *iti* *u-mefanek* *b-a-tanur,*
most important cooking slow and-indulging in-the-oven

kama *še-yoter* *zman* *yece* ***yoter***
how.much that-more time will.come.out **more**

xaval al hazman.
it's a waste of time → amazing

'The most important thing is slow and indulging cooking in the oven, the more time [it remains in the oven - IB] the **more amazing** it comes out.'

(tinyurl.com/4m792yr7)

b. *ha-diskotek* ***haxi*** ***xaval al hazman***
the-discotheque **the.most** **it's a waste of time → amazing**

'**The most amazing** discotheque'

(tinyurl.com/muj3vzve)

This state of affairs testifies to a high lexicalization status on this parameter.

5.2.1.2 Xaval al hazman is an attributive NP

I spotted quite a number of examples where *xaval al hazman* fills the slot of the first NP in an NP of an NP genitive construction (also known as a *binomial noun-phrase* or a *quality pseudo-partitive*).[68] The first NP is a modifier which amplifies a pragmatically inferred property of the second NP, as in Example (5.5).

[68] For the Hebrew version of this construction see Halevy 2001; for the English version see, e.g., Aarts 1998; for other languages see Halevy 2001: 73-74 and Aarts 1998: footnote 3.

(5.5) *ra'ita* *et* *ha-ona* *ha-rišona?*
you.saw ACC the-season the-first

xaval al hazman ***šel*** ***ktiva*** ***misxak*** ***u-vimuy.***
it's a waste of time → wonder **of** **writing** **playing** **and-staging**

ve-gam *ma'avarey^* *scena* *še-lo* *me-ha-olam* *ha-ze.*
and-also transitions scene that-not from-the-world the-this

'Have you watched the first season? **It's a wonder of writing, playing and staging**. And also scene-transitions out of this world.'

(tinyurl.com/2m2u93yv)

The syntactic flexibility (see Section 4.3.6) of *xaval al hazman* could account for filling the (formally) NP slot. In addition, the fact that *xaval al hazman* fills this NP slot testifies to its conventionality as a lexical item. This is due to the status of the NP of an NP construction in Hebrew. It is considered standard — not sub-standard — (Shatil, 2015), where the first NP slot hosts *conventional* abstract nouns (conveying canonical attributes such as size, beauty, charm etc.; Halevy 2001). This state of affairs testifies to a high lexicalization status on this parameter.

5.2.1.3 Intensifier xaval al hazman in a sequence of other intensifiers

According to Méndez-Naya (2003), an intensifier (e.g., *very*) which cannot modify or be modified by another intensifier is fully lexicalized, while intensifiers which have not completely lexicalized can (see aslo Klein, 1998: 140-145).[69] Indeed, I failed to find any instance of intensifier *xaval al hazman* which either modifies or is modified by another intensifier. This state of affairs testifies to a high lexicalization status on this parameter.

5.2.1.4 The position of intensifier xaval al hazman with respect to intensified adjectives and verbs

Hebrew intensifiers, such as *meod* and *nora* both denoting 'very', and *kol-kax* 'so', can assume either a pre- or a post-adjectival/verbal position. This means that the scaffoldless intensifier *xaval al hazman* should, in principle, be found in both positions. Tables 5.1a and 5.1b show that this is indeed the case. However, whereas the results of pre- and post-adjectival/verbal *meod* and *nora* 'very', and *kol-kax* 'so' are on the same order of magnitude (see the ratio columns in both tables),[70] in the case of *xaval al hazman*, the post-adjectival/verbal position is much more frequent than the pre-adjectival/verbal position. The post-adjectival/verbal position is 19 and 27 times more

[69] This is in line with Ariel (2008: 260-264), who argued that expressions of the very same semantic and pragmatic status, regardless of their word class, cannot be used when they have the same scope.

[70] These results are in line with the results presented in Bar-Ziv Levy (2017: 141-146).

occupied than the pre-adjectival/verbal position, respectively; see the ratio columns in Tables 5.1a and 5.1b.

Intensifier	# Pre-adjectival	# Post-adjectival	Ratio
***nora* 'very'**	13147	3488	3.8 : 1
***kol-kax* 'so'**	143066	45667	3.1 : 1
***meod* 'very'**	277,511	495,638	1 : 1.8
***xaval al hazman* 'extremely'**	5	95	1 : 19

Table 5.1a: The frequencies of pre- and post-adjectival intensifiers, and their ratios. Data extracted from HeTenTen corpus.

Intensifier	# Pre-verbal	# Post-verbal	Ratio
***nora* 'very'**	9804	4778	2 : 1
***kol-kax* 'so'**	50859	33496	1.5 : 1
***meod* 'very'**	126,150	152,538	1 : 1.2
***xaval al hazman* 'so much'**	3	80	1 : 27

Table 5.1b: The frequencies of pre- and post-verbal intensifiers, and their ratios. Data extracted from HeTenTen corpus.

The fact that the vast majority of intensifier *xaval al hazman* is post-adjectival/verbal precludes full lexicalization on this parameter. Had *xaval al hazman* undergone full lexicalization, it would have more often preceded the adjectives/verbs.

5.2.1.5 The polarity of the adjectival and verbal collocates of intensifier xaval al hazman

Lorenz (2002) suggested that a specifically negative (or affirmative) polarity of the adjectival collocates of an intensifier attests to its degree of lexicalization. For English *terribly* he showed that its 20 most frequent collocates are of negative polarity (e.g., *sorry, upset, sad*). These data indicate that *terribly* has not (yet?) shaken off its negative connotation and is therefore not a full-fledged intensifier. Had it been completely lexicalized, it should have preceded any adjectival collocate, regardless of polarity.

I examined all 155 instances of intensifier *xaval al hazman* modifying adjectives in IsraBlog corpus: 83 of them amplify adjectives of positive polarity; 53 amplify adjectives of negative polarity (the polarity of 19 additional adjectives was unclear). I also examined all 141 instances of intensifier *xaval al hazman* modifying verbs: 46 of them amplify verbs of positive polarity; 83 amplify verbs of negative polarity (the polarity of 12 additional verbs was unclear). These distributional findings, which are not biased for a single polarity, testify that intensifier *xaval al hazman* is at an advanced stage of lexicalization on this parameter.

5.2.1.6 Negated adjective and intensifier xaval al hazman

While thus far I have shown how advanced *xaval al hazman* is in its lexicalization process, I suggest that with respect to negation, the process is not yet complete. To test

this intuition, I compared the acceptability of negated *xaval al hazman* resonating a prior mention, or not. Resonance is said to license innovative uses by way of analogy (Du Bois, 2014). So if a negated *xaval al hazman* is judged acceptable only when resonating a prior mention, then it has not quite lexicalized on this parameter (i.e., negation).

I ran 2 *exploratory* tests (December 2018-January 2019) to test this intuition. One test examined the acceptability of a negated adjectival *xaval al hazman*. The other one examined the acceptability of a negated intensifying *xaval al hazman*. I describe them in turn.

Test 1: negated adjectival *xaval al hazman*?

Participants. All 32 participants were students of linguistics, all native speakers of Hebrew (20 women and 12 men), aged 23–29 (*M* = 25.94, *SD* = 4.88).

Stimuli. The stimuli were the two alternatives of the dialogue in (5.6a) and the two alternatives of (5.6b). The difference between the two alternatives of (5.6a) is the adjective used — either *xaval al hazman* or the conventionalized *meragešet* 'moving'. This difference applies to the two alternatives of (5.6b). Importantly, the difference between (5.6a) and (5.6b) is the resonance (or lack thereof) between A's utterance and B's response, respectively. In (5.6a), the adjectives are first asserted by A and then repeated under negation by B. This is not the case in (5.6b).

Each participant was presented with only one of the four alternatives.

(5.6) a. A: *muzika* *kubanit* *hi*
music Cuban is

***xaval al hazman* / *meragešet*.**
it's a waste of time → outstanding / moving

'Cuban music is **outstanding/moving**.'

B: *hi* *lo* ***xaval al hazman* / *meragešet.***
she not **it's a waste of time → outstanding / moving**

hi *benonit.*
she so-so

'it is **not outstanding/moving**. It is so-so.'

(5.6) b. A: *ma* *at* *xoševet* *al* *muzika* *kubanit?*
what you think on music Cuban

'What do you think about Cuban music?'

B: *le-da'ati* *hi* *lo*
to-my.opinion she not

xaval al hazman / ***meragešet.***
it's a waste of time → outstanding / moving

'In my opinion, it's **not outstanding/moving**.'

Procedure. The exploratory tests took place in the first 5 minutes of class on pragmatics, just before class began. Participants got the stimulus printed on a sheet of paper, and were asked to determine how acceptable B's response is on a 7-point Likert scale (1 = not acceptable al all, 7 = highly acceptable). If they did not want to complete the task, they could have turned in the paper without responding (as two of them did).

Results. A Kruskal-Wallis test (Kruskal-Wallis H test) showed that there is a statistically significant difference between the groups ($H(3, N = 32) = 17.89$, $p < 0.001$) with mean rank of 19.65 for the rating of a resonating *xaval al hazman* response in (5.6a), 17.43 for the rating of a resonating *meragešet* 'moving' response in (5.6a), 5.25 for the rating of a non-resonating *xaval al hazman* response in (5.6b), and 23.5 for the rating of a non-resonating *meragešet* 'moving' response in (5.6b). A strong effect size was detected, $\varepsilon^2(H) = 0.58$, 95% CI [0.36, 0.81].

Additionally, post-hoc pairwise Mann-Whitney tests were conducted to compare all group pairs, using a Bonferroni-adjusted α level of 0.0083 (0.05/6). These tests showed that only the non-resonating *xaval al hazman* response ($Mdn = 2$) was significantly different from all other groups — the resonating *xaval al hazman* response ($Mdn = 6.5$), the resonating *meragešet* 'moving' response ($Mdn = 6$) and the non-resonating *meragešet* 'moving' response ($Mdn = 7$). The difference between the resonating *xaval al hazman* and the non-resonating *xaval al hazman* was significant, $U(n_{res\_xaval\ al\ hazman} = 10, n_{non\text{-}res\_\ xaval\ al\ hazman} = 8) = 75.5$, $p = 0.0015$, 95% CI [2.0, 5.0]. The difference between a resonating *meragešet* 'moving' and a non-resonating *xaval al hazman* was significant, $U(n_{res\_meragešet} = 7, n_{non\text{-}res\_xaval\_al\_hazman} = 8) = 55$, $p = 0.0018$, 95% CI [3.0, 5.0]. The difference between the non-resonating *meragešet* 'moving' and the non-resonating *xaval al hazman* was also significant, $U(n_{non\text{-}res\_\ meragešet} = 8, n = 8\ n_{non\text{-}res\_xaval\ al\ hazman}) = 55.5$, $p = 0.0013$, 95% CI [3.0, 6.0]. None of the other comparisons were found significant after Bonferroni adjustment (all ps > 0.087). The results are presented in Figure 5.2.

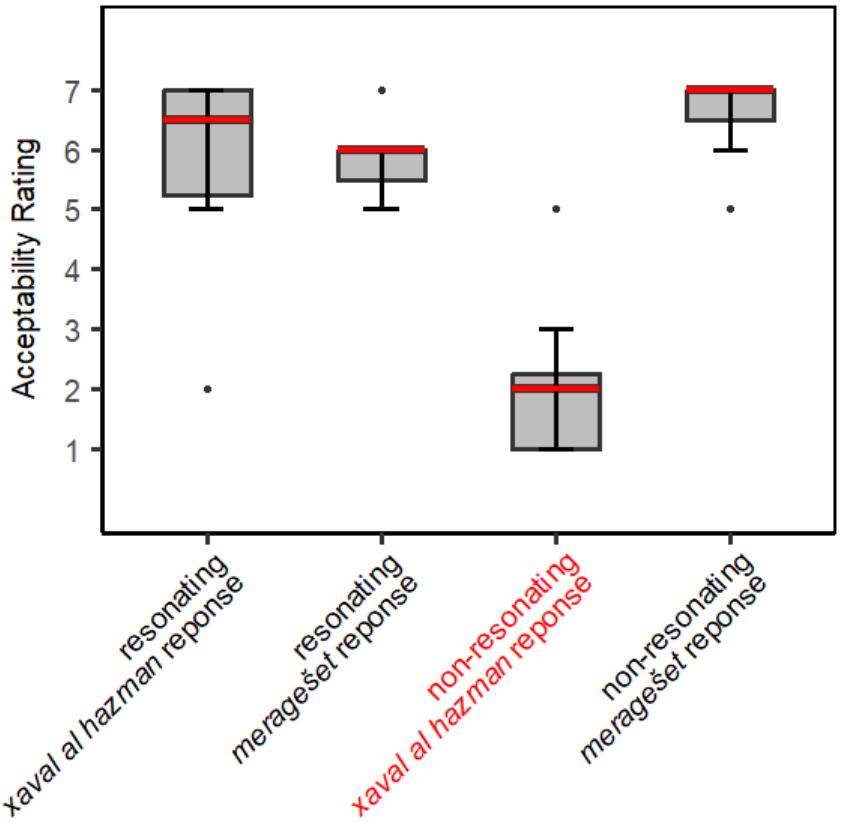

Figure 5.2: Box plot representing the acceptability of responses which include negated adjectival *xaval al hazman* and *meragešet* 'moving', whether resonating with a prior mention, or not

The results of this test show that adjectival *xaval al hazman* is not yet free to occur under negation, and needs a prior (resonating) mention.

Test 2: negated intensifier *xaval al hazman*?

Participants. All 34 participants were students of linguistics, all native speakers of Hebrew (21 women and 13 men), aged 23–44 (M = 27.15, SD = 6.77).

Stimuli. The stimuli were the two alternatives of the dialogue in (5.7a) and the two alternatives of (5.7b). The difference between the two alternatives of (5.7a) is the (negated) intensifier used — either *xaval al hazman* or the conventionalized *nora* 'awfully'. This difference applies to the two alternatives of (5.7b). Importantly, the difference between (5.7a) and (5.7b) is the resonance (or lack thereof) between A's utterance and B's response, respectively. In (5.7a) the intensifiers are first asserted by A and then repeated under negation by B. This is not the case in (5.7b).

Each participant was presented with only one of the four alternatives.

(5.7) a. A: *muzika* *kubanit* *hi* ***meragešet***
music Cuban is **moving**

***xaval al hazman* / *nora*.**
it's a waste of time → extremely / awfully

'Cuban music is **extremely/awfully moving**.'

B: *hi* ***lo*** ***meragešet***
she **not** **moving**

***xaval al hazman* / *nora*.**
it's a waste of time → extremely / awfully

hi *benonit.*
she so-so.

'it is **not extremely/awfully moving**. It is so-so.'

(5.7) b. A: *ma* *at* *xoševet* *al* *muzika* *kubanit?*
what you think on music Cuban

'What do you think about Cuban music?'

B: *le-da'ati* *hi* ***lo***
to-my.opinion she **not**

xaval al hazman* / *nora meragešet.
it's a waste of time → extremely / awfully moving

'In my opinion, it's **not extremely/awfully moving**.'

Procedure. As in test 1 above. All participants completed the assignment.

Results. A Kruskal-Wallis test (Kruskal-Wallis H test) showed that there was a statistically significant difference between the groups ($H(3, N = 34) = 14.35$, $p = 0.0025$) with mean rank of 19.55 for the rating of a resonating *xaval al hazman* response in (5.7a), 21.17 for the rating of a resonating *nora* 'awfully' response in (5.7a), 6.69 for the rating of a non-resonating *xaval al hazman* response in (5.7b) and 21.21 for the rating of a non-resonating *nora* 'awfully' response in (5.7b). A strong effect size was detected, $\varepsilon^2(H) = 0.43$, 95% CI [0.22, 0.71].

Additionally, post-hoc pairwise Mann-Whitney tests were conducted to compare all group pairs, using a Bonferroni-adjusted alpha level of 0.0083 (0.05/6). These tests showed that only the non-resonating *xaval al hazman* response (*Mdn* = 3) was

significantly different from all other groups, the resonating *xaval al hazman* response (*Mdn* = 7), the resonating *nora* 'awfully' response (*Mdn* = 7) and the non-resonating *nora* 'awfully' response (*Mdn* = 7). The difference between the resonating *xaval al hazman* response and the non-resonating *xaval al hazman* response was significant, $U(n_{res\_xaval\ al\ hazman} = 10, n_{non\text{-}res\_xaval\ al\ hazman} = 7) = 70$, $p = 0.0072$, 95% CI [1.0, 5.0]. The difference between the resonating *nora* 'awfully' response and the *non*-resonating *xaval al hazman* response was significant, $U(n_{res\_nora} = 9, n_{non\text{-}res\_xaval\ al\ hazman} = 7) = 68$, $p = 0.0017$, 95% CI [1.0, 5.0]. The difference between the *non*-resonating *nora* 'awfully' response and the *non*-resonating *xaval al hazman* response was significant, $U(n_{non\text{-}res\_nora} = 7, n_{non\text{-}res\_xaval\ al\ hazman} = 7) = 52.5$, $p = 0.0044$, 95% CI [1.0, 5.0]. None of the other comparisons were found significant after Bonferroni adjustment (all *p*s > 0.566). The results are presented in Figure 5.3.

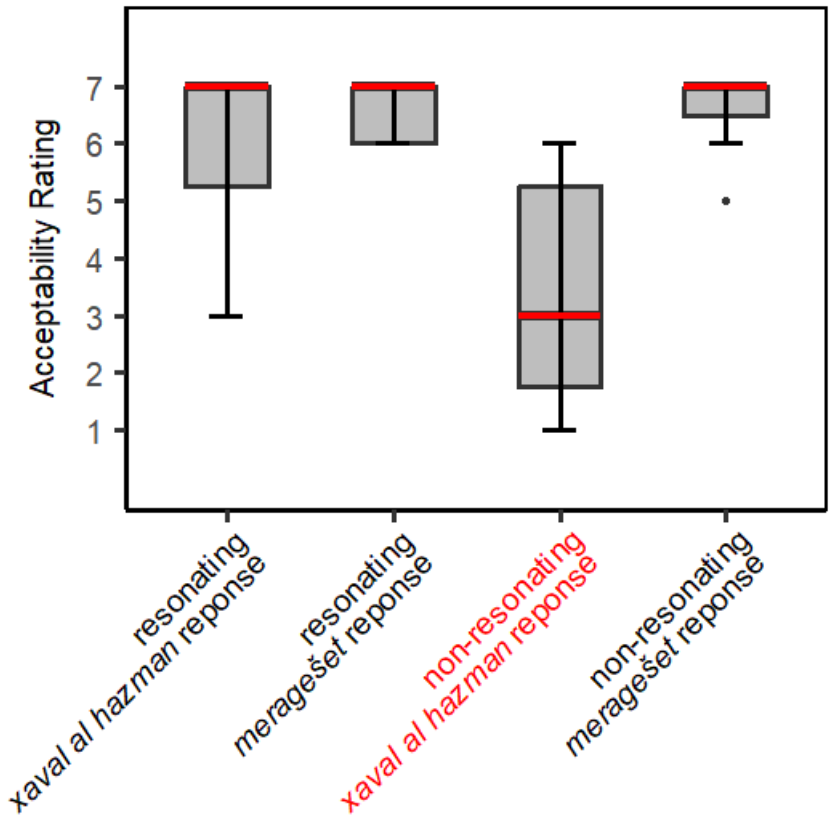

Figure 5.3: Box plot representing the acceptability of responses which include negated intensifying *xaval al hazman* and *nora* 'awfully', whether resonating with a prior mention, or not

The results of the second test support the same conclusion as the first one. *Xaval al hazman* does not yet freely occupy the role of a negated constituent, unless resonating a prior mention.

Discussion of the results of Test 1 and Test 2. These results attest to the role of dialogic resonance in the acceptability of novel constructions, lexemes included. Whereas the conventionalized negated adjective *meragešet* 'moving' and intensifier *nora* 'awfully' are acceptable regardless of dialogic resonance between interlocutors, negated *xaval al hazman*, whether an adjective of an intensifier, is acceptable only when resonating a prior mention. In fact, such resonance, Du Bois (2014: 364) suggested, "[…] yields a broad range of impacts on […] creativity, grammaticization, ultimately contributing to the self-organization of new linguistic structure."[71] The fact that only prior mention licenses the negated *xaval al hazman*, whether an adjective or an intensifier, implies that *xaval al hazman* does not score high on degree of lexicalization on this specific parameter.

Since the acceptability of negated *xaval al hazman* is conditioned on prior mention, I assumed that no instance of a negated *xaval al hazman* should be found even in the rather informal, yet not dialogic IsraBlog corpus, HeTenTen corpus, or the web. This was indeed the case.

In sum, *xaval al hazman* does not display a perfectly homogenous picture of lexicalization. It scores high on some parameters, but low on the negation one.

5.2.2 The acronym XVL(")Z

As mentioned above, *xaval al hazman* has also evolved an acronym, *XVL"Z*, pronounced /xavláz/. I here analyze its morphological and syntactic patterning showing that it is more deeply lexicalized than the fully spelled *xaval al hazman*. I suggest that this is due to its shortness and semantic opacity, two features which are typical of words.

An acronym is an orthography-originated word which conflates the initial letters of a sequence of words which constitute a phrase (Brinton & Traugott, 2005: 42; Pawley, 1986: 106), and pronounced as a word rather than as a sequence of letters (Bauer, 1983: 237-238; Blank, 2001: 1605) (e.g., *REM* = *R*apid *E*ye *M*ovement is an acronym, whereas *IBM* = *I*nternational *B*usiness *M*achines is not). Hebrew acronyms are conventionally marked as such by a quotation mark between the ultimate and the penultimate letters (i.e., *XVL"Z*).

Another acronym-related phenomenon is the Hebrew-unique *ex acronym*. An ex acronym is an acronym written without a quotation mark. The absence of a quotation mark indicates that speakers no longer conceive of the word as an acronym (Tadmor, 1988), which points to an even higher degree of lexicalization.

5.2.2.1 The association of XVL(")Z with scaffoldless intra-sentential positions

If, as suggested above, *XVL"Z* is more word-like, then it should occur in scaffoldless intra-sentential positions (stage IV) more than in extra-sentential and scaffolding-assisted intra-sentential ones (stages I-III). This is indeed the case. I found 523 instances of the acronym *XVL"Z* and 97 instances of the ex acronym *XVLZ* in IsraBlog corpus.

[71] See also Ariel (2008: 176-177, 251-252) for the role of dialogic syntax in facilitating innovative constructions, such as the "illogical" *for the whole part* when resonating with a contextually salient *for the most part*.

Regardless of their word class, both occur in any of the I-IV contexts (see Section 4.3.1 above) without exception. Interestingly, however, once comparing the distribution of the orthographic variants (i.e., *XVL"Z* / *XVLZ* and the fully spelled *xaval al hazman*), a Chi-square test of homogeneity shows that the acronyms are associated more significantly with scaffoldless intra-sentential uses (stage IV) (371/620=60%) than the fully spelled *xaval al hazman* 608/1543 = 40%), χ^2 (1, N = 2163) = 73.72, p = 5.9×10^{-18} (φ = 0.18, a small effect size), as shown in Figure 5.4.[72]

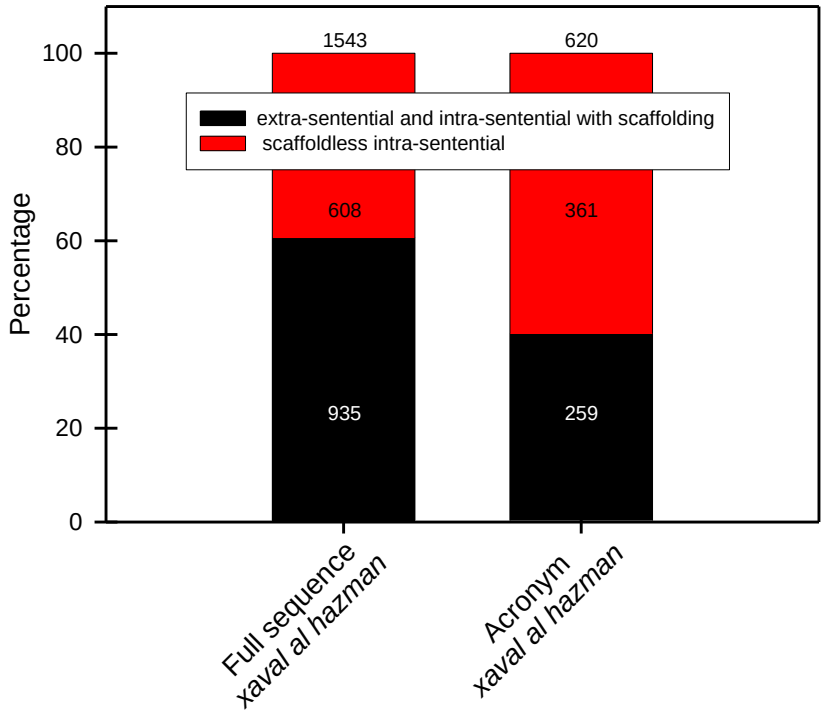

Figure 5.4: A snapshot (as of July 2017) of the distribution of extra-sentential and scaffolding-assisted intra-sentential (items of stages I-III), and scaffoldless intra-sentential *xaval al hazman* (items of stage IV) as a function of the orthographic variant – fully spelled and acronymic. Data extracted from IsraBlog corpus.

Example (5.8a) is a representative example of an adjectival use of *XVL"Z* (note the conjunction *ve* 'and' and the definite article *ha-* which substantiate the adjectivity of

[72] The acronym and the ex acronym were collapsed because a more refined analysis had shown that there is no significant difference between the use of the acronym (313/523 = 60%) and the ex acronym (58/97 = 60%) as a scaffoldless intra-sentential elements, χ^2 (1, N = 620) = 9.64×10^{-5}, p = .99.

XVL"Z). Example (5.8b) is a representative example of an adjectival use of the ex acronym *XVLZ*.

(5.8) a. *bni* *ha-matok* *ve-he-xamud*
my.son the-sweet and-the-cute

*ve-ha-**XVL"Z*** *hit'orer*.
and-the-**it's a waste of time → amazing** woke.up

'My sweet, cute and **amazing** son has just woken up.'
(IsraBlog)

b. *ha-yom* *hexanu* *oxel* *im* *ima* *šel* *xavera*
today we.prepared food with mother of girlfriend

šeli *še-hi* *mexina* oxel ***XVLZ!!!!***
my that-she prepares food **it's a waste of time → outstanding**

'Today we prepared food with my girlfriend's mother who prepares **outstanding** food!!!!'
(IsraBlog)

5.2.2.2 Further morphological development of XVL(")Z (derivation and inflection)
"Any expression which can serve as the base for inflected or derived formations can be regarded as lexicalized to some degree" (Pawley, 1986: 107). This is true for adjectival acronymic *XVL"Z* and ex acronym *XVLZ*.

Concatenative derivation is the most typical way to derive Hebrew adjectives from nouns (Nir, 1993: 109-120). Data show that the default strategy is suffixation with *–i*.[73] No wonder, then, that I found 6 instances of *XVL(")Z-**i*** SG.M (one of which is Example 5.9a) and 1 instance of *XVL(")Z-**it*** SG.F in IsraBlog corpus. A Google search of the web (as of 16 June 2022) returned several examples of *XVL(")Z-**im*** PL.M. Example (5.9b) is one of them.

(5.9) a. *ha-tekes* *acmo* *haya* *mehamem* *u-merageš*
the-ceremony itself was stunning and-moving

*ve-**XVL"Z-i***
and-**it's a waste of time → amazing-SG.M**

[73] For corpus-based evidence see Ravid and Shlesinger (1987), Fisherman (1994) and Muchnik (2000), and for experimental evidence see Bolozky (1999: 87-89).

‘the ceremony itself, was stunning and moving and **amazing**…’
(IsraBlog)

b. *ha-efektim* *šel* *vin dizel* ***XVLZ-im***
the-effect of Vin Diesel **it’s a waste of time → amazing-PL.M**

‘Vin Diesel’s effects are **amazing** […]’
(tinyurl.com/bp7dcfvm)

Importantly, the fact that the adjective *XVL(“)Z-i* is derived from the base *XVL”Z* (or *XVLZ)* suggests that this base has psychological reality, namely, “an independent existence in the lexicon of the average Hebrew speaker” (Ravid & Shlesinger, 1987: 59-60; translation mine). In other words, *XVL”Z* / *XVLZ* seems to be entrenched in the lexicon.

But then, one may wonder why Hebrew speakers have an urge to derive the adjective *XVL(“)Z-i* from the adjectival *XVL(“)Z.* Such a process is quite rare (Bolozky, 1999: 86; Ravid & Shlesinger, 1987), for there seems to be no semantic difference between the two. However, if speakers do feel that adjectival *XVL(“)Z-i* and adjectival *XVL(“)Z* are not semantically identical, that is, *XVL(“)Z* is not adjectival “enough”, then they may produce *XVL(“)Z-i* in order to underscore the adjectival status of *XVL(“)Z.*

Also note that both *XVL”Z* and *XVL”Z(-i)* in Examples (5.8a) and (5.9a), respectively, are preceded by the conjunction *ve-* ‘and’ (and in Example 5.8a also the definite article *ha-*). No instance of the fully spelled *xaval al hazman* preceded by the conjunction *ve* ‘and’ was detected. This may suggest that the acronymic *XVL”Z(-i)* is “more” adjectival than the fully spelled *xaval al hazman*.

5.2.2.3 Intensified adjectival XVL(“)Z

I surmise that it is the length of the fully spelled *xaval al hazman*, or rather, its complex morphology, that hinders treating it as an unmarked adjective. *XVL(“)Z-(i)*, on the other hand, is treated as a simple, opaque word. This difference is reflected in the intensifiers that the fully spelled *xaval al hazman* and the acronym team up with. While the fully spelled adjectival *xaval al hazman* can be modified only by *mamaš* ‘really’ (see Example 5.3 above and the relevant discussion about the interpretation of *mamaš*), the derived *XVL”Z-i* is modifiable by the more lexicalized *meod*, *nora*, and *be-yoter*, all denoting ‘very’. Example (5.10) is a representative example. This may suggest a more lexicalized status for the derived adjectival *XVL(“)Z-i*.

(5.10) *šavu’a* ***XVL”Z-i*** ***meod!***
a.week **it’s a waste of time → amazing-SG.M** **very**

‘A **very amazing** week!”
(tinyurl.com/yc7kc4fj)

5.2.2.4 Negated adjective XVL(“)Z

As in the case of the fully spelled *xaval al hazman*, not a single instance of *XVL(“)Z* was found in the scope of negation in IsraBlog and HeTenTen corpora, for reasons explained in Section 5.2.1.6. However, a Google search of the web (as of 16 June 2021) returned one derived acronym in the scope of negation, *XVL”Z-i*, see Example (5.11). Again, this may suggest a more lexicalized status for the inflected *XVL”Z-i* than for the non-inflected *XVL”Z*, all the more so for the fully spelled (and fully pronounced) *xaval al hazman*.

(5.11) *ha-meser* *xad* *ve-kole’a,* *le-da’ati* *yaxol*
the-message sharp and-to.the.point to-my.opinion can

la’avor *l-a-šalav* *ha-ba…* *aval*
to.pass to-the-stage the-next but

lo ***XVL”Z-i***
not **it’s a waste of time → outstanding-SG.M**

‘The message is sharp and to the point, and in my opinion it can pass on to the next stage…but it’s **not outstanding**.’

(tinyurl.com/3vuv5x66)

In sum, the acronym *XVL”Z* and the ex acronym *XVLZ*, both adjectives (not intensifiers), seem to score higher than their fully spelled counterpart on the relevant lexicalization parameters presented above. This is, most probably, due to their shortness and semantic opacity which make them more word-like than their fully spelled counterpart.

5.2.3 Further developments *of xaval al hazman*

5.2.3.1 Xaval al hazman as a strong agreement marker

Early 2002 saw a dialogic, utterance-level *xaval al hazman* which implicates an emphatic affirmative answer to a question, similar to ‘totally’ (and Hebrew *legamrey* ‘totally’; see Shaviv 2018). Example (5.12) is the earliest example I detected. The context is a journal article about people whose life is all about soap operas. In this (representative) example, the interviewer wonders about the effect the addiction to soap operas has on the life of one of the interviewees.

This new use follows from the upgrade (Pomerantz, 1984) it offers to the content and/or stance implicated by the question, amplifying the (here) concealed affirmative answer of the interviewee. Interestingly, this use testifies that *xaval al hazman* did not lose its original status as an independent utterance (i.e., an idiomatic sentence), despite its rather high degree of lexicalization, as described above (see Hopper’s 1991 Layering principle). This is why it can still give rise to new developments, to new constructional changes. This development is in line with Bardenstein’s (2021) Persistence principle

which builds on Hopper's (1991) Persistence principle. Once the initial implicature of 'remarkability' was entrenched for conveying a strong stance, this strong stance persisted for later evolved uses, here an emphatic affirmative answer to a question.

(5.12) Interviewer: *ve-ze* *lo* *mištalet* *lax* *al* *ha-xa'im?*
and-this not takes.over to.you on the-life

'Doesn't it take over your life?'

Interviewee: ***xaval al hazman.*** *betax.* *at* *margiša*
it's a waste of time → totally sure you feel

kvar *xelek* *me-ha-alila.*
already part from-the-plot

'**Totally**. Sure. You feel already part of the plot [of the soap opera - IB].'

(January 2002, Yedioth Ahronoth)

5.2.3.2 A further reinforced xaval al hazman

A further development of *xaval al hazman* is the more complex construction *xaval* ***laxem*** *al hazman* 'it's a waste of **your** time', as in Example (5.13). It incorporates an additional dative participant, which invariably refers to the addressees. *Xaval* ***laxem*** *al hazman* is still an intensifier, likely to reinforce the potentially weakening *xaval al hazman*, which may be losing its emotive force over time (see, e.g., Hopper & Traugott, 2003 [1993]: 122; Klein, 1998: 26; Méndez-Naya, 2003). *Xaval* ***laxem*** *al hazman* was first spotted in IsraBlog corpus in 2003, about two years after this web-site had been launched (and some ten years after the dativeless *xaval al hazman* had been first documented).

(5.13) *yeš* *oxel* *ta'im* ***xaval*** ***laxem*** ***al***
there.is food delicious **it's.a.pity** **to.you.2PL.M** **on**

hazman.
the-time → amazingly

'There's **amazingly** good food.'

(IsraBlog)

While both *xaval al hazman* and *xaval* ***laxem*** *al hazman* are intensifiers, the explicature (or 'what-is-said' content) that triggers the intensification is different in the

two cases. Whereas *xaval al hazman* ‘it’s a waste of time’ “says” that the ‘wasted’ time would be the *speaker’s* attempt to describe an extremely remarkable state of affairs, *xaval **laxem** al hazman* ‘it’s a waste of **your** time’ “says” that the *addressees’* attempt to contest the speaker’s extreme evaluation of a specific state of affair is futile, because the said event is unequivocally remarkable.

All in all, in this section I have presented a wide gamut of constructional changes undergone by the newly evolved *xaval al hazman*, attesting to its depth of lexicalization. I argued above that this state of affairs could come about only once the contextual constructional scaffolding was deemed unnecessary. The prediction falling from this argument is that the other members of the Ultimate construction family would show a somewhat different picture of further constructional changes. I begin with *en dvarim ka’ele/u*.

5.3 Further developments of *en dvarim ka’ele/u*, originally, ‘there are no such things’

I propose that although *en dvarim ka’ele/u* has not yet fully “dismantled” the contextual constructional scaffolding, it does show some early signs of doing so, as witnessed by a few stage IV cases (marked yellow) in Figure 5.1 above. It therefore does show further constructional changes. Close inspection of the data which pertain to *en dvarim ka’ele/u* along the parameters considered in the previous sections reveals the following:

The two instances of adjectival *en dvarim ka’ele/u* found in IsraBlog corpus indicate that *en dvarim ka’ele/u* is not yet a ‘central’ adjective, because both are predicative. One of them is modified by the intensifier *kol-kax* ‘so’, as exemplified in (5.14), which is the concluding line of a raving journal article about the set of *Game of Thrones*, an American fantasy drama television series.

(5.14)

kol-kax	***en dvarim ka’ele***		*še-ze*	*nir’e*
so	**there are no such things → amazing**		that-it	looks.like
be’emet	*kmo*	*fantazia.*		
truly	like	fantasy		

‘**So amazing** that it truly looks like a fantasy.’

(IsraBlog)

I spotted no instances of the adjectival *en dvarim ka’ele/u* in the comparative or superlative forms. A Google search of the web yielded similar results (the latest of which is 18 June 2022).

No instances of negated *en dvarim ka’ele/u* was spotted either, but this is not surprising in view of the negative particle *en* ‘there’s.no’ which is part of the sequence and could clash with the negator *lo* ‘not’.

The three instances of the intensifier *en dvarim ka’ele/u* found in IsraBlog corpus, scoping over a verb, were only found in post-verbal positions. One of them collocates

with a positive adjective and one with a negative one. This is in line with the *še* 'that' (version) data (stage III; also from IsraBlog corpus) — *še-en dvarim ka'ele/u* — for which I found 23 instances following adjectives. Here too, 11 adjectives were positive and 10 were negative (the polarity of the remaining two was unclear).

A web search for evidence of more constructional changes and deeper lexicalization confirmed that no acronym was formed for *en dvarim ka'ele/u*. However, a few attributive NP examples were detected in the NP of an NP construction. Example (5.15) is representative.

(5.15) ***en dvarim ka'ele*** *šel* *rofe!!!*
there are no such things → a.wonder of a.doctor

oman *amiti,* *mikco'i* *ve-adiv.*
artist real professional and-kind

'**A wonder** of a doctor!!! A real artist, professional and kind.'
(tinyurl.com/2p88jrvu)

The fact that *en dvarim ka'ele/u* can be hosted by the NP of an NP construction, as well as the even distribution between positive and negative adjectival collocates, seem to indicate that this member of the Ultimate construction family too is "shedding" its contextual constructional scaffolding. Nonetheless, the overall distributional picture shows that *en dvarim ka'ele/u* is not as advanced on the lexicalization path as *xaval al hazman*.

5.4 Further developments of *ba livkot*, originally, 'it feels like crying' and *en milim*, originally, 'there are no words'

What about *ba livkot* 'it feels like crying' and *en milim* 'there are no words'? Both *ba livkot* and *en milim* haven't reached the stage where they can function as *bona fide* adjectives, adverbs and intensifiers in the absence of any constructional scaffolding (stage IV), as shown in Figure 5.1. Accordingly, no further constructional changes similar to those undergone by *xaval al hazman* (Section 5.2) and *en dvarim ka'ele/u* (Section 5.3) were detected. Interestingly, however, the reasons for not reaching stage IV differ for the two expressions.

Ba livkot, I suggest, hasn't reached stage IV simply because it hasn't managed to "shed" the contextual constructional scaffolding. This is due to the iconic status of the entire source syntagma *yafot še-ba livkot* 'beautiful up to a point that it feels like crying', as mentioned above at the end of Section 4.3.5.2. Indeed, *yafot* 'beautiful' is often replaced by other adjectives, but the entire syntagma has conventionalized to such an extent that it thwarts the shedding of the contextual constructional scaffolding, and consequently the transition to stage IV.

En milim, in contrast, is not a case of "not shedding" the contextual constructional scaffolding, but rather a case of not incorporating into this very same scaffolding (see

Section 4.3.5 for a detailed analysis as for the reason for this state of affairs). *En milim* hasn't reached stage IV simply because it hardly reached stage II, let alone III.

In sum, by applying a set of parameters, I showed that the depth of lexicalization, reflected by the extent of further constructional changes, depends on the removal of the constructional scaffolding context. Indeed, different parameters may point to opposite directions, e.g., the evenly distributed polarity of adjectival collocates of the intensifier *xaval al hazman* (Section 5.2.1.5) as opposed to the total lack of negated contexts (Section 5.2.1.6). But all in all, *xaval al hazman* seems to be the most advanced member of the Ultimate construction family on the cline of lexicalization, followed by *en dvarim ka'ele/u*. Both b*a livkot* and *en milim* lag behind. The former, because it has not (yet) managed to "shed" the contextual constructional scaffolding. The latter because it hardly managed to incorporate into this very same scaffolding (and in fact not motivated to do so, as accounted for in Section 4.3.5.2).

5.5 Putting everything together

Recall that Chapter 3 analyzed the preconditions that full sentences must meet in order to become idiomatic sentences (i.e., unanalyzable and semantically opaque), and full-fledged words later on (i.e., relational). These preconditions explain the infrequency of this phenomenon. The focus of Chapter 4, however, was the actual process of change in the grammatical status of these idiomatic sentences, first from independent sentences — already *functionally* idiomatic sentences — into dependent clauses (with *(ad) še* '(up.until) that…' Correlative endpoint resultant-state clause), and later, from dependent *syntactic* clauses (again, *functionally* idiomatic sentences) to full-fledged words (an adjective, an adverb or an intensifier). I emphasized the role of the constructional scaffolding (the Anaphoric degree-adverb exclamative sentence construction and the Correlative endpoint resultant-state clause) in this process.

The model that I proposed for the change in grammatical status builds on Lehmann's (1988) continua of clause linkage. Lehmann, however, focused on the nominalization of subordinate clauses (see the very many papers that cite his work on this topic), where "the [subordinate] clause becomes a nominal or adverbial constituent of a matrix clause" (p. 193) while losing elements which denote mood, tense, aspect, followed by loss of verbal inflection and finally – loss of subject. This nominalization process differs from the one I propose in this dissertation in two respects: In Lehmann's model, (i) no semantic change is necessarily involved and (ii) although the subordinate clauses start out as full sentences, the resulting words constitute just a fraction of these (full) sentences. This is not the case in the lexicalization process studied here.

The key point of my analysis, however, is the function of the idiomatic sentences, all members of the Ultimate construction family. They convey a strong stance, in line with the strong stance already conveyed by the proposition in the preceding sentence. Now, these idiomatic sentences could very well remain syntactically independent, occupying a separate evaluative utterance. As such, their expressive amplification function would have enjoyed high discourse prominence. But this expressive motivation meets a competing, economical motivation, which calls for the production

of a single complex unit instead of two. As such, they counter the natural loss of emotive force in that (preceding) sentence. In order for this incorporation to take place, there should be a (perfect) match between the newly evolved idiomatic sentences and the preceding sentence in terms of function and grammatical status. Thus, the preceding sentence can function as contextual constructional scaffolding which accommodates the newly evolved idiomatic sentences.

I showed that the Anaphoric degree-adverb exclamative sentence construction is exactly such constructional scaffolding. It is a very expressive construction, and hence, just like any other expressive proposition, it is prone to weakening. Reinforcement is then called for in order to compensate for the natural loss of emotive force. This construction, however, does not permit additional reinforcement by repetition of the intensifier (as is the case with many other intensifiers). Rather, a Correlative endpoint resultant-state clause has to be added on, emphasizing the extreme nature of the proposition conveyed by the (bare) exclamative. This provided a perfect fit with the members of the Ultimate construction family. Note that the Correlative endpoint resultant-state clause is, of course, a syntactic clause, and thus can incorporate the newly evolved idiomatic sentences (the members of the Ultimate construction family) just because despite their *functional* status as idiomatic sentences, they have not yet lost their *grammatical* status as clauses.

The match between the newly evolved idiomatic sentences and the constructional scaffolding of the Anaphoric degree-adverb exclamative sentence construction has another—just as important—aspect. The foci of this construction are of various word classes — a noun, a verb or an adjective. As independent utterance modifiers, the newly evolved idiomatic sentences, on their part, are potential flexible modifiers (by virtue of being newcomers to the lexicon, as well as mono-morphemic and therefore semantically opaque). This is why they can be embedded in the Correlative endpoint resultant-state clause slot and equally modify any of these focal elements (of the Anaphoric degree-adverb exclamative sentence construction) regardless of its word class. They are eventually reanalyzed as adjectives, adverbs or intensifiers, depending on the word class of the modified (i.e., focal) element. This match also accounts for the simultaneous emergence of the new adjective(s), adverb(s) and intensifier(s). This finding is in line with Croft's (2001) suggestion that it is the construction—rather than the word— that constitutes the primitive unit of language. In fact, it is the construction that determines the word class of the elements that it includes.

If the constructional scaffolding is no longer required, then "wordification" is complete and the linguistic forms have reached a well-defined grammatical phase. In that case, the possibility of deeper lexicalization opens up, as reflected by further constructional changes. I have shown this in the present chapter with respect to *xaval al hazman* and *en dvarim ka'ele/u*. However, If the constructional scaffolding is impossible to get rid of, then no full-fledged adjectives, adverbs and intensifiers emerge. I have shown this in the present chapter with respect to *ba livkot*.

I also showed that this whole lexicalization process can take place only in the absence of any constructional competitor, an allosentence, which just like the

Anaphoric degree-adverb exclamative hosting the members of the Ultimate construction family, makes up a single complex unit instead of two. I have shown this in Chapter 4 with respect to *en milim*.

In sum, for the drastic change in grammatical status described (in the last two chapters) to "go all the way", what is required is (i) the "right" constructional scaffolding, (ii) the lack of any constructional competitor to the "right" constructional scaffolding, and (iii) easy dismantling the constructional scaffolding, once not needed anymore.

All in all, I showed that construction-based motivations *alone*, and the network links between the constructions can explain the typologically rare grammatical change of independent full sentences into full-fledged words. This has been done by outlining a parsimonious—yet exhaustive—model for this lexicalization process. This analysis provides (yet another) piece of evidence for the claim that "there is no discrete cut off point between grammar and lexicon" (Bybee, 1998: 429). They may therefore belong to the same level of representation, regardless of level of syntactic complexity.

5.6 What's next?

Naturally, the analysis presented in this chapter relies on the availability of diachronic data. These data, however, were not available to me at the early stages of my research. The (temporary) lack of diachronic data prompted me to propose several methods to substantiate semantic change when only synchronic data are at hand. These methods are described and applied in the next chapter.

Chapter 6: How to detect semantic change in the absence of a diachronic corpus – Getting around a methodological problem by using the notion of EVALUATION

תֶּבֶן אֵין נִתָּן לַעֲבָדֶיךָ וּלְבֵנִים אֹמְרִים לָנוּ עֲשׂוּ -- שמות ה', ט"ז
(There is no straw given unto thy servants, and they say to us: Make brick;
-- Exodus, 5, 16)

"There is nothing like first-hand evidence", he remarked.
-- Arthur Conan Doyle, *A Study in Scarlet*, 1887: Ch. 4

In this chapter I introduce three methods I devised in order to substantiate semantic change in the absence of a diachronic corpus or a synchronic corpus tagged for speakers' age. All three rely on evidence from speakers' metalinguistic activity. They are, in fact, applications of Weinreich, Labov and Herzog's (1968) notion of EVALUATION which implies that "[…] changes [can] be evaluated in terms of their effects […] upon communicative efficiency (as related, e.g., to functional load), and on the wide range of nonrepresentational factors involved in speaking" (p. 101).

6.1 The problem and a proposed solution

Semantic change involves meaning change of syntagmas. Establishing the semantic change of a syntagma requires many instances of that syntagma in a diachronic corpus. Each instance is paired with a meaning and associated with a specific period of time. A distribution of meanings as a function of time can be drawn, allowing the researchers to decide whether semantic change has occurred, or not.

But what if some semantic change is evidently in progress, but linguists only have a synchronic corpus at their disposal? If the available synchronic corpus is tagged for speakers' age, then running an apparent time analysis (Bailey, Wikle, Tillery, & Sand, 1991; Labov, 1963; 1994: Ch. 3) is a possible solution. But such a corpus is not always available. How, then, can one substantiate the presence of semantic change and its direction, as well as pinpoint the stage of change, given only a synchronic corpus not tagged for speakers' age?

In this chapter, I introduce three methods to accomplish this task, given such challenging data. The proposed methods examine the semantic polysemy (as manifested in a synchronic corpus) created by the semantic change from two perspectives: The cognitive perspective, which examines the salience of simultaneous coded meanings in the minds of speakers; and the sociopragmatic perspective, which examines the conventionalization status of simultaneous coded meanings across different speech communities. One method can detect semantic change, its direction and its stage. The second can detect semantic change and its direction. The third, ancillary method, can detect (under certain conditions) the stage of change, whether initial or advanced.

My proposal is based on speakers' *metalinguistic activity*. Each of the methods builds on a different type of metalinguistic activity, but all alike "let the speakers do the talking", thus allowing the researchers to tap into the speakers' minds. These methods can thus reduce the need to rely solely on researchers' potentially subjective interpretive interventions otherwise needed in identifying speakers' intentions.

I will apply these three methods to the semantic change of *xaval al hazman* (originally, 'it's a waste of time'), discussed in detail in the previous chapters. *Xaval al hazman* was chosen because it offers many examples illustrating the three methods. In order to support the effectiveness of these methods, while simultaneously pointing to their limitations, each of these methods will also be applied to three other semantically changed syntagmas, *en dvarim ka'ele/u* (originally, 'there are no such things') and *ba livkot* (originally, 'it feels like crying'), also discussed in the previous chapters, and *sof haderex* (originally, 'the end of the road').

I start by introducing the cognitive and the sociopragmatic perspectives on polysemy and semantic change.

6.2 Two perspectives on polysemy and semantic change

6.2.1 The cognitive perspective: Salience reversal

The cognitive perspective "view[s] the word in the minds of the speakers with regard to its entrenchment in the individual mental lexicons of the speakers and the conceptual status it has achieved there" (Schmid, 2016 [2011]: 71). Degree of entrenchment affects degree of cognitive salience (Schmid, 2007), which refers to the mental accessibility of concepts.

I am here concerned only with context-*in*dependent lexical salience, and adopt Giora's (1997, 2003) *Graded Salience Hypothesis* in defining lexical entrenchment.[74] The Graded Salience Hypothesis focuses on synchronic polysemy, where the various coded meanings of a syntagma (a single word or a phrase) lie along the salience/non-salience continuum. The more salient a meaning is, the more highly it is ranked for prominence in the mental lexicon. Such prominence depends on degree of

[74] Context-dependent salience which is the outcome of concept activation induced by the prior context (e.g., Ariel, 1990; Chiarcos, Claus, & Grabski, 2011; Gibbs, 1986, 1994, 2002; Jaszczolt & Allan, 2011) is not relevant to the present study.

entrenchment, a combination of cognitive factors (such as prototypicality or individual relevance) and degree of exposure (resulting from experiential familiarity, frequency, or conventionality). The salient meaning enjoys speed superiority of processing over the corresponding low-salience meaning. The former springs to mind unconditionally when the syntagma is encountered, whereas the latter is activated more slowly, regardless of prior context (for similar views, see Duffy, Morris, & Rayner, 1988; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Williams, 1992).

As is well-known, linguistic expressions undergoing semantic change do not shift their original meaning, X, to an innovative meaning, Y, abruptly. Typically, there is a period of polysemy between the coded X and Y (e.g., Blank, 2001: 1597, 1603; Bréal, 1964 [1899]: Ch. 14-15; Brinton & Traugott, 2005: 21; Detges, 2010; Hopper, 1991; Hopper & Traugott, 2003 [1993]; Traugott & Dasher, 2002) which can even last hundreds of years (Traugott & Dasher, 2002). During this period of co-existing meanings, "[the] relationship [of X and Y] to each other in terms of *saliency* may change" (Traugott and Dasher 2001: 12; emphasis mine). In other words, when speakers notice semantic change, what they experience is the emergence of polysemy, accompanied by changes in the relative salience of competing coded meanings. Accordingly, the process of semantic change can be described in terms of *salience reversal*, where the innovative Y gradually "takes over" the syntagma, "pushing aside" the original X.

Indeed, the mere co-existence of X and Y— different simultaneous coded meanings of the same syntagma — in a synchronic corpus is suggestive of semantic change. It cannot indicate, however, the direction of change, nor its stage. When only a synchronic corpus is available, explicit evidence for semantic change, its direction and its stage, I suggest, can be deduced using alternative tools derived from the Graded Salience Hypothesis (Giora, 1997, 2003), which attest to salience reversal. These tools take into account the metalinguistic activity of speakers who are linguistically sensitive to the salience reversal of simultaneous coded meanings, as I will show in Sections 6.5 and 6.6 below.

6.2.2 The sociopragmatic perspective: Contrastive lexical choices

The sociopragmatic perspective "view[s] the word in the speech community with regard to the extent of its spread and diffusion, i.e. the degree of use and familiarity among the members of the speech community" (Schmid, 2016 [2011]: 71).

Much the same way 'salience reversal' can gauge semantic change in the minds of the speakers (when semantic change is examined from the cognitive perspective), 'degree of conventionality' of coded meanings across speech communities can also gauge semantic change (when semantic change is examined from the sociopragmatic perspective). A 'linguistic convention', according to Schmid (2020: 88), is a regularity of linguistic behavior, norms which the members of a speech community conform to and expect each other to conform to.

In the absence of data tagged for date of production (i.e., lack of a diachronic corpus), the well-known apparent time analysis can be applied to a corpus tagged for speakers'

age in order to substantiate a sematic change (Bailey et al., 1991; Labov, 1963; 1994: Ch. 3). The contrast between older and younger speakers, lexically-rigid versus lexically-flexible, respectively, in terms of monosemy versus polysemy, may reflect semantic change and its direction (but see Petré & Van de Velde, 2018: 869 for criticism).

I propose that even a synchronic corpus which is not tagged for speakers' age can provide explicit evidence for semantic change and its direction. This builds on (further) exploiting the notion of lexical flexibility, by *contrasting the lexical choices* of conservative versus flexible speech communities. This approach takes into account the metalinguistic activity of speakers who are linguistically sensitive to the effect of new coded meanings on (their) social identity, as I will show in Section 6.7 below.

Note that while there are models that incorporate both aspects of semantic change — the cognitive and the sociopragmatic (e.g., Baxter and Croft (2016); Petré and Van de Velde (2018); Schmid (2015, 2020) — this is not my goal in this chapter. Rather, I wish to highlight a specific facet of each of them, a facet that appears to be useful for detecting semantic change — speakers' metalinguistic activity.

The specific aspects of speakers' metalinguistic activity relevant to the methods I here propose are presented in the next section.

6.3 Metalinguistic activity

In the present study, the term *metalinguistic activity* does not refer to any formal tools that logicians use in order to compute the truth value(s) of sentences in natural languages (e.g., Carnap, 1970 [1939]; Tarski, 1944), nor does it refer to any cognitive model *à la* Culioli (e.g., 1990: 177-213; 1995) used by linguists to sketch schematic representations of utterances.[75] In the present study, the term metalinguistic activity refers to certain everyday language uses made by laypersons. After all, "[l]inguistics is not an arrogant discipline that does not care about the layman's opinion" (Kabatek, 2015: 224) and would therefore consider laypersons' intuitions invaluable.[76]

---

[75] Note, however, that Culioli (1995) acknowledged that

> "[a metalinguistic set of representations] can mean a great many extremely varied things, such as using the *gloss* speakers produce when, given a text, we ask them to make utterances or equivalent commentaries. Roughly speaking, we say: "I don't understand. Could you please reformulate your statement? What do you mean by that?" […] Language activity, hence languages, has the potential of being used for metalinguistic purposes." (p. 24; original emphasis)

Nevertheless, he uses the term 'metalinguistic' in connection with his formal cognitive model rather than in certain uses of everyday language.

[76] See also Traugott and Trousdale (2013: 22) who note that "[t]ypically, language-users may not be aware of the change having occurred (Keller 1994), but sometimes there are metatextual comments made by grammarians or *others* who observe change." [emphasis mine]

Jakobson (1960) considered this kind of metalinguistic activity one of the six functions of language as a system of communication. A metalinguistic message, he suggested, conveys information about the lexical code rather than about objects in the real world. The interaction in Example (6.1) (taken from Jakobson) illustrates metalinguistic messages, specifically, glossing.

(6.1) *A: The sophomore was plucked.*
B: But what is plucked?
A: Plucked means the same as flunked.
B: And flunked?
A: To be flunked is to fail an exam.
B: And what is sophomore?
A: A sophomore is (or means) a second-year student.

(Jakobson 1960: 356)

A metalinguistic message can also explain, predicate, or comment on the meaning of propositions (Hübler & Bublitz, 2007: 2), as exemplified in (6.2).

(6.2)

ani	*xoševet*	*še-yeš*	*li*	*peruš*	*xadaš*
I	think	that-there.is	to.me	interpretation	new

l-a-bituy	*"emor*	*li*	*mi*	*xaverexa*
for-the-idiom	tell	to.me	who	your.friends

ve-omar	*lexa*	*mi*	*ata."*
and-I.will.tell	to.you	who	you

'I think I have a new interpretation for the idiom "Tell me who your friends are and I will tell you who you are." '

(tinyurl.com/43cbutn7)

Just like Jakobson, Weinreich (1966: 162-163) considered metalinguistic activity at the level of everyday language use. He suggested a vocabulary of metalinguistic operators that can attest to metalinguistic activity of the semantic type, for example, *real*, *so-called*, *strictly speaking* (cf. Reichenbach, 1947: 9ff., 344-346; Schiffrin, 1980), all intended to resolve ambiguity.[77]

But metalinguistic activity is not limited to talking about the linguistic code. Metalinguistic activity is an umbrella term referring to any activity of the interlocutors, involving "*conscious* management (reflection or an intentional control over) of the language objects, either as objects per se or in terms of the use to which they are put"

[77]And see also Bateson (1972: 183-198) for an anthropological context.

(Gombert, 1992: 4; emphasis mine; see also Verschueren, 2004 and many references in both).

Accordingly, wordplay also necessarily involves metalinguistic activity (in this case implicit),[78] because speakers must consciously recruit their metalinguistic understanding of linguistic signs in order to produce wordplay of various kinds (Attardo, 1994: Ch. 4, and p. 329; 2018; Delabastita, 2001; Gombert, 1992: 114-119; MacLaren, 1989; Yaguello, 1998; Zirker & Winter-Froemel, 2015a; and see also various authors in Zirker & Winter-Froemel, 2015b).

Metapragmatics is also subsumed under the heading of metalinguistic activity (Verschueren, 2004) or, at least, considered a cognate concept (Jaworski, Coupland, & Galasiński, 2004). Metapragmatics is the reflexive awareness of language users regarding the language they use and its potential interpretation in context (Culperer & Haugh, 2014: Ch. 8). It should be noted that the term metapragmatics is used to describe many related phenomena (see, Caffi, 1994; Hübler & Bublitz, 2007: Ch. 1) which are far beyond the scope of my study. Relevant to this chapter, however, is the effect metapragmatics has on consistent conscious linguistic behaviors leading to "normative ideas about language use that are shared across *particular social groups*" (Culperer & Haugh, 2014: 255; emphasis mine), that is, linguistic behaviors affected by the sociocultural background of the interlocutors.

I will here show that,

(i) metalinguistic comments can testify to the semantic change of a given syntagma, its direction and its stage, in terms of relative cognitive salience;
(ii) in some cases of semantic change, wordplay can attest to the relative cognitive salience of simultaneous coded meanings (of the same syntagma), thus potentially corroborating the results of method (i) with regard to the stage of change;
(iii) language use of particular social groups (here ideological, linguistically conservative speakers) can provide further support for the results of method (i) with regard to the semantic change and its direction.

As already noted in Section 6.1, the proposed methods will be applied to four specific syntagmas claimed to have undergone semantic change. I will, therefore, precede each analysis (of each syntagma) with evidence from a diachronic corpus attesting that semantic change has indeed occurred. I start with *xaval al hazman*, which has already been shown to have undergone semantic change in the previous chapters. See Figure 4.4, repeated here for convenience, as Figure 6.1.

[78] Jakobson (1960) considers wordplay a manifestation of the *poetic* — not the *metalinguistic* — function of language. But these two functions are closely connected (as suggested by Jakobson himself and also by Lyons, 1977: 55).

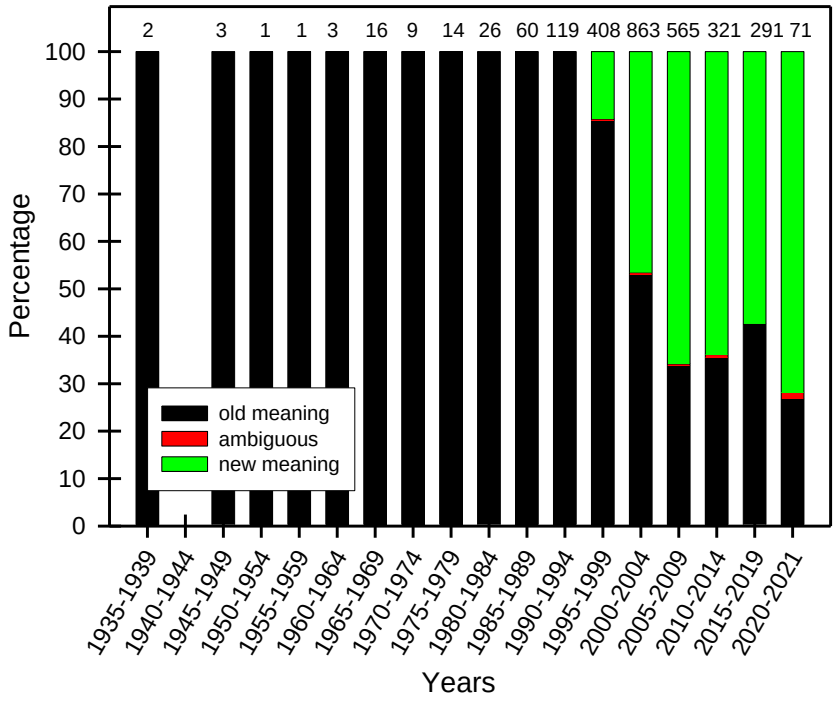

Figure 6.1: The distribution of the old versus the newly evolved meaning(s) of *xaval al hazman* as a function of time. The number of counts for every five-year interval is marked on top of each bar.[79] Note that items of different grammatical statuses were considered *en bloc*. Items classified as either accompanied by metalinguistic comments or simply names (of books, newspaper columns, and the like) were filtered out. Data extracted from Yedioth Ahronoth corpus.

I can now move on to discussing to what extent the synchronic corpus from which I extracted all the data (used to demonstrate the proposed methods for detecting semantic change) is up to the task(s).

6.4 The web-based corpus used in this chapter

Due to its size (1.0×10^9 tokens; and see Section 2.5), HeTenTen corpus is expected to provide very many instances of any neologism (Leech, 2007), including, of course, those considered in this chapter.

But as evident from Figure 6.1, the semantic change undergone by *xaval al hazman* seems to have occurred during the nineties of the twentieth century, prior to the birth of Web-2.0, which was/is fed by ordinary people's contributions (see, Blank & Reisdorf, 2012; O'Reilly & Battelle, 2009). By that time, the innovative positive (and

[79] I assume that the change in total counts is possibly a result of change in the overall number of tokens. Since the overall number of tokens is not available, this assumption is based on the changed number of articles printed in Yedioth Ahronoth over the years.

intensifying) *xaval al hazman* (see Examples 1.2) has already made it into the lexicon, downgrading the salience of the original, negative *xaval al hazman* (see Example 1.1). In light of this state of affairs, HeTenTen corpus is seen as unsuitable for detecting semantic change, for it is not only synchronic, but it also postdates the semantic change. Nonetheless, I here propose three methods to overcome these shortcomings.

In the next sections (6.5-6.7), I present the proposed methods, in turn, and apply each of them to *xaval al hazman*. I start with the metalinguistic comment-based method.

6.5 The metalinguistic comment-based method

The method introduced in this section is derived from Giora's (1997, 2003) Graded Salience Hypothesis (Section 6.5.1) which, I claim, allows to detect semantic change, its direction, as well as the stage of the change (Section 6.5.2). This method is related to prior research on metalinguistic comments in the context of semantic change (Section 6.5.3), although the research goals are different.

6.5.1 Metalinguistic comments as a means to activate a low-salience meaning

As noted above (in Section 6.2), the Graded Salience Hypothesis maintains that access to the various coded meanings of a given syntagma is ordered: Salient meanings are activated instantly when a stimulus is encountered; low-salience meanings, albeit coded, are activated more slowly because they are less cognitively prominent. Still, speakers *do* sometimes intend the low-salience meanings. How, then, do they get the low-salience meaning across?

Givoni (2020), and Givoni et al. (2013) introduced the *Low-Salience Marking Hypothesis*, proposing that speakers explicitly alert addressees to the need to access low-salience meanings by using low-salience markers, such as (the Hebrew) *be'emet* ('really', 'truly'), *literali* ('literally'), *bimlo muvan hamila* ('in the full sense of the word') and *tartey mašma* ('double entendre') (cf. Katz & Ferretti, 2003; Nerlich & Clarke, 2001; Norén & Linell, 2007). Givoni and Givoni et al. ran off-line rating experiments and on-line reading-time experiments where participants were presented with syntagmas of multiple meanings, such as *it's all down in black and white* (either literally, 'written in black ink on white paper' or idiomatically 'the message is stated in the clearest terms'). Results showed that in the absence of low-salience markers, it was the salient meaning (here, the idiomatic one) that was activated; However, when the same syntagmas were followed by low-salience markers, those markers induced a meaning shift, evoking the low-salience meaning too (here, the literal one).[80] These *markers* constitute *metalinguistic comments*, in fact.

The existence of metalinguistic comments begs the following question: Why would speakers chose a lexical item known to be polysemous (thus violating Grice's, 1975 *Manner Maxim*), in the first place, and then resolve the entanglement by means of metalinguistic comments? I suggest that the juxtaposition of an element of an emotive

[80] Note, that low-salience marking is not necessarily intended to suppress the salient meaning, but rather to highlight the low-salience meaning.

nature — here, *xaval al hazman* — and a metalinguistic comment serves the (rather obvious) rhetorical goal of the speaker/writer to convey a strong(er), highly expressive and extravagant message. As such, the use of *xaval al hazman* is an instance of strategic ambiguity (according to the model presented in Winkler, 2015; see a detailed parameter-based model in Winter-Froemel & Zirker, 2015), from which not only the addressee but also the researchers may benefit, as I show in the next sections.

6.5.2 Metalinguistic comments as a means to spot semantic change, its direction and stage of change

I extracted all 2954 instances of *xaval al hazman* from HeTenTen corpus, and identified 142 tokens accompanied by some metalinguistic comment about their meaning. These 142 metalinguistic comments can be divided into two different subsets.

(a) 109 metalinguistic comments accompanying the syntagma undergoing semantic change (i.e., the neologism), marking that syntagma by word-pointers, such as "the word", "the term", "the concept" and "the expression". These word-pointers are intended to indicate that the neologism is *mentioned* rather than used (Lyons, 1977: 5 ff.) and is therefore the subject matter (Svanlund, 2018). Most often, these are cases where speakers share their attitude towards the new *xaval al hazman* (among other neologisms) with their addressees. Most speakers take a prescriptive stand, as in Example (6.3), originally in Hebrew.

(6.3) Within the last year or two, I studied negative idioms and their psycholinguistic aspects […] Although I was careful to use ***xaval al hazman*** only in the old (and literal) sense [here, 'it's a waste of time' – IB], I eventually gave up, and soon ***xaval al hazman*****s** came out of my mouth, every now and then.
(HeTenTen, tinyurl.com/2p8m7ncz)[81]

(b) 33 Metalinguistic comments accompanying the syntagma undergoing semantic change (i.e., the neologism), which is *used* naturally, here as a modifier of modifiable elements (see Examples 1.2 above), rather than just mentioned (thus constituting the subject matter, as in (a) above). Kerremans (2015: 20) dubbed this kind of usage "object-linguistic usage". The text in bold in Example (6.4), originally in Hebrew, illustrates this kind of usage of metalinguistic comments, while the subsequent underlined text illustrates the other kind of usage, mentioned in (a) above.[82]

[81] This is an example from HeTenTen corpus complemented by a direct link to the relevant web-site.

[82] The distinction between these two different subsets (i.e., (a) and (b)) can be seen as a *de dicto/de re* distinction, where "*de dicto* […] refers to parts of the discourse as linguistic forms rather than to the semantic content of the forms (*de re*)" (Hopper & Traugott, 2003 [1993]: 185).

(6.4) My blog also looks ***xaval al hazman*** **(but in the version of 1994** [here, 'it's a waste of time' – IB], <u>one second before the meaning of this expression, *xaval al hazman*, changed</u>).

(HeTenTen, tinyurl.com/f89t8d48)

Metalinguistic comments of type (b) are what Givoni (2020), and Givoni et al. (2013) dubbed *low-salience markers* and are the focus of my analysis.

A close inspection of the contents of these 33 instances enables the description of this semantic change in more detail. Examples (6.5a-g), are representative. In each of them, the speaker indicates explicitly the meaning of *xaval al hazman* she wishes to communicate (boldfaced), thus revealing that semantic change has indeed occurred. Example (6.5a) indicates that *xaval al hazman* had a different meaning in the past. Example (6.5b) indicates that it has a current, innovative meaning (used by youngsters). Example (6.5c) is ambiguous. It alludes to the idiomatic nature of either the new or the old *xaval al hazman* (or both?). Example (6.5d) attests to a contrast between two meanings, "in slang" (i.e., idiomatic) and "for real" (i.e., literal). Example (6.5e) indicates that the old meaning of *xaval al hazman* is not slangy, and Example (6.5f) indicates that the current meaning is slangy. Taken together, Examples (6.5e) and (6.5f) resolve the ambiguity of Example (6.5c). The new meaning is the idiomatic one, simply because it is slang. If the old meaning is literal (as implied by the combination of Examples 6.5d and 6.5e), namely of a negative meaning, then the new meaning must be the positive one, as indicated by Example (6.5g).

(6.5) a. *xaval al hazman* ***(b-a-muvan*** ***he-atik)***.
it's a waste of time **in-the-sense** **the-ancient**

'It's a waste of time **(in the ancient sense)**.'

(HeTenTen, tinyurl.com/5na839w6)

b. ***ex*** ***ha-dor*** ***he-xadaš*** ***haya*** ***megiv?***
how **the-generation** **the-new** **would.have** **reacted?**

xaval al hazman!
it's a waste of time!

'**How would the young generation have reacted?** It's a waste of time!'

(HeTenTen, tinyurl.com/4sekhufp)

c. *xaval al hazman* ***b-a-muvan*** ***he-xadaš*** ***šel*** ***ha-bituy***.
it's a waste of time **in-the-sense** **the-new** **of** **the-expression**

'It's a waste of time **in the new sense of the expression**.'
(HeTenTen, tinyurl.com/yk4k88de)

d. *xaval al hazman* ***(ve-ha-pa'am*** ***lo*** ***be-sleng***
it's a waste of time **and-this-time** **not** **in-slang**

ela ***be'emet)***.
but **for.real**

'It's a waste of time **(and this time not in slang but for real)**.'
(HeTenTen)[83]

e. *"xaval al hazman"* **—** ***b-a-muvan*** ***ha-yašan***
it's a waste of time **in-the-sense** **the-old**

ve-ha-lo-slengy…
and-the-not-slangy

' "It's a waste of time" **in the old and non-slangy sense**…'
(HeTenTen, tinyurl.com/4dbya9hm)

f. *"xaval al hazman"* ***(b-a-mašma'ut*** ***ha-slengit***
It's a waste of time **in-the-meaning** **the-slangy**

ve-ha-axšavit).
and-the-current)

'it's a waste of time (**in the slangy and current meaning)**.'
(HeTenTen, tinyurl.com/mt5jvb4w)

[83] This is an example from HeTenTen corpus, but unfortunately a direct link to the relevant web-site is broken.

g. *"xaval al hazman",* ***ve-lo*** ***b-a-muvan*** ***ha-xiyuvi***
it's a waste of time **and-not** **in-the-sense** **the-positive**

šel ***ha-inyan.***
of **the-matter**

' "It's a waste of time", **and not in the positive sense of the matter**.'
(HeTenTen)

The set of 33 items (of type b) is further divided into two subsets:

(bi) 21 items in which speakers use metalinguistic comments to invite the old (negative) *xaval al hazman*, and
(bii) 12 items in which speakers use metalinguistic comments to invite the new (positive and intensifying) *xaval al hazman*.

A comparison between (bi) and (bii) shows that there are more negative (n = 21) than positive (and intensifying) (n = 12) inviting items for *xaval al hazman*. The difference is marginally significant (binomial test, p = 0.081), and may suggest that currently, the old meaning is the non-salient one, while the new meanings are the salient ones.

For the sake of reproducibility, I repeated this procedure with another web-corpus, *Seret* (see Section 2.4), where I detected no instances of *xaval al hazman* marked by word-pointers (as in (a) above). I did, however, detect 17 items in which *xaval al hazman* was used naturally as a modifier (as in (b) above), rather than as the subject matter (as in (a) above). In all these cases, the speakers used metalinguistic comments to invite only the negative *xaval al hazman*, thus testifying to its current status as a low-salience meaning, and by implication the current high salience of the positive (and intensifying) *xaval al hazman*.

6.5.3 The metalinguistic comments discussed in this chapter in light of the previous literature on metalinguistic comments

The scarce literature on metalinguistic comments in reference to semantic change has only considered metalinguistic comments of subset (a) in Section 6.5.2 above, those used as word pointer to neologisms which are mentioned rather than naturally used. This literature considered metalinguistic comments (i) as cues to speakers' awareness of neologisms (Schmid, 2008: 11, 16-17) and therefore (ii) as cues to the stage of conventionalization (Fischer, 1998: 176-178; Svanlund, 2018). It also examined (iii) the effect comments have on the frequency of conventionalization (Kerremans, 2015; Svanlund, 2018). Although different from each other, these studies share the same conclusion: Metalinguistic comments which serve as word pointers accompany neologisms during the early stages of lexicalization, but hardly ever during advanced stages.

The metalinguistic comments considered in this chapter, those termed "object-linguistic usage" (see subset (b) in Section 6.5.2 above), however, are different in the

sense that they are not intended to explicate the new meaning of neologisms as in Fischer's, Schmid's, Kerremans' and Svanlund's works, but to highlight the intended meaning. As I have shown above, these comments testify to the non-salient status of the old meaning of *xaval al hazman* and to the salient status of the new *xaval al hazman*, a neologism at an advanced stage of lexicalization, perhaps not even a neologism anymore.

Crucially, both types of comments, those considered in previous literature and the ones considered in this chapter, have the very same goal. Both invite/clarify the low-salience meaning, thereby implying what the salient meaning is, and consequently what the stage of change is.

In sum, the metalinguistic comment-based method (which considers comments of the "object-linguistic usage" kind) has been shown to indicate semantic change, its direction and its stage — a new idiomatic meaning for *xaval al hazman* with a positive flavor, alongside an old meaning of a negative flavor, in line with the results presented in the previous chapters and illustrated in Figure 6.1. The current salient meaning appears to be the new meaning.

In the next section I introduce an ancillary method intended to examine whether the current salient meaning of *xaval al hazman* is indeed the one indicated by the metalinguistic comment-based method.

6.6 The wordplay-based method

The method introduced in this section is also derived from Giora's (1997, 2003) Graded Salience Hypothesis (see Section 6.2.1). It is intended to determine the current salient meaning of a syntagma from among several competing coded meanings. It is based on a specific type of wordplay, the *Optimal Innovation* (Giora et al., 2004), described in the next section.

6.6.1 What is an Optimal Innovation?

Wordplay results from any intentional manipulation of linguistic material, phonetically, semantically, or grammatically, and its core function is to produce a humorous effect (among several other functions; see, for example, Attardo, 1994; Thaler, 2016; Winter-Froemel, 2016).

Optimal innovation is a kind of wordplay. Giora et al. (2004) defined an optimal innovation of a syntagma as an innovation that constitutes a qualitative—not quantitative—variation on an underlying formally-close syntagma. By "qualitative variation", Giora et al. implied a modification to the underlying syntagma which is not trivial as pluralization is, for example, but rather one that constitutes a meaning contrast between the novel and the underlying syntagma. In the conventional terms of wordplay research, an optimal innovation is *paronymy in absentia* (Winter-Froemel, 2016). Crucially, the "underlying formally-close syntagma" (i.e., the one that underlies the surface form) must be a collocation, most often an idiom.

An optimal innovation — unlike a pure innovation (see Appendix F) — allows for the automatic recoverability of the underlying syntagma, which means that the meaning

of the underlying syntagma is its salient meaning (see Section 6.2.1). The meaning contrast between the optimal innovation (i.e., the surface form) and the underlying syntagma is what makes addressees rate the optimal innovation(s) as highly enjoyable. For instance, *Body and Sole*, as the name of a shoe store,[84] is an optimal innovation, for it automatically activates the salient meaning of the underlying *body and soul*, the relevant idiomatic meaning.

6.6.2 Optimal Innovation as an indication of the salient meaning of a polysemous underlying syntagma

Now, imagine that the underlying covert syntagma of the optimal innovation is in itself polysemous, where both meanings are collocations. Following Giora's (1997, 2003) Graded Salience Hypothesis (see Section 6.2.1), the evoked meaning of the underlying syntagma must be its salient meaning (from among other potential meanings), regardless of context. In fact, according the Graded Salience Hypothesis, this is the case with exposure to any syntagma (whether involved in wordplay, or not).

So, if a synchronic corpus displays a collection of optimal innovations (cases of *paronymy in absentia*) based consistently on one out of several potential meanings of a specific underlying syntagma (all of which are collocations), then this meaning of the underlying syntagma is its current salient meaning. This is exactly how this wordplay can reveal the salient meaning of a given syntagma in a synchronic corpus.

Appendix F presents the criteria for identifying an optimal innovation, and their application to *xaval al hazman*. I submitted queries to HeTenTen corpus implementing these criteria and spotted 21 instances of optimal innovation based on *xaval al hazman*. *Xašmal al hazman* (*xašmal* 'electricity'), for example, is the catchy name of an on-line electrical appliance store[85], based, no doubt, on the positive idiomatic *xaval al hazman*. *Xalav al hazman* (*xala*v 'milk'), on the other hand, is the title of an article which lists the shortcomings of consuming cow milk, and it is based on the (old) negative collocation *xaval al hazman*.

An informant, versed in detecting instances of optimal innovation, was presented with all 21 examples of optimal innovations based on *xaval al hazman*. I asked her to determine whether they are indeed cases of optimal innovation, and to further determine for each case, which of the two underlying meanings of *xaval al hazman* serves as the substrate for innovation. Indeed, there are more cases where the syntagma underlying the optimal innovations is the positive (and intensifying) idiomatic *xaval al hazman* (n = 14) than the original, negative counterpart (n = 7). This difference is only marginally significant (binomial test, p = 0.055). This state of affairs testifies to the current higher

[84] https://tinyurl.com/23z63vu3

[85] https://tinyurl.com/hsezydwz

salience of the positive (and intensifying) *xaval al hazman* over the negative *xaval al hazman*, in line with the results of the metalinguistic comment-based method above.[86]

Note that this wordplay-based method is applicable in the case of *xaval al hazman* because *both* meanings, the old and the new one, are collocations. Obviously, this in not always the case, as I will show below.

All in all, taking a cognitive approach, I relied on the Graded Salience Hypothesis to substantiate the claim about semantic change from a negative to a positive (and intensifying) *xaval al hazman*. This change is at an advanced stage, where the new positive (and intensifying) meaning is becoming the salient meaning, "taking over" the syntagma and downgrading the salience of the original negative meaning.

In the next section, I take a different approach, a sociopragmatic approach to detect semantic change (again, when only a synchronic corpus is available).

6.7 The conservative speaker-based method

Apparent time analysis assumes sociolectal differences between speakers of the same language. So do I, which is why I can, after all, use synchronic data not tagged for speakers' age. I suggest comparing the lexical choices of a speech community known to be relatively lexically-conservative (and therefore lexically-rigid) with the lexical choices of the general population, the *adopters* and the *laggards*, respectively, in terms of Rogers' theory of diffusion of innovations (2003 [1962]: Ch. 7; see also Section 6.7.2 below). Given that the general population is ahead of the lexically-conservative community in adopting neologisms (as I will show in Section 6.7.2), a difference in choices between the two groups may indicate semantic change, as well as its direction.

As it happens, Hebrew-speaking religiously devout ultra-orthodox Jews in Israel constitute a lexically-conservative community (with respect to the general, mostly secular population of Hebrew speakers). In the next sections I explain who the ultra-orthodox are (Section 6.7.1), I focus on their attested lexical choices in Hebrew (Section 6.7.2), and compare the frequency of *xaval al hazman* in their speech to that of the general population of Hebrew speakers in Israel (Section 6.7.3).

6.7.1 Who are the ultra-orthodox Jews?

The religious observance of Jews in Israel is ordered on a scale ranging "from secular, moderately observant to orthodox-religious and ultra-orthodox" (Henkin, 2020: 61). The ultra-orthodox Jews, who constitute 16.5% of the Jewish population in Israel (Cahaner & Malach, 2021), voluntarily adopted a policy of cultural separatism from the surrounding secular majority, which is manifested via their separate residential areas, their unique-traditional dress code and their independent educational system (e.g., Baumel, 2006; Fridman et al., 2011; Friedman, 1991; Perry-Hazan, 2013). Many group members and leaders are wary of outsiders who might induce acculturation (e.g., Baumel, 2006: 6-7; Isaacs, 1999b; Spiegel, 2011: 19-20, 29; Tannenbaum & Abugov,

[86] My attempt to also detect optimal innovations in Seret corpus (see Section 2.4) failed. This is not at all surprising. If HeTenTen corpus of 1.0×10^9 tokens produced 21 instances of optimal innovation, then Seret corpus of 1.5×10^6 tokens would be expected to produce less than one instance.

2010: 80, 85 and endnote 2), which explains why research that requires direct contact with ultra-orthodox, whether of a linguistic nature, or not, is scarce. It should be noted, however, that the ultra-orthodox community is not monolithic with respect to cultural separatism.[87]

As part of the cultural separatism and the wish to form a unique group identity, some of the ultra-orthodox leaders encouraged the use of Yiddish as the community L1 (Assouline, 2015, 2017, 2018; Munro, 2022), but as a result of the economic dependence of the ultra-orthodox on the secular majority and the need to communicate with the general population, some sects adopted Hebrew as their L1, despite the decree of their leaders (e.g., Glinert & Shilhav, 1991; Isaacs, 1999a).[88]

In the next section I summarize the literature which examined the actual Hebrew used by the ultra-orthodox community in light of the voluntary cultural separatism and language policy.

6.7.2 The attested lexical choices of Hebrew-speaking ultra-orthodox Jews in Israel

The lexical choices of ultra-orthodox Hebrew speakers, as described in the literature, are (almost) always compared to their secular counterparts (Bar-Asher, 2012: 84-86, 90-91; Ben-Rafael, 2002: 72-74; Schwarzwald, 2002: 153-154).

Comprehensive evidence for the differences between ultra-orthodox and secular Hebrew was provided by Sela (2004) and Cohen (2008), who examined various linguistic aspects of the written language used by ultra-orthodox, opinion journalism of male publicists and fiction written by female authors, respectively. Both Sela and Cohen conclude that the language used by ultra-orthodox writers, whether men or women, is different from their secular counterparts on the lexical, morpho-syntactic and rhetorical aspects.[89] With regard to the lexical aspect, Sela showed that the lexical choices of ultra-orthodox publicists draw on early layers of Hebrew — Mishnaic and Talmudic exegeses. She clearly stated that "[t]here isn't the slightest doubt that a colloquial idiom like *xaval al hazman* will be found [in ultra-orthodox dailies]" (p. 23; translation mine).

---

[87] The ultra-Orthodox community is not homogenous. It is divided roughly into Hassidim and Mitnagdim/Litvish (historically 'opponents' of the Hassidic movement). The Hassidim are considered more separatist than the Mitnagdim. Each of the two factions is divided into further sects (e.g., Friedman, 1991: Introductory Ch.; Heilman & Friedman, 1991; Loewenthal, 2013; Perry-Hazan, 2013: Ch. 3).

[88] The declared motivation behind using Yiddish instead of modern Hebrew and (most probably) behind the hesitancy to adopt neologisms for those ultra-orthodox who do use modern Hebrew is the following Talmudic rabbinic *midraš* (=an exegesis related to — or an interpretation of — biblical themes):

Israel were redeemed from Egypt on account of four things; because they did not change their names, **they did not change their language**, they did not go tale-bearing, and none of them were found to have been immoral." -- Leviticus Rabbah 32 (http://tinyurl.com/ym9f4tuf).

The Hebrew referred to in this specific *midraš* is pre-modern Hebrew as well as Aramaic.

[89] Note that both genres — opinion journalism (Sela, 2004) and fiction (Cohen, 2008) — are cases of planned formal speech, although the former is, as argued by Shlesinger (2000: 189) with regard to Hebrew, a somewhat lower register than the latter.

The lexical choices of secular publicists, however, draw on modern Hebrew and borrowings, and they often use colloquial Hebrew too. Similarly, Cohen showed that ultra-orthodox writers use more collocations/idioms from early layers of Hebrew than secular writers. Moreover, the latter are more innovative in that they often produce intended playful deviations from the original (archaic) collocations/idioms. These deviations, in fact, can be deemed optimal innovations.

Focusing on informal speech, Karni (2004) provided further support for these findings. Karni ran an experiment with a psycholinguistic flavor. In her study, high school adolescents distinguishable by their degree of religious observance ― secular, national religious (which are moderately religious) and ultra-orthodox ― were asked to write down the meanings of polysemous lexical items they were presented with, according to the order in which the meanings sprang into their minds (easily activated, in psycholinguistic terms, and therefore salient; see Section 6.2). Each item had an old, traditional, religiously related denotation alongside a modern secular one. Karni's results attest to three different sociolects influenced by levels of religious observance. The secular adolescents were only aware of the secular denotation of lexical items. All religious adolescents (whether ultra-orthodox or national religious) were aware of the two meanings of the (polysemous) lexical items. But in the case of the ultra-orthodox adolescents, the first meaning that sprang into their minds was the old, traditional one, whereas in the case of the national religious adolescents, results were mixed. In sum, these results indicate that the lexical choices of religious Jews draw on Jewish classical texts to a significantly greater extent than their secular counterparts. They attest to the relative salience of meanings (of the relevant lexical items) for each group of speakers on the community level (for a similar observation, see also Baumel, 2006: 87).

Similar evidence with respect to *non*-planned informal speech, was provided by Baumel (2006) who documented the spontaneous speech of Hebrew-speaking ultra-orthodox in the domestic sphere, the education system and the public domain. One of his research interests was the issue of *omission*, that is, the absence of certain words, speech patterns and dialectal expressions from the ultra-orthodox vocabulary. He noted the relative command of ultra-orthodox speakers of colloquial Hebrew, including slang, which they use on occasion. Even the leaders of the community use colloquial Hebrew and slang on occasion, in order to play to their audience, often born-again Jews (p. 34, 149). So, it looks as if there's a difference between the *de jure* language policy and *de facto* praxis. But Baumel went on to say that ultra-orthodox also know what "proper" language is, and they would not cross "the invisible linguistic line" (p. 78) that would remove them from the ultra-orthodox community. For example, he mentioned ultra-orthodox parents who were exposed to slang at work, but would not use it at home, so that the children would be brought up hearing only "proper language" (p. 95, 98, 153). Another example he cited is of a young man who had made an effort to refrain from using—what he called—"street language", including *xaval al hazman* (p. 98).

Baumel's finding are similar to those of Oryan (1997) who suggested that Hebrew-speaking ultra-orthodox women reject vulgar language, as well as linguistic innovations, such as slang, for "it is street talk. Therefore, anyone who uses such

language indicates that she has a defective character" (p. 12; translation mine). Oryan also noted that avoiding a certain linguistic register, here slang (as well as vulgar language) which is considered inferior, is part of the "counter culture" of the ultra-orthodox community (see Section 6.7.1 above), although many slangy items are neither vulgar nor pejorative. Fader (2009: 162-164) suggested the same for bilingual ultra-orthodox Jews in the US who speak both Yiddish and English, regardless of age.

It appears, then, that the Hebrew vernacular of ultra-orthodox Jews reflects their unique metapragmatic considerations with respect to colloquial Hebrew and slang. Example (6.6), originally in Hebrew, written by an ultra-orthodox (as the nickname *Nefeš Yehudi* 'a Jewish soul' implies), states this attitude explicitly.

(6.6) This street talk **"*xaval al hazman*"** [here, 'amazing/amazingly/extremely' – IB] struck roots even among us. At times, one hears it from "our" people every other sentence. It clings to you; it's a pain in the neck and it becomes a linguistic routine. How can one get rid of it?

(https://tinyurl.com/a342b9yn)

Overall, all the literature shows that the ultra-orthodox are aware of colloquial Hebrew and slang, and they use it. But they don't use it as often as their secular counterparts, in order to maintain their separate identity.

Taking a theoretical perspective, the low adoption frequency of lexically innovative items by the ultra-orthodox, or even the reluctance to adopt them altogether, can be accounted for by Rogers' (2003 [1962]: Ch. 6) theory of diffusion of (any) innovation. Roger proposed a five-factor model to account for the diffusion of an innovation. Noteworthy in the present context are two attributes subsumed under the first factor — the perceived attributes of innovation — which can explain why the ultra-orthodox block the spread of linguistic innovations:

(i) "*Relative advantage* is the degree to which an innovation is perceived as being better than the idea it supersedes." (p. 311)[90]

(ii) "*Compatibility* is the degree to which an innovation is perceived as consistent with the existing values, past experiences, and needs of potential adopters. […] An innovation can be compatible or incompatible (1) with sociocultural values and beliefs, […]." (p. 324)

In light of the declared policy of cultural separatism, adopting lexical innovations originating in the secular surrounding would be *incompatible* with the sociocultural values and beliefs of the ultra-orthodox community. The consequent *relative advantage* would be negative, possibly some kind of a social distancing (in a socially tight-knit community). The ultra-orthodox community is then a community of linguistic *laggards*

[90] In the 1983 edition of Rogers' book, he argued that "[r]elative advantage, in one sense, indicates the strength of the reward or punishment resulting from adoption of an innovation" (p. 217).

(as opposed to the secular *adopters*), those "who are the last in a social system to adopt an innovation" (Rogers, 2003 [1962]: 377).

6.7.3 The lexical choices of ultra-orthodox Jews versus those of the general population as an indication of semantic change and its direction

In light of the attitude of the ultra-orthodox towards colloquial Hebrew, which affects their lexical choices, I predicted that the distribution of the old versus the new meanings of *xaval al hazman* among them would be different from the distribution in the general population. If the ultra-orthodox population shows a different distribution, then semantic change has occurred (at least in the secular community). Specifically, if the ratio between the meaning suspected to be the new one and the meaning suspected to be the old one among the ultra-orthodox is smaller than it is in the general population, then the meaning suspected as the new meaning is indeed the new one, and, by implication, the other meaning is the old one.

I had an expert from the ultra-orthodox community tag the web-sites from which *xaval al hazman* was extracted, for the degree of religious observance of their writers (2373 out of the 2954 cases), as described in Appendix G. The distribution of *xaval al hazman* as a function of the religious observance of speakers is presented in Table 6.1. Only two factions of Hebrew speakers are considered, diametrically opposed from the aspect of religious observance, because all I wish to show is a contrast.

	Positive and intensifying *xaval al hazman*	**Negative *xaval al hazman***
General secular Jewish population	1066	1030
Ultra-orthodox Jews	59	151

Table 6.1: The distribution of *xaval al hazman* extracted from HeTenTen corpus by the religious observance of the speakers. (The figures include no instance of *xaval al hazman* accompanied by metalinguistic comments.)

The results of a two-tailed Fisher exact test are significant, $p = 1.90\times10^{-10}$, *Odds Ratio* = 2.65 (95% CI [1.92, 3.69], which is equivalent to a medium effect size; Chen, Cohen, & Chen, 2010) . These results indicate a difference in usage of *xaval al hazman* between the two populations and consequently point to semantic change. The lower ratio between the positive (and intensifying) *xaval al hazman* and the negative one among the ultra-orthodox ($^{59}/_{151} = 0.39$) as compared to the general population ($^{1066}/_{1030} = 1.03$) indicates that the ultra-orthodox Jews are less likely to use the positive (and intensifying) than the negative *xaval al hazman*. Given the lexical conservatism of the ultra-orthodox community, the former is then the new meaning of this syntagma, whereas the latter is the old one. These results are in line with the results of the metalinguistic comment-based method above (Section 6.5).

6.7.4 In support of the conservative speaker-based method

Admittedly, taking the ultra-orthodox speech community as a lexically monolithic group is misleading. After all, "a linguistic community is never homogeneous and hardly ever self-contained" (André Martinez in his preface to Weinreich, 1979 [1953]: vii). The individual idiolects of speakers (of the same speech community) are affected by the speakers' unique personality, gender and register, among other variables (e.g., Barlow, 2013; Guy, 2013; Zenner, Kristiansen, & Geeraerts, 2016, *inter alia*). These findings challenge claims about the collective linguistic behavior of any community. But this drawback is offset by the fact that the ultra-orthodox speakers examined here make up a specific faction of this conservative community, a rather permissive one who has access to the web, and it is therefore as well-defined as possible on the lexical aspect.

In addition, ultra-orthodox speakers active on web are expected to be more familiar with the lexical habits of the general secular population. If such permissive speakers still adhere to specific lexical choices, which are different from those of the general population, then the attested contrast (between them and the general population) is quite likely indicative of semantic change.[91]

6.8 Application of the new methods to other syntagmas that have undergone semantic change

One might argue that the methods I proposed above have been proven useful due the wealth of data available for *xaval al hazman*. This amount of data can be claimed to have resulted from the unusual attention *xaval al hazman* attracted from Hebrew speakers as a result of the stark meaning contrast between the new and the old meanings. In order to address this potential criticism, I will apply the proposed methods to three additional syntagmas: *En dvarim ka'ele/u*, originally, 'there are not such things', *sof haderex*, originally, 'the end of the road', and *ba livkot*, originally, 'it feels like crying' (two of which belong the Ultimate construction family — *en dvarim ka'ele/u* and *ba*

[91] The method here proposed to detect semantic change can be naturally tested on languages used by ultra-orthodox communities elsewhere — the English of the ultra-orthodox community in North-America and England, and the Flemish of the ultra-orthodox community in Belgium.

A similar community which is worth examining due to commitment to religious life and maintaining a separate identity is the conservative tight-knit Amish community. Indeed, the accepted view is that the members of this community speak a special German dialect among themselves, mostly the one known as Pennsylvania Dutch (but other dialects exist), while American English is the language of school, of printed communications and the language used in commerce and interactions with non-Amish (see, e.g., Hostetler, 1993: 241-244; Kraybill, 2001 [1989]: 55-57). But this is not quite the case, as argued by Thompson (2006: 275): "English has become not just the language of outsiders, but also one of the languages *internal* to the community" [emphasis mine]. Thompson enumerated several lexical differences between the English of Amish and non-Amish speakers. The Amish, then, seem to be another potential population to test the conservative speaker-based method, focusing on the kind of English that they speak within the community, rather than their special German dialect.

livkot). Each one of these items is somewhat different form *xaval al hazman* either in the amount of data available or in the semantic contrast between the old and the new meanings. In the following sections I will show that the methods proposed above are still useful for detecting semantic change.

6.8.1 *En dvarim ka'ele/u*, originally, 'there are no such things'

The syntagma *en dvarim ka'ele/u* means literally 'there are no such things', referring to an object which does not exist, either locally or globally, as exemplified in (6.7a) and (6.7b), respectively. The denial of the Rock concert in Example (6.7b) can be interpreted more loosely, as a hyperbole implying that it is 'highly rare', rather than nonexistent. This can explain the evolution of the expression into a positive-flavor idiomatic adjective denoting 'amazing' as in Example (6.8a), an adverb denoting 'so much' as in Example (6.8b), and an intensifier denoting 'extremely' as in Example (6.8c).

(6.7) a. *ha-munax "mekubalim" kayam rak b-a-yesodi.*
the-term popular.guys exists only in-the-elementary.school.

b-a-xativa u-v-a-tixon ***en***
in-the-middle.school and-in-the-high.school **there.are.no**

dvarim ka'ele.
things like.these

'The term "popular guys" exists only in elementary school. **There are no such things** in middle school and in high school.'

(HeTenTen, tinyurl.com/cfy3ud4u)

b. *mofa rok amiti* **—** ***en dvarim***
performance rock genuine **there.are.no things**

ka'ele *yoter. pašut en. hem lo kayamin.*
like.these anymore. simply there.are.no they not exist

halxu le'ibud b-a-misxari'ut še-pašta b-a-muzika.
went lost in-the-commerciality that-spread in-the-music

'A genuine Rock concert **— there are no such things** anymore. There just aren't. They no longer exist. They got lost due to the spreading commercialization in the field of music.'

(HeTenTen, tinyurl.com/umszhfs7)

(6.8) a. *ha-yexolet* *šela* *laxdor* *l-a-lev* *hayta* *mašehu*
the-ability her to.penetrate to-the-heart was something

meyuxad, *bešeket* *bešeket,* *im* *xiyux*
special quietly quietly with a.smile

*še-**en dvarim ka'ele***.
that-**there are no such things → amazing**

'Her ability to get into your heart was something special, very quietly, with an **amazing** smile.'

(HeTenTen, tinyurl.com/mr224r6u)

b. *nehenenu* *mi-kol* *rega* *b-a-makom* *ha-kasum*
we.enjoyed from-every moment in-the-place the-magical

ha-ze *še-**en dvarim ka'ele***.
the-this that-**there are no such things → so much**

'We enjoyed every single moment in this magical place **so much**.'

(HeTenTen, tinyurl.com/3ddmtwxs)

c. *aval* *ze* *haya* *ta'im*
but this was delicious

*še-**en dvarim ka'ele***.
that-**there are no such things → extremely**

'But it was **extremely** delicious.'

(HeTenTen, tinyurl.com/2pv8vnwj)

Data from Yedioth Ahronoth corpus, presented in Figure 4.5 and repeated here, for convenience, as Figure 6.2, have already shown that *en dvarim ka'ele/u* has undergone semantic change during the second half of the first decade of the twenty-first century.

Note that the meaning contrast (at least) between the global reading of *en dvarim ka'ele/u* (6.7b) and the new meaning (6.8a-c) is rather debatable. In fact, one can substitute one for the other without any apparent clash with the relevant contexts. This was not the case with *xaval al hazman* analyzed above.

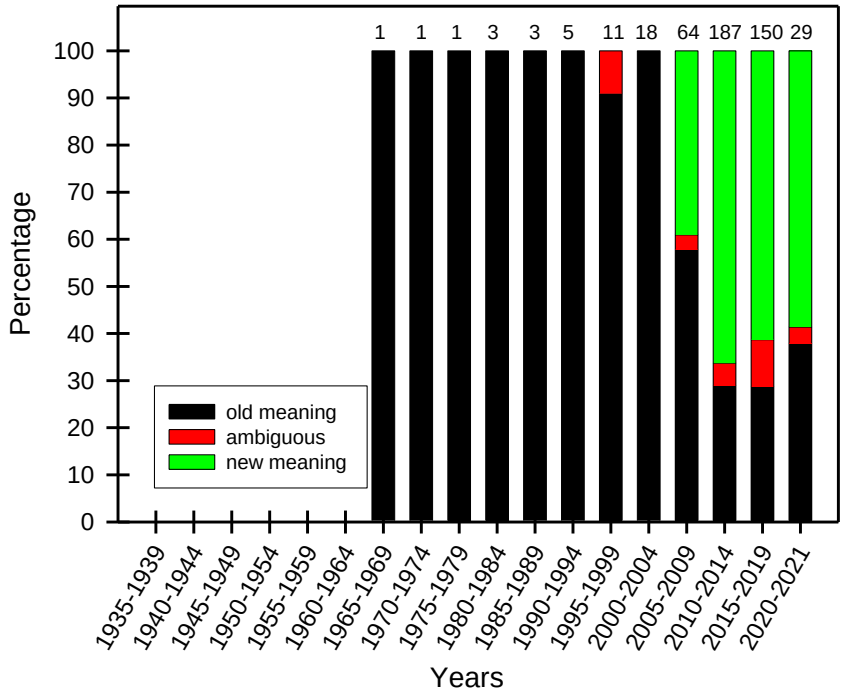

Figure 6.2: The distribution of the old versus the newly evolved meaning(s) of *en dvarim ka'ele/u* 'there are no such things' as a function of time. The number of counts for every five-year interval is marked on top of each bar.[92] Note that items of different grammatical statuses were considered *en bloc*. Items classified as either accompanied by metalinguistic comments or simply names (of books, newspaper columns, and the like) were filtered out. Data extracted from Yedioth Ahronoth corpus.

In the synchronic HeTenTen corpus 1359 instances of *en dvarim ka'ele/u* were detected. That's about a half (~46%) of the *xaval al hazman* tokens.

6.8.1.1 Comment-based method

A close examination of the contents of the 53 instances of *en dvarim ka'ele/u*, accompanied by metalinguistic comments, supports my analysis of the semantic change undergone by this syntagma. The following Examples (6.9a-d), are representative examples of comments which accompany the *use* (rather than mention) of *en dvarim ka'ele/u* (see Section 6.5.2). In each of them, the speaker indicates explicitly the meaning of *en dvarim ka'ele/u* she wishes to communicate (boldfaced), thus revealing that semantic change has indeed occurred. Example (6.9a) indicates that *dvarim ka'ele/u* is an idiom which has a prior ("original") meaning. One of the meanings is considered slang, and therefore necessarily idiomatic (6.9b). Taken together, Examples (6.9a) and (6.9b) testify to a current idiomatic meaning. The intended meaning in

[92] I assume that the change in total counts is possibly a result of change in the overall number of tokens. Since the overall number of tokens is not available, this assumption is based on the changed number of articles printed in Yedioth Ahronoth over the years.

Example (6.9c) is specifically the literal one which implies that it is the "original" meaning mentioned in Example (6.9a). The idiomatic, slangy meaning must be the current meaning, for it is favored by youngsters (6.9d). Note that (6.9d) is a mixed example which highlights also the literal meaning ("truly").

(6.9) a. *en dvarim ka'ele,* ***b-a-mašma'ut*** ***ha-mekorit***
there are no such things **in-the-meaning** **the-original**

šel ***ha-bituy***.
of **the-idiom**

'There are no such things**, in the original meaning of the idiom**.'
(HeTenTen, tinyurl.com/4n5dku9s)

b. ***ex*** ***omrim*** ***be-sleng?*** *en dvarim ka'ele!*
how **they.say** **in-slang** there are no such things

'**How do they say in slang?** There are no such things!'
(HeTenTen)

c. *en dvarim ka'ele,* ***pšuto*** ***ke-mašma'o***.
there are no such things **its.face.value** **as-its.meaning**

'There are no such things, **literally**.'
(HeTenTen, tinyurl.com/2s42v2u8)

d. *o* ***kmo*** ***še-omer*** ***ha-no'ar*** ***eclenu:*** *en dvarim ka'ele.*
or **like** **that-says** **the-youth** **at.us** there are no such things

be'emet *še-en.*
for.real that-there.are.no

'Or **as our youth says:** There are no such things. **Truly**, there are no such things.'
(HeTenTen)

There are 11 such examples, 6 of which invoke the old, literal meaning of *en dvarim ka'ele/u* and 5 invoke the new, idiomatic meanings. These data point to equal salience for the two meanings.

Taken together, Examples (6.9a-d) testify to new idiomatic meaning(s) of *en dvarim ka'ele/u*, alongside an old literal one, in line with the results of the diachronic corpus presented in Figure 6.2 above. The old and the new meanings seem to be equally salient.

Importantly, in the case of *en dvarim ka'ele/u* (unlike the case of *xaval al hazman*) there's hardly any meaning contrast between the old and the new meanings (i.e., no negative versus positive polarity, respectively). And still, there are quite some metalinguistic comments that attest to semantic change. *En dvarim ka'ele/u* thus refutes a possible claim that only a sharp meaning contrast would trigger speakers' use of (metalinguistic) comments (in order to clarify which meaning they wish to communicate).

6.8.1.2 Wordplay-based method

This method is not applicable here because only one meaning is a collocation, the new idiomatic one. It is not surprising, then, that the few instances of optimal innovation found in HeTenTen corpus are based on the new idiomatic meaning only. Nevertheless, they cannot testify to the relative salience of this new meaning over the other, old one.

6.8.1.3 Conservative speaker-based method

Religious people (in general, not just ultra-orthodox) are sensitive to the (secular) origin of *en dvarim ka'ele/u*. Example (6.10), originally in Hebrew, is an excerpt from a user account of a young woman on a dating web-site. She explains why she would never date a secular man, thus implying that she is religious. And in doing so, she reveals that she knows that *en dvarim ka'ele/u* is used by secular Jews.

(6.10) I find it funny and somewhat odd to cite, without even paying attention to it, something from the Mishna, that (where I come from) is equivalent to **the secular *en dvarim ka'ele*** [here, 'amazing/amazingly/extremely' – IB], and see a confused look of total lack of understanding on the face of my date.

(HeTenTen)

The results of the classification of the 1269 classifiable instances of *en dvarim ka'ele/u* (extracted from the synchronic HeTenTen corpus) according to the religious observance of the speakers who produced them are summarized in Table 6.2. As in the case of *xaval al hazman*, only data of secular and ultra-orthodox Jews are considered.

	Positive and Intensifying *en dvarim ka'ele/u*	**Literal *en dvarim ka'ele/u***
General secular Jewish population	962	283
Ultra-orthodox Jews	13	11

Table 6.2: The distribution of *en dvarim ka'ele/u* extracted from HeTenTen corpus by the religious observance of the speakers. (The figures include no instance of *en dvarim ka'ele/u* accompanied by metalinguistic comments.)

The results of a two-tailed Fisher exact test are significant, $p = 0.012$, *Odds Ratio* = 2.92 (95% CI [1.17, 7.15], which is equivalent to a medium effect size; Chen, Cohen, & Chen, 2010). The lower ratio between the positive (and intensifying) *en dvarim ka'ele/u* and the literal one among the ultra-orthodox ($^{13}/_{11} = 1.18$) as compared to

the general population ($^{962}/_{283} = 3.40$) indicates that the ultra-orthodox are less likely to use the positive (and intensifying) *en dvarim ka'ele/u* over the literal one, as compared to the general secular population. The positive (and intensifying) meaning is therefore the new meaning of this syntagma whereas the literal one is the old meaning. These results are in line with those of the metalinguistic comment-based method above.

In sum, results of two methods attest to semantic change from the literal *en dvarim ka'ele/u* to the positive (and intensifying) one. This change is at a stage where both meanings seem to be equally salient. The third method it not applicable.

6.8.2 *Sof haderex*, originally, 'the end of the road'

Sof haderex is a construct phrase which means literally 'the end of the road' (*sof* 'end'; *derex* 'road'), an actual physical road, as exemplified in (6.11a), alongside a metaphorical road, as exemplified in (6.11b). And quite like *xaval al hazman* and *en dvarim ka'ele/u*, *sof haderex* too has turned into a positive-flavor idiomatic adjective denoting 'amazing' (6.12a), an adverb denoting 'amazingly' (6.12b), as well as an intensifier denoting 'extremely' (6.12c).

(6.11) a.

yeš	*le'hagi'a*	*le-taxana*	*česington*	*darom*	*lifnot*
one.should	to.get	to-station	Chessington	South	to.turn

yamina	*ve-lalexet*	*ad*	***sof***	***ha-derex.***
to.the.right	and-to.go	until	**the.end.of**	**the-road**

'One should get to Chessington South station, turn right and go up until the **end of the road**.'

(tinyurl.com/3drs3y2a)

b.

ha-seret	*"predot"*	*osek*	*be-"**sof***	***ha-derex**",*
the-movie	departures	is.about	in-**the.end.of**	**the-road**

be-tipul	*b-a-metim.*
in-taking.care.of	in-the-dead

'The movie "Departures" is about **the end of the road**, about taking care of the dead.'

(http://www.psy.org.il/)

(6.12) a. *ha'iti* *b-a-seret* *be-yom* *revi'i.* *ha-rikudim* *šam*
I.was in-the-movie in-day Wednesday the-dances there

madhimim *ve-ha-širim* ***sof haderex***.
astonishing and-the-songs **the end of the road → amazing**

'I watched this movie on Wednesday. The dancing there is astonishing and the songs are **amazing**.'

(tinyurl.com/czehaa8)

b. *seret* *meratek.* *mevuyam* ***sof haderex***.
movie fascinating directed **the end of the road → amazingly**

'A fascinating movie. **Amazingly** directed.'

(tinyurl.com/zvkcdzj3)

c. *yofi* *be-eney* *ha-mitbonen* **—** *ma* *še-nir'e*
beauty in.the.eyes.of the-beholder what that-seems

lexa *yafe* *u-meyuxad* *le-axat* *ze* *yafe*
to.you beautiful and-unique to-one it.is beautiful

sof haderex, *le-axeret* *mexo'ar* *le-haxrid.*
the end of the road → extremely to-another ugly terribly

'Beauty is in the eyes of the beholder — what seems to you beautiful and unique, for one person it is **extremely** beautiful, for another it is terribly ugly.'

(tinyurl.com/2sue5hc)

Examples (6.11a) and (6.11b) are emotively neutral, so the connotative contrast between them and (6.12a-c) is, no doubt, less stark than the contrast between the old (negative) and the new (positive and intensifying) meaning of *xaval al hazman* (presented in Section 6.5-6.7), and similar to the case of *en dvarim ka'ele/u* (presented in Section 6.8.1 above).

The data from Yedioth Ahronoth corpus indicate that the semantic change of *sof haderex* has occurred during the late nineties of the twentieth century (see Figure 6.3 below), somewhat later than the change undergone by *xaval al hazman*. In fact, I found a call to linguists posted in Yedioth Ahronoth corpus, from June 1999: "Linguists! For your information: *Sof haderex* is the official replacement of *xaval al hazman*" [translation mine]. This observation is corroborated by Triger (2007) who noted that when he had left Israel for New-York in 1999, the new *sof haderex* (6.12a-c) did not

yet exist. However, in one of his visits to Israel (between 1999-2002), he realized that the new *sof haderex* couldn't be avoided.

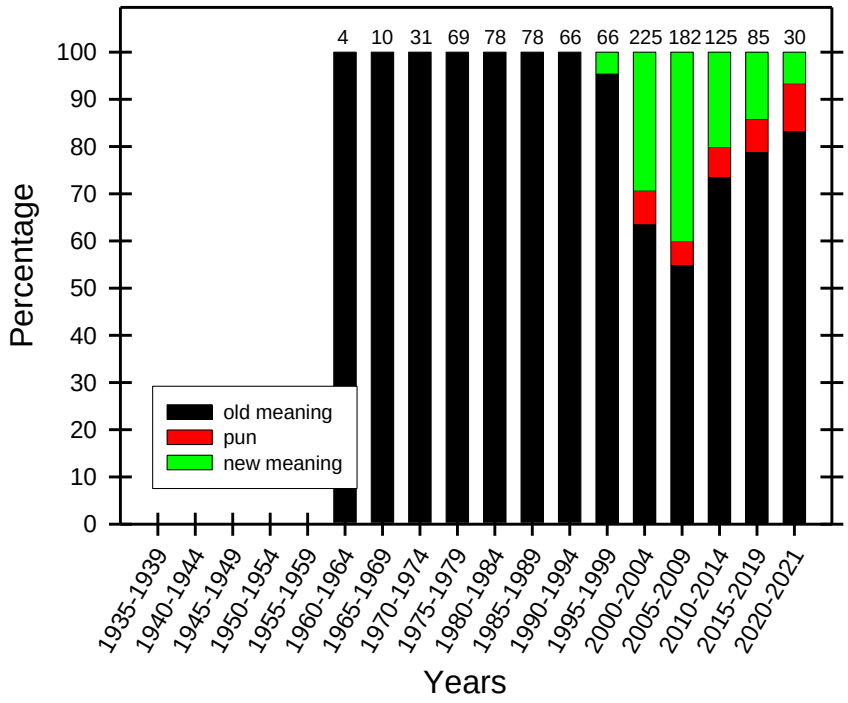

Figure 6.3: The distribution of the old versus the new meanings of *sof haderex* as a function of time. The number of counts for every five-year interval is marked on top of each bar.[93] Note that items of different grammatical statuses were considered *en bloc*. Items classified as either accompanied by metalinguistic comments or simply names (of books, newspaper columns, and the like) were filtered out. Data extracted from Yedioth Ahronoth corpus.

In the synchronic HeTenTen corpus, 1306 instances of *sof haderex* were found, a little less than half (~45%) of *xaval al hazman* tokens. I next apply each of the three proposed methods to *sof haderex*.

6.8.2.1 Comment-based method

Examples such as (6.13) indicate that *sof haderex* has changed its meaning, while roughly noting the approximate date of change (compatible with Triger's, 2007 observation above).

[93] I assume that the change in total counts is possibly a result of change in the overall number of tokens. Since the overall number of tokens is not available, this guess is based on the changed number of articles printed in Yedioth Ahronoth over the years.

(6.13) *avira* *šel* *mesiba* *sof haderex,* ***kmo*** ***še-hayu*** ***omrim***
an.air of a.party the end of the road **like** **that-were** **saying**

b-a-naintiz, *rak* *še-hapa'am* *ani* *be'emet* *mitkaven*
in-the-ninetees only that-this.time I truly mean

la-ze.
to-this

'An air of a *sof haderex* party, **as they used to say in the nineties**, only this time I really mean it.'

(HeTenTen, tinyurl.com/2nrddb7e)

A close inspection of the contents of the 50 instances of *sof haderex,* accompanied by some comment about its meaning, enables a more detailed description of this semantic change. The following (6.14a-d) are the examples found in HeTenTen corpus where *sof haderex* is *used* naturally rather than mentioned (which is also the case with Example (6.13); and see Section 6.5.2). In each of them, the speaker indicates explicitly the meaning of *sof haderex* she wishes to communicate (boldfaced), thus revealing that semantic change has indeed occurred. Example (6.14a) attests to a literal meaning. Example (6.14b) indicates that there's a meaning which is not used by youngsters, the literal meaning, and it is therefore an old meaning. Note, that in the context of Example (6.14b), this old meaning is not quite literal, contrary to what the speaker indicates. It is, rather, metaphorical. In certain contexts, this metaphorical meaning could be negative, as evidenced by Example (6.14c). The new meaning must therefore be the positive one, as indicated by Example (6.14d). Taken together, Examples (6.14a-d) highlight the non-salient meaning which must be explicitly invited. This meaning is the literal and/or the metaphorical one. The new meaning is, by implication, the positive (and intensifying) meaning. In sum, Examples (6.14a-d) attest to semantic change, all in line with the data from Yedioth Ahronoth corpus presented in Figure 6.3 above.

(6.14) a. *ha-kfar* *turtuk* *hu* ***pšuto*** ***ke-mašma'o***
the-village Turtuk is **its.face.value** **as-its.meaning**

sof haderex.
the end of the road

'The village Turtuk is **literally** the end of the road.'

(HeTenTen, tinyurl.com/5n6uxxww)

b. *dimuy* *ha-70* *nir'e* *mamaš* *sof haderex.*
the.image^ the-70 seems really the end of the road

ve-lo *sof haderex* ***b-a-muvan***
and-not the end of the road **in-the-sense**

še-ha-ce'irim ***mištamšim*** ***b-o*** ***ela*** ***mamaš***
that-the-young.people **use** **in-it** **but** **really**

— *sof haderex* — *lefaxot* *karov* *le-sof haderex.*
the end of the road at.least close to-the end of the road

'The image of [age] 70 seems like the end of the road, **and not** the end of the road **in the sense used by young people, but really** the end of the road or at least close to the end of the road.'

(HeTenTen)

c. *ha-xa'im* *nir'im* *"sof haderex",* ***aval*** ***b-a-muvan***
the-life seems the end of the road **but** **in-the-sense**

ha-šlili ***šel*** ***ha-musag***.
the-negative **of** **the-term**

'Life seems "the end of the road", **but in the negative sense of the term**.'

(HeTenTen, tinyurl.com/2h8mws3s)

d. *"sof haderex",* *ve-lo* ***b-a-muvan*** ***ha-xiyuvi***
the end of the road and-not **in-the-sense** **the-positive**

šel ***ha-munax***.
of **the-term**

' "The end of the road", **and not in the positive sense of the term**.'

(HeTenTen)

Note that unlike in the case of *xaval al hazman*, where the comments invoke both meanings (although the old one somewhat more so), in the case of *sof haderex*, all 5 examples (6.13 and 6.14a-d) invoke only the old meanings, indicating that *sof haderex* is presumably at a more advanced stage of semantic change than *xaval al hazman*.

6.8.2.2 Wordplay-based method

Similar to *xaval al hazman*, some of the meanings of *sof haderex* too are collocations and therefore constitute substrates for optimal innovations. One of the old meanings is metaphorical, and all the new meanings are hyperboles. I found 3 instances of optimal innovation based on *sof haderex* in HeTenTen corpus: Two instances of *xof haderex* 'beach of the road' in the context of promoting the Israeli version of Burning Man festival (HeTenTen), and as the title of a blog post celebrating the twenty-fifth anniversary of the hit song "I'm walking on sunshine" (HeTenTen, tinyurl.com/4xs48p3v); and *šef haderex* 'chef of the road' as a brand name of pre-cooked meals for field trips (HeTenTen, tinyurl.com/234y42mu). *Xof haderex* observes the phonological criterion used to spot an optimal innovation (see Appendix G), and *šef haderex* constitutes a somewhat weaker observation of this criterion. Both observe the semantic criterion. All three are based on the new meaning of *sof haderex* (6.12a-c). No example of optimal innovation based on the earlier, metaphorical meaning of *sof haderex* (6.11b) was detected. This distribution attests to the salience of the new meaning of *sof haderex*. A few more examples were found on a targeted Google search (as of 11 December 2021). All, but one, are based on the new meaning of *sof haderex*, again attesting to the new meaning being the current salient meaning, in line with the results of the metalinguistic comment-based method.

6.8.2.3 Conservative speaker-based method

The results of the classification of the 1075 classifiable instances of *sof haderex* by the religious observance of the speakers who produced them are summarized in Table 6.3. As in the case of *xaval al hazman*, only data of secular and ultra-orthodox Jews are considered.

	Positive and Intensifying *sof haderex*	**Literal & Metaphoric *sof haderex***
General secular Jewish population	441	568
Ultra-orthodox Jews	9	26

Table 6.3: The distribution of *sof haderex* extracted from HeTenTen corpus by the religious observance of the speakers. (The figures include no instance of *sof haderex* accompanied by metalinguistic comments.)

The results of a two-tailed Fisher exact test are significant, $p = 0.037$, *Odds Ratio* = 2.24 (95% CI [1.0052, 5.50], which is equivalent to a medium effect size; Chen et al. 2010). The lower ratio between the positive (and intensifying) *sof haderex* and the literal and metaphoric ones among the ultra-orthodox ($^{9}/_{26} = 0.35$) as compared to the general population ($^{441}/_{568} = 0.78$), indicates that the ultra-orthodox are less likely to prefer the positive (and intensifying) *sof haderex* over the literal and metaphorical ones. The positive (and intensifying) meaning is therefore the new meaning of this syntagma, whereas the literal and metaphorical meanings are the old ones. These results are in the line with the results of the metalinguistic comment-based

method above. A closer look at the confidence Interval of the results indicates that the lower bound is quite close to 1.0 (specifically, 1.0052). This implies that the distribution of the old and the new meanings of *sof haderex* are not quite different for the general secular Jewish population and the ultra-orthodox Jews. Hence a change of meanings among the ultra-orthodox is, quite possibly, underway.

In sum, the results of the three methods here proposed attest to the semantic change of *sof haderex*, and its direction — from the literal and metaphorical (perhaps, but not necessarily, negative) meanings to the positive and intensifying one. They also attest to the current salience of the positive (and intensifying) meaning over the two other old meanings (for the general population).

6.8.3 *Ba livkot,* originally, 'it feels like crying'

In the previous two sections I have shown that it is not necessarily the stark meaning contrast between the old and the new meanings (as in the case of *xaval al hazman*), that attracts speakers' attention and therefore yields a wealth of data. Even syntagmas where the contrast is debatable, as in the cases of *en dvarim ka'ele/u* and *sof haderex*, evoke speakers' metalinguistic interest, and hence yield a fair amount of metalinguistic data. I further suggest that syntagmas that are much less frequent, such as *ba livkot* 'it feels like crying' — 206 instances in HeTenTen corpus altogether — may also benefit from the proposed methods.

Ba livkot bears two simultaneous meanings exemplified in (6.15) and (6.16). Example (6.15) displays the literal meaning. In Example (6.16) *ba livkot* can be interpreted as either 'exceedingly' (intensifier) or 'amazing' (adjective), depending on the lexeme it modifies, 'full' (adjective) or 'juice' (noun), respectively.

(6.15)

zo	*avira*	*mamaš*	*apokaliptit.*	*kol*	*ha-nof*	*nir'e*
this	atmosphere	really	apocalyptic	all	the-landscape	looks

afor-kaxol,	*pašut*	***ba***	***livkot.***
grey-blue	simply	**feels.like**	**to.cry**

'There's an apocalyptic atmosphere. the entire landscape is grey-blue, **it** simply **feels like crying.**'

(HeTenTen, tinyurl.com/4dpbpvjr)

(6.16)

ve-ha-klemantinot	*mele'ot*	*asis*
and-the-tangerines	full.with	juice

še-ba livkot
that-it feels like crying → exceedingly/amazing

'[…] and the tangerines are **(exceedingly)** full with **(amazing)** juice.'

(HeTenTen)

The data from Yedioth Ahronoth corpus indicate that the new, positive and intensifying *ba livkot* appeared on the language scene at the early seventies of the twentieth century (see Figure 6.4 below).

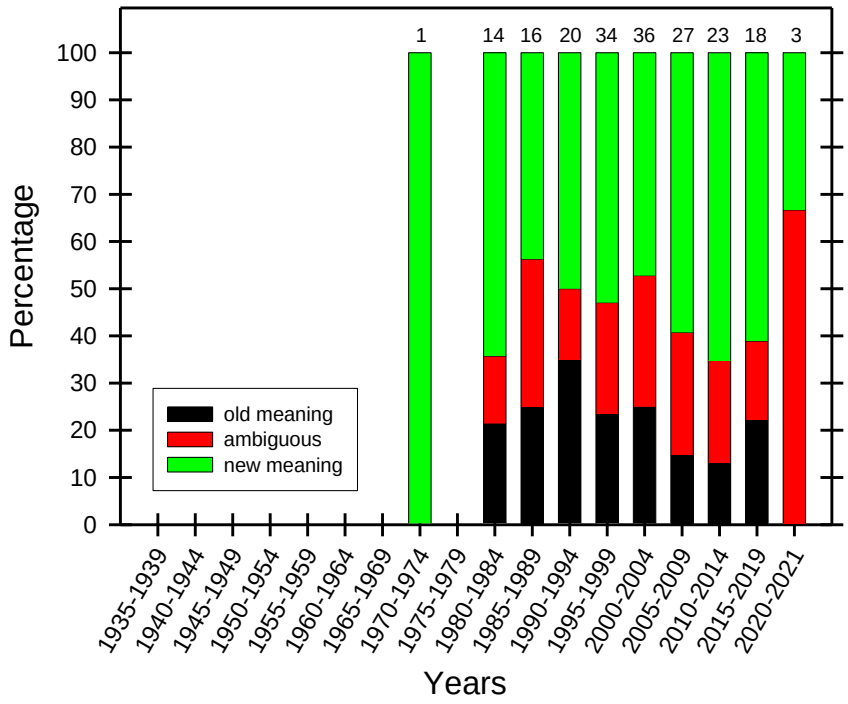

Figure 6.4: The distribution of the old versus the new meanings of *ba livkot* as a function of time. The number of counts for every five-year interval is marked on top of each bar.[94] Note that items of different grammatical statuses were considered *en bloc*. Items classified as either accompanied by metalinguistic comments or simply names (of books, newspaper columns, and the like) were filtered out. Data extracted from Yedioth Ahronoth corpus.

Having established that semantic change has indeed occurred, I now examine the 206 instances of *ba livkot* in the synchronic HeTenTen corpus.

6.8.3.1 Comment-based method

I detected two instances of *ba livkot* accompanied by metalinguistic comments. Both are of the type where *ba livkot* is *used* naturally rather than mentioned (see subset (b) in Section 6.5.2 above). One of them is Example (6.17), where the comment "and I really cried", invokes the low-salience meaning, here the literal meaning of actual crying. The other example (not cited here) is similar. Both instances imply that the low-salience meaning of *ba livkot* is the literal one, and by implication that the salient

[94] I assume that the change in total counts is possibly a result of change in the overall number of tokens. Since the overall number of tokens is not available, this guess is based on the changed number of articles printed in Yedioth Ahronoth over the years.

meaning of *ba livkot* is the idiomatic intensifying one denoting 'extremely' or 'amazing'. But note the paucity of data.

(6.17) *ha-sefer* *madhim,* *katuv* *kol-kax* *yafe*
the-book amazing written so beautifully

še-ba livkot *(ve-ani* ***be'emet*** *baxiti).*
that-it feels like crying → extremely and-I **truly** cried

'The book is amazing, so beautifully written that it feels like crying (and I **truly** cried).'

(HeTenTen)

Another interesting example, not quite a comment, but perhaps a creative instance, requires some background. *Ba livkot* is mentioned for the first time in a popular song from the mid-sixties of the twentieth century. In that song, a womanizer confesses that he is very much emotionally moved by the presence of attractive women, so much that he feels like bursting into tears: *yafot še-ba livkot* 'beautiful up to a point that it feels like crying'.[95] In fact, the entire syntagma, not just *ba livkot*, has become idiomatic for speakers. Now, in Example (6.18), a review about highly expensive eye shadows, the speaker uses the adjective *yafot* which immediately invokes the missing continuation of the intensifier *še-ba livkot*. But instead of using the original version of the intensifier, the speaker uses it creatively, to agonize over high prices. This creative use invokes the low-salience, literal meaning. It then shows the intensifying meaning as the one that needs to be actively rejected, for it is already the salient automatic meaning (at the time of speech).

(6.18) *lo* *yaxolti* *limco* *l-a-post* *ha-ze* *koteret* *yoter* *meduyeket*
not I.could to.find to-the-post the-this title more accurate

avuran — *yafot,* *yafot!* *aval* *im* *hen* *kol-kax* *yafot,*
for.them beautiful beautiful but if they so beautiful

az *lama* ***bexol-zot*** ***ba*** ***livkot?***
so why **in.spite.of.it** **feels.like** **to.cry**

'I couldn't have found a more accurate title for this post – beautiful, beautiful! But if they [eye shadows – IB] are so beautiful, **why is it that one still feels like crying**?'

(tinyurl.com/2p8m7ncz)

[95] The lyrics of *yafot, yafot*: http://tinyurl.com/3898mhpf

Note that in the specific case of *ba livkot*, the limited selection of metalinguistic comments allows to establish semantic change and salience, but not the direction of change.

6.8.3.2 Wordplay-based method

As in the case of *en dvarim ka'ele/u* (see Section 6.8.1.2), here too, this method is not applicable. Only the new meaning of *ba livkot* is collocational/idiomatic and can constitute the substrate for wordplay, not the old one.

6.8.3.3 Conservative speaker-based method

The results of the classification of the 185 classifiable instances of *ba livkot* by the religious observance of the speakers who produced them are summarized in Table 6.4. Again, only data of secular and ultra-orthodox Jews are considered.

	Positive and Intensifying *ba livkot*	**Negative** *ba livkot*
General secular Jewish population	137	38
Ultra-orthodox Jews	1	6

Table 6.4: The distribution of *ba livkot* extracted from HeTenTen corpus by the religious observance of the speakers. (The figures include no instance of *ba livkot* accompanied by metalinguistic comments.)

The results of a two-tailed Fisher exact test are significant, $p = 8.7\times10^{-4}$, *Odds Ratio*= 21.18 (95% CI [2.4587, 996.2954], which is equivalent to a large effect size; Chen et al. 2010). The lower ratio between the intensifying *ba livkot* and the literal one among the ultra-orthodox ($^{1}/_{6} = 0.167$) as compared to the general population ($^{137}/_{38} = 3.61$) attests to semantic change — the negative *ba livkot* gave way to the positive (and intensifying) one.

Noteworthy is the fact that the channel by which *ba livkot* was introduced into the language scene (as noted above) is a popular song, which may have contributed to the spread of *ba livkot* among speakers. However, the specific provocative content of this song makes it entirely unacceptable to ultra-orthodox Jews, which may explain why they have barely adopted the new, intensifying *ba livkot* (thus drawing a linguistic line between them and the general Jewish population in Israel).

Taken together, the results of the relevant methods attest to semantic change from the compositional, negative *ba livkot* to the idiomatic, positive and intensifying one. It could be that this change is at an advanced stage where the new meaning is the salient meaning, but keep in mind that data are scarce.

It seems, then, that applying the (relevant) new methods here proposed even to a small amount of data — ~200 items as opposed to 1000-3000 items in the three previous cases — can still provide some support for a claim about semantic change and the possible stage of change.

6.9 Limitations of the methods proposed in this chapter

The proposed methods I have introduced (and exemplified) above seem straightforward and effortless. One can then wonder how come they have not been used by historical linguists as a matter of course. Obviously, these methods are not without limitations.

The comment-based method is probably applicable only if speakers entertain both (all?) meanings of a given syntagma in mind. This is not always the case.

The wordplay-based method is applicable only if both (all?) meanings constitute a substrate for attested wordplay. This is not always the case, either because one of the two does not serve as a substrate for wordplay or just because wordplay requires substantial cognitive resources (as compared to utterances which are not wordplay; See longer reading times for optimal innovations in Giora et al. 2004: Exp. 4; Giora, Givoni, Heruti, & Fein 2017: Exp. 2).

The conservative speaker-based method depends on the existence of a lexically-conservative community among the native speakers of a language, which rejects neologisms as a matter of ideology. But such communities are not always available or accessible to the researchers.

6.10 A note about sensitivity to changes

This chapter focused on semantic change. But then the syntagmas here considered have also undergone change in grammatical status. However, not a shred of evidence (neither explicit nor implicit) for speakers' sensitivity to change in grammatical status has been spotted. This implies that speakers are much more aware of semantics, but much less so of syntax (except in the case of careful speech, which is obviously not the case here). The semantic level of representation, regardless of level of syntactic complexity, may imply that speakers regard all syntagmas, or constructions, whether they "belong" to the lexicon or to the syntax as part of the same inventory, with no division between lexicon and syntax.

6.11 Summary and conclusions

I this chapter I proposed three methods that allow linguists to argue for the reality of semantic change in the absence of diachronic data, or when the only available corpus is a synchronic corpus which is not tagged for speakers' age. Such a state of affairs is rather challenging. But the difficulties can be overcome, as I have argued above, by using alternative methods — the metalinguistic comment-based method, the wordplay-based method and the conservative speaker-based method. These methods take advantage of speakers' (conscious) metalinguistic activity, from the cognitive, as well as the sociopragmatic perspectives. They can attest to a change in the status of simultaneous meanings of a given syntagma, where a previous coded salient meaning gives way to a new coded meaning, within and across speech communities. Speakers who "provide" the data for the analyses must be linguistically sensitive, that is, aware of the linguistic system in general and of the semantic change in particular.

I applied these methods to four syntagmas in Hebrew for which semantic change has been established independently (by using a diachronic corpus). *Xaval al hazman*

provided a large data set for the analyses, probably due to the stark semantic contrast between the old and new meaning(s), which attracted speakers' attention. But *ba livkot*, also of a stark contrast, provided a rather small amount of data. *En dvarim ka'ele/*u and *sof haderex* provided a fair amount of data, despite the much smaller contrast between their various meanings (if at all).

Obviously, these methods (just like any other method) are not without limitations, but they can fill a gap when a diachronic corpus is not available. But even when a diachronic corpus is available, collecting psycholinguistic and sociolinguistic evidence of a metalinguistic nature attesting to sematic change is worthwhile, because it adds *first-hand* evidence for the change, based on a glimpse into the speakers' minds. Such evidence can also reduce the researchers' reliance on the interpretative interventions required in pinpointing speakers' communicative intentions. I therefore advocate the use of such data at least as self-monitoring for researchers.

Chapter 7: Concluding remarks

"Come, children, let us shut up the box and the puppets, for our play is played out."
-- William Makepeace Thackeray, *Vanity Fair*, Ch. 67, 1847-1848

In this dissertation I set out to provide a full explanation along the lines of Construction Grammar for a cross-linguistically rare phenomenon attested in Hebrew — full sentences lexicalizing into full-fledged words. The proposed model of the lexicalization process emphasized the role of a multidimensional Construct-i-con — defined by the various kinds of *links* obtaining between various constructions (e.g., Diessel, 2020, 2023; Schmid, 2020; Sommerer & Smirnova, 2020) — in accounting for this phenomenon.

I examined a set of sentences that seem to show similar distributional behavior. I showed that this set of sentences is linked to higher-level, more abstract constructions via (i) *vertical inheritance links*, thus making a taxonomic family – the Ultimate construction family. In this specific family — a family of sentences that contain no subject — there are especially strong links between the predicate and an adjacent NP or an infinitive. Each of these sentences is interpreted as one conceptual unit (similar to VPs) and therefore tends to undergo semantic change to become a semantically opaque idiomatic sentence. Being evaluative, the newly evolved idiomatic sentences are relational or semantically incomplete, and therefore in search of a modifiable element in prior discourse. But these idiomatic sentences are not just evaluative. The evaluation that they convey is a highly intense evaluation. They are therefore used to reinforce an emotively bleaching intensifier in the preceding sentence via (ii) *filler-slot links* between them and their respective context. This context is a context that allows such reinforcement (almost) exclusively by incorporating these idiomatic sentences (which have indeed changed semantically, but are still syntactic sentences) as subordinate clauses, rather than by reduplication of the emotively bleaching intensifier. This process ends up by the idiomatic sentences being reanalyzed as full-fledged modifying words. I then showed that once "wordification" has been completed, further developments on the lexicalization path are conditioned by the removal of the special context involved in this lexicalization process. I also showed that this lexicalization process, specifically, the change in the grammatical status of the members of the Ultimate construction

family, is affected by the presence (or lack of) (iii) *horizontal links* between the newly evolved idiomatic sentences incorporated into the special context that enables the change in their grammatical status and competing constructions, such that may block this whole process.

In sum, I showed that the abovementioned links, which form a multidimensional network, allow to provide a parsimonious—yet exhaustive—account for this lexicalization process along the lines of Construction Grammar *exclusively*.

Indeed, in retrospect and measured by the results reported in the previous chapters, the choice of Construction Grammar as a theoretical framework for the phenomenon here studied has proven to be an excellent choice. But this choice was, by no means, opportunistic. Construction Grammar has never dismissed peripheral phenomena as negligible. In fact, it has undertaken to account for every possible linguistic phenomenon, as peripheral as it may be, without any exception.

Apart from the resolution of the lexicalization process here studied, the research provided support for the assumption that all constructions belong to the same representation level, compatible with the claim that there is no division between syntax and lexicon. Firstly, I showed that not just phrases, which are intra-sentential elements (by definition), can turn into words, but so can sentences, which are independent extra-sentential elements. Secondly, this lexicalization process can be appreciated as a case of *embedded productivity* (Booij & Audring, 2018) where "the productive use of morphology cannot be analyzed in isolation, without taking its syntactic context into account" (p. 227).

Beyond the contribution to Construction Grammar, the phenomenon here studied provided additional support for the consensual notion that semantic change in lexicalization (as opposed to grammaticalization) cannot happen across phrasal or even clause boundaries, but only within the boundaries of a phrase (Lehmann, 2020). As long as no boundaries are crossed, whether a linguistic sequence is considered a phrase or a full sentence (as is the case in this dissertation), it makes up a potential candidate for undergoing semantic change and then lexicalization to become a content word.

I complemented my analysis with a solution to a practical problem I had faced at the early stages of my research ― the lack of a diachronic corpus which is necessary to substantiate the presence of semantic change. Indeed, it has been suggested that "the speakers of a language […] are generally not aware of language change taking place" (Keller, 1994: 77). But I relied on data produced by speakers who are sensitive to semantic change, and proposed alternative (quantitative) methods to substantiate this change in the absence of a diachronic corpus. Speakers' exclusive sensitivity to the semantic—not the grammatical—aspect of (polysemous) syntagmas also provided support (if unconsciously) for the assumption that all constructions, regardless of level of complexity, belong to the same level of representation compatible with the claim that there is no division between syntax and lexicon, a thread running through the entire dissertation.

References

Aarts, Bas. (1998). Binominal noun phrases in English. *Transactions of the Philological Society, 96*(1), 117-158.

Allerton, David J., & Cruttenden, Alan. (1979). Three reasons for accenting a definite subject. *Journal of Linguistics, 15*(1), 49-53.

Ariel, Mira. (1990). *Accessing Noun-Phrase Antecedents*. London: Routledge.

Ariel, Mira. (2008). *Pragmatics and Grammar*. New York: Cambridge University Press.

Ariel, Mira. (2017). What kind of pragmatics? *Xaval al hazman*. Handout from a talk in Ben-Gurion University, Be'er Sheba, Israel. [in Hebrew].

Assouline, Dalit. (2015). ha-hitnagdut ha-kanaʔit la-ʕivrit ha-xadaša ba-mciʔut ha-lšonit be-Israʔel [The opposition of ultra-orthodox "zealots" to Modern Hebrew in contemporary Israel]. *Carmillim: lexeker ha-lašon ha-ʕivrit u-lšonot smuxot [For the Study of Hebrew and Related Languages], 11*, 123-132. [in Hebrew].

Assouline, Dalit. (2017). *Contact and Ideology in a Multilingual Community: Yiddish and Hebrew among the Ultra-Orthodox*. Boston: De Gruyter Mouton.

Assouline, Dalit. (2018). Haredi Yiddish in Israel and the United States. In B. Hary & S. Bunin Benor (Eds.), *Languages in Jewish Communities, Past and Present* (pp. 472-488). Boston & Berlin: De Gruyter Mouton.

Attardo, Salvatore. (1994). *Linguistic Theories of Humor*. Berlin & Boston, MA: De Gruyter.

Attardo, Salvatore. (2018). Universals in puns and humorous wordplay. In E. Winter-Froemel & V. Thaler (Eds.), *Cultures and Traditions of Wordplay and Wordplay Research* (pp. 89-110). Berlin & Boston: De Gruyter.

Audring, Jenny, & Booij, Geert. (2016). Cooperation and coercion. *Linguistics, 54*(4), 617-637.

Bailey, Guy, Wikle, Tom, Tillery, Jan, & Sand, Lori. (1991). The apparent time construct. *Language Variation and Change, 3*(3), 241-264.

Bar-Asher, Moshe. (2012). ʕal ribuy paneha šel ha-ʕivrit bat yamenu (A multi-faceted language). *pirkey ʕiyun ba-ʕivrit ha-xadaša u-va-ʕasiya ba (Studies in Modern Hebrew)* (pp. 70-92). Jerusalem: The Academy of the Hebrew Language. [in Hebrew].

Bar-Ziv Levy, Miri. (2017). *me-racuy le-macuy: yicug lešon ha-dibur be-sirtey ʕalila be-ʕivrit 1932-1988 [The Cinematic Representation of Hebrew Speech 1932-1988]*. Jerusalem: the Hebrew University in Jerusalem dissertation. [in Hebrew].

Barðdal, Jóhanna, & Gildea, Spike. (2015). Diachronic Construction Grammar: Epistemological context, basic assumptions and historical implications. In L. Sommerer, J. Barðdal, S. Gildea, & E. Smirnova (Eds.), *Diachronic Construction Grammar*. Amsterdam: John Benjamins.

Bardenstein, Ruti. (2021). Persistent argumentative discourse markers: The case of Hebrew rectification-marker *be-ʕecem* (‘actually’). *Journal of Pragmatics, 172*, 254-269.

Bardenstein, Ruti, & Ariel, Mira. (2022). The missing link between truth and intensification. *Studies in Language, 46*(2), 285-322.

Bardenstein, Ruti, & Ariel, Mira. (2023). mamaš: xizuk tiʕuni matmid [Mamaš: Persistent argumentative strengthening]. *Lěšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 30*, 298-327. [in Hebrew].

Barlow, Michael. (2013). Individual differences and usage-based grammar. *International Journal of Corpus Linguistics, 18*(4), 443-478.

Bateson, Gregory. (1972). *Steps to an Ecology of Mind. Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. Northvale, NJ: Jason Aronson Inc.

Bauer, Laurie. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.

Baumel, Simeon D. (2006). *Sacred Speakers: Language and Culture among the Ultra-Orthodox in Israel*. New-York, NY & Oxford: Berghahn Books.

Baxter, Gareth, & Croft, William. (2016). Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change, 28*(2), 129-173.

Beltrama, Andrea, & Bochnak, M. Ryan. (2015). Intensification without degrees cross-linguistically. *Natural Language & Linguistic Theory, 33*(3), 843-879.

Ben-Rafael, Eliezer. (2002). rav-tarbutiyut ve-rav-lešonyiut be-Israʔel [Multiculturalism and multilingualism in Israel]. In S. Izre'el & M. Mendelson (Eds.), *medabrim ʕivrit: le-xeker ha-lašon ha-mduberet ve-ha-šonut ha-lšonit be-Israʔel [Speaking Hebrew: Studies in the Spoken Language and in Linguistic Variation in Israel]* (pp. 67-84). Tel Aviv: Tel Aviv Univerisy Press. [in Hebrew].

Berg, Thomas. (2014). Competition as a unifying concept for the study of language. *Mental Lexicon, 9*(2), 338-370.

Bergs, Alexander, & Diewald, Gabriele. (2008). *Constructions and Language Change*. Berlin & Boston, MA: De Gruyter Mouton.

Berman-Aronson, Ruth. (1978). *Modern Hebrew Structure*. Tel Aviv: University Publishing Projects.

Berman, Ruth A. (1980). The case of an (S)VO language: Subjectless constructions in Modern Hebrew. *Language, 56*(4), 759-776.

Biber, Douglas, & Gray, Bethany. (2016). *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge: Cambridge University Press.

Blank, Andreas. (2001). Pathways of lexicalization. In W. Raible, M. Haspelmath, E. König, & W. Oesterreicher (Eds.), *Language Universals and Language Typology* (pp. 1596-1608). Berlin: De Gruyter Mouton.

Blank, Grant, & Reisdorf, Bianca C. (2012). The participatory web: A user perspective on Web 2.0. *Information, Communication & Society, 15*(4), 537-554.

Blau, Joshua. (1999). ha-gvul ha-dak ben mišpat zika le-mišpat šeʔela ʕakifa [Between indirect questions and relative clauses in Modern Hebrew] *Lěšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 62*(1/2), 87-92. [in Hebrew].

Boas, Hans C. (2013). Cognitive Construction Grammar. In G. Trousdale & T. Hoffmann (Eds.), *The Oxford Handbook of Construction Grammar Oxford*. Oxford: Oxford University Press.

Bolozky, Shmuel. (1999). *Measuring Productivity in Word Formation: The Case of Israeli Hebrew*. Leiden: Brill.

Booij, Geert. (2009). Lexical integrity as a formal universal: A constructionist view. In S. Scalise, E. Magni, & A. Bisetto (Eds.), *Universals of Language Today* (pp. 83-100). Dordrecht: Springer.

Booij, Geert, & Audring, Jenny. (2018). Category change in construction morphology. In K. Van Goethem, M. Norde, E. Coussé, & G. Vanderbauwhede (Eds.), *Category Change from a Constructional Perspective* (pp. 209-228). Amsterdam & Philadelphia: John Benjamins.

Börjars, Kersti, Vincent, Nigel, & Walkden, George. (2015). On constructing a theory of grammatical change. *Transactions of the Philological Society, 113*(3), 363-382.

Botha, Rudolf P. (1981). A base rule theory of Afrikaans synthetic compounds. In M. Moortgat, H. van der Hulst, & T. Hoekstra (Eds.), *The Scope of Lexical Rules* (pp. 1-77). Dordrecht: Foris.
Boucher, Jerry, & Osgood, Charles E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior, 8*, 1-8.
Bréal, Michel. (1964 [1899]). *Semantics: Studies in the Science of Meaning* (N. Cust, Trans.). New York, NY: Henry Holt & Company.
Brinton, Laurel J., & Traugott, Elizabeth Closs. (2005). *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
Brown, Cheryl. (1983). Topic continuity in written English narrative. In T. Givón (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-language Study* (pp. 313-342). Amsterdam: John Benjamins.
Bruening, Benjamin. (2018). The lexicalist hypothesis: Both wrong and superfluous. *Language, 94*(1), 1-42.
Bybee, Joan L. (1985). *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam: John Benjamins.
Bybee, Joan L. (1998). The emergent lexicon. In M. C. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *CLS* (Vol. 34, pp. 421-435). Chicago, IL: Chicago Linguistic Society.
Bybee, Joan L. (2003). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The New Psychology of Language* (Vol. 2, pp. 145-167). New York, NY: Psychology Press.
Bybee, Joan L. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*(4), 711-733.
Bybee, Joan L. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
Caffi, Claudia. (1994). Metapragmatics. In R. E. Asher & J. M. Y. Simpson (Eds.), *The Encyclopedia of Language and Linguistics* (pp. 2461-2466). Oxford: Pergamon Press.
Cahaner, Lee, & Malach, Gilad. (2021). *Statistical Report on Ultra-Orthodox Society in Israel*. Retrieved from the web-site of the Israel Democracy Institue: https://en.idi.org.il/haredi/2021/?chapter=38439. [in Hebrew].
Cappelle, Bert. (2006). Particle placement and the case for 'allostructions'. *Constructions, special volume 1: Constructions all over: Case studies and theoretical implications*. Retrieved from https://hal.science/hal-01495786.
Carnap, Rudolf. (1970 [1939]). *Foundations of Logic and Mathematics* (13th ed.). Chicago, IL: The University of Chicago Press.
Chafe, Wallace L. (1974). Language and consciousness. *Language, 50*(1), 111-133.
Chen, Henian, Cohen, Patricia, & Chen, Sophie. (2010). How big is a big Odds Ratio? Interpreting the magnitudes of Odds Ratios in epidemiological studies. *Communications in Statistics - Simulation and Computation, 39*(4), 860-864.
Chiarcos, Christian, Claus, Berry, & Grabski, Michael (Eds.). (2011). *Salience: Multidisciplinary Perspectives on its Function in Discourse*. Berlin & Boston, MA: De Gruyter Mouton.
Chomsky, Noam. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.
Chomsky, Noam. (1970). Remarks on nominalization. In R. Jacobsand & P. Rosenbaum (Eds.), *Readings in English Transformational Grammar* (pp. 184-221). Waltham, MA: Ginn and Company.
Cohen, Tirza. (2008). *kavey lašon yixudiyim le-sofrot datiot ve-xarediot be-hašvaʔa le-sofrot xiloniot ba-siporet ha-isreʔelit bat zmanenu [Unique Linguistic Features of Religiously Observant Women Authors Compared to Secular Women Authors in Contemporary Israeli Literature]*. Ramat Gan: Bar-Ilan University dissertation. [in Hebrew].
Comrie, Bernard. (1989). *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago, IL: University of Chicago Press.

Coussé, Evie, Andersson, Peter, & Olofsson, Joel. (2018). *Grammaticalization Meets Construction Grammar*. Amsterdam: John Benjamins.
Croft, William. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. New York: Oxford University Press.
Culioli, Antoine. (1990). *Pour une Linguistique de L'énonciation, Tome I. Opérations et Représentations*. Paris: Ophrys.
Culioli, Antoine. (1995). *Cognition and Representation in Linguistic Theory*. Amsterdam: John Benjamins.
Culperer, Jonathan, & Haugh, Michael. (2014). *Pragmatics and the English Language*. London & New York, NY: Palgrave Macmillan.
Culperer, Jonathan, & Kytö, Merja. (2010). *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.
Dahl, Östen, & Fraurud, Kari. (1996). Animacy in grammar and discourse. In T. Fretheim & J. K. Gundel (Eds.), *Reference and Referent Accessibility* (pp. 47-64). Amsterdam: John Benjamins.
Daneš, František. (1966). A three-level approach to syntax. In F. Daneš, K. Horálek, V. Skalička, P. Trost, & J. Vachek (Eds.), *L'École de Prague D'aujourd'hui. Travaux Linguistiques de Prague.* (Vol. 1, pp. 225-240). Alabama: University of Alabama Press.
Danet, Brenda. (2001). *Cyberpl@y: Communicating Online*. New York: Routledge.
Dattner, Elitzur. (2019). The Hebrew dative: Usage patterns as discourse profile constructions. *Linguistics, 57*(5), 1073-1110.
De Smet, Hendrik. (2012). The course of actualization. *Language, 88*(3), 601-633.
De Smet, Hendrik, D'hoedt, Frauke, Fonteyn, Lauren, & Van Goethem, Kristel. (2018). The changing functions of competing forms: Attraction and differentiation. *Cognitive Linguistics, 29*(2), 197-234.
Delabastita, Dirk. (2001). Aspects of interlingual ambiguity: Polyglot punning. In P. Bogaards, J. Rooryck, & P. J. Smith (Eds.), *Quitte ou Double Sens* (pp. 45-64). Amsterdam: Rodopi.
Detges, Ulrich (2010). Computed or entrenched? The French imparfait de politesse. In H.-J. Schmid & S. Handl (Eds.), *Cognitive Foundations of Linguistic Usage Patterns* (pp. 195-224). Berlin & New York, NY: De Gruyter Mouton.
Diessel, Holger. (2015). Usage-based construction grammar. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 296-322). Berlin, München & Boston, MA: De Gruyter Mouton.
Diessel, Holger. (2020). A dynamic network approach to the study of syntax. *Frontiers in Psychology, 11*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2020.604853.
Diessel, Holger. (2023). *The Constructicon: Taxonomies and Networks.* Elements in Construction Grammar, Retrieved from https://www.cambridge.org/core/elements/constructicon/ECD7CC294373BC24CF5C2D97068F1F82.
Diewald, Gabriele. (2006). Context types in grammaticalization as constructions. *Constructions, special volume 1: Constructions all over: Case studies and theoretical implications*. Retrieved from https://constructions.journals.hhu.de/article/view/443.
Divjak, Dagmar. (2010). *Structuring the Lexicon: A Clustered Model for Near-Synonymy*. Berlin & Boston, MA: De Gruyter Mouton.
Divjak, Dagmar, & Janda, Laura A. (2008). Ways of attenuating agency in Russian. *Transactions of the Philological Society, 106*(2), 138-179.
Dixon, Robert M. W., & Aikhenvald, Alexandra Y. (2002). Word: A typological framework. In R. M. W. Dixon & A. Y. Aikhenvald (Eds.), *Word: A Cross-linguistic Typology* (pp. 1-41). Cambridge: Cambridge University Press.
Doyle, Arthur Conan. (2000 [1887]). *Study in Scarlet*. London: ElecBook.

Doyle, Arthur Conan (1892). *The Adventures of Sherlock Holmes*. London: George Newnes Ltd.
Du Bois, John W. (2003). Argument structure: Grammar in use. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Eds.), *Preferred Argument Structure: Grammar as Architecture for Function* (pp. 11-60). Amsterdam: John Benjamins.
Du Bois, John W. (2014). Towards a dialogic syntax. *Cognitive Linguistics, 25*(3), 359-410.
Dubnov, Keren. (2005a). šem ha-toʔar be-tafkid HGM ba-ʕivrit ha-xadaša be-rešita [Adjectives functioning as Impersonals (Ḥagam) in early Modern Hebrew]. In T. Alexander, J. Tobi, D. Laor, Z. Amishai-Maisels, & O. Schwartzwald (Eds.), *Iggud: Selected Essays in Jewish Studies* (pp. 31-40). Jerusalem: World Union of Jewish Studies. [in Hebrew].
Dubnov, Keren. (2005b). *tirgumey šeʔila mivniyim ba-roved ha-mukdam šel ha-ʕivrit ha-xadaša [Structural Loan Translations in Early Modern Hebrew]*. Jerusalem: The Hebrew University in Jerusalem dissertation. [in Hebrew].
Duffy, Susan A., Morris, Robin K., & Rayner, Keith. (1988). Lexical ambiguity and fixation times in reading. *Journal of memory and language, 27*(4), 429-446.
Englebretson, Robert. (1997). *Genre and grammar: Predicative and attributive adjectives in spoken English.* Paper presented at the the 23rd Annual Meeting of the Berkeley Linguistics Society, UC Berkeley, CA.
Epstein, Ruth. (1971). *ha-mišpatim ha-musaʔiyim ba-mikra [The Object Clauses in the Old Testement]*. Tel Aviv: Tel Aviv University thesis. [in Hebrew].
Evans, Nicholas, & Wilkins, David. (2000). In the mind's ear: The semantic extensions of perception verbs in Australian languages. *Language, 76*(3), 546–592.
Fader, Ayala. (2009). *Mitzvah Girls: Bringing Up the Next Generation of Hasidic Jews in Brooklyn*. Princeton, NJ: Princeton University Press.
Ferris, D. Connor. (1993). *The Meaning of Syntax : A Study in the Adjectives of English*. London: Longman.
Fillmore, Charles J., Kay, Paul, & O'Connor, Mary Catherine. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Langauge, 64*(3), 501-538.
Finkbeiner, Rita, & Meibauer, Jörg. (2016). Boris "Ich bin drin" Becker ('Boris I am in Becker'). Syntax, semantics and pragmatics of a special naming construction. *Lingua, 181*, 36-57.
Firbas, Jan. (1974). Some aspects of the Czechoslovak approach to problems of functional sentence perspective In F. Daneš (Ed.), *Papers of Functional Sentence Perspective* (pp. 11-37). Prague: Academia/The Hague, Mouton.
Fischer, Roswitha. (1998). *Lexical Change in Present-Day English: A Corpus-Based Study of the Motivation, Institutionalization, and Productivity of Creative Neologisms*. Tübingen: G. Narr.
Fisherman, Haya. (1994). kivuney šinuyim be-šimuš be-milim zarot be-ʕivrit bat-yamenu [The direction of changes of foreign words used in contemporary Hebrew]. In M. Muchnik (Ed.), *Foreign Influences on Contemporary Hebrew* (pp. 57-71.). Tel Aviv: The Open university Press. [in Hebrew].
Flach, Susanne. (2020). Constructionalization and the Sorites Paradox: The emergence of the into-causative. In L. Sommerer & E. Smirnova (Eds.), *Nodes and Networks in Diachronic Construction Grammar* (pp. 45-67). Amsterdam & Philadelphia: John Benjamins.
Flach, Susanne. (2021). From movement into action to manner of causation: changes in argument mapping in the into-causative. *Linguistics, 59*(1), 247-283.
Fodor, Jerry A., & Bever, Thomas G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior, 4*(5), 414-420.
Fridman, Israela, Shaul-Mena, Nava, Fogel, Nir, Romanov, Dmitri, Feldman, Mark, Sehayek, Ruth, Schifris, Gustavo, & Portnoy, Haim. (2011). *šitot medida va-ʔamidat godla šel ha-ʔuxlusia ha-xaredit be-Israʔel [Measurement and estimates of the population of*

ultra-orthodox Jews in Israel]. *Israel Central Bureau of Statistics.* Retrieved from the web-site of the Israeali Central Bureau of Statistics: https://www.cbs.gov.il/he/publications/DocLib/tec/tec25/tec25.pdf. [in Hebrew].
Fried, Mirjam. (2009). Construction Grammar as a tool for diachronic analysis. *Constructions and Frames, 1*(2), 262-291.
Friedman, Menachem. (1991). *ha-xevra ha-xaredit — mekorot, megamot ve-tahalixim [The Haredi (Ultra-Orthodox) Society — Sources, Trends and Processes]*. Jerusalem: The Jerusalem Institute for Israel Studies. [in Hebrew].
Gaeta, Livio. (2015). Lexeme formation in a conscious approach to the lexicon. In L. Bauer, L. Körtvélyessy, & P. Štekauer (Eds.), *Semantics of Complex Words* (pp. 115-140). Heidelberg: Springer.
Gaeta, Livio, & Ricca, Davide. (2009). *Composita solvantur*: Compounds as lexical units or morphological objects? *Rivista di Linguistica, 21*(1), 35-70.
Gibbs, Raymond W. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General, 115*(1), 3-15.
Gibbs, Raymond W. (1994). *The Poetics of Mind: Figurative Thought, Language, and Understanding*. New York, NY: Cambridge University Press.
Gibbs, Raymond W. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics, 34*(4), 457–486.
Giora, Rachel. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 8*(3), 183-206.
Giora, Rachel. (2003). *On Our Mind: Salience, Context, and Figurative Language*. New York, NY: Oxford University Press.
Giora, Rachel, Fein, Ofer, Kronrod, Ann, Elnatan, Idit, Shuval, Noa, & Zur, Adi. (2004). Weapons of mass distraction: Optimal innovation and pleasure ratings. *Metaphor and Symbol, 19*(2), 115-141.
Giora, Rachel, Givoni, Shir, Heruti, Vered, & Fein, Ofer. (2017). The role of defaultness in affecting pleasure: The optimal innovation hypothesis revisited. *Metaphor and Symbol, 32*(1), 1-18.
Givón, Talmy. (1976). Topic, pronoun and grammatical agreement. In C. N. Li (Ed.), *Subject and Topic* (pp. 151-188). New York, NY: Academic Press.
Givón, Talmy. (2009). *The Genesis of Syntactic Complexity: Diachrony, Ontogeny, Neuro-cognition, Evolution*. Amsterdam & Philadelphia, PA: John Benjamins.
Givoni, Shir. (2020). *Marking Multiple Meanings.* Tel Aviv: Tel Aviv University dissertation.
Givoni, Shir, Giora, Rachel, & Bergerbest, Dafna. (2013). How speakers alert addressees to multiple meanings. *Journal of Pragmatics, 48*(1), 29-40.
Glinert, Lewis. (1989). *The Grammar of Modern Hebrew*. Cambridge: Cambridge University Press.
Glinert, Lewis, & Shilhav, Yosseph. (1991). Holy land, holy language: A study of an ultraorthodox Jewish ideology. *Language in Society, 20*(1), 59-86.
Goldberg, Adele E. (1995). *Constructions : A Construction Grammar Approach to Argument Structure*. Chicago, IL: The University of Chicago Press.
Goldberg, Adele E. (2003). Constructions: A new theoretical approach to language. *Trends in cognitive sciences, 7*(5), 219-224.
Goldberg, Adele E. (2006). *Constructions at Work: The Nature of Generalization in Language*. New York, NY: Oxford University Press.
Goldshtein, Maria. (2014). *A Discourse Based Typology of Intensifiers*. Tel Aviv: Tel Aviv University thesis.
Gombert, Jean Émile. (1992). *Metalinguistic Development*. Chicago, IL: University of Chicago Press.
Grice, Herbert Paul. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics: Speech Acts* (Vol. 3, pp. 41-58). New York, NY: Academic Press.
Guy, Gregory R. (2013). The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables? *Journal of Pragmatics, 52*, 63-71.

Halevy, Rivka. (2001). tmurot be-maʕamada šel ha-smixut ha-pruda ba-ʕivrit šel yamenu [Functional changes of *šel* phrases in contemporary Hebrew]. *Lĕšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 63*(1-2), 61-80. [in Hebrew].
Harris, Alice C. (2000). Where in the word is the Udi clitic? *Language, 76*(3), 593-616.
Haspelmath, Martin. (2001). Non-canonical marking of core arguments in European languages. In A. Y. Aikhenvald, R. M. W. Dixon, & M. Onishi (Eds.), *Non-Canonical Marking of Subjects and Objects* (pp. 53-83). Amsterdam: John Benjamins.
Haspelmath, Martin. (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica, 45*(1), 31-80.
Haspelmath, Martin. (2019). What is the difference between a clause and a sentence? Retrieved from Diversity Linguistics Comment: Language Structures Throughout the World web-site: https://dlc.hypotheses.org/1725.
Haspelmath, Martin. (2022). Two senses of "lexicon": The inventorium and the lexemicon. Retrieved from Diversity Linguistics Comment: Language Structures Throughout the World web-site: https://dlc.hypotheses.org/2720.
Heaney, Seamus. (1966). *Death of a Naturalist*. London: Faber & Faber.
Heilman, Samuel C., & Friedman, Menachem. (1991). Religious fundamentalism and religious Jews: The case of the Haredim. In M. E. Marty & R. S. Appleby (Eds.), *Fundamentalisms Observed* (pp. 197-264). Chicago, IL: University of Chicago Press.
Heine, Bernd. (2002). On the role of context in grammaticalization. In I. Wischer & G. Diewald (Eds.), *New Reflections on Grammaticalization* (pp. 83-101). Amsterdam & Philadelphia: John Benjamins.
Heine, Bernd, & Kuteva, Tania (2002). *World Lexicon of Grammaticalization*. New York, NY: Cambridge University Press.
Henkin, Roni. (1994). kinuyey remez mašvim ve-šimušam le-derug [Comparative demonstratives as superlatives in Modern Hebrew]. *Lĕšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 54*(2), 119-143. [in Hebrew].
Henkin, Roni. (2020). Sociolinguistics of Modern Hebrew. In R. A. Berman (Ed.), *Usage-based Studies in Modern Hebrew: Background, Morpho-lexicon, and Syntax* (pp. 51-96). Amsterdam & Philadelphia, PA: John Benjamins.
Hetterle, Katja. (2015). *Adverbial Clauses in Cross-linguistic Perspective*. Berlin, München & Boston: De Gruyter Mouton.
Hilpert, Martin. (2013). *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.
Hilpert, Martin. (2014). *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.
Hilpert, Martin. (2018). Three open questions in Diachronic Construction Grammar. In E. Coussé, P. Andersson, & J. Olofsson (Eds.), *Grammaticalization Meets Construction Grammar* (pp. 21-39). Amsterdam & Philadelphia: John Benjamins.
Hopper, Paul J. (1991). On some principles of grammaticalization. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization: Theoretical and Methodological Issues* (Vol. 1, pp. 17-35). Amsterdam: John Benjamins.
Hopper, Paul J., & Traugott, Elizabeth Closs. (2003 [1993]). *Grammaticalization*. Cambridge: Cambridge University Press.
Hostetler, John A. (1993). *Amish Society* (4th ed.). Baltimore, MD: The John Hopkins University Press.
Hübler, Axel, & Bublitz, Wolfram. (2007). *Metapragmatics in Use*. Amsterdam: John Benjamins.
Huddleston, Rodney, & Pullum, Geoffrey K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
Hundt, Marianne, Nesselhauf, Nadja, & Biewer, Carolin. (2007). *Corpus Linguistics and the Web*. Leiden: Brill | Rodopi.

Irwin, Patricia. (2015). Expressive meaning in an AAE attributive construction. *Language Sciences, 50*, 12-29.
Isaacs, Miriam. (1999a). Contentious partners: Yiddish and Hebrew in Haredi Israel. *International Journal of Sociology and Language, 1999*(138), 101-122.
Isaacs, Miriam. (1999b). Haredi, haymish and frim: Yiddish vitality and language choice in a transnational, multilingual community. *International Journal of Sociology and Language, 1999*(138), 9-30.
Izre'el, Shlomo. (2002). le-tahalixey ha-hithavut šel ha-ʕivrit be-Israʔel [The emergence of Spoken Israeli Hebrew]. In S. Izre'el & M. Mendelson (Eds.), *medabrim ʕivrit: le-xeker ha-lašon ha-mduberet ve-ha-šonut ha-lšonit be-Israʔel [Speaking Hebrew: Studies in the Spoken Language and in Linguistic Variation in Israel]* (pp. 217-238). Tel Aviv: Tel Aviv Univerisy Press. [in Hebrew].
Izre'el, Shlomo. (2012). mispax o mišpat — le-mivne ha-mišpat ba-ʕivrit ha-mduberet [Hebrew sentence structure: Old habits die hard]. In M. Muchnik & T. Sadan (Eds.), *mexkarim be-ʕivrit u-ve-lšonot ha-yehudim lixvod Ora Schwarzwald [Studies in Hebrew and in Jewish languages in honor of Ora Schwarzwald]* (pp. 399-416). Jerusalm: Carmel. [in Hebrew].
Jackendoff, Ray. (1977). *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: The MIT Press.
Jakobson, Roman. (1960). Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in Language* (pp. 350-377). Cambridge: The MIT Press.
Janda, Laura A., & Divjak, Dagmar. (2015). The role of non-canonical subjects in the overall grammar of a language: A case study of Russian. In M.-L. Helasvuo & T. Huumo (Eds.), *Subjects in Constructions — Canonical and Non-canonical* (pp. 293-318). Amsterdam: John Benjamins.
Jaszczolt, Kasia M., & Allan, Keith (Eds.). (2011). *Salience and Defaults in Utterance Processing*. Berlin & Boston: De Gruyter Mouton.
Jaworski, Adam, Coupland, Nikolas , & Galasiński, Dariusz (2004). Metalanguage: Why now? In A. Jaworski, N. Coupland, & D. Galasiński (Eds.), *Metalanguage: Social and Ideological Perspectives* (pp. 3-8). Berlin & New York, NY: De Gruyter Mouton.
Ježek, Elisabetta, & Ramat, Paolo. (2009). On parts-of-speech transcategorization. *Folia Linguistica, 43*(2), 391–416.
Jing-Schmidt, Zhuo. (2007). Negativity bias in language: A cognitive affective model of emotive intersifiers. *Cognitive Linguistics, 18*(3), 417-443.
Johnson, Neal F. (1965). The psychological reality of phrase-structure rules. *Journal of Verbal Learning and Verbal Behavior, 4*(6), 469-475.
Kabatek, Johannes. (2015). Wordplay and discourse traditions. In A. Zirker & E. Winter-Froemel (Eds.), *Wordplay and Metalinguistic / Metadiscursive Reflection: Authors, Contexts, Techniques, and Meta-reflection*. Berlin & Boston, MA: De Gruyter.
Kanerva, Jonni M. (1987). Morphological integrity and syntax: The evidence from Finnish possessive suffixes. *Language, 63*(3), 498-521.
Karni, Tsofia. (2004). ha-memad ha-dati ke-meʔavxen lešoni [The dimension of religiosity as a linguistic feature]. *derex Efrata - bitʔon bet ha-midraš le-morim ve-le-gananot Efrata [Derech Efrata - Teachers College Yearbook], 11*, 129-147. [in Hebrew].
Katz, Albert N., & Ferretti, Todd R. (2003). Reading proverbs in context: The role of explicit markers. *Discourse Processes, 36*(1), 19-46.
Kavafis, Konstantinos Petrou. (2007). *The Collected Poems* (E. Sachperosglou, Trans.). New York, NY: Oxford University Press.
Keller, Rudi. (1994). *On Language Change: The Invisible Hand in Language* (B. Nerlich, Trans.). London & New York, NY: Routledge.
Kerremans, Daphné. (2015). *A Web of New Words: A Corpus-based Study of the Conventionalization Process of English Neologisms*. Bern: Peter Lang.

Kilgarriff, Adam, Baisa, Vít, Bušta, Jan, Jakubíček, Miloš, Kovář, Vojtěch, Michelfeit, Jan, Rychlý, Pavel, & Suchomel, Vít. (2014). The Sketch Engine: Ten years on. *Lexicography, 1*(1), 7-36.
Klein, Henny. (1998). *Adverbs of Degree in Dutch and Related Languages*. Amsterdam: John Benjamins.
Kraybill, Donald B. (2001 [1989]). *The Riddle of Amish Culture*. Baltimore, MD: Johns Hopkins University Press.
Kronrod, Ann, & Engel, Orit. (2001). Accessibility theory and referring expressions in newspaper headlines. *Journal of Pragmatics, 33*, 683-699.
Kuno, Susumu. (1972). Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry, 3*(3), 269-320.
Kuroda, Sige-Yuki (1972). The categorical and the thetic judgment: Evidence from Japanese syntax. *Foundations of Language, 9*(2), 153-185.
Kuzar, Ron. (1989). *mivne ha-meser šel ha-mišpat ba-ʕivrit ha-Isreʔelit [Message structure of the sentence in Israli Hebrew]*. Jerusalem: the Hebrew University in Jerusalem dissertation. [in Hebrew].
Kuzar, Ron. (1992a). he-x.g.m ― xelek diber o ʕemda taxbirit? [The Nominal Impersonal: A part of speech or a syntactic construction?]. *Lěšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 56*(3), 241-248. [in Hebrew].
Kuzar, Ron. (1992b). psukiot nominalizacia ba-ʕivrit ha-isreʔelit [Nominal clauses in Israeli hebrew]. *balšanut ʕivrit [Hebrew Linguistics: A Journal for Hebrew Descriptive, Computational and Applied Linguistics], 36*, 71-89. [in Hebrew].
Kuzar, Ron. (2001). *Hebrew and Zionism: A Discourse Analytic Cultural Study* Berlin & Boston, MA: De Gruyter Mouton.
Kuzar, Ron. (2002). tavnit he-x.g.m ha-pšuta ba-lašon ha-meyuceget ki-mduberet [The simple impersonal construction spoken-like langauge]. In S. Izre'el & M. Mendelson (Eds.), *medabrim ʕivrit: le-xeker ha-lašon ha-mduberet ve-ha-šonut ha-lšonit be-Israʔel [Speaking Hebrew: Studies in the Spoken Language and in Linguistic Variation in Israel]* (pp. 329-352). Tel Aviv: Tel Aviv University. [in Hebrew].
Kuzar, Ron. (2006a). digmey ha-mišpat šel ha-ʕivrit ha-Isreʔelit ʕal-pi Rosén: ʕiyun bikorti ve-hacaʕot tikun [Sentence patterns of Israeli Hebrew according to Rosén: A critical review and a corrected model]. *ha-ʕivrit ve-ʔaxyoteha [Hebrew and its Sisters], 6-7*, 269-294. [in Hebrew].
Kuzar, Ron. (2006b). tavnit mišpat ha-kiyum ke-toremet be-mašmaʕut ha-kiyum [The contribution of the pattern of the existential sentence to existential meaning]. In A. Maman & S. E. Fassberg (Eds.), *mexkarim be-lašon [Language Studies]* (Vol. 10, pp. 101-112). Jerusalem: The Hebrew University. [in Hebrew].
Kuzar, Ron. (2012). *Sentence Patterns in English and Hebrew*. Amsterdam: John Benjamins.
Labov, William. (1963). The social motivation of a sound change. *WORD, 19*(3), 273-309.
Labov, William. (1994). *Principles of Linguistic Change* (Vol. 1: Internal Factors). Cambridge: Blackwell.
Labov, William. (2017). Uriel Weinreich: Builder on empirical foundations. *Journal of Jewish Languages, 5*(2), 253-266.
Lambrecht, Knud. (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
Lambrecht, Knud. (2000). When subjects behave like objects: An analysis of the merging of S and O in Sentence-Focus constructions across languages. *Studies in Language, 24*(3), 611–682.
Lambrecht, Knud, & Polinsky, Maria. (1997). *Typological variation in Sentence-Focus constructions*. Paper presented at the 33rd Regional Meeting of the Chicago Linguistic Society, Chicago, IL.
Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar. Theoretical Prerequisites* (Vol. 1). Stanford, CA: Stanford University Press.

Langacker, Ronald W. (1991). *Foundations of Cognitive Grammar. Descriptive Application* (Vol. 2). Stanford, CA: Stanford University Press.
Langacker, Ronald W. (2008). *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
Leech, Geoffrey. (2007). New resources, or just better old ones? The Holy Grail of representativeness. In M. Hundt, N. Nesselhauf, & C. Biewer (Eds.), *Corpus Linguistics and the Web* (pp. 133-149). Amsterdam: Rodopi.
Lehmann, Christian. (1988). Towards a typology of clause linkage. In J. Haiman & S. A. Thompson (Eds.), *Clause Combining in Grammar and Discourse.* (pp. 181-226). Amsterdam & Philadelphia: John Benjamins.
Lehmann, Christian. (2020). Univerbation. *Folia Linguistica, 54*(41), 205-252.
Lipka, Leonhard. (1992). *An Outline of English Lexicology: Lexical Structure, Word Semantics, and Word-Formation*. Tübingen: Walter de Gruyter.
Loewenthal, Naftali. (2013). The Hasidic ethos and the schisms of Jewish society. *Jewish History, 27*(2/4), 377-398.
Lorenz, Gunter. (2002). Really worthwhile or not really significant? A corpus-based approach to the delexicalization and grammaticalization of intensifiers in Modern English. In I. Wischer & G. Diewald (Eds.), *New Reflections on Grammaticalization* (pp. 143-161). Amsterdam: John Benjamins.
Lyons, John. (1977). *Semantics* (Vol. 1). Cambridge: Cambridge University Press.
MacLaren, Richard I. (1989). The distinction between linguistic awareness and metalinguistic consciousness: An applied perspective. *Rassegna Italiana di Linguistica Applicata, 21*(1-2), 5-18.
Marantz, Alec. (1984). *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.
Marx, Karl, & Engels, Friedrich. (1848). *Manifest der Kommunistischen Partei*. London: J. E. Burghard.
McNabb, Yaron. (2012). Cross-categorial modification of properties in Hebrew and English. *Semantics and Linguistic Theory, 22*, 365-382.
Meibauer, Jörg. (2007). How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction. *Morphology, 17*(2), 233-259.
Melnik, Nurit. (2002). *Verb-initial Constructions in Modern Hebrew*. Berkeley, CA: University of California dissertation.
Melnik, Nurit. (2006). A constructional approach to verb-initial constructions in Modern Hebrew. *Cognitive Linguistics, 17*(2), 153-198.
Melnik, Nurit. (2014). tmurot ba-lašon ha-ʕivrit ha-dvura: kinyan ve-kiyum [Language change in spoken Hebrew: Existence and possession]. *balšanut ʕivrit [Hebrew Linguistics: A Journal for Hebrew Descriptive, Computational and Applied Linguistics], 68*, 67-82. [in Hebrew].
Méndez-Naya, Belén. (2003). On intensifiers and grammaticalization: The case of *SWIÞE*. *English Studies, 84*(4), 372-391.
Meyer, David, Zeileis, Achim, & Hornik, Kurt (2008). Visualizing Contingency Tables. In C.-h. Chen, W. Härdle, & A. Unwin (Eds.), *Handbook of Data Visualization* (pp. 590-616). Berlin: Springer-Verlag.
Michaelis, Laura A. (2001). Exclamative constructions. In W. Raible, M. Haspelmath, E. König, & W. Oesterreicher (Eds.), *Language Universals and Language Typology* (pp. 1038-1050). Berlin: Walter De Gruyter.
Michaelis, Laura A. (2004). Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics, 15*(1), 1-67.
Michaelis, Laura A., & Lambrecht, Knud. (1996). The exclamative sentence type in English. In A. E. Goldberg (Ed.), *Conceptual Structure, Discourse and Language* (pp. 375-389). Chicago, IL: CSLI Publications.
Mithun, Marianne. (2006). Polysynthesis in the Arctic. In M.-A. Mahieu & N. Tersis (Eds.), *Variations on Polysynthesis: The Eskaleut Languages* (pp. 3–17). Amsterdam: John Benjamins.

Mithun, Marianne. (2014). Syntactic and prosodic structures: Segmentation, integration, and in between. In T. Raso & H. Mello (Eds.), *Spoken Corpora and Linguistics Studies* (pp. 297-330). Amsterdam: John Benjamins.

Mithun, Marianne. (2020). Grammaticalization and polysynthesis: Iroquoian. In W. Bisang & A. Malchukov (Eds.), *Grammaticalization Scenarios from Africa, the Americas, and the Pacific* (Vol. 2, pp. 943-976). Berlin & Boston, MA: De Gruyter Mouton.

Mor, Galila. (1992). ben lašon ktuva li-mduberet: ba-me še-ben ha-teʔur ha-meʔayex ha-maʕacim la-muʕacam al-yadav [Between written and spoken language: The interaction between the intensifier and the intensified element]. In U. Ornan, R. Ben-Shahar, & G. Toury (Eds.), *ha-ʕivrit safa xaya [Hebrew: A Living Language. Studies on the language in social and cultural contexts]* (Vol. 1, pp. 129-139). Haifa: Haifa University Press. [in Hebrew].

Mor, Uri, & Pat-El, Naama. (2016). The development of predicates with prepositional subjects in Hebrew. *Journal of Semitic Studies, 61*(2), 327-346.

Muchnik, Malka. (2000). darxey tecuratam šel šmot-ha-toʔar bi-lšon ha-slang [Adjective formation in Hebrew slang]. In O. Schwartzwald, S. Blum-Kulka, & O. Elite (Eds.), *sefer Raphael Nir: mexkarim be-tikšoret, be-valšanut u-ve-horaʔat ha-lašon [The Raphael Nir Jubilee Book: Studies in Communication, Linguistics and Language Teaching]* (pp. 377-389.). Jerusalem: Carmel Publishing House. [in Hebrew].

Munro, Heather L. (2022). The Politics of language choice in Haredi communities in Israel. *Journal of Jewish Languages*, *10*(2), 169-199.

Nerlich, Brigitte, & Clarke, David D. (2001). Ambiguities we live by: Towards a pragmatics of polysemy. *Journal of Pragmatics*, *33*(3), 1-20.

Neuman, Yishai. (2017). *psukit zika ʕacmaʔit bli šin ha-zika ba-ʕivrit ha-xadaša [Independent relative clause lacking relative that in modern Hebrew]*. Paper presented at the The 44th conference of the Israel Association of Applied Linguistics, Achva Academic College, Israel. [in Hebrew].

Nir, Raphael. (1993). *darxey ha-yecira ha-milonit be-ʕivrit bat-zmanenu [The Ways of Lexical Formation in Contemporary Hebrew]* Tel Aviv: The Open University. [in Hebrew].

Noël, Dirk. (2007). Diachronic construction grammar and grammaticalization theory. *Functions of Language, 14*, 177-202.

Norén, Kerstin, & Linell, Per (2007). Meaning potentials and the interaction between lexis and contexts: An empirical substantiation. *Pragmatics, 17*(3), 387-416.

O'Grady, William. (1998). The syntax of idioms. *Natural Language & Linguistic Theory, 16*(2), 279-312.

O'Reilly, Tim, & Battelle, John. (2009). Web squared: Web 2.0 five years on. Retrieved from http://dgroups.org/file2.axd/45127812-6239-4448-8bb6-508dc1b1e204/web2009_websquared-whitepaper.pdf.

OED, online. (2019). Oxford University Press.

Oryan, Shlomit. (1997). lašon cnuʕa - nefeš cnuʕa: defusej tikšoret miluliyim be-kerev banot ve-našim xarediot [Modest language — modest soul: Verbal communication among ultra-orthodox Jewish girls and women]. *balšanut ʕivrit [Hebrew Linguistics: A Journal for Hebrew Descriptive, Computational and Applied Linguistics], 41-42*, 7-19. [in Hebrew].

Ovid. (1998). *Metamorphoses (trans. Melville, A. D.)*. Oxford: Oxford University Press. [Original work published 8 AD].

Pat-El, Naama. (2018). The diachrony of non-canonical subjects in Northwest Semitic. In B. Jóhanna, N. a. Pat-El, & S. M. Carey (Eds.), *Non-canonically Case-marked Subjects : The Reykjavík-Eyjafjallajökull Papers* (pp. 159-184). Amsterdam: John Benjamins.

Pawley, Andrew. (1986). Lexicalization. In D. Tannen & J. Alatis (Eds.), *Language & Linguistics: The Interdependence of Theory, Data & Application. Georgetown University Round Table on Languages* (pp. 98-120). Washington, DC: Georgetown University Press.

Perek, Florent. (2012). Alternation-based generalizations are stored in the mental grammar: Evidence from a sorting task experiment. *Cognitive Linguistics, 23*(3), 601-635.
Perry-Hazan, Lotem. (2013). *ha-xinux ha-xaredi be-Israʔel: ben mišpat, tarbut u-folitika [The Ultra-Orthodox Education in Israel: Law, Cultura and Politics]*. Srigim-LiOn: Nevo Publishing House. [in Hebrew].
Petré, Peter, & Van de Velde, Freek. (2018). The real-time dynamics of the individual and the community in grammaticalization. *Language, 94*, 867-901.
Pomerantz, Anita. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 57-101). Cambridge: Cambridge University Press.
Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey, & Svartvik, Jan. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
Rabin, Chaim. (1985). miškeley ha-šem [Noun-patterns]. In B.-Z. Luria (Ed.), *sefer Avraham Even-Shoshan [Book in honor of Avraham Even-Shoshan]* (pp. 255-265). Jerusalem: Kiryat-Sefer. [in Hebrew].
Ramat, Paolo. (2016). What's in a word? *SKASE Journal of Theoretical Linguistics, 13*(2), 106-121.
Ramat, Paolo, & Ricca, Davide. (1998). Sentence adverbs in the languages of Europe. In J. van der Auwera & D. P. Ó. Baoill (Eds.), *Adverbial Constructions in the Languages of Europe* (pp. 187-273). Berlin & New York, NY: Mouton de Gruyter.
Ravid, Dorit. (1990). Internal structure constraints on new-word formation devices in Modern Hebrew. *Folia Linguistica, 23*(3-4), 289-346.
Ravid, Dorit. (1999). *šmot ha-pʕula ba-ʕivrit ha-xadaša: ʕiyun morfologi* [On the essence of derived nominals in Modern Hebrew]. *balšanut ʕivrit [Hebrew Linguistics: A Journal for Hebrew Descriptive, Computational and Applied Linguistics], 45*, 61-78. [in Hebrew].
Ravid, Dorit, & Shlesinger, Yitzhak. (1987). ʕal miyunam ve-darxey gziratam šel šmot-toʔar baʕaley sofit *-i* ba-ʕivrit ha-xadaša [On the classification and structure of *–i* suffixed adjectives]. *balšanut ʕivrit [Hebrew Linguistics: A Journal for Hebrew Descriptive, Computational and Applied Linguistics], 25*(25), 59-70. [in Hebrew].
Reichenbach, Hans. (1947). *Elements of Symbolic Logic*. New York, NY: The Macmillan Company.
Rogers, Everett M. (2003 [1962]). *Diffusion of Innovations* (5th ed.). New York, NY: The Free Press.
Ronneberger-Sibold, Elke. (2006). Lexical blends: Functionally tuning the transparency of complex words. *Folia Linguistica, 40*(1-2), 155-181.
Rosén, Haiim B. (1977). *Contemporary Hebrew*. Paris: Mouton.
Rosenthal, Ruvik. (2005/2018). *milon ha-slang ha-makif [The Comprehensive Dictionary of Israeli Slang]*. Jerusalem: Keter Books. [in Hebrew].
Rubinstein, Aynat. (2019). Historical corpora meet the digital humanities: the Jerusalem Corpus of Emergent Modern Hebrew. *Language Resources and Evaluation, 53*(4), 807-835.
Rubinstein, Eliezer. (1968). *ha-mišpat ha-šemani: ʕiyunim be-taxbir yamenu [The Nominal Sentence: A Study in the Syntax of Comtenporary Hebrew]*. Tel Aviv: HaKibbutz HaMeuchad. [in Hebrew].
Rudashevski, Ilana. (2022). *Taska*. Tel Aviv: 2Sfarim. [in Hebrew].
Sag, Ivan A. (2012). Sign-Based Construction Grammar: An informal synopsis. In H. C. Boas & I. A. Sag (Eds.), *Sign-Based Construction Grammar* (pp. 69-202). Stanford, CA: CSLI Publications.
Salazar-García, Ventura. (2010). Degree words, intensification, and word class distinctions in Romance languages. In U. Ansaldo, J. Don, & R. Pfau (Eds.), *Parts of Speech: Empirical and Theoretical Advances* (pp. 201-226). Amsterdam: John Benjamins.

Sasse, Hans-Jürgen. (1987). The thetic/categorical distinction revisited. *Linguistics, 25*(3), 511-580.
Schiffrin, Deborah. (1980). Meta-talk: Organizational and evaluative brackets in discourse. *Sociological Inquiry, 50*(3-4), 199-236.
Schmid, Hans-Jörg. (2007). Entrenchment, salience and basic levels. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 117-138). Oxford: Oxford University Press.
Schmid, Hans-Jörg. (2008). New words in the mind: Concept-formation and entrenchment of neologisms. *Anglia, 126*(1), 1-36.
Schmid, Hans-Jörg. (2015). A blueprint of the Entrenchment-and-Conventionalization model. *Yearbook of the German Cognitive Linguistics Association, 3*(1), 3-26.
Schmid, Hans-Jörg. (2016 [2011]). *English Morphology and Word-formation: An Introduction* (3rd ed.). Berlin: Erich Schmidt.
Schmid, Hans-Jörg. (2020). *The Dynamics of the Linguistic System: Usage, Conventionalization, and Entrenchment*. Oxford: Oxford University Press.
Schönefeld, Doris (Ed.) (2006). *Constructions, special volume 1: Constructions All Over: Case Studies and Theoretical Implications* (Vol. 1).
Schwarzwald, Ora. (2002). goney lašon ba-ʕivrit bat zmanenu [Language varieties in contemporary Hebrew]. In S. Izre'el & M. Mendelson (Eds.), *medabrim ʕivrit: le-xeker ha-lašon ha-mduberet ve-ha-šonut ha-lšonit be-Israʔel [Speaking Hebrew: Studies in the Spoken Language and in Linguistic Variation in Israel]* (pp. 141-175). Tel Aviv: Tel Aviv Univerisy Press. [in Hebrew].
Seidenberg, Mark S., Tanenhaus, Michael K., Leiman, James M., & Bienkowski, Marie. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*(4), 489-537.
Sela, Pnina. (2004). *ha-reka ha-sociolingvisti bi-lšona šel ha-publicistika ba-ʕitonut ha-xaredit [Sociolinguistic Factors in Ultra-Orthodox Newspapers]* Ramat Gan: Bar-Ilan University dissertation. [in Hebrew].
Shatil, Nimrod. (2006). maʕarexet miškeley ha-šem ba-ʕivrit bat yamenu be-reʔiya kognitivit [The Hebrew noun system: A structural cognitive perspective]. *Lĕšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 68*(3-4), 243-282. [in Hebrew].
Shatil, Nimrod. (2015). The nature and diachrony of Hebrew quality pseudo-partitives: Are they a calque from the contact languages? *Journal of Jewish Languages, 3*(1-2), 301-308.
Shaviv, Tamar. (2018). legamrey – gilgulu šel maʕacim [The evolution of an Intensifier]. *Helkat Lašon - A Journal for Theoretical and Applied Linguistics, 51*, 152 - 174. [in Hebrew].
Shibatani, Masayoshi, & Kageyama, Taro (1988). Word formation in a modular theory of grammar: Postsyntactic compounds in Japanese. *Langauge, 64*(3), 451-484.
Shirtz, Shahar, & Goldberg, Adele E. (submitted for publication). When a phrase masquerades as a word, people play along: The English Phrase-as-Lemma Construction.
Shlesinger, Yitzhak. (2000). *lešonot ha-ʕitonut: meʔafyenim signoniyim bi-lšon ha-ʕitonut ha-yomit ha-katuva ʕal medoreha ha-šonim [Jounalistic Hebrew: Stylistic Aspects of Israeli Newspaper Sections]*. Beer Sheva: Ben-Gurion University of Negev Press. [in Hebrew].
Smirnova, Elena, Mailhammer, Robert, & Flach, Susanne. (2019). The role of atypical constellations in the grammaticalization of German and English passives. *Diachronica, 36*(3), 384-416.
Smirnova, Elena, & Sommerer, Lotte. (2020). The nature of the node and the network – Open questions in Diachronic Construction Grammar. In L. Sommerer & E. Smirnova (Eds.), *Nodes and Networks in Diachronic Construction Grammar* (pp. 1-42). Amsterdam & Philadelphia: John Benjamin.

Smith, Philip L., & Little, Daniel R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin & Review, 25*(6), 2083-2101.
Sommerer, Lotte. (2020). Constructionalization, constructional competition and constructional death: Investigating the demise of Old English POSS DEM constructions. In L. Sommerer & E. Smirnova (Eds.), *Nodes and Networks in Diachronic Construction Grammar* (pp. 69-103). Amsterdam & Philadelphia: John Benjamins.
Sommerer, Lotte, Gildea, Spike, Barðdal, Jóhanna, & Smirnova, Elena (Eds.). (2015). *Diachronic Construction Grammar*. Amsterdam: John Benjamins.
Sommerer, Lotte, & Smirnova, Elena (Eds.). (2020). *Nodes and Networks in Diachronic Construction Grammar*. Amsterdam & Philadelphia: John Benjamins.
Spiegel, Ehud. (2011). *"ve-talmud tora keneged kulam": xinux xaredi le-vanim bi-yrušalayim ["Talmod Tora is Equivalent to All": The Ultra-Orthodox (Haredi) Education System for Boys in Jerusalem]*. Jerusalem: The Jerusalem Institue for Israel Studies. [in Hebrew].
Stern, Naftali. (1983). x.g.m. u-dmuy poʕal + šem poʕal ba-ʕivrit ha-isreʔelit [Impersonals (x.g.m) and verboids + infinitives in Israeli Hebrew]. *Lěšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 47*(3-4), 248-263. [in Hebrew].
Svanlund, Jan. (2018). Metalinguistic comments and signals: What can they tell us about the conventionalization of neologies? *Pragmatics & Cognition, 25*(1), 122-141.
Svensson, Maria Helena. (2008). A very complex criterion of fixedness: Non-compositionality. In F. Meunier & S. Granger (Eds.), *Phraseology : An Interdisciplinary Perspective* (pp. 81-93). Amsterdam: John Benjamins.
Tadmor, Uri. (1988). ha-notrikin ba-ʕivrit ha-xadaša [The acronym in modern Hebrew]. *Lěšonénu: A Journal for the Study of the Hebrew Language and Cognate Subjects, 39*(9-10), 225-257. [Hebrew].
Tannenbaum, Michal, & Abugov, Netta. (2010). The legacy of the linguistic fence: Linguistic patterns among ultra-orthodox Jewish girls. *Heritage Language Journal, 7*(1), 74-90.
Tarski, Alfred. (1944). The semantic conception of truth: And the foundations of semantics. *Philosophy and Phenomenological Research, 4*(3), 341-376.
Thackeray, William Makepeace. (1847-1848). *Vanity Fair*. London: Bradbury & Evans.
Thaler, Verena. (2016). Varieties of wordplay. In S. Knospe, A. Onysko, & M. Goth (Eds.), *Crossing Languages to Play with Words* (pp. 47-62). Berlin & Boston, MA: De Gruyter.
Thompson, Chad L. (2006). The English of the Swiss Amish of Northeastern Indiana. In T. E. Murray & B. L. Simon (Eds.), *Language Variation and Change in the American Midland: A New Look at 'Heartland' English* (pp. 275–292). Amsterdam & Philadelphia: John Benjamins.
Thompson, Sandra A. (1990). A discourse approach to the cross-linguistic category 'adjective'. In J. A. Hawkins (Ed.), *Explaining Language Universals* (pp. 167-185). Oxford: Basil Blackwel.
Titone, Debra A., & Connine, Cynthia M. (1994). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity, 9*(4), 247-270.
Tomlin, Russell S. (1986). *Basic Word Order: Functional Principles*. London: Croom Helm.
Traugott, Elizabeth Closs. (2003). Constructions in Grammaticalization. In B. D. Joseph & R. D. Janda (Eds.), *The Handbook of Historical Linguistics* (pp. 624-647). Malden, MA: Blackwell.
Traugott, Elizabeth Closs, & Dasher, Richard B. (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
Traugott, Elizabeth Closs, & Trousdale, Graeme. (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
Triger, Zvi. (2007). sof haderex/xaval al hazman [The end of the road/it's a waste of time]. In A. Rosenblum & Z. Triger (Eds.), *lelo milim: ha-tarbut ha-isreʔelit bi-reʔi ha-safa*

[Speechless: How Contemporary Israeli Culture is Reflected in Language] (pp. 228). Or Yehuda: Kinneret, Zmora-Bitan, Dvir. [in Hebrew].
Trips, Carola. (2012). *Empirical and theoretical aspects of phrasal compounds: Against the "syntax explains it all" attitude.* Paper presented at the the 8th Mediterranean Morphology Meeting, Cagliari.
Trips, Carola, & Kornfilt, Jaklin. (2015). Typological aspects of phrasal compounds in English, German, Turkish and Turkic. *STUF - Language Typology and Universals, 68*(3), 281-321.
Ungerer, Tobias. (2021). Using structural priming to test links between constructions: English caused-motion and resultative sentences inhibit each other. *Cognitive Linguistics, 32*(3), 389-420.
Ungerer, Tobias. (2024). Vertical and horizontal links in constructional networks: Two sides of the same coin? *Constructions and Frames, 16*(1), 30-63.
Vaisman, Carmel, & Gonen, Illan. (2011). *ʕivrit ʔinternetit [Hebrew On-line]*. Jerusalem: Keter Publishing House. [in Hebrew].
Van de Velde, Freek. (2014). Degeneracy: The maintenance of constructional networks. In R. Boogaart, T. Colleman, & G. Rutten (Eds.), *Extending the Scope of Construction Grammar* (pp. 141-179). Berlin: De Gruyter.
Van den Eynde, Karel. (1995). Methodological reflections on descriptive linguistics. Knud Togeby's principles and the Pronominal Approach. In L. Schøsler & M. Talbot (Eds.), *Studies in Valency* (Vol. 1, pp. 111-130). Odense: Odense University Press.
Van den Eynde, Karel, Kirchmeier-Andersen, Sabine, Mertens, Piet, & Schøsler, Lene. (2002). Distributional syntactic analysis and valency: Basic notions, procedures, and applications of the Pronominal Approach. In S. M. Johnson & B. E. Nevin (Eds.), *The Legacy of Zellig Harris: Language and Information Into the 21st Century* (Vol. 2, pp. 163-202). Amsterdam: John Benjamins.
Van Goethem, Kristel, Vanderbauwhede, Gudrun, & De Smet, Hendrik. (2018). The emergence of a new adverbial downtoner: Constructional change and constructionalization of Dutch [*ver van* X] and [*verre van* X] 'far from X'. In K. Van Goethem, M. Norde, E. Coussé, & G. Vanderbauwhede (Eds.), *Category Change from a Constructional Perspective* (pp. 179-205). Amsterdam & Philadelphia: John Benjamins.
Van Lier, Eva, & Rijkhoff, Jan. (2013). Flexible word classes in linguistic typology and grammatical theory. In J. Rijkhoff & E. Van Lier (Eds.), *Flexible Word Classes: Typological Studies of Underspecified Parts of Speech* (pp. 1-30). New York, NY: Oxford University Press.
Verschueren, Jef. (2004). Notes on the role of metapragmatic awareness in language use. In A. Jaworski, N. Coupland, & D. Galasiński (Eds.), *Metalanguage: Social and Ideological Perspectives* (pp. 53-74). Berlin & New York, NY: De Gruyter Mouton.
Vilkuna, Maria. (1989). *Free Word Order in Finnish: Its Syntax and Discourse Functions*. Helsinki: Suomalaisen Kirjallisuuden Seura.
Weinreich, Uriel. (1966). On the semantic structure of language. In J. H. Greenberg (Ed.), *Universals of Language* (pp. 142-216). Cambridge: MIT Press.
Weinreich, Uriel. (1979 [1953]). *Languages in Contact* (9th ed.). Paris: Mouton de Gruyter.
Weinreich, Uriel, William, Labov, & Herzog, Marvin I. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for Historical Linguistics: A Symposium* (pp. 95-195). Austin, TX: University of Texas Press.
Wexler, Paul. (1990). *The Schizoid Nature of Modern Hebrew: A Slavic Language in Search of a Semitic Past*. Wiesbaden: Otto Harrassowitz.
Wierzbicka, Anna. (1988). *The Semantics of Grammar*. Amsterdam: John Benjamins.
Wiese, Richard. (1996). Phrasal compounds and the theory of word syntax. *Linguistic Inquiry, 27,* 183-193.

Williams, John N. (1992). Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research, 21*(3), 193-218.

Winkler, Susanne. (2015). Exploring ambiguity and the ambiguity model from a transdisciplinary perspective. In S. Winkler (Ed.), *Ambiguity: Language and Communication* (pp. 1-26). Berlin & Boston, MA: De Gruyter.

Winter-Froemel, Esme. (2016). Approaching wordplay. In S. Knospe, A. Onysko, & M. Goth (Eds.), *Crossing Languages to Play with Words: Multidisciplinary Perspectives* (pp. 11-46). Berlin & Boston, MA: De Gruyter.

Winter-Froemel, Esme, & Zirker, Angelika. (2015). Ambiguity in speaker-hearer-interaction: A parameter-based model of analysis. In S. Winkler (Ed.), *Ambiguity: Language and Communication* (pp. 283-340). Berlin & Boston, MA: De Gruyter.

Yaguello, Marina. (1998). *Language Through the Looking Glass: Exploring Language and Linguistics*. Oxford: Oxford University Press.

Zadka, Yitzhak. (1991). "Yeš ma le-hašiv" al "ein el mi lifnot" ["There is how to respond" to "there's no one to turn to"]. *balšanut ʕivrit [Hebrew Linguistics: A Journal for Hebrew Descriptive, Computational and Applied Linguistics], 31-32*, 75-77. [in Hebrew].

Zehentner, Eva, & Traugott, Elizabeth Closs. (2020). Constructional networks and the development of benefactive ditransitives in English. In L. Sommerer & E. Smirnova (Eds.), *Nodes and Networks in Diachronic Construction Grammar* (pp. 167-211). Amsterdam & Philadelphia: John Benjamins.

Zenner, Eline, Kristiansen, Gitte, & Geeraerts, Dirk. (2016). Individual differences and in situ identity marking: Colloquial Belgian Dutch in the reality TV show “Expeditie Robinson“. In M. Romano & M. D. Porto (Eds.), *Exploring Discourse Strategies in Social and Cognitive Interaction: Multimodal and Cross-linguistic Perspectives* (pp. 39-77). Amsterdam & Philadelphia, PA: John Benjamins.

Zirker, Angelika, & Winter-Froemel, Esme. (2015a). Wordplay and its interfaces in speaker-hearer interaction: An introduction. In A. Zirker & E. Winter-Froemel (Eds.), *Wordplay and Metalinguistic/Metadiscursive Reflection: Authors, Contexts, Techniques, and Meta-Reflection* (pp. 1-22). Berlin & Boston, MA: De Gruyter.

Zirker, Angelika, & Winter-Froemel, Esme (Eds.). (2015b). *Wordplay and Metalinguistic/Metadiscursive Reflection: Authors, Contexts, Techniques, and Meta-Reflection*. Berlin & Boston, MA: De Gruyter.

Zwicky, Arnold M., & Pullum, Geoffrey K. (1983). Cliticization vs. Inflection: English N'T. *Language, 59*(3), 502-513.

Appendices

Appendix A. The inferential steps and mechanisms involved in the semantic change of the members of the *Ultimate* construction family

Originally, when *xaval al hazman* 'it's a waste of time' ('waste' left for inference) was used to convey an opinion, say, about a movie, the enriched proposition developed from the compositional meaning was that the 'wasted time' refers to the time that would be spent on *watching the movie*. If watching the movie is a waste of time, then the speaker must be implicating a negative evaluation of it. However, as suggested by Ariel (2017) the 'wasted time' need not necessarily refer to the projected state of affairs. In a metalinguistic use (initiated in the mid-nineties of the twentieth century), this 'wasted time' could refer to the speaking time required to do justice to the message the speaker wanted to convey. It would be a waste of time on the speaker's part to try and find the appropriate (strong) words to express her stance on some stance-object. The implicature is then that the speaker holds an extremely strong stance about the stance-object (explicated in the preceding utterance). This extreme stance usage of *xaval al hazman* later specialized for specifically positive contexts, probably due to the propensity for positive rather than for negative contexts (the *Pollyanna effect*, namely the natural tendency "to look on (and talk about) the bright side of life"; Boucher & Osgood 1969: 1). Recurrently derived implicatures may semanticize (Grice, 1975; Traugott & Dasher, 2002), and this is how *xaval al hazman* 'it's a waste of time' came to encode a general amplifying meaning.

Similar to *xaval al hazman* 'it's a waste of time', *xaval al hamilim* 'it's a waste of words', *en milim* 'there are no words' and *en ma ledaber/lehagid* 'there's nothing to say/speak' are also cases of metalinguistic use. The speaker uses them when she cannot find strong enough words to express her amazement in the face of a remarkable state of affairs, thus conveying an extreme stance.

En dvarim ka'ele/u 'there are no such things' is a hyperbolic utterance, an exaggeration referring to a stance-object which the speaker takes as not real. If the stance-object is (as-if) not real, then it must be extremely remarkable (see Goldshtein, 2014 for the 'unbelievability' category).

Ba livkot/lamut 'it feels like crying/dying' and *efšar lehištage'a*/*lamut* 'it's possible to go crazy/die' express the speaker's emotional reaction to a stance-object. This negative response is metonymically reduced to the high emotional intensity which, in turn, metaphorically maps onto semantic intensity (see Jing-Schmidt, 2007 for a cognitive-affective model; see also Heine & Kuteva, 2002: 50).

Appendix B. The details of the acceptability experiment

This experiment is a replication of that of Divjak and Janda (2008, and Janda and Divjak 2015) applied to Hebrew. It is intended to decide whether the Hebrew predicates *efšar* 'it's possible' and *ba* 'it feels like' form a 'complex event' with their respective infinitives. The decision is based on the results of two tests: (i) One examines the conceptual subordination of the infinitive, i.e., its status as a (possible) subject, and (ii) the other examines the temporal separation between the predicate and the respective infinitive.

To this end, six native Hebrew speakers between the ages 26-44 (3 women and 3 men) participated in a small-N design experiment. They were interviewed over Zoom on two separate occasions, one week apart.

(a) In the first Zoom meeting, the participants were presented with *pro-forms* of constructions along the Pronominal Approach (Van den Eynde, 1995; Van den Eynde, Kirchmeier-Andersen, Mertens, & Schøsler, 2002). A pro-form is the schematic representation of a construction which contains a specific verb. But the slots of the construction, which are usually filled with lexical items, are here replaced with pronouns representing the referents of the construction (see, for example, (B1b) below). This approach is used to determine the valency of verbs, while minimizing the mutual interaction of actual lexical items (i.e., the referents). Such an interaction could influence the (non)acceptability of the construction.

The participants were asked to decide whether specific pro-form constructions, among which were constructions hosting *efšar* 'it's possible' and *ba* 'it feels like', are acceptable. If they judged these constructions to be acceptable, then they were asked to produce an example of their own for each pro-form.

(b) In the second Zoom meeting, the participants were asked to judge whether constructed examples produced by other participants and by myself were acceptable.

In both meetings, conceptual subordination, or lack thereof (see (i) above), was tested by Divjak's (2010: Ch. 2) *thing-test*. This test aims to determine whether an infinitive following a predicate occupies the argument slot of the predicate (a slot usually occupied by an NP). If not, then the infinitive—rather than the predicate—is the "anchor point of the construction" (p. 41), and the predicate is but a modifier. Together, the predicate and the infinitive produce a 'complex event'. (B1a) is an example of a pro-form used for the thing-test; (B1b) is its gloss; (B1c) is an example cast in the form of (B1a); and (B1d) is its gloss. The marginal acceptability of (B1c) is indicated by a question mark.

(B1)		a.	*ba*	*le-mi*	*ma*
		b.	it.feels.like	to-who	what
	?	c.	*ba*	*l-i*	*rica*
		d.	it.feels.like	to-me	run_N

(B2) is another example of a pro-form. It is a dialogue between Π and Σ, where the only difference is Π‘s question. In (B2b), Π‘s question includes the verb *la'asot* ‘to do’, whereas in (B2a) it doesn't. If the predicate *ba* ‘it feels like’ projects an argument slot, then given Σ‘s response, only Π‘s question in (B2a) is acceptable. However, if the predicate *ba* ‘it feels like’ doesn't project an argument slot, then given Σ‘s response, only Π‘s question in (B2b) is acceptable.

(B2)	a.	Π: *ma*	*ba*	*lexa?*	
		what	it.feels.like	to.you	
		Σ: *laruc*	*b-a-park.*		
		to.run	in-the-park		
	b.	Π: *ma*	*ba*	*lexa*	*la'asot?*
		what	it.feels.like	to.you	to.do
		Σ: *laruc*	*b-a-park.*		
		to.run	in-the-park		

The same logic applies to (B3).

(B3)	a.	*laruc*	*b-a-park.*	*ze*	*ma*	*še-ba*	*li.*
		to.run	in-the-park	that.is	what	that-it.feels.like	to.me
	b.	*laruc*	*b-a-park.*	*ze*	*ma*	*še-ba*	*li*
		to.run	in-the-park	that.is	what	that-it.feels.like	to.me
		la'asot.					
		to.do					

In the second meeting, I examined also time separability or lack thereof (see (ii) above), using Divjak's (2010: Ch. 2) *time-test*. This test is intended to determine whether the predicate and the infinitive that follows permit conflicting temporal modifiers, thus occupying different points on a timeline. If not, then the two overlap temporally to produce a single unit, a ‘complex event’. (B4a) is an example of a pro-form used for the time-test; (B4b) is its gloss; (B4c) is an example cast in the form of (B4a); and (B4d) is its gloss. The unacceptability of (B4c) is indicated by an asterisk.

(B4)	a.	*ba*	*le-mi*	***hayom***	*la'asot*		***maxar***
	b.	it.feels.like	to-who	today	to.do		tomorrow
*	c.	*ba*	*l-i*	***hayom***	*lalexet l-a-sifriya*		***maxar***
	d.	it.feels.like	to-me	today	to.go	to-the-library	tomorrow

In both meetings, participants were presented with the two predicates of interest (*efšar* ‘it's possible’ and *ba* ‘it feels like’), an additional predicate associated only with P1 S-pattern thetic propositions (*keday* ‘worth’) and 3 other predicates associated with

Verbal S1 S-pattern categorical propositions (*lomed* ‘learns’, *mavti’ax* ‘promises’ and *me’ašer* ‘authorizes’).

Each participant was presented with all the pro-forms and the constructed examples, but in a unique pseudo-random order generated just for her/him. Here, I report speakers’ judgments regarding only the predicates relevant to this dissertation, *efšar* ‘it’s possible’ and *ba* ‘it feels like’.

In the first meeting, where participants were presented with pro-form constructions like (B1), one participant determined that *efšar* ‘it’s possible’ cannot be followed by a noun. The others produced examples followed by nouns, specific kinds of food (e.g., ice-cream, pizza, and coffee). However, they all spontaneously pointed out that such examples sound like “children’s-talk”, “something is missing”, “it’s the way my young niece would speak”, “ellipsis”, “it’s being used in special contexts, where one already knows what to do with this kind of object” (i.e., ‘receive’ for *efšar* ‘it’s possible’) etc. *Ba* ‘it feels like’ evoked the same judgements, examples and comments. The results regarding *efšar* ‘it’s possible’ are in line with Kuzar’s (2012: 107) observation that “[s]ome conventionalized situation types are associated with the predicate *efšar* ‘it’s possible’”, and the event described in those sentences is “[…] metonymically reconstructed around the NP”. Put differently, the situation is easily inferred from the NP. The results of the experiment testify that the same holds for *ba* ‘it feels like’ as well.

Additionally, in the second meeting, each participant was asked to determine which of the examples (a) and/or (b) in (B2) and (B3) were acceptable. Note that the infinitival VPs used in those examples were extracted from the examples that the remaining participants produced in the first meeting.

Results for *efšar* ‘it’s possible’ were clear cut — participants preferred the examples with *la’asot* ‘to do’ over the examples without *la’asot* ‘to do’ (10 out of 12). However, the results for *ba* ‘it feels like’ were less conclusive (6 out of 12). Still, the participants pointed out that the decision regarding the preference of examples with *la’asot* ‘to do’ over example without *la’asot* ‘to do’ is context-dependent. As for pro-forms like (B4), with *efšar* ‘it’s possible’, no time separation between the predicate and the infinitive is possible; with *ba* ‘it feels like’, two participants determined that no time separation is possible. The remaining four participants came up with “weird” and “awkward” examples (in their own words) that allowed *ba* ‘it feels like’ to be followed by an infinitive in contrastive contexts only. For example, *Yesterday I felt like flying tomorrow to London, but today I feel like flying tomorrow to Paris*. When asked whether such sentences are acceptable out of a contrastive context, participants ruled them out completely.

All in all, the results of this experiment attest to the status of *efšar* ‘it’s possible’ and the following infinitive as a ‘complex event’. This is not surprising in light of the fact that *efšar* ‘it’s possible’ is a full-fledged modal. *Ba* ‘it feels like’ and the following infinitive also form a ‘complex event’, though to a lesser degree.

Appendix C. More examples

This appendix is a collection of relevant examples of every member of the Ultimate constructions family. While reading them, the perceptive reader may notice that *xaval al hazman* (in C.1) is in a more advanced stage of lexicalization than any of its counterparts (C.2-C.10), as evidenced from the fact that *xaval al hazman* does not need *še* 'that' in order to modify modifiable elements, while the other members of the Ultimate constructions family do. A detailed analysis of this state of affairs is provided in Chapter 4

C.1 *Xaval al hazman*, originally, 'It's a waste of time'

(C1) a. *zo* *hayta* *mesiba* ***[xaval al hazman]***$_{\text{ADJECTIVE}}$.
this was a.party **it's a waste of time → amazing**

'This was an **amazing** party.'

(tinyurl.com/ymc2juu4)

b. *hi* *šara* *kol-kax* *be-vitaxon* *ve-ocma* *ve-hi*
she sings so with-confidence and-intensity and-she

rokedet ***[xaval al hazman]***$_{\text{MANNER ADVERB}}$.
dances **it's a waste of time → amazingly**

'She sings with such confidence and so intensely and she dances **amazingly**.'

(tinyurl.com/4rfj98ua)

c. *ha-herayon* *šeli* *hirgiš* *kmo* *maxala,* *savalti*
the-pregnancy my felt like an.illness I.suffered

[xaval al hazman]$_{\text{ADVERB}}$.
it's a waste of time → so much

'My pregnancy period felt like an illness, I suffered **so much**.'

(tinyurl.com/35zfz23y)

d. *anaxnu* *karega* *b-a-malon* *haxi haxi* *[…]*
we at.the.moment in-the-hotel most most

nof *yafe* ***[xaval al hazman]***$_{\text{INTENSIFIER}}$ *[…]*
a.view beautiful **it's a waste of time → extremely**

‘At the moment, we are in the very best hotel […] an **extremely** beautiful view […]’

(tinyurl.com/yvpxs32e)

e. *sirton* ***[xaval al hazman]***$_{\text{INTENSIFIER}}$ *yafe*
a.clip **it’s a waste of time → extremely** beautiful

knisa *xova:*
an.entrance must

‘An **extremely** beautiful clip, entering (the link) is a must […]: [here comes a link – IB]’

(tinyurl.com/2p8cfc6r)

C.2 *Xaval al hamilim*, originally, ‘It’s a waste of words’

(C2) a. *Leksus RX 350* *jip* *im* *netunim*
Lexus RX 350 a.jeep with specifications

*še-**[xaval al hamilim]***$_{\sim\text{ADJECTIVE}}$.[96]
that-**it’s a waste of words → amazing**

‘Lexus RX 350 – a jeep with **amazing** specifications’

(tinyurl.com/3n9nzbhr)

b. *ha-madrix^*[97] *košer* *šeli* *menašek* *et* *mixal*
the-instructor fitness my kisses ACC Michal

*še-**[xaval al hamilim]*** $_{\sim\text{MANNER ADVERB}}$.
that-**it’s a waste of words → amazing manner**

‘My fitness instructor is kissing Michal in an **amazing manner**’

(tinyurl.com/2p82xu66)

c. *makom* *madhim* *spa* *še-en* *kmo-to*
a.place amazing a.spa that-there’s.no like-it

ba-a-arec *[…]* *yeš* *camud* *l-a-makom*
in-the-country there.is right.next to-the-place

---

[96] The notation ~ indicates that had the *še* ‘that’ been dropped, the member of the Ultimate construction family that follows it would have been reanalyzed as a full-fledged word belonging to the relevant word class.

[97] The notation ^ stands for a construct state.

ulam *anak* *še-**[xaval al hamilim]*** ~INTENSIFIER *[…]*
a.hall huge that-**it's a waste of words → amazingly**

'An amazing place, there's no spa like this in the whole country […], right next door there's an **amazingly** huge hall […]'

(tinyurl.com/3rth4m58)

C.3 *Ba livkot*, originally, 'It feels like crying'

(C3) a. *ma'afim* *me'ulim,* *lexem* *še-**[ba livkot]*** ~ADJECTIVE
pastries excellent bread that-**it feels like crying → amazing**

ve-granola *me-ha-agadot* *mexakim* *laxem* *b-a-makom.*
and-granola from-the-fairy.tales waiting to.you in-the-place

'High-quality pastries, **amazing** bread and legendary granola are waiting for you in this place.'

(tinyurl.com/2hy8jnyx)

b. *[…]* *ve-mixal* *mekateret* *še-**[ba livkot]*** ~ADVERB.
and-Michal complains that-**it feels like crying → so much**

'[…] and Michal complains **so much**'

(tinyurl.com/2p9fnue8)

c. *haya* *mehamem* *še-**[ba livkot]*** ~INTENSIFIER.
it.was stunning that-**it feels like crying → extremely**

'It was **extremely** stunning.'

(tinyurl.com/bdzfp9rhf)

C.4 *Ba lamut*, originally, 'It feels like dying'

(C4) a. *Jumira bič rezidens:* *aruxot* *šel* *koxvej^* *mišlen,*
Jumeirah Beach Residence meals of stars Michelin

xofim *zehubim* *ve-nof* *še-**[ba lamut]*** ~ADJECTIVE.
beaches golden and-a.view that-**it feels like dying → stunning**

'Jumeirah Beach Residence: Michelin-rated meals, golden beaches and a **stunning** view'

(tinyurl.com/2p84d3at)

b. *lexem^* *kemax^* *šipon* *tari* *[…]* *male* *sumsum*
bread flour rye fresh a.lot sesame

mi-lema'ala *ve-ta'im* *še-**[ba lamut]*** ~INTENSIFIER.
from-top and-tasty that-**it feels like dying → extremely**

'Fresh rye bread, lots of sesame seeds on top and **extremely** tasty.'
(tinyurl.com/yppw3adc)

C.5 *Efšar lehištage'a*, originally, 'It's possible to go crazy'

(C5) a. *[…]* *ve-ha-xodeš* *haxi* *yafe* *b-a-šana* *marhiv*
and-the-month most beautiful in-the-year decorates

et *arcenu* *bi-cva'im* *u-fraxim* *ve-rexot*
ACC our.country in-colors and-flowers and-fragrances

*še-**[efšar lehištage'a]*** ~ADJECTIVE *[…]*
that-**it's possible to go crazy → astounding**

'[…] and the most beautiful month of the year decorates our country with **astounding** colors, flowers and fragrances…'
(tinyurl.com/55pu8xyh)

b. *simlat^* *kleopatra* *be-orex* *maksi* *mi-bad* *na'im*
dress Cleopatra in-length maxi from-fabric pleasant

*še-**[efšar lehištage'a]*** ~INTENSIFIER*!*
that-**it's possible to go crazy → extremely**

'A maxi Cleopatra-style dress from an **extremely** soft-to-the-touch fabric!'
(tinyurl.com/3fjce39e)

C.6 *Efšar lamut*, originally, 'It's possible to die'

(C6) a. *[…]* *anaxnu* *mamlicim* *al* *mis'edet^* *ha-gag* *šel*
we recommend on restaurant the-roof of

*malon^* *mamila,* *mušlemet* *le-kvucot* *ktanot,* *trendit*
hotel Mamilla perfect to-groups small trendy

ve-im *nofey^* *ha-ir* *ha-atika*
and-with views the-city the-old

*še-**[efšar lamut]*** ~ADJECTIVE.
that-**it feels like dying → stunning**

'[…] we recommend the restaurant on the roof of the Mamilla hotel, perfect for small visitor groups, trendy and with the **stunning** views of the old city [of Jerusalem – IB].'

(tinyurl.com/3vwyd74p)

b. *be-gil* *45* *(ve-yafa* *še-**[efšar lamut]*** ~INTENSIFIER)
in-age 45 and-beautiful that-**it feels like dying → extremely**

sandi *bar* *xozeret* *ledagmen* *[…]*
Sandy Bar goes.back to.model

'At the age of 45 (and **extremely** beautiful) Sandy Bar goes back to modeling […]'

(tinyurl.com/4cv6unhf)

C.7 *En ma lehagid*, originally, 'There's nothing to say'[98]

(C7) a. *en* *ve-lo* *yihye* *kmo* *ha-ta'am* *ha-ze*
there's.no and-not will.be like the-taste the-this

b-a-olam *naki* *ve-tari* *ve-šerut*
in-the-world clean and-fresh and-service

*še-**[en ma lehagid]*** ~ADJECTIVE *ve-haxi* *xašuv*
that-**there's nothing to say → wonderful** and-most important

ha-be'alim *im* *lev^* *zahav.*
the-owner with heart gold

'There isn't and there will never be anything like this taste in the whole world, clean and fresh and **wonderful** service and most importantly — the owner has a heart of gold.'

(tinyurl.com/nnsj3kp9)

[98] Note that *en ma lehagid* can be alternatively interpreted as 'beyond dispute' or 'incontestable'. Clearly, this interpretation and the intensifying interpretation are related. If a stance-object is of high quality, then it's clear that there's no point arguing about this fact. This alternative interpretation, 'beyond dispute' or 'incontestable', is made salient in contexts where a speaker wishes to eliminate, in advance, any criticism as for the validity of an assertion she is about to make. To this end she may use *en ma lehagid* as an introduction to her assertion.

b. *Šani […] hi metumtemet še-**[en ma lehagid]*** ~INTENSIFIER *[…]*
Shani is stupid that- **there's nothing to say → exceedingly**

'Shani […] is **exceedingly** stupid [..] '

(tinyurl.com/yckw39vp)

C.8 *En ma ledaber*, originally, 'There's nothing to speak'[99]

(C8) a. *ani xayav lehagid ulam mispar 1 b-a-cafon*
I must to.say banquet.hall No. 1 in-the-north

*oxel še-**[en ma ledaber]*** ~ADJECTIVE *icuv šel*
food that-**there's nothing to speak → amazing** design of

ha-beyoker ve-haxi xašuv šerut ve-yaxas […]
the-costly and-most important service and-attitude

'I must admit, it's the No. 1 banquet hall in the north [of Israel – IB], **amazing** food, expensive design, and most importantly – excellent service and attitude […]'

(tinyurl.com/3tm595w3)

b. *hi hayta yalda muclaxat, šketa, yešara, tova*
she was girl fine quiet honest good

va-xaxama ve-yafa – yafa
and-smart and-beautiful – beautiful

*še-**[en ma ledaber]*** ~INTENSIFIER.
that-**there's nothing to speak → extremely**

'She was a fine, quiet, honest, good and smart girl, and beautiful – **extremely** beautiful.'

(tinyurl.com/2p83n24n)

C.9 *En milim*, originally, 'There are no words'

(C9) a. *[…] kol ha-tov ha-ze cofe le-nof*
all the-good the-this watches to-a.view

*še-**[en milim]*** ~ADJECTIVE*! mamlicim be-xom!*
that-**there are no words → wonderful** recommend in-warmth

[99] See fn. 98.

'[…] All this good [an outstanding hotel – IB] looks over a **wonderful** view! Highly recommended!'

(tinyurl.com/2p9fa24s)

b. *[…] ha-mikum lo mašehu aval na'im šam ecel*
the-location not something but pleasant there at

*ha-puštakim ve-ta'im še-**[en milim]*** ~INTENSIFIER***!***
the-Puštakim and-tasty that-**there are no words → extremely**

'[…] The location is not the best, but it's nice there, at Puštakim [a name of a restaurant – IB], and **extremely** tasty!'

(tinyurl.com/3k9d5urv)

C.10 *En dvarim ka'ele/u*, originally, 'There are no such things'

(C10) a. *ani osa dag marokai*
I make fish Moroccan

*še-**[en dvarim ka'ele]*** ~ADJECTIVE.
that-**there are no such things → amazing**

'I prepare an **amazing** Moroccan-style fish dish.'

(tinyurl.com/y9rjr5da)

b. *[…] hem sixaku kaduregel*
they played football

*še-**[en dvarim ka'ele]*** ~MANNER ADVERB *[…]*
that-**there are no such things → amazingly**

'[…] They played football **amazingly** [...] '

(tinyurl.com/2p87ady5)

c. *[…] ze xomer^ kri'a le-anašim kamoni še-san'u*
this material reading to-people like.me that-hated

*et ha-sefer še-**[en dvarim ka'elu]*** ~ADVERB *[…]*
ACC the-book that-**there are no such things → immensely**

'[…] This is relevant reading material for people like me who hated the book **immensely** […]'

(tinyurl.com/yckmc2rm)

d. *ata* *tarxan* *še-**[en dvarim ka'ele]*** ~INTENSIFIER
you tedious.person that-**there are no such things → extremely**

'You are an **extremely** tedious person.'

(tinyurl.com/yc6c3vy4)

Appendix D. Alternative analyses and representations of the data in Figures 4.6 and 4.9

I here present an alternative analysis for the datasets of Figures 4.6 and 4.9 in Chapter 4. The results of the analysis are identical to those presented in Sections 4.3.4 and 4.3.6, respectively.

D.1 An alternative analysis and representation of the data in Figure 4.6

Since the size of the dataset is small and "larger bins hence safe-guard against a disproportionate influence of fluctuation in the data and facilitate the detection of developmental patterns by sensible data aggregation" (Flach, 2021: 259), I applied the Variability-based Neighbor Clustering algorithm to the data (Hilpert, 2013: 32-45). The resulting dendrogram produced four *non*-arbitrary periods of time. Table D1 is the corresponding contingency table. A Fisher test confirms that the distribution shows a statistically significant interaction between stage/context and period ($p = 1.51 \times 10^{-4}$).

	Period				
Stage/Context	**2006-2007**	**2008**	**2009-2012**	**2013-2018**	**Sum**
Stage I	21	91	97	33	242
Context IIa	0	3	5	4	12
Stage III	0	7	25	17	49
Sum	21	101	127	54	303

Table D1: Tokens of stage I, context IIb and stage III *en dvarim ka'ele/u* by time period

This clustering procedure enabled me to draw a Pearson residual-based association plot (see Meyer, Zeileis, & Hornik, 2008 for a detailed description of associoation plots), presented in Figure D1. Each tile in this plot represents data in the respective cell in Table D1. Tiles above the dotted line represent tokens which are observed more than expected. Tiles below the dotted line represent tokens which are observed less than expected. Their colors and shades indicate whether the association between the observed and expected number of tokens is significant (dark blue or red shades), or just a trend (light blue or red shades). The grey color represents no trend. The width of each tile is proportional to the square root of the expected frequency — the higher the frequency the wider the tile (Smirnova, Mailhammer, & Flach, 2019). Tile height is proportional to Pearson residual. Thus, tile area is proportional to the difference between observed and expected frequency.

The relevant part of the association plot which provides support for issue (iv) — the timeline of Anaphoric degree-adverb exclamatives (IIa) and their counterparts lacking the anaphoric degree-adverb (III) — is the left half of the plot, 2006-2009. During this time period, stage I *dvarim ka'ele/u* is observed more than expected while context IIa and stage III *en dvarim ka'ele/u* are observed less than expected. This suggests that

stage I *en dvarim ka'ele/u* showed up chronologically earlier than both context IIa and stage III *en dvarim ka'ele/u*. If we focus on the last two, we can see that context IIa *en dvarim ka'ele/u* is observed less than expected but not significantly so, while stage III *en dvarim ka'ele/u* is also observed less than expected, this time significantly. These results suggest that context IIa *en dvarim ka'ele/u* showed up chronologically earlier than stage III *en dvarim ka'ele/u*. This time evolution of the three stages/contexts may be applied to *xaval al hazman* too.

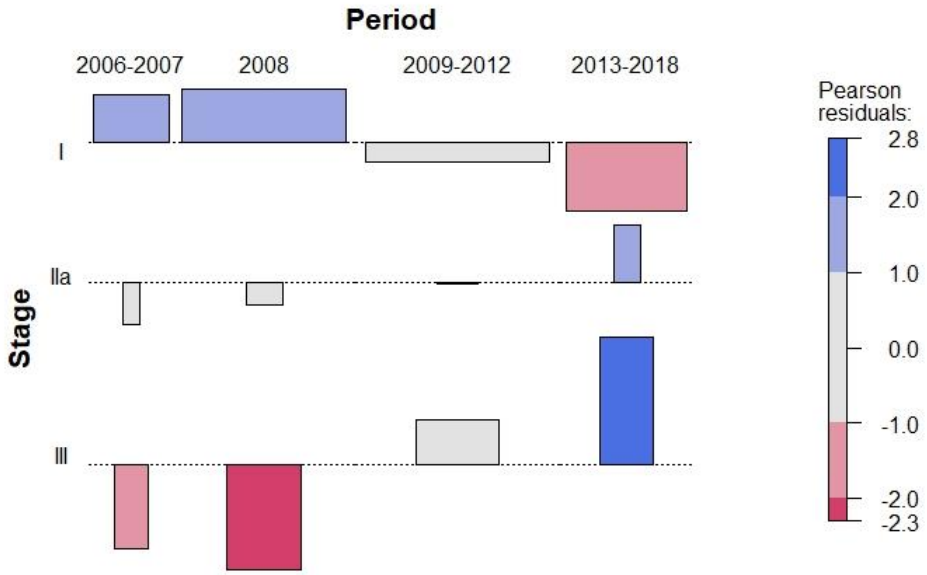

Figure D1: Residual-based association plot suggesting the timeline of evolution of *en dvarim ka'ele/u*: from an independent idiomatic sentence (I), through an idiomatic sentence integrated into the preceding Anaphoric degree-adverb exclamative by means of a Correlative endpoint resultant-state clause (IIa), and then an idiomatic sentence integrated into the preceding, non-exclamative sentence with no degree-adverb, again by means of a Correlative endpoint resultant-state clause (III). Data extracted from IsraBlog corpus.

D.2 An alternative analysis and representation of the data in Figure 4.9

I repeated the entire procedure described in Appendix D.1 for the data presented in Figure 4.9.

Again, the resulting dendrogram produced four *non*-arbitrary periods of time. Table D2 is the corresponding contingency table. A Fisher test confirms that the distribution shows a *non*-significant interaction between category and period ($p = 0.55$).

	Period				
Category	**2008-2012**	**2013**	**2014**	**2015-2018**	**Sum**
Modifiers of nouns	10	1	5	0	16
Modifiers of adjectives	15	4	3	1	23
Modifiers of verbs	7	0	3	0	10
Sum	32	5	11	1	49

Table D2: Tokens of the three categories of modifiers *en dvarim ka'ele/u* by time period

This clustering procedure enabled me to draw a residual-based association plot presented in Figure D2. Not only are most tiles grey, which implies a non-significant association (and may be the result of the scarce dataset), but there does not seem to be any distinct pattern like the diagonal pattern of positive associations in Figure D1 from top left to bottom right, which indicates a diachronic evolution. The associations in Figure D2 appear random, thus implying no distinct line of evolution, and therefore simultaneous inception of the three categories of stage III *en dvarim ka'ele/u*.

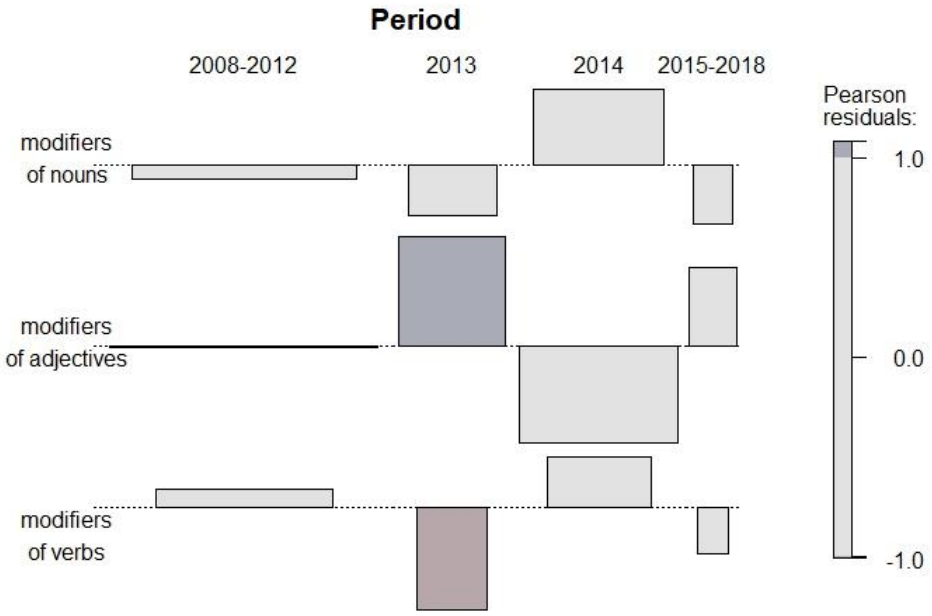

Figure D2: Residual-based association plot suggesting the simultaneous inception of the three categories of stage III *en dvarim ka'ele/u*. Data extracted from IsraBlog corpus.

Appendix E. A more comprehensive analysis of the data pertaining to *en milim* 'there are no words' in Section 4.3.5

In Figure E1 below, I present an analysis parallel to the one I presented in Figure 4.8, where I compared the frequencies of the bare *en milim* and *en milim (befi) leta'er* over the years. Note, however, that I here consider not just *en milim (befi) leta'er*, but a whole set of similar verbs which are interchangeable with *leta'er* 'to describe' — *levate* 'to express', *lehabi'a* 'to convey', *lehasbir* 'to explain', *lehagdir* 'to define', *lehagid* and *lomar* 'to say'. Just like in Figure 4.8, no single time interval shows a statistically significant higher frequency of the bare *en milim* as compared with *en milim (befi) leta'er* and similar verbs. In no case is *en milim* (red bars) favored by speakers more than *en milim (befi) leta'er* and similar verbs (black bars).

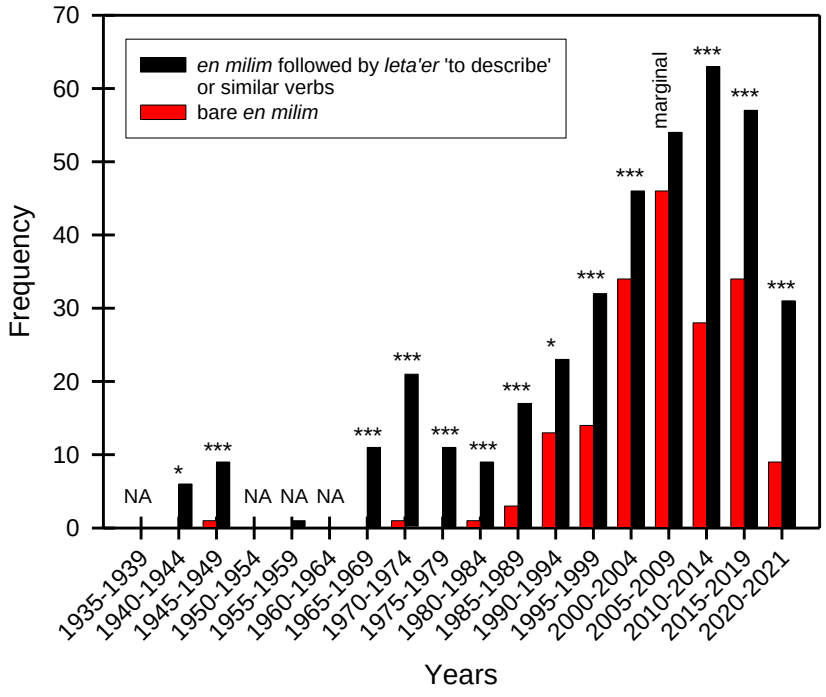

Figure E1: The frequencies of bare *en milim* and *en milim (befi) leta'er* and similar verbs over the years. The level of significance for each five-year interval (which is the outcome of a binomial test) is marked above each pair of bars. '*' = $p < 0.05$, '**' = $p < 0.01$; '***' = $p < 0.001$; 'NA' = Not Applicable. Data extracted from Yedioth Ahronoth corpus.

Appendix F. Spotting optimal innovations

Giora et al. (2004) suggested that it is the amount of formal modification of the familiar syntagma underlying the innovation that determines whether a given instance of innovation is an instance of optimal innovation, or not. Consider the two variants of *body and soul — body and sole* and *Bobby and Saul*. Both variants "preserve the rhythm and the form [i.e., syntactic structure – IB] of the original [expression]" (p. 121). However, of the two variants — *body and sole* and *Bobby and Saul* — only *body and sole*, "allows for a salient response to get through" (p. 126), as it makes no more than a single modification to the familiar underlying *body and soul*. In contrast, *Bobby and Saul* introduces two formal modifications to the familiar underlying *body and soul*, and it is therefore a pure innovation.[100] In short, one criterion for an optimal innovation is recognizability of the substrate. The more changes are made, the less recognizable the substrate is.

The requirement for a single modification (of the innovation) to the underlying familiar syntagma is a necessary condition for the innovation to be considered optimal. But it is not a sufficient condition. Take, for example, *body and souls*, which also introduces only a single modification to *body and soul*. Is *body and souls*, then, also an instance of optimal innovation? Giora et al. ruled out *body and souls* as an instance of optimal innovation, since this modification is trivial, that is, it is of a *quantitative*, rather than a *qualitative*, nature, and therefore produces no meaning contrast between the innovation and the underlying familiar syntagma.

The requirement for a single formal modification and the requirement for a semantic contrast between the surface form and the underlying syntagma are exactly what specifies the optimal innovation a case of paronymy *in absentia*.

In order to extract optimal innovations for *xa'val al haz'man* from the part-of-speech tagged and morphologically annotated HeTenTen corpus, I sought out two kinds of sequences: (i) sequences of any noun or adjective preceding *al haz'man* (roughly, 'on the time'), which potentially replace *xa'val* (roughly, 'it's a waste'); (ii) sequences in which *xa'val al* (roughly, 'it's a waste of') is immediately followed by any noun, potentially taken as replacing *haz'man* (roughly, 'the time').[101] Three types of optimal innovations, classified by their rhyming — which is an element of rhythm — are listed in Table F1 below.

(i) The first query produced 21 syntagmas which can be counted as optimal innovations. All of them were of the type in which a noun was immediately followed by *al haz'man*. Each of these 21 syntagmas was evaluated by an informant versed in

[100] This criterion which applies to phrases has an equivalent in the field of morphology which must be satisfied in order to produce a *contour blend*, such as *Chinglish* 'English with Chinese grammar mixed together' < *Chinese* + *English* (the Urban Dictionary; tinyurl.com/2p8wbr5n) (Ronneberger-Sibold, 2006): The matrix word "can be traced, so to speak, by several phonological features of high importance for its recoverability. These are firstly its overall rhythmical contour defined by its number of syllables and the place of its main stress […]" (p. 170).

[101] The transcription here includes stress because the rhythmical criterion necessarily involves stress.

detecting instances of optimal innovation. The informant was asked to determine whether the syntagmas are indeed cases of optimal innovation (by carefully observing the criteria specified above), and to further determine for each case, which of the two underlying meanings of *xa'val al haz'man* serves as the substrate for innovation.

It should be noted that the surface forms of several instances of optimal innovations are the same. However, these items were produced by different speakers in different contexts. See, for example, the two instances of *xa'lal al haz'man* (*xalal* 'outer space') listed in Table F1: The first example of *xa'lal al haz'man* is the title of a blogpost about a TV series covering the Roswell UFO incident, and the second example is the title of a positive report about an exhibition focusing on outer space.

(ii) The second query produced no hits that observe the criteria (specified above) for innovations to be considered optimal innovations.

Query	Optimal innovations		Broadly-transcribed examples	The rhyme	Gloss	Meaning in context
(i) ***N/ADJ*** ***al haz ˈman***	***N/ADJ*** rhymes with *xa ˈval*	***N/ADJ*** rhymes with *haz ˈman*				
	a **perfect** rhyme		***na ˈmal*** *al haz ˈman*	***na ˈmal —*** *xa ˈval*	***na ˈmal*** ‘harbour’	The slogan of the 80th birthday celebration to Tel Aviv harbor
			xa ˈlal *al haz ˈman*	***xa ˈlal —*** *xa ˈval*	***xa ˈlal*** ‘outer space’	The title of a blogpost praising a TV series covering the Roswell UFO incident
			xa ˈlal *al haz ˈman*	***xa ˈlal —*** *xa ˈval*	***xa ˈlal*** ‘outer space’	The title of a positive report about an exhibition focusing on outer space
	a **general** rhyme		***xa ˈva*** *al haz ˈman*	***xa ˈva —*** *xa ˈval*	***xa ˈva*** ‘farm’	The title of a positive report about an agricultural farm opened for family visits on weekends
			xa ˈlav *al haz ˈman*	***xa ˈlav —*** *xa ˈval*	***xa ˈlav*** ‘milk’	The title of a newspaper article encouraging parents to feed their children dairy products
		a **perfect** rhyme	***a ˈšan*** *al haz ˈman*	***a ˈšan —*** *haz ˈman*	***a ˈšan*** ‘smoke’	The title of an enthusiastic report about a smoky jazz club
			sa ˈtan *al haz ˈman*	***sa ˈtan —*** *haz ˈman*	***sa ˈtan*** ‘Satan’	A movie about the son of Satan who repents of his sins
(ii) ***xa ˈval al ha-*** ***N***	*No Relevant Results*					

Table F1: Examples of optimal innovations based on *xa’val al haz’man* classified by the type of rhyme used. Data extracted from HeTenTen corpus.

Appendix G. Classifying web-sites according to religious observance

I excluded the 142 instances of *xaval al hazman* accompanied by metalinguistic comments (see Section 6.5) from the list of 2954 *xaval al hazman* items (extracted from HeTenTen corpus), ending up with 2812 items. I was able to tag 2762 items for their specific function — 1389 instances of the positive (and intensifying) *xaval al hazman* and 1373 instances of the negative *xaval al hazman*. I failed to tag the remaining 50 items for lack of sufficient context.

From the resulting list of 2762 items, I extracted all 865 unique web addresses. I presented this list of web addresses to a former ultra-orthodox informant (aged 28) who is highly active on the web. He was asked to identify the community associated with each web-site — as either ultra-orthodox, or not — by examining the web addresses one by one, that is, browsing through each web-site. He was not informed of the precise purpose of this classification, nor did he have any access to the 2762 items.

Only 778 web addresses were accessible. Not only did my informant tag each web address as related to (and fed by) ultra-orthodox population (or not), he provided a more refined tagging of the exact religious observance of the religious population associated with each web address, see Table G1 below. Then, each web address was associated with the original list of 2762 items, producing 2373 items tagged for the religious observance of their writers. The analysis reported in Section 6.7.3 includes only items extracted from web-sites associated distinctly with either secular or ultra-orthodox populations.

The same procedure was repeated for *sof haderex* (originally, ‘the end of the road’), *en dvarim ka’ele/u* (originally, ‘there are no such things’) and *ba livkot* (originally, ‘it feels like crying’), presented in Section 6.8.

Table G1: Web-sites classified according the religious observance of their contributors. The full list comprises of several hundred web-sites. The background colors indicate the fine classification provided by the informant. UO = Ultra-Orthodox. Data extracted from HeTenTen corpus.

Further comments about the web-site	Ultra-orthodox or not?	Web-site
ליטאים וחסידים = Hassidic and Litvish UO	כן = yes	pashkevil.co.il
אתר כיכר השבת – בעיקר ליטאים אך יש גם חסידים = Mostly Litvish but also Hassidic UO	כן = yes	kikarhashabat.co.il
ליטאים ; האתר בתחזוקה, אך ארגון אש התורה שייך לחרדים ליטאים = This site is under maintenance but this organization is associated with Litvish UO	כן = yes	aish.co.il
הכותבת חרדית, ככל הנראה ליטאית מודרנית = The writer is UO but most likely a modern Litvish UO woman	כן = yes	yaelzals.co.il
כיכר השבת לנשים דתיות= *kikar hashabbat* for women	כן = yes	mame.co.il
חסידים, ליטאים ומזרחיים - Hassidic, Litvish and Sephardi UO	כן = yes	kolhazman.co.il
חרדים מעט מודרניים, אין שם הבחנה לפי זרמים ועדות, האתר דומה מעט לכיכר השבת. = Somewhat modern UO. Cannot be distinguished by factions	כן = yes	jdn.co.il
יש באתר נוכחות של כלל ארגוני ההחזרה בתשובה החרדיים – מכל הזרמים = An organization encouraging secular Jews to join the UO community	כן = yes	haemet.net
קהילתו של אמנון יצחק – מעין כת חרדית מזרחית עצמאית = The community of Amnon Itzkhak – an independent Sephardi UO community	כן = yes	shofar.tv
בעיקר חרדים מזרחיים = Mostly Sephardi UO	כן = yes	radio2000.co.il

jewish-education.info	כן = yes	חב"ד = Chabad community
col.org.il	כן = yes	חב"ד = Chabad community
besimcha.org	כן = yes	חב"ד = Chabad community
ch10.co.il	כן = yes	בעיקר חב"ד = mostly Chabad community
hageula.com	כן = yes	חב"ד = Chabad community
neshei.com	כן = yes	חב"ד = Chabad community
yardbirdsil.info	חלקית =partially	בעמוד http://yardbirdsil.info/judaism/judaism.htm יש קישורים לדפים שנכתבו על ידי חרדים – לא ברור מאלו זרמים = In this webpage there are links to other webpages written by UO, not clear from which factions exactly.
telechofesh.co.il	חלקית =partially	האתר מיועד לציבור הדתי והחרדי = A religious web-site for UO and national religious
aminadav.org.il	חלקית =partially	אגודה העוסקת בגיוס בנות דתיות וגם חרדים (בנים ובנות) לשירות לאומי = An organization which recruits religious teenagers and also UO to civil service.
rotter.net	חלקית =partially	אתר רוטר הוקם על ידי חרדים אך כיום הוא בעל אוריינטציה דתית-לאומית-ימנית = A web-site established by UO but has recently gained a more nuanced orientation
hakolhayehudi.co.il	חלקית =partially	דתי כללי, יש שם גם חרדים אשכנזים (לא חסידים) ומזרחיים = A web-site where one can find religious Jews of various observance, including Ashkenazi Litvish and Sephardi UO
hagabay.net	לא ברור = not clear	מכיל גם תוכן הקשור בחרדים אך האתר הוא דתי כללי = A web-site which includes context related to UO, but is oriented to any kind of religious Jews
uziyaron.com	לא ידוע = unknown	בלוג סגור למורשים בלבד = A private blog

otiyot-sukaryot.com	לא, אבל דתי לאומי = no, but national religious	דתיים לאומיים = National religious
vbm-torah.org	לא, אבל דתי לאומי = no, but national religious	דתיים לאומיים = National religious
kayama.co.il	לא, אבל דתי לאומי = no, but national religious	הכותבת – הדס דודסון – היא דתייה = The writer is a religious woman
yhy.co.il	לא, אבל דתי לאומי = no, but national religious	אתר של ישיבת הסדר = A web-site of a military Yeshiva
hebron.org.il	לא, אבל דתי לאומי = no, but national religious	יהודים המתגוררים בחברון = Jews living in Hebron
kamoha.org.il	לא, אבל דתי = no, but religious	הומואים דתיים = National religious gays
makshivim.org.il	לא, אבל דתי = no, but religious	דתיים = Religious
laitman.co.il	לא, אבל דתי = no, but religious	הכותב דתי אך ככל הנראה איננו חרדי = the writer is religious but not UO
inn.co.il	לא, אבל דתי = no, but religious	ערוץ 7 = Channel 7, a religious new channel
yba.org.il	לא, אבל דתי = no, but religious	בני עקיבא = A national religious youth group
musaf-shabbat.com	לא, אבל דתי = no, but religious	מוסף מקור ראשון = The Shabbat supplement of a religious daily
agenda.co.il	לא, אבל יש שם גם דתיים = no, but you can find there religious	האתר מכיל פורום בשם דתיות גאות שגם חרדים כותבים בו: https://www.agenda.co.il/266/ = This web-site includes a forum by the name religious gays where some of the participants are UO
halachabrura.org	לא בהכרח, אבל מאוד דתי = not necessarily, but very religious	
nitsanem.com	לא ידוע = unknown	אתר פרטי - A private web-site

chabad770.org	לא ברור, כנראה חסר משמעות = unclear	לא נראה שיש שם הרבה מעבר לכתובת דומיין = Nothing more than the name of a domain
chabad8.org	לא ברור, כנראה חסר משמעות = unclear	לא נראה שיש שם הרבה מעבר לכתובת דומיין = Nothing more than the name of a domain
israelhayom.co.il	לא = no	
headlines.co.il	לא = no	
urbanbridesmag.co.il	לא = no	
lametayel.co.il	לא = no	
hakshev.co.il	לא = no	
ramkol.co.il	לא = no	
wind-surf.co.il	לא = no	

הפקולטה למדעי הרוח ע"ש לסטר וסאלי אנטין
ביה"ס לפילוסופיה, בלשנות ולימודי מדע

תבניות לשוניות "בתנועה": מתבניות משפטיות עצמאיות לתבניות לקסיקליות

חיבור זה הוגש לשם קבלת התואר
"דוקטור לפילוסופיה"

על-ידי
ישראלה בקר

העבודה הוכנה בהדרכת
פרופ' מירה אריאל ופרופ' רחל גיורא

ספטמבר 2023

תקציר

חבל על הזמן הוא משפט שהיה משך שנים ארוכות אמצעי לבטא את הערכתם השלילית (העקיפה) של הדוברים ביחס למושא הערכה מסוים בשיח. כוונת הדוברים הייתה שמושא ההערכה האמור הוא חסר ערך, ולכן העיסוק בו איננו אלא בזבוז זמן (1). והנה, לפני כשלושים שנה, פרץ אל הזירה הלשונית העברית הביטוי *חבל־על־הזמן*. *חבל על הזמן* שינה את משמעותו מן הקצה על הקצה, ואף שינה את המעמד התחבירי שלו. לא עוד משפט עצמאי בלבד, אלא מילה מן־המניין שיכולה לשמש בתפקידים תחביריים שונים — שם־תואר (2א), תואר־פועל (2ב) ומגביר (2ג-ד) — שהמאפיין המשותף להם הוא המסר הרגשי המפליג שהם מבטאים. בדוגמה (2א), *חבל־על־הזמן* משמעותו 'מדהים'; בדוגמה (2ב) משמעותו 'באופן מדהים'; ובדוגמות (2ג) ו־(2ד) – 'באופן קיצוני'.

(1) אחד הסרטים הגרועים שראיתי בתקופה האחרונה!!!! [...] הַקִיצֶר, תישארו בבית, **חבל על הזמן**!!!
(tinyurl.com/38syphbc)

(2) א. אחד הטובים! סרט **חבל־על־הזמן**! מומלץ בחום!
(tinyurl.com/mpw32bjc)

ב. סרט סבבה לגמרי!... סיליאן מרפי משחק **חבל־על־הזמן**! ורייצ'ל מק'אדמס [...] שחקנית מעולה פשוט.
(tinyurl.com/ycyvcnvv)

ג. סרט מצחיק **חבל־על־הזמן**, [...] סרט מפתיע, קליל ומאוד מצחיק. מומלץ בחום.
(tinyurl.com/5ymk7m9n)

ד. כדאי ללכת לסרט [...] אני צחקתי **חבל־על־הזמן**, ובתי בת ה־17 גם היא נהנתה.
(tinyurl.com/2t3jkvp5)

חבל־על־הזמן משך תשומת־לב יוצאת דופן מצד דוברי העברית. זאת, ככל הנראה, משום השינוי הסמנטי הרדיקלי שהוא חָווה, וייתכן גם משום שמדובר במשפט מלא ההופך למילה מן־המניין. זהו שינוי תחבירי פריפריאלי שאיננו רווח בעברית, ובכלל.

דיסרטציה זו בוחנת את התופעה הפריפריאלית הזאת — לקסיקליזציה של משפט מלא לכדי מילה — באמצעות שימוש בנתונים (כמותיים) מִקוֹרפורה כתובים. **שאלת המחקר העיקרית שלי היא: באלו תנאים יעבור משפט מלא לקסיקליזציה לכדי מילה מן־המניין?**

המסגרת התיאורטית שאימצתי לצורך הצעת מודל המספק הסבר חסכני אך ממצה לתופעה זו היא *דקדוק התבניות*. דקדוק התבניות היא תיאוריה שתופשת כל אלמנט לשוני, בכל רמה שהיא — מורפמה, מילה, צירוף, פסוקית ומשפט רב־פסוקיות — כְּתַבְנִית (למשל, Croft 2001; Goldberg 1995; Langacker 1987, 1991). תבנית היא צמד מבנה־משמעות. כל הידע הלשוני של הדוברים כולו הוא אוסף של תבניות המהוות רשת (למשל, Goldberg 2003). כל תבנית קשורה לתבניות אחרות באמצעות 'קשרים' מסוגים שונים, כך שהרשת האמורה היא רשת רב־ממדית (למשל, Diessel 2023; Schmid 2020). הקשרים הללו מעידים על תכונות משותפות לתבניות

הקשורות בקשרים אלה, תכונות פורמליות ו/או תכונות סמנטיות. לא פחות חשובה מכך, ורלוונטית לניסיון שלי לענות על שאלת המחקר העיקרית, היא הנחה נוספת העומדת בבסיס תיאוריה זו. כל התבניות כולן מסודרות על רצף שכולל בתוכו גם את הלקסיקון וגם את התחביר, כלומר הלקסיקון והתחביר אינם ישויות מובחנות זו מזו (Goldberg 2006: 220). מכאן נובע שכל תבנית חופשית לנוע מן הקצה המשפטי — ולכן גם המורכב — של הרצף האמור אל הקצה הפשוט והאידיוסינקרטי, כלומר לחוֹת לקסיקליזציה.

כדי להבין את מהות השינוי הלשוני שחווה *חבל על הזמן*, אני בוחנת את מקומו של *חבל על הזמן* בתוך רשת התבניות הרב־ממדית (המהוות את הידע הלשוני של הדוברים) משתי נקודות מבט המשלימות זו את זו. תחילה אני מתבוננת בשינוי הלשוני מנקודת מבטו של מושא השינוי הלשוני (כאן, *חבל על הזמן*), ואחר־כך אני מתבוננת בשינוי זה מנקודת מבטו של הֶקשר המאפשר את השינוי הלשוני. כל זאת, תוך שאני מדגישה את יחסי הגומלין בין התבניות השונות — הן המשפט החווה שינוי לשוני והן ההקשר המאפשר שינוי זה. (יחסי הגומלין הללו הם חלק מן הקשרים היוצרים את הרשת הרב־ממדית.)

כדי להציג ניתוח איתן מבחינה תיאורטית, יש צורך, כמובן, בהכללה. ולכן אני מראה ש־*חבל על הזמן* איננו יחיד במינו. יש משפטים אחרים (בעברית) שחוו (ועדיין חווים) לקסיקליזציה בדיוק כמוהו. המדובר ב״משפחה״ של משפטים היצוקים בדיוק באותה תבנית שבה יצוק *חבל על הזמן*. חברי המשפחה הזאת, שאני מכנה *משפחת המבנים האולטימטיביים* (על שום המסר הרגשי המפליג שהם מבטאים), קשורים ב׳קשרי הורשה׳ לתבנית משפטית מופשטת יותר שממנה הם יורשים את תכונותיהם. מבחינה צורנית, תבנית זו היא תבנית משפטית שבראשה עומד פרדיקאט (ולא נושא); מבחינה פונקציונלית, תבנית זו מבטאת הערכה ביחס למושא הערכה מסוים בשיח (Kuzar 2012). מלבד *חבל על הזמן*, נמנים, למשל, על משפחה הזאת גם המשפטים *אין דברים כאלה*, *אין מילים*, ו־*בא לבכות*.

אני טוענת שהמשפטים הללו אינם שונים מהותית מצירופים פעליים. מרכיביהם של משפטים אלה רלוונטיים מאד זה לזה. רלוונטיים עד כדי כך שהם יוצרים ׳יחידה הדוקה מבחינה פרשנית׳, ממש כמו צירופים פעליים. וכשם שצירופים פעליים נוטים לשנות את משמעותם ולהפוך אטומים מבחינה סמנטית (בעברית, ובכלל), כך גם חברי משפחת המבנים האולטימטיביים. ולכן אין תימה ש־*חבל על הזמן* ובני משפחתו נוטים להפוך אטומים מבחינה סמנטית, בעודם משפטים עצמאיים. זאת ועוד, מעצם היותם של המשפטים הללו מבעים שמבטאים הערכה מפליגה, הם מהווים יחידות בעלות זיקה ליחידות אחרות, כלומר הם ׳בלתי־שלמים מבחינה סמנטית׳. ולכן, אין זה פלא שהם תרים אחר אלמנטים בני־הערכה, בדרך־כלל בשיח המקדים, כדי לאייכם. כלומר, טבועה בהם מוטיבציה להסתפח אל משפט קודם להם ולאייך אלמנט בר־הערכה. נוסף על כך, משום היותם מונר־מורפיים ולכן גם אטומים מבחינה סמנטית, חברי משפחת המבנים האולטימטיביים מהווים מועמדים מצוינים לאייך מגוון אלמנטים תחביריים. במילים אחרות, הם מהווים מועמדים פוטנציאליים להוות *מאייכים גמישים*.

אני גם מציגה מודל שמאפשר להאיר את תפקידו הקריטי של ההקשר בשינוי הלשוני האמור. עצם העובדה שהתופעה הנחקרת מערבת בתוכה משפט מלא שהופך למילה מן־המניין, מצביעה על הצורך לאמץ מודל של קישור בין פסוקיות לכדי משפט מורכב, כזה המתאר מעבר מרצף פאראטאקטי לרצף היפּוֹטאקטי, ספציפית, כזה המתאר כיצד שתי פסוקיות עצמאיות נקשרות זו לזה, כך שאחת מהן היא משפט עיקרי החולש על השנייה שהופכת משפט משועבד. מודל כזה הוצע על־ידי Lehmann (1988). אבל המודל של Lehmann התמקד בנומינליזציה של פסוקיות משועבדות *שאיננה* כרוכה בהכרח בשינוי סמנטי. נוסף על כך, המילה הנוצרת איננה כוללת בתוכה את כל חלקי המשפט המלא, אלא *רק* שבר ממנו. לעומת זאת, המודל שאני מציעה מספק הסבר גם לשינוי הסמנטי שחווה משפט מלא — שינוי סמנטי שהוא קריטי להמשך תהליך הלקסיקליזציה — וגם לשינוי המעמד התחבירי של המשפט המלא על כל חלקיו.

ההקשר המאפשר את השינוי הסמנטי של משפחת המבנים האולטימטיביים הוא כמובן כל הֶקשר שמבטא עוצמה רגשית גבוהה. אבל מבין סוגי ההקשרים שעונים על הקריטריון הזה, רק תבנית משפטית אחת — משפטי קריאה — היא כזאת שמאפשרת את שינוי המעמד התחבירי של המשפט המלא (כלומר, משפט הנמנה על משפחת המבנים האולטימטיביים). אך אין מדובר בכל משפטי הקריאה, אלא במשפטי קריאה הכוללים מעצים שהוא כינוי רמז מְדָרג־מפליג — *כזה/כזאת/כאלה* ו־*כל־כך*, כמו בדוגמה (3). משפטים אלו, מעצם הווייתם, מאפשרים קישור

פסוקיות לכדי משפט מורכב. זאת, משום שמעצימים שהם כינויי רמז מדרגים־מפליגים אינם נוטים לחזור על עצמם כדי לפצות על האובדן הטבעי של עוצמתם הרגשית (כמו *מאד מאד* במקום *מאד*, או *ממש ממש* במקום *ממש*). האסטרטגיה היחידה שלהם לפצות על אובדן זה היא המשך בדמות משפט תוצאה משועבד שגם הוא בעל מסר רגשי מפליג, כמו בדוגמה (4).

(3) זה סרט **כזה** מצחיק.

(https://tinyurl.com/59eu6uyy)

(4) זה סרט **כזה** מצחיק, (עד) ש**חבל על הזמן**.

חברי משפחת המבנים האולטימטיביים מתמקמים בנוחות בתוך משבצת משפט התוצאה המשועבד (שהוא בעל מסר רגשי מפליג), משום שלמרות השינוי הסמנטי שהם חוו, הם עדיין נתפשים מבחינה תחבירית כמשפטים. זהו קשר מסוג 'מילוי־משבצת'. כך אני מספקת מוטיבציה פרגמטית ליצירת המשפט המורכב.

נוסף על שינוי המעמד התחבירי, משפטי הקריאה הללו מאפשרים גם להסביר את הבשלתם של חברי משפחת המבנים האולטימטיביים לכדי מאייכים גמישים. משפטי הקריאה הכוללים מעצים שהוא כינוי רמז מדרג־מפליג יצוקים בתבנית משפטית שבראשה עומד נושא (ולא פרדיקאט). חלקם הם משפטים שמניים שבהם הפוקוס של המשפט הוא שם־עצם או שם־תואר. אחרים הם משפטים פעליים שהפוקוס שלהם הוא פועל. משפט התוצאה המשועבד שהוא בעל מסר רגשי מפליג מאייך, בעצם, את הפוקוס של משפטי הקריאה הללו, יהא אשר יהא פוקוס זה. ולכן התפקיד התחבירי של חברי משפחת המבנים האולטימטיביים מוכתב על־ידי האלמנט המאוייך — שם־עצם, שם־תואר או פועל. אין גם לשכוח, כמובן, שבחברי משפחת המבנים האולטימטיביים עצמם חבוי הפוטנציאל להפוך מאייכים גמישים משום שהם מונו־מורפמיים ואטומים סמנטית. לאחר שחברי משפחת המבנים האולטימטיביים חוברים למשפט הקריאה, כמו בדוגמה (4), אין עוד צורך בכינוי הרמז המדרג־מפליג (כאן, *כזה*) שעוצמתו קהתה ממילא, והוא מושמט. כמו כן, מושמט גם סמן השעבוד ש, כך שבסופו של דבר חברי משפחת המבנים האולטימטיביים נתפשׂים כמאייכים ישירות את הפוקוס של משפט הקריאה, כמילים מן־המניין.

גם העובדה שחברי משפחת המבנים האולטימטיביים אינם חווים לקסיקליזציה באותה מידה זוכה להסבר במסגרת המודל שאני מציעה. חלקם מקיימים יחסי גומלין עם תבניות אחרות — ספציפית 'קשרים אופקיים' בין אַלו־משפטים — באותה רשת רב־ממדית. קשרים אלו משרים תחרות בין האַלו־משפטים ומגבילים את נטייתם של חלק מחברי משפחת המבנים האולטימטיביים לחבור למשפטי הקריאה הרלוונטיים ולחוות שינוי במעמדם התחבירי.

בהמשך אני גם מראה שההקשר הספציפי האחראי לשינוי התחבירי — קרי, משפטי הקריאה — חייב להעלם (או לפחות להפוך אופציונלי) כדי לאפשר שינויים תבניתיים נוספים למי מבין חברי משפחת המבנים האולטימטיביים שהפך למילה מן־המניין (למשל, גזירה ונטייה במקרה של שמות־תואר). במידה שהֶקשר זה איננו נעלם, לא יתרחשו שינויים תבניתיים נוספים.

לא זו בלבד שהמודל שאני מציעה מספק הסבר חסכני וממצה *לכל* הפנים השונים של תהליך הלקסיקליזציה האמור תוך הנחת רשת רב־ממדית של תבניות הקשורות זו לזו בקשרים מסוגים שונים, הוא גם מספק תמיכה לטענה (של דקדוק התבניות) שהתחביר והלקסיקון אינם ישויות נפרדות לחלוטין, אלא אוסף תבניות ברמות שונות של מורכבות השייכות לאותה רמה של ייצוג לשוני. ולכן, אין זה פלא שתבנית מורכבת יכולה להפוך באופן טבעי ביותר לתבנית פשוטה יותר.

מלבד הניסיון לענות על שאלת המחקר העיקרית שלי — באלו תנאים יהפוך משפט מלא לכדי מילה מן־המניין? — נאלצתי בתחילת המחקר שלי להתמודד עם העדר קורפוס לשוני גדול, דָבוּר, דיאכרוני ונגיש, שהוא, כידוע, חיוני למחקר העוסק בשינוי לשוני. כדי לעקוף בעיה משמעותית זו, לפחות באשר לביסוס העובדה שחברי משפחת המבנים האולטימטיביים אכן חוו שינוי סמנטי, נאלצתי להציע שיטות חלופיות לזיהוי שינוי סמנטי. **שאלת המחקר המשנית שלי היא איפוא: כיצד אפשר לבסס את קיומו של שינוי סמנטי בהעדר קורפוס דיאכרוני?**

אני מראה שבחינת הפעילות המטא־לשונית של דוברים, בין אם היא גלויה או סמויה, מאפשרת לבסס את קיומו של שינוי סמנטי. השיטות שאני מציגה מסתמכות על בחינת הפן הקוגניטיבי והפן הסוציופרגמטי (Schmid 2016 (2011)) בפעילותם המטא־לשונית של דוברים החשים בשינוי סמנטי.

הפן הקוגניטיבי מסתמך על שלוש תיאוריות פסיכובלשניות: *תיאורית הבולטות המדורגת* (Giora 1997, 2003), *תיאורית סימון רב משמעות* (Givoni 2020; Givoni et al. 2011) ו-*תיאורית החידוש האופטימלי* (Giora et al. 2004). אני מראה שדוברים שרגישים לשינוי סמנטי מסמנים אותו במפורש. תוכן הסימונים מעיד על כיוון השינוי הסמנטי ועל הבולטות היחסית של המשמעויות השונות. אני מראה גם שאפשר לתמוך בממצאים ביחס לבולטות היחסית של המשמעויות השונות באמצעות בחינת משחקי מילים שיוצרים הדוברים באופן ספונטני.

הפן הסוציופרגמטי מסתמך על ההנחה שדוברים מסוימים, שרגישים לשינוי סמנטי שהתרחש, נמנעים במודע, מתוך שיקולי שמרנות, משימוש במשמעות החדשה שנתפשת בעיניהם כערעור על זהותם החברתית הייחודית. ספציפית, אני משווה את המוכנות של דוברים מן הקהילה החרדית, הידועה בשמרנותה הלשונית, לאמץ חידושים לשוניים, בהשוואה לדוברים מן הקהילה הכללית של דוברי העברית. אני מראה שההבדל בין שתי הקהילות מעיד על שינוי סמנטי.

למעשה, גם החלק הזה של המחקר, מלבד המטרה המוצהרת שלו למלא חסר מתודולוגי, מספק תמיכה לכך שאין חיץ ברור בין התחביר לבין הלקסיקון, לפחות מן הבחינה הסמנטית. זאת, משום שפעילות הדוברים מתמקדת אך ורק בפן הסמנטי של השינוי הלשוני, והם אינם מתייחסים כלל לפן התחבירי שלו.

לסיכום, המטרות של שני חלקי הדיסרטציה שונות זו מזו, ולכן גם שאלות המחקר הנשאלות שונות זו מזו. אבל מה שעולה מתוך המודל המסביר את תהליך הלקסיקליזציה שחווה משפט מלא לכדי מילה מן־המניין, כמו גם מתוך רגישותם המפורשת של דוברים לפן הסמנטי בלבד ולא התחבירי של תהליך זה, הוא שהלקסיקון והתחביר אינם ישויות נפרדות, אלא אוסף תבניות ברמות שונות של מורכבות, אבל, ככל הנראה, באותה רמה של ייצוג לשוני.