

Beyond Regularization:
The Role of Native Learners in
Shaping Modern Hebrew
Relative Clauses

Aviad Eilam, University of Pennsylvania
GALANA III Sept. 4-6, 2008

Outline

2

- General background
- The genesis of Modern Hebrew
- Resumptive pronouns in Yiddish and present-day Hebrew
- Resumptive pronouns throughout Modern Hebrew: Quantitative analysis
- Theoretical implications
- Conclusions and open questions

General Background

3

- When faced with inconsistent linguistic input, children have a tendency to impose regularity on the data rather than to simply reproduce it (Hudson Kam & Newport 2005).
- Some research on ASL (Singleton & Newport 2004) , and artificial languages in an experimental setting (Hudson Kam & Newport 2005, Wonnacott & Newport 2005).
- Sparse data available from real situations of creole formation.
- Modern Hebrew (MH) emerged in a creole-like situation, and so might provide the type of data we're looking for.

The Genesis of Modern Hebrew

4

- Like creole languages, MH emerged when a group of adults shifted to speaking a language which they had not natively acquired.
- We thus predict that they produced the type of irregular variation relevant to this line of research.
- Unlike creole situations, the language they attempted to speak had been in existence previously, but was preserved only as a written language (Chomsky 1957, Izre'el 2003, Zuckermann 2006, a.o.).

The Genesis of Modern Hebrew

5

- The founding population of MH consisted primarily of native speakers of Yiddish, which thus heavily influenced their Hebrew and whose impact on the language is pervasive to this day (Zuckermann 2006, Eilam 2008).
- But both the language of the first MH speakers and present-day MH are not simply a replica of Yiddish (pace the relexification hypothesis of Horvath & Wexler 1994, 1997, Wexler 1990, 2002):
 - Lexical items and morphology
 - Structural distinctions

The Genesis of Modern Hebrew

6

- Perhaps the first speakers' language can be best described as an interlanguage (Selinker 1972), i.e., an L2 learner grammar consisting of elements from Classical Hebrew and Yiddish, but also involving changes which are internal to it and regulated by various factors.
- Examine one grammatical property distinguishing Yiddish from present-day MH: **resumptive pronouns in relative clauses.**

Resumptive Pronouns in Yiddish and Present-day MH Relative Clauses

7

□ Yiddish (Lowenstamm 1977, Jacobs 2005):

- RCs can be formed with either a relative pronoun, inflected for number, gender and case (*vel/x* or *vemen*), or an uninflected complementizer (*vos*).
- In the latter strategy, resumptive pronouns are obligatory in oblique and genitive positions:
 - 1) a šmole kladke vos oyf *(ir) me hot koym gekent geyn
a narrow footbridge that on it one AUX barely able walk
'a narrow footbridge one could barely walk on'
 - Resumptive pronouns are optional in the direct object and subject position, including the highest subject:
 - 2) der dokter vos ix ken (im) iz haynt ongekomen.
the doctor that I know (him) is today arrived
'The doctor that I know arrived today.'
 - 3) der yid vos (er) geyt, iz mayn xaver.
the man that (he) walks is my friend
'The man who walks is my friend.'

Resumptive Pronouns in Yiddish and Present-day MH Relative Clauses

8

- (3) violates the robust crosslinguistic generalization known as the Highest Subject Restriction (McCloskey 1990, 2006).
- Attempts to derive the Highest Subject Restriction in the generative literature primarily appeal to properties of the complementizer (e.g., Otsuka 2006, Suñer 1998) or to binding theory (McCloskey 1990, 2006).
- **The Highest Subject Restriction is a statistical generalization, not a universal statement.** It is violated both by some adult grammars, and among children: Günzberg-Kerbel, Shvimer & Friedmann (in press), report 4.4% subject resumptives among 3;7-5;0 year old MH speakers (see also Finer 1992).

Resumptive Pronouns in Yiddish and Present-day MH Relative Clauses

9

- **Present-day MH (Doron 1982, Borer 1984, Shlonsky 1992):**
 - RCs are formed with an uninflected complementizer (še- or ašer).
 - Resumptive pronouns are obligatory in oblique and genitive positions:
- 4) ha-iš še-xašavti al-*(av)
the-man that-I.thought about-him
'the man that I thought about'
- Resumptive pronouns are optional in the direct object position and the non-highest subject position, and banned in the highest subject position:
- 5) ha-iš še-ra'iti (oto)
the-man that-I.saw (him)
'the man that I saw'
- 6) ha-iš še-rak al kesef (hu) xošev
the-man that-only about money (he) thinks
'the man who only thinks about money'
- 7) **ha-iš še-(*hu) ohev et Rina**
the-man that-(*he) loves ACC Rina
'the man who loves Rina'

Resumptive Pronouns in Yiddish and Present-day MH Relative Clauses

10

- **Nonrestrictive RCs** pattern the same in categorical terms:
- 8) dani, še-(*hu) hitkonen la-meruc bemešex šana, siyem axaron.
Danny that-he prepared to.the-race for year finished last
'Danny, who prepared for the race for a year, finished last.'
- In fact, the Highest Subject Restriction seems to hold equally of nonrestrictive RCs across languages, a fact which has gone unnoticed in the literature until now.
 - In Northern Italian dialects, one finds resumptive clitics for the object in nonrestrictive relatives, but not for the subject (Cecilia Poletto, p.c.).
 - In Greek nonrestrictive relatives, resumption is obligatory in all positions, but barred in the highest subject position (Alexopoulou 2006).
 - In Palestinian Arabic restrictive relatives, resumptive pronouns are obligatory in all positions but the highest subject, where they are disallowed (Shlonsky 1992). The same is true of nonrestrictives.

Quantitative Findings

11

□ The Corpora

- **GEN1:** 3 native Yiddish speakers, who were born in Eastern Europe and immigrated to Palestine in the late 19th/early 20th century. Formal written texts and written speeches (the “average formal style” of Blanc 1968) available online at the Project Ben-Yehuda website (<http://benyehuda.org/>).
- **GEN2:** 4 native MH speakers born to native Yiddish-speaking parents in Palestine in the late 19th/early 20th century. Same type of linguistic data as GEN1.
- **GEN3:** Grammaticality judgments of present-day speakers & a random sample of 555 RCs from the Hebrew Treebank Version 2.0 (<http://www.mila.cs.technion.ac.il>). The database is similar to those used for GEN1 and GEN2. Reasonable to assume that the authors of these texts are one generation removed from GEN2.

Quantitative Findings

12

□ GEN1: Subject RCs

Gap/RP Position	# tokens	Percent
Subject	587	56.5%
Object	170	16.4%
Preposition	222	21.4%
Possessor	60	5.8%
Total	1039	100%

Relativization Positions

Restrictiveness	# tokens	Percent
Restrictive	378	64.4%
Nonrestrictive	209	35.6%
Total	587	100%

Subject Relatives

Dependency type	# tokens	Percent
Gap	376	99.5%
RP	2	0.5%
Total	378	100%

Restrictives

Dependency type	# tokens	Percent
Gap	180	86.1%
RP	29	13.9%
	(18)	(8.6%)
Total	209	100%

Nonrestrictives

- The 2 tokens of resumptives in restrictive RCs are embedded subjects.
- **We find 18/209 (8.6%) unambiguous resumptives in the highest subject position in nonrestrictive RCs.**

Quantitative Findings

13

□ GEN1: Subject RCs

➤ Examples of highest subject resumptives:

- 9) “nifgašti gam im cemax, še-**hu** amad az beroš agudat ‘habima’”
I.met also with Tsemakh that-he stood then at.head society Habima
‘I also met with Tsemakh, who headed the ‘Habima’ Society at the time’
(Bialik: http://benyehuda.org/bialik/dvarim_shebeal_peh54.html)
- 10) “ve-lešon ha-kodeš šel ha-xalfan, še-**hu** haya yehudi mizraxi”
and-tongue the-sacred of the-money.changer that-he was jew Oriental
‘and the holy tongue [Hebrew] of the money changer, who was an Oriental Jew’
(Ben-Yehuda: http://benyehuda.org/by/haidan_harishon.html)
- 11) “hamonim yehudiyim, še-**hem** ya'asu ve-hem yeyacru tarbut klalit”
masses Jewish that-they will.produce and-they will.create culture general
‘Jewish masses, who will produce and create a general culture’
(Brenner: http://benyehuda.org/brenner/lama_ragzu.html)

Quantitative Findings

14

□ GEN2: Subject RCs

Gap/RP Position	# tokens	Percent
Subject	303	57.4%
Object	92	17.4%
Preposition	96	18.2%
Possessor	37	7%
Total	528	100%

Relativization Positions

Restrictiveness	# tokens	Percent
Restrictive	208	68.6%
Nonrestrictive	95	31.4%
Total	303	100%

Subject Relatives

Dependency type	# tokens	Percent
Gap	206	99%
RP	2	1%
Total	208	100%

Restrictives

Dependency type	# tokens	Percent
Gap	95	100%
RP	0	0%
Total	95	100%

Nonrestrictives

- One of the resumptives in a restrictive RC is an embedded subject, and the other appears in a kind-sentence, which is also possible in English (Prince 1997).
- **There are no highest subject resumptives in restrictive or nonrestrictive RCs.**

Quantitative Findings

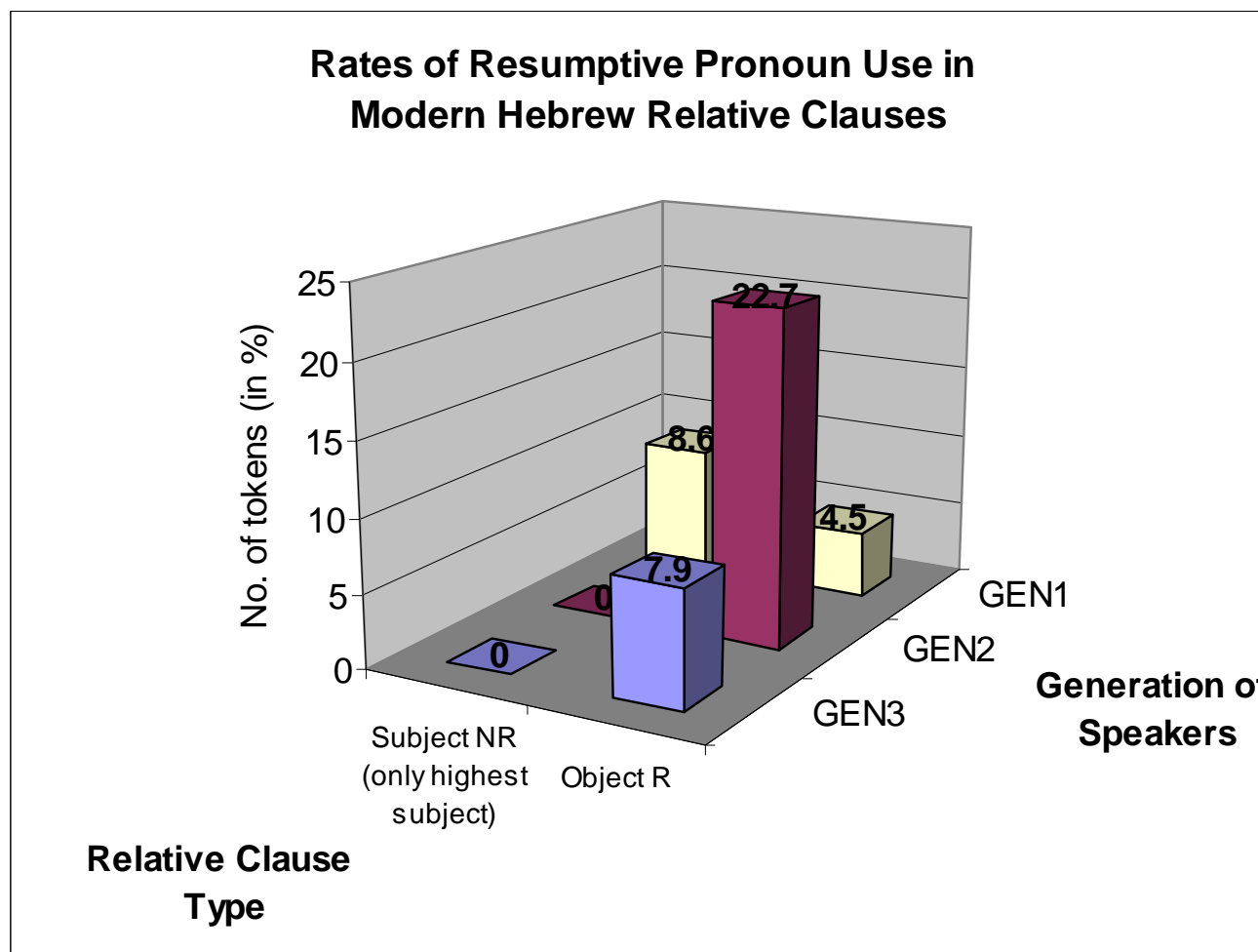
15

□ **GEN3: Subject RCs**

- There are no resumptives in the highest subject position, regardless of whether the RC is restrictive or nonrestrictive. They are also judged to be ungrammatical by present-day speakers (cf. 7-8).

Quantitative Findings

16



Theoretical Implications

17

- To what extent were the RCs of MH modeled after Classical Hebrew and Yiddish?
 - Hebrew basis, with varying degrees of Yiddish interference.
 - Not identical to the patterns of Classical Hebrew: (1) Classical Hebrew freely allows bare relatives, while MH bans them from subject position, and (2) subject resumptive pronouns do not seem to be part of the grammar of Classical Hebrew (Peretz 1967, Holmstedt 2002).
 - Pace Belikova (2007), not wholesale transfer of Yiddish patterns: (1) MH does not have an unambiguous relative pronoun, (2) MH exhibits RC types which are absent from Yiddish but found in Classical Hebrew.

Theoretical Implications

- Why were highest subject resumptives lost?
 - This type of structural change in the grammar of MH could only have been introduced by GEN2 speakers during the critical period (Newmeyer 1998).
 - Highest subject resumptives are a marked grammatical option, as reflected both in acquisition and in their typological status (i.e., the Highest Subject Restriction).
 - In *typical* cases of competition between an unmarked option and a marked alternative found in the adult grammar, children eventually converge on the target adult grammar (Kroch 2005).
 - The first generation of native MH speakers acquired their language in an *atypical* environment, so that we might predict the acquisition of an unmarked non-target grammatical option to be possible.

Theoretical Implications

19

- Why were highest subject resumptives lost?
 - Two specific questions remain:
 1. What is the underlying cause of the marked status of highest subject resumptives?
 2. What exactly allowed the unmarked grammatical option to surface? In other words, what distinguishes the input to first generation MH speakers from that to Yiddish speakers, given that only the latter use highest subject resumptives?

Theoretical Implications

- What allowed the unmarked option to surface?
 - Input frequency: 8.6% of subject nonrestrictive RCs
 - What would be a minimum threshold of frequency below which some grammatical option is not acquired?
 - Children do acquire constructions which are rare in their input: English subject-auxiliary inversion in non-question contexts, Hebrew expletive negation (4 tokens in a corpus containing 341,526 tokens and 16,030 types; Eilam 2007).
 - Object resumptives were found in 4.5% of restrictive RCs, and yet they are grammatical in present-day Hebrew.
 - Resumptives are apparently not very common in Yiddish either: Prince (1990) reports 28/164 (17.1%) resumptives; 24/151 (15.9%) in restrictive RCs and 4/13 (30.8%) in nonrestrictives.

Theoretical Implications

21

- What allowed the unmarked option to surface?
 - The sociolinguistic status of GEN1 (Bar-Adon 1977)
 - In a typical environment for acquisition, parents fulfill a monitoring function, even without any explicit corrective action (Versteegh 1993).
 - MH was characterized by an entire generation whose parents were not native speakers of the language.
 - As in dialect contact situations (Kerswill 1996), GEN2 speakers turned to their peers as primary linguistic role models. The lack of powerful adult norms facilitated a change in which a developmental form failed to be suppressed.

Conclusions and Future Work

22

- Much of the research on the genesis of MH revolves around the question of its genetic classification (e.g., Wexler 1990, Zuckermann 2006).
 - Deciding if MH is Semitic or not will tell us very little about the cognitive mechanisms and social processes which played a role in its formation.
 - Usually based on a small number of recurring examples, without any quantitative, in-depth analysis of the phenomena in question.
- I suggest an alternative perspective, which can provide greater insight into broader, crosslinguistic questions in the domain of language acquisition.
- General learning mechanisms, which lead to regularization of probabilistic input, are only one facet of the multidimensional picture of acquisition in an atypical environment. This study shed light on other factors which may guide children in such situations.

Conclusions and Future Work

23

□ **Future work:**

- Obtain measures of frequency in a Yiddish corpus.
- Examine other possible case studies in the diachrony of MH.
- Apparent time study among different generations.
- Why do certain languages allow highest subject resumptives?

Thanks to Tony Kroch and Charles Yang for their input,
and

THANK YOU!