

**SUBJECT PRO-DROP IN ISRAELI HEBREW:
MORPHOSYNTACTIC/PRAGMATIC VARIATION**

Similar to verbal systems of other Semitic languages which maintain the prefix-conjugation/suffix-conjugation dichotomy, Israeli Hebrew verbs vary in the extent to which they permit the omission of an overt subject. Non-variationist accounts of Israeli Hebrew pro-drop have come up with the following descriptive generalizations:

- (1) In the past and future tenses, pro-drop is optional in the 1st and 2nd persons.
- (2) In the past and future tenses, an overt subject is obligatory in the 3rd person.
- (3) In the present tense, an overt subject is obligatory in all three persons.

These generalizations appear in works such as Borer 1986 and 1989, Vainikka & Levy 1995 and 1999, and Levy & Vainikka 1999-2000. Berman 1990 offers a more “flexible” account, providing a scale ranging from “most obligatory” to “most optional” contexts for subject pro-drop. Berman takes into account not only formal morphosyntactic features such as tense, person, number and gender, but also such discourse-related criteria as “speech-event, situational”, “subordinate clauses, anaphoric”, and “adjacent utterances, narrow topic”, to name a few.

Some examples:

- (4) axal-ti
ate I
- (4') ani axal-ti
I ate I
- (5) axal-t- ∅
ate you-F
- (5') at-∅ axal-t- ∅
you-F ate you-F
- (6) hem/*∅ axl-u
they ate they
- (7) ani/at-∅/hi/*∅ oxel- et
I / you-F/she/AMBIG eats/is eating F

It is acknowledged in the literature that the future tense behaves somewhat differently in *spoken* Hebrew, namely, that there is more ambiguity, presumably leading to a decrease in null subject usage. In all varieties of Hebrew, written/formal and spoken/casual, forms with a *t-* prefix can be either 2nd person masculine singular or 3rd person feminine singular:

- (8) t-oxal
'you-M/she will eat'.
 (8') at-a t-oxal
you-M (you-M) will eat
 (8'') hi t-oxal
she (she) will eat

More and more do speakers of Hebrew use the prefix conjugation not only for declarative future marking, but also as a new imperative form. The disambiguating pronoun does not typically precede such verbs:

- (8''') Ø/*ata toxal!
'eat-M-Sg (already)!'

Another problem the prefix conjugation poses has to do with the 1st person future having been historically marked by a prefixed glottal stop, whose place in the phonemic inventory of contemporary Israeli Hebrew is questionable. In written texts, it is still apparent, as spelling is rather archaic and maintains the grapheme א. However, in casual speech it is often realized as a phonetic zero. Since the independent 1st person singular pronoun *ani* ends with a high front vowel, vowel initial verbs following it are likely to be realized with an intervening palatal glide. A palatal glide is also the prefix that marks 3rd person singular in future verbs, causing the two forms to merge.

- (9) ?ani ?-oxal → ani Ø-oxal → ani^yoxal → ani yoxal
I I will eat
 (10) hu y-oxal
he he will eat

What the phonological process in (9) suggests is that pronoun realization in at least the 1st person singular future preceded the merger of the two prefixes. Given that other prefixed forms still allow pro-drop (e.g., *n-oxal* 'we will eat'), it is hard to accept the syntacticians' argument, that "[a]s expected, the result is that null-subject in 1st-sg. Future tense in this dialect is not possible" (Borer 1989:94). This chicken-and-egg problem seems somewhat more straightforward than its ancient evolutionary predecessor.

For the current study, data were extracted from three sources:

1. Newspaper articles: a news item and an op-ed from the online edition of the daily *Haaretz* (<http://www.haaretz.co.il>). Both articles are dated April 18, 2003.
2. An excerpt from a recording transcribed in Izre'el 2002, consisting of a three-way conversation between three Israelis, all native speakers of Israeli Hebrew, as they were driving home from a wedding celebration. The recording was made on September 4, 2000 as part of the pilot stage of the compilation of the Corpus of Spoken Israeli Hebrew CoSIH).
3. Three online postings on Israeli Internet bulletin boards on the LGBT web site <http://www.gogay.co.il>: one from the general forum, one from the “girltalk” forum and the third from the “bears” forum. They were posted between April 15-17, 2003.

These three sources were chosen, hoping that they would yield grammatical forms more representative of various registers of contemporary Hebrew, which would be less monolithic than described and analyzed in most of the theoretical literature. The notion of linguistic variation being inherent to language itself is not only adopted here for the sake of preferring one linguistic approach over others. Variable rules in a grammar can have theoretical implications. Borer used the supposedly categorical pro-drop licensing conditions in Hebrew to make a point about Agreement being a feature of anaphora under Control Theory. Vainikka & Levy's work suggests that Hebrew (and Finnish) dropped subjects stem from a different node in the syntactic tree than overt pronouns and full NPs. They extend their analysis to account for learnability of the two structures in first language acquisition. While neither syntactic universals nor acquisition are at the focus of the current study, its results may offer a new angle to tackle them.

Following work by Dimitriadis (1995, 1996) and Prince (1999), I offer a multivariate analysis of pro-drop in Israeli Hebrew. 234 tokens were extracted from the sources mentioned above. They were coded for 13 factors:

1. Dependent variable: Subject

Full NP, pronoun, null, impersonal, clause, right dislocation, wh- word.

2a. Independent variable for style: Text

News article, op-ed, GoGay general forum, bears forum, girltalk (lesbian) forum, car dialogue.

2b. Independent variables for morphosyntax:

Person: 1st, 2nd, 3rd

Is person explicitly marked in the morphology? Yes, no, in between.

Gender: Feminine, masculine, neuter (for 1st person and plural forms).

Number: Singular, plural

Verb form: Suffixed (“past”), prefixed (“future”), participle/nominal (“present”), ambiguous.

Word order: Subject-initial, verb-initial, neutralized (subject dropped), V2, quote-initial.

“Positiveness”: Positive, negative, interrogative, imperative, rhetorical negative question.

Subordination/embedding: relative clause, complement, conjunctive/coordinate, conditional, none.

2c. Independent variables: Centering¹

Subj(U_i)=Cp(U_{i-1}): Yes, no.

Subj(U_i)=Cb(U_i): Yes, no, no Cb

Transition type: Continue, retain, smooth shift, rough shift, no prior Cb.

The data were run in the multivariate analysis program Varbrul in two ways. First, with null pronouns constituting an application of the pro-drop rule and any other variant of the dependent variable constituting a non-application. The second run only took into account overt pronouns (but not, e.g., full NPs) as non-applications of the rule. To eliminate knockouts, several factors were recoded or ignored. Most important were the following changes:

- Relative clauses were excluded (null pronouns occurred in 78% of relative clauses when full NPs were included; 91% when they were excluded).
- The word order factor group was eliminated.
- All three LGBT bulletin board texts were recoded as one text-type.
- Zero Cb was recoded as “no”.
- Transition with zero Cb was subsequently recoded as rough shift.

As the tables below show, both 3rd person verbs and participles (“present tense” verbs) are subject to prop-drop, contrary to the generalizations made by Borer and Vainikka & Levy. The results reflected here are for all utterances except relative clauses (where pro-drop is rampant).

¹ Centering Theory (Grosz, Joshi and Weinstein 1995) is briefly introduced on the PowerPoint slides.

First, for all subject types:

Group		Apps	Non- apps	Total	%
Person					
3	N	17	87	104	53
	%	16	84		
1	N	41	23	64	33
	%	64	36		
2	N	13	15	28	14
	%	46	54		
Total	N	71	125	196	
	%	36	64		
Verb form (tense)					
Pref	N	26	28	54	28
(FUT)	%	48	52		
Partic	N	24	32	56	29
(PRES)	%	43	57		
Suff	N	21	65	86	44
(PAST)	%	24	76		
Total	N	71	125	196	
	%	36	64		

And for null and pronominal subjects only:

Group		null	overt pro	Total	%
Person					
3	N	17	15	32	26
	%	53	47		
1	N	41	22	63	51
	%	65	35		
2	N	13	15	28	23
	%	46	54		
Total	N	71	52	123	
	%	58	42		

Verb form (tense)					
Pref	N	26	13	39	32
(FUT)	%	67	33		
Partic	N	21	34	55	45
(PRES)	%	38	62		
Suff	N	24	5	29	24
(PAST)	%	83	17		
Total	N	71	52	123	
	%	58	42		

The factor groups chosen by Varbrul as significant were as follows:

With all tokens counted:

Person	Varbrul weight	% null subject
3rd	0.300	16%
1st	0.802	64%
2nd	0.585	46%
Explicitness		
in-between	0.300	21%
yes	0.724	51%
no	0.294	24%
Subj=Cp(U_{i-1})		
no	0.366	27%
yes	0.771	57%
Transition		
rough shift	0.382	27%
continue	0.819	66%
retain	0.126	7%
smooth shift	0.561	35%

Thus, on the grammatical level, pro-drop is more likely in 1st and 2nd persons, in particular when person is explicit in the morphology, supporting what might be considered a functional explanation. This is in line with the assumptions made in the syntactic literature about the stronger tendency of such verbs to license pro-drop, but it is not as categorical as some of the literature makes it seem.

As far as discourse coherence is concerned, a subject is more likely to be dropped when it is the preferred center of the preceding utterance, and when it is in either a Continue transition or a smooth shift.

With null and overt pronoun subjects only:

Text-type	Varbrul weight	% null subject
news	0.922	88%
GoGay.co.il	0.581	73%
op-ed	0.356	44%
dialogue	0.375	43%
Explicitness		
in-between	0.152	38%
yes	0.713	80%
no	0.329	38%
Gender		
feminine	0.275	36%
masculine	0.508	65%
neuter	0.722	87%
Subj=Cp(U_{i-1})		
no	0.371	52%
yes	0.663	65%

When comparing only the variation between null subjects and overt pronouns, similarly significant weights are given to explicitness of person in the morphology and for the subject being the Cp of the preceding utterance. Two of the texts (but mostly one, the news article) prefer pro-drop as well. While both the news text and the GoGay bulletin board are written texts, they represent quite different writing styles. One can come up with several assumptions as to this particular grouping of the texts, but clearly more research is needed, with a more diverse corpus, to reach any meaningful conclusions on this point. Finally, the morphologically marked gender of the verb appears to have a significant role in promoting pro-drop as well. Neuter subjects, which are either first person or plural (but not in the present), show a strong preference for being realized with a null subject. Again, further research is needed to verify this feature and explain it.

REFERENCES:

Berman, Ruth A. 1990. On acquiring an (S)VO language: Subjectless sentences in children's Hebrew. *Linguistics* 28:1135-1166.

- Borer, Hagit. 1989. Anaphoric AGR. In: Osvaldo Jaeggli and Kenneth J. Safir (eds.). *The Null Subject Parameter*. (Studies in Natural Language and Linguistic Theory; 15). Dordrecht, Boston and London: Kluwer Academic Publishers. 69-109.
- Dimitriadis, Alexis. 1995. When pro-drop languages don't: On overt pronominal subjects in Greek. *PWPL* 2.2:45-600.
- Dimitriadis, Alexis. 1996. When pro-drop languages don't: Overt pronominal subjects and pragmatic inference. *CLS* 32:33-47.
- Grosz, Barbara J., Aravind K. Joshi and Scott Weinstein. 1995 Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21:203-225.
- Izre'el, Shlomo. 2002. The Corpus of Spoken Israeli Hebrew (CoSIH): Textual examples. *Lěšonénu* 64:289-314. [in Hebrew].
- Levy, Yonata and Anne Vainikka. 1999-2000. The development of a mixed null subject system: A cross-linguistic perspective with data on the acquisition of Hebrew. *Language Acquisition* 8:363-384.
- Prince, Ellen F. 1999. Subject-prodrop in Yiddish. In: Peter Bosch and Rob van der Sandt (eds.). *Focus: Linguistic, Cognitive, and Computational Perspectives*. (Studies in Natural Language Processing). Cambridge: Cambridge University Press. 82.104.
- Prince, Ellen F. 2003. The Yiddish Impersonal Pronoun *men* 'one' in Discourse. Paper presented at the annual meeting of the LSA, January 2003.
- Vainikka, Anne & Yonata Levy. 1999. Empty subjects in Finnish and Hebrew. *NLLT* 17:613-671.
- Vainikka, Anne and Yonata Levy. 1995. Empty subjects in Finnish and Hebrew. IRCS Report 95-31, University of Pennsylvania Institute for Research in Cognitive Science.
- Walker, Marilyn, Masayo Iida and Sharon Cote. 1994. Japanese discourse and the process of centering. *Computational Linguistics* 20:1-37.