University of Massachusetts Amherst

From the Selected Works of Shmuel Bolozky

1999

On the Special Status of the Vowels a And e in Israeli Hebrew

Shmuel Bolozky, *University of Massachusetts - Amherst*

Available at: https://works.bepress.com/shmuel_bolozky/42/

ON THE SPECIAL STATUS OF THE VOWELS *a* AND *e* IN ISRAELI HEBREW*

Shmuel Bolozky

University of Massachusetts/Amherst

As the “minimal” vowel of Israeli Hebrew, *e* is the vowel most likely to break unpronounceable consonant clusters, and to be affected by casual vowel deletion. Both *e*-insertion and casual deletion are automatic, which suggests that *e* may be characterized as phonetically unmarked. In contrast, *a* is the most prominent vowel in Israeli Hebrew. It has the highest sonority, is the least marked phoneme in the five-vowel system, and is the most frequent vowel in the language by far. Consequently, *a* functions as the default choice in acronym formation, which is a conscious, non-automatic process. It is thus the natural, most expected vowel of Israeli Hebrew from a phonological point of view that is not *automatic-phonetic*, but rather “*psychologically-based*.” Its status and frequency have at least one important application to the teaching of reading to beginners and to vowel marking in glossaries and dictionaries.

The vowel system of Israeli Hebrew is simpler than those in earlier phases of the language. It contains only five phonemes: *i*, *e*, *a*, *o*, *u*.¹ This article argues that in determining which of them is “unmarked,” that is, the most natural, or most expected, it is useful to posit a dichotomy between the automatic-phonetic and conscious-psychological domains.

I. PHONETIC VS. PSYCHOLOGICALLY-BASED MARKEDNESS

Factors underlying phonological and morphological change are complex, but for simplicity of presentation, one could characterize the major ones by means of a precarious, though viable balance between ease of articulation, the pull towards paradigm uniformity, and the word (or formative) frequency effect on the one hand, and the need to maintain auditory and semantic distinctiveness on the other. Processes motivated by ease-of articulation, and to some extent by paradigm uniformity pressures, are generally automatic; so too is the effect of word frequency which through repeated usage decreases attention that may result in assimilation, reduction, etc. Maintenance of distinctiveness, however, is essentially conscious, and thus psychologically based. Whereas the purpose of

* This article is based on a paper read at the 12th World Congress of Jewish Studies, Jerusalem, July 29–August 5, 1997, entitled “On the ‘centrality’ of the vowel /a/ in Israeli Hebrew.” I wish to thank two anonymous reviewers of *Hebrew Studies*, as well as Ora Schwarzwald and Robert Rothstein, for their useful comments.

¹ Vowel length is not phonemic in Israeli Hebrew. Stress is word-final unless marked otherwise

dissimilation is to maintain distinctions, and it is thus psychologically based, assimilation and reduction processes are driven by ease-of-articulation.

The following concrete examples illustrate these factors.

(i) In Black English, *t* or *d* is very likely to be elided at the end of a word when preceded by another consonant (preferably a sonorant), particularly when the next word begins with a consonant as well, as in *sand castle*, *don't go*, *just now*, *fast car*.² Its elision facilitates articulation. Also, *don't*, *just*, and *fast* are frequent words, and *sand castle*, *just now*, and *fast car* are frequent collocations. Frequent items, which promote casualness in the way they are produced, draw little attention to themselves, and consequently do not inhibit assimilation and reduction. At the same time, their high frequency also guarantees their recoverability (in comprehension) from the context even when reduced. It has also been pointed out that when *d/t* manifests the past tense morpheme, as in *burned coal* or *work[t]* *very hard*, speakers are less likely to elide it, since it marks the past tense, distinguishing between *work* and *worked*.³ This is a psychologically-based consideration.

(ii) The common colloquial tendency in Israeli Hebrew to merge the first person masculine singular future form with the third (m.sg.) person, as in *ani adaber ito* 'I will speak with him' > *ani yedaber ito* (cf., *hu yedaber* 'he will speak'), may be argued to stem from the expected tendency to align the first person future stem with the rest of the conjugation (the *eCaCeC* stem in this case, i.e., *edaber*). The merger unifies the future paradigm and thus simplifies it. The concomitant automatic insertion of *y*, resulting in the merger with *yedaber* 'he will speak,' is phonetically conditioned. It is hard to articulate a vowel sequence whose first member is *i* without a facilitating *y* glide in between. The change is due to an automatic, ease-of-articulation type of process. As a consequence of the changes, however, the presence of the first person pronoun, *ani* 'I,' becomes obligatory in *ani yedaber* 'I will speak,' in order to distinguish the first person from the third. This is a psychologically-based development.

(iii) Either the *CaCaC* pattern (from historical *CaC_iC_iaC*, e.g., *nagar* 'carpenter,' *tabax* 'cook') or *+an* (e.g., *kablan* 'contractor,' *xalilan* 'flutist') can be used to designate agent nouns. Clark and Berman, and Boložky, show that the latter is preferred in child as well as in adult

² E.g., R. W. Fasold, *Tense Marking in Black English: A Linguistic and Social Analysis* (Washington, DC: Center for Applied Linguistics, 1972).

³ E.g., W. Wolfram, *A Sociolinguistic Description of Detroit Negro Speech* (Washington, DC: Center for Applied Linguistics, 1969). The past tense morpheme is realized as [t] after a voiceless consonant.

neologisms.⁴ The prominence of the suffix facilitates its association with the agent (or “performer,” incorporating agents and instruments into a single category), as well as with agent attributes, for example, *ragzan* ‘bad tempered person (N)’ or ‘bad tempered (Adj.)’. *CaCaC* from *CaC_iC_iaC*, however, is only associated with occupations, and is homonymous with other *CaCaC* patterns having nothing to do with occupations. This contributes to opacity. So the saliency of *+an* and its broad but well-defined semantic domain attract an innovator looking for a pattern in which to realize some agentive target meaning. The considerations involved are psychological.

Thus, naturalness of segments or morphemes, and the degree to which they are expected as neutral, or ‘default’ realizations, depends on whether naturalness is characterized in terms of factors such as ease of articulation and possibly paradigmatic leveling and word frequency, manifest in essentially automatic processes, or at the phonological/psychological domain, where distinctiveness, prominence, and other psychological considerations play a role.

II. THE UNMARKED STATUS OF *e* AT THE PHONETIC LEVEL

Phonetically, the vowel *e*, which incorporates what used to be biblical *cere*, *segol*, and *šva mobile*, is the “minimal” vowel of Israeli Hebrew—minimal in that it is the vowel most likely to split consonant clusters that are “impermissible” phonotactically, and the first to undergo elision facilitating “ease of pronunciation.” Thus, from a purely phonetic point of view, it is unmarked. When an unpronounceable consonant cluster is liable to arise, *e* is automatically inserted to split it up, as in the cases of *wanted* and *prodded* in English:

- (1) *avád+ti* ‘I worked’ > *avádeti*, *šavát+ti* ‘I was on strike’ > *šaváteti*.

The same happens when geminates are broken (the assumption that this *e* is epenthetic is based on pattern comparison):

- (2) *zalelan* ‘glutton,’ *xatetan* ‘meddler,’ cf. *kamcan* ‘miser’;
noxexut ‘presence,’ *holelut* ‘folly, hilarity,’ cf. *roxxut* ‘pharmacology.’

It occurs also when the elision of a vowel separating identical consonants is blocked (although reduction to *e* still takes place):

⁴ E. V. Clark and R. A. Berman, “Structure and Use in the Acquisition of Word Formation,” *Language* 6:3 (1984) 542–559; S. Bolzky, *Measuring Productivity in Word Formation: the Case of Israeli Hebrew* (Leiden: E. J. Brill, 1999).

- (3) *xagag* 'he celebrated' ~ *xagega* 'she celebrated'
 (cf. *katav* 'he wrote' ~ *katva* 'she wrote');
kucac 'it was cut' ~ *kucecu* 'they were cut'
 (cf. *šupac* 'it was overhauled' ~ *šupcu* 'they were overhauled');
hitpalel 'he prayed' ~ *hitpalela* 'she prayed'
 (cf. *hitlabeš* 'he got dressed' ~ *hitlabša* 'she got dressed').

e also prevents formation of clusters that would have violated the sonority hierarchy:

- (4) *yladim* 'children' > *yeladim* (cf. *klavim* 'dogs');
mtuka 'sweet, (f.s.)' > *metuka* (cf. *ptuxa* 'open, [f.s.]');
lvanim 'white (m.pl.)' > *levanim* (cf. *ktanim* 'small, [m.pl.]');
nmuxim 'short (m.pl.)' > *nemuxin* (cf. *gnuvin* 'stolen, [m.pl.]');
t'ufa 'aviation' > *te'ufa* > *teufa* (cf. *tsuva* 'answer');⁵
tilmdu 'you (pl.) will learn' > *tilmedu* (cf. *tiknu* 'you [pl.] will buy');
metalfnim 'phone (v, m.pl.)' > *metalfenim*
 (cf. *medabrim* 'speak, [m.pl.]').

At the same time, Boložky and Schwarzwald show that the minimal vowel *e* is the vowel most susceptible to casual/fast speech elision, reduction, and assimilation.⁶ In particular, reduction tends to apply under various conditions.

It occurs (i) when *e* belongs to an affix, or a cliticized function word, as in *te+* and *et* below:

- (5) *te-šév* *be-šéket* 'sit down quietly!' ~ *tšév bešéket* ~ *čév bešéket*
 you (m.) will sit in quiet

te-sadrí *miyád* *et ha-xéder* ~ *tsadrí miyád ta xéder*;
 you (f.) will-arrange immediately acc. the-room
 'tidy up (f.sg.) the room right away!'

(ii) in the context of an appended vowel which removes the sonority-sequence violation by allowing consonant redistribution through re-syllabification:

- (6) **yladim*, but *ha-yeladím* 'the children' ~ *hayladím*
 **lvana*, but *xulcá levaná* 'a white shirt' ~ *xulcá lvaná*

⁵ Even though the historical 'ayin is not realized, it is useful to regard it as a "consonant slot."

⁶ S. Boložky and O. R. Schwarzwald, "On vowel assimilation and deletion in casual Modern Hebrew," *Hebrew Annual Review* 12 (1990) 23-48.

(iii) when both conditions apply, that is, when the item involved is an affix or a cliticized grammatical word, *and* an appended vowel facilitates re-syllabification:

(7) *hu me-laméd ivrít* 'he teaches Hebrew' ~ *hu mlaméd ivrít*
 he pres.-teach Hebrew

hi te-xaké li ba-pina ~ *hi txaké li bapina*
 she (f.s.) will-wait for me at the-corner
 'she'll wait for me at the corner'

hi sogéret et a-délet 'she is closing the door' ~ *i sogért a délet*
 she close-f acc. the-door

bá lexa la-léxet le-séret ~ *bá lxa laléxet leséret*
 come to you to go
 'do you feel like going to a movie?'

ani rocé še-ta-vo a-érev ~ *ani rocé štavó aérev*
 I want that-you will-come the-night
 'I'd like you to ('that you') come tonight'

In very fast or very casual speech, even the *e* separating between two identical consonants may be elided:

(8) 'od *ló raíti zalelán kazé* 'I've never seen such a glutton' ~ *od ló raíti zallán kazé*

hi šatetá šalóš kosót 'she drank three cups' ~ *i šattá šalóš kosót*

hi xagegá et yóm ha-hulédet šela etmól 'she celebrated her birthday yesterday' ~ *i xaggá ta yomulédet šela etmól*

The tendency to elide such "minimal" vowels is phonetically natural. Stress and frequency facilitate reduction when the vowel is in an affix, a clitic, or a function word. Clitics, affixes, and most function words (shorter than three syllables) usually do not carry lexical stress. Whatever stress they bear is normally a consequence of their syntactic or morphological position. Since they are often unstressed, or the stress they carry can be shifted almost unnoticeably through changes in syllable configuration, it is easier to reduce the vowels within them. Also, because of their frequency, they are easily recoverable from the residue of reduction. Optimality of consonant sequences also plays a role. There are universal preferences for the ordering of consonants within the syllable, as well as across syllable boundaries. Within a syllable, consonants must rise in sonority from the beginning of the onset towards the nucleus, and decline in sonority from the nucleus to the end of the coda. The sonority slope is not obligatory

across syllable boundaries, but sequences that maintain it are favored: the "Contact Law" prefers syllable sequences in which the onset-initial consonant of the syllable is lower in sonority than the coda-final consonant of the preceding syllable.⁷ Note the difference between the realizations of *sret* and *ktem* in fast/casual speech:

- (9) *ra'íta kvar et a-séret a-ze?* 'Have you (already) seen this movie?'
 ~ *raíta kvar t asért aze;*
ra'íta kvar et a-kétem a-ze? 'Have you (already) seen this stain?'
 ~ ?*raíta kvar t akétm aze*

The second *e* of *ktem* is less likely to be reduced because *m* is more sonorous than *t*.

The Contact Law can also be used to account for the application of *e*-deletion the *ha-yeladim* > *hayladim* group above (*l* is less sonorous than *y*, etc.). Other phonological considerations, not directly relevant to this discussion, are involved.

e from *cere*, which is rarer, is not as readily available to elision/reduction. It cannot be elided, for instance, in

- (10) *ha-šrutím* 'the services/bathroom' ~ **hašrutím*;
ha-terucím 'the excuses' > **hatrucím*.

Assuming syllable division that would have split the consonant sequence, that is, *haš-rutim*, *hat-rucim* (though *ha-šrutim* and *ha-trucim* would have been equally acceptable), Landau attributes the resistance of such *e* to elision to the Contact Law (*r* being more sonorous than *t* or *š*).⁸ The Contact Law, however, cannot always account for *maCeCa* (with *e* from *cere*) not undergoing *e*-reduction when preceded by a vowel. Its deletion would create the configuration that the Contact Law defines as non-optimal in some cases, for example,

- (11) *ha-maceva* 'the tombstone' > **hamacva*;
ha-magefa 'the plague' > **hamagfa*,

but not in others. In

- (12) *ha-masexa* 'the mask' > **hamasxa* and
ha-xašexa 'the darkness' > **haxašxa*,

for instance, *x* is not more sonorant than either *s* or *š*.

⁷ T. Vennemann, *Preference Laws for Syllabic Structure and the Explanation of Sound Change, with Special Reference to German, Germanic, Italian and Latin* (Berlin/New York/Amsterdam: Mouton de Gruyter, 1988).

⁸ I. Landau, "Suicidal vowels and Syllables that hate them: quantity sensitivity and feature persistence in Modern Hebrew" (MIT ms.).

In other words, speakers may still require either reference to a former *cere*, or simply to two morphological patterns, *CeCuC(+im)* and *maCeCa*, which are immune to *e*-deletion. The latter is probably more realistic, since the former would require a more complicated phonology.⁹ When the item or collocation is sufficiently frequent, however, reduction of *e* from *cere* may occasionally apply in affixes, as to *mevin* 'understand' in

- (13) *ani ló mevín áf milá mimá še-hu omér* 'I don't understand a word of what (that) he says' ~ *an ló mvín áf milá mimá šu omér*,

or to the future-used-imperatively in

- (14) *tešev*¹⁰ *bešéket* 'sit quietly!' > *tšev bešéket* (> *šev bešéket*¹¹).

Note that the hypothesis requiring reference to specific *miškalim* (rather than to ex-*cere e*) is supported by the fact that regardless of how frequent *hašerutim* 'the services' or *masexa* 'mask' might be, their *e* never undergoes deletion.¹² In any event, ex-*cere e*'s are marginal, and the few cases that are protected from reduction hardly affect the general "reducibility" of *e*.

i may be elided/reduced as well, particularly in affixes/clitics/function words, but to a lesser degree. In the verb, reduction occurs mostly in the future/imperative, but it may also apply in common clitics:

- (15) *tiškáv bešéket* 'lie down quietly!' ~ *tškáv bešéket* ~ *čkáv bešéket*;
tistalék mipó 'get out of here!' ~ *tstalék mipó* ~ *ctalék mipó*;
tiza(h)ér miménu 'watch out for him!' ~ *dza(h)ér miménu*;
ani ló medabér ito 'I don't talk to him' ~ *an ló mdabér ito*.

It is hardly ever reduced, however, in regular lexical items such as

- (16) *ha-tinok-ot* 'the babies' > **hatnokot*, or *ha-šir-im* 'the songs' > **hašrim*.

Thus, *e* is a very likely choice for breaking impermissible clusters, and is easily elided in casual speech. *i* is also often elided in similar

⁹ Except for older non-native speakers, particularly those of Ashkenazi origin, who may still maintain the pronunciation of /ey/ for *e* from *cere*.

¹⁰ Elision of such *e* from *cere* may be restricted to this item, where articulation is facilitated by reinterpretation of *tš* as a single affricate, cf. S. Boložky, "On the Monophonemic Interpretation of Modern Hebrew Affricates," *Linguistic Inquiry* 11:4 (1980) 793–799. Similar forms like *telex* 'go!' are not reduced to **ilex*.

¹¹ In S. Boložky, "On the new imperative in colloquial Hebrew," *Hebrew Annual Review* 3 (1979) 17–24, it is argued that in Israeli Hebrew, forms like *šev*, though identical to the normative imperative, are actually derived by "chopping off" the *iV+* prefix, to make the command shorter and thus more urgent or forceful.

¹² Pointed out to me by Ora Schwartzwald (personal communication).

registers, though not so often. The other three vowels, *u*, *o*, and *a*, are generally more resistant to casual deletion. *o* may be reduced in clitics and in some function words, as in *oto/ota/otam* 'him/her/them,'

(17) *hu kibél oto miyád* 'he received him immediately' > *u kibél to miyád*.

It occurs also in common lexical items, such as *roce* 'want,' which owing to their high frequency are easy to recover from their reduced form:

(18) *ani ló rocé ledabér ito* 'I don't want to speak to him' ~ *an ló rce ldabér ito*.

In general, however, rounded vowels are more resistant to reduction. It is hard to determine whether or not this resistance is inherent, since they rarely occur in affixes. *u* is rarely elided, for instance in

(19) *ani ló muxaná* 'I am not ready' > *an(i) ló ???mxaná*.

This example may be used to prove that *u* is inherently resistant to deletion, particularly since the Contact Law would have favored deletion. There are, however, few such cases in affixes, and none with *o*.

a is reducible in some clitics/function words, like *ata/atem* 'you,' for example,

(20) *ata mevín oti* 'do you understand me?' ~ *ta mvín oti*,

and occasionally in affixes, as in

(21) *ani menatáxat et hamišpat aze káxa* 'I (f.) analyze this sentence thus' ~ *ani mnatáxt et a mišpat aze káxa*.

Despite these examples, *a* is less reducible than either *i* or *e*. This becomes clear by comparing reducibility in similar, potentially-reducible environments. *a* is resistant to a number of casual reduction processes.¹³

(i) Its elision is blocked in imperative reduction, as in *takum* 'get up!':

(22) *takúm miyád* 'get up immediately!' > **tkúm*¹⁴ *miyád*.

¹³ Note that "deletions" such as */davar+im/* 'things' > *dvarim* or */katav+a/* 'she wrote' > *katva* are frozen morphophonological alternations that have nothing to do with "real" casual deletion. The same applies, of course, to *e* in */diber+a/* 'she spoke' > *dibra*, to *o* in */tipol+i/* 'you (f.) will fall' > *tipli* and */tixtov+u/* 'you (pl.) will write,' > *ixtevu*, etc.

¹⁴ The whole prefix may be "chopped off," as noted in n. 11, resulting in *kum*, which is identical to the normative imperative, but belongs to a different register. But that is a separate process. *a*'s resistance to deletion only refers to the vowel reduction process producing variants such as *tšev* 'sit down!' and *tškav* 'lie down!' above.

(ii) It is less likely to be affected by post-stress deletion. For example, *pérax* 'flower' in

(23) *tén li leharíax et hapérax haze* 'let me smell this flower'
 ~ *tén li laríax t a ??pérx aze*,

though not impossible, is less likely to become *pérx* in casual speech than *séret* 'movie' is to be reduced to *sert* in (9) above.

(iii) It is unaffected by (substandard) *hif'il* centralization of prefixal *i* to *e*, for example,

(24) *hisbir* 'explained' > *hesbir* and *hidgiš* 'emphasized' > *hedgiš*,

but not in

masbir 'explain' > **mesbir* or *madgiš* 'emphasize' > **medgiš*.

Hif'il centralization is described in some detail by Boložky, who argues that, while the *hiCiC* > *heCiC* and *heCiC* > *hiCiC* shifts, as in

(25) *hipil* 'dropped' > *hepil*, *mapil* 'drop' > *mepil*,
hicig 'presented' > *hecig*, *macig* 'present' > *mecig*,
hevin 'understood' > *hivin*, *hekim* 'raised' > *hikim*,

exemplify analogy or paradigmatic leveling, the *hisbir* 'explained' > *hesbir* centralization is a phonetically motivated process.¹⁵ The fact that it does not affect *a* — **mesbir* is impossible — constitutes another piece of evidence for the stability of *a* compared with *e* or *i*.

One might argue that automatic insertion is not limited to *e*. Where pattern similarity suggests there should have been a consonant cluster, and the first segment (or consonantal slot) in a cluster used to be a guttural, a process of *a*-insertion may be postulated, as in the following:

(26) /šo'el+im/ 'ask (m.pl.)' > *šoalim* (cf. *kotvim* 'write, [m.pl.]', suggesting intermediate /šo'lim/),
 /meša'mem/ 'bore (m.sg.), boring' > *mešaamem* (cf. *metargem* 'translate, [m.sg.]', suggesting intermediate /meša'mem/),
 /mitnahel+im/ 'conduct self (m.pl.)' > *mitna(h)alim* (cf. *mitlabšim* 'get dressed, [m.pl.]', suggesting intermediate /mitnahlim/),
 /mefaxed+im/ 'fear (m.pl.)' > *mefaxadim* (cf. *medabrim* 'speak, [m.pl.]', suggesting intermediate /mefaxdim/) ~ *mefaxdim*
 'even 'stone' ~ *avanim* 'stones' (cf. *kélev* 'dog' ~ *klavin* 'dogs', suggesting intermediate /vanim/),

¹⁵ S. Boložky, *Measuring Productivity in Word Formation: the Case of Israeli Hebrew*.

'eved 'slave' ~ *avadim* 'slaves' (cf. *kélev* 'dog' ~ *klavin* 'dogs,' suggesting intermediate /'vadim/),
harug 'dead, (m.sg.)' ~ *harugim* 'dead, killed (m.pl.)' (cf. *sagur* 'closed, [m.sg.]' ~ *sgurim* 'closed, [m.pl.]', suggesting intermediate /hrugim/),
xašuv 'important, (m.sg.)' ~ *xašuvim* 'important (m.pl.)' (cf. *sagur* 'closed, [m.sg.]' ~ *sgurim* 'closed, [m.pl.]', suggesting intermediate /xšuvim/).

My claim here, however, is that regularities involving *a* in the neighborhood of ex-gutturals have been morphologized in Israeli Hebrew, and that the *a* of forms like *šo'alim* or *avanim* is either lexically acquired for each item, or is derived by a morphological (rather than phonetic) process. One may argue that such *a* is not less motivated phonetically than is the *e* of *xagega* 'she celebrated' or of *yeladim* 'children.' However, there is an important difference between *e* and *a* in these two apparently-similar groups. While the former may often be elided in fast/casual speech, as in *xagega* > *xagga*, *hayeladim* 'the children' > *hayladim*,¹⁶ the latter usually is not. *šo'alim* does not reduce to **šolim*; *haavanim* 'the stones' never becomes **havanim*; *haxašuvim* does not reduce to **haxšuvim*, etc. If the first segment in a cluster is *x*, insertion is optional: the *mefaxadim* 'fear (m.pl.)' type does have a *mefaxdim* variant. But it results not from casual reduction, but rather from analogical restructuring as a regular form (cf. *medabrim* 'speak, (m.pl.)', since *x* is no longer a guttural.¹⁷ Thus, when the same vowel is both the automatic default choice inserted to break clusters, and the first one to be elided when the environment (e.g., *hayeladim* 'the children' > *hayladim*) or the register (e.g., *xagega* 'she celebrated' > *xagga*) allows, it is clearly the phonetically "minimal" vowel in the language, and thus unmarked in the automatic phonetic sense.

¹⁶ This is not the case, though, when the resulting string would violate phonotactic constraints, as in *tilmedu* 'you (pl.) will learn' > **tilmdu*.

¹⁷ When *x* affects adjacent vowels, it is usually *x* from historical *xet*, but not necessarily, as can be seen in sub-standard *oráxai din(it)* 'lawyer (f.sg.)' for standard *oréxet din*, *šaxaxu* 'they forgot' for standard *šaxexu*, *betoxaxem* 'within you (m.pl.)' for standard *betoxexem* (data provided by Ora Schwarzwald, personal communication). Schwarzwald feels that the sub-standard preference for *a* over *e* in such cases proves that *a* is at least as unmarked as *e* at the automatic phonetic domain. My own view is that these are merely consequences of the merger of historical *xet* with *xaf*, and subsequent broadening of the domain of morphophonemic rules applying to *x* from historical *xet*.

III. THE UNMARKED STATUS OF *a* AT THE CONSCIOUS-PSYCHOLOGICAL DOMAIN

While *e* is the “minimal” vowel of Israeli Hebrew, *a* is the most prominent one at the conscious-psychological domain.

(i) It is the most sonorant vowel of the five. Since the mouth is wide open, and little obstruction is involved, minimal energy is expended on production, and maximal volume of sound comes out.

(ii) In the phonological system of Israeli Hebrew there are two high vowels, two mid ones, and only one low vowel, *a*. Although phonetically, *a* is probably sort-of mid-to-back, it is the only low vowel in the system, and if considered back, it is the only non-round back vowel. Regardless of how one looks at it, it is the least marked vowel phonologically within the five-vowel system.

(iii) *a* can be shown, by means of vowel counts like N. Plada's, to be the most frequent vowel in Modern Hebrew, and thus could be regarded as “statistically unmarked.”¹⁸ Plada's analysis was based on an early frequency list, normalized so as to also account for construct state forms, with 150,000 words surveyed, incorporating 586,827 phonetic segments (vowels as well as consonants).¹⁹ A pilot study reported on by Boložky, intended to check whether Plada's conclusions still hold today, examined a small written corpus of mixed registers, containing 6,159 vowels.²⁰ Below are the respective findings, in percentages (of the total number of vowels in each survey/sample):

(27) Two Israeli Hebrew vowel counts

Vowel	% in Plada (1958/59)	% in Boložky (1990)
<i>a</i>	42.12	39.90
<i>e</i>	28.99	23.50
<i>i</i>	12.83	16.20
<i>o</i>	11.14	12.90
<i>u</i>	4.92	7.50
Total:	100.00	100.00

¹⁸ N. Plada, “*Šexiḥut ha-hagayim be-ivrit* (phoneme frequency in Hebrew) (with an introduction by Chaim Rabin),” *Lešonenu* 23 (1958/59) 235–242.

¹⁹ E. Rieger, *Ocar milot ha-yesod šel ha-lašon ha-ivrit ha-šimušit* [A Thesaurus of Basic Words in Everyday Hebrew] (Jerusalem, 1935/36).

²⁰ S. Boložky, “*Al simun ha-tnu‘ot a ve-e ve-al simun he‘der tnu‘a ba-ktiv šel ha-ivrit ha-ḥadaša* [On marking the vowels *a* and *e* and on marking the absence of a vowel in Modern Hebrew orthography]” *Lašon Ve-ivrit* 5 (1990) 34–37.

Although more comprehensive updated counts need to be conducted, including recorded spoken samples, it is very unlikely that the distribution of vowels today is any different. Even if the ratios of other vowels have shifted, *a* is still as prominent and as frequent today as it was thirty years ago. Note that a distinction is drawn between the frequency of words (or formatives) and of individual phonemes. Word frequency, or formative frequency—frequency of clitics or other morphemes—is regarded here as a cause of decrease in attention, resulting in increased assimilation and reduction characteristic of the more casual registers. Reduction is tolerated in meaningful frequent items, since their frequency makes them easy to recover even in reduced form. Phoneme frequency, however, has nothing to do with casualness and consequent reduction; it merely increases the general prominence of frequently occurring segments.

Thus, the combination of inherent sonority, position within the five-vowel system, and statistical predominance makes *a* the most prominent vowel in Israeli Hebrew. As such, it is the vowel chosen, at the conscious-psychological domain, as the default vowel in acronym formation.

IV. THE UNMARKED STATUS OF *a* AS MANIFEST IN ACRONYM FORMATION

Pronounceable acronyms, e.g., NATO, RADAR in English, require vowels to make them pronounceable. So why do those who form them in Hebrew rarely use the phonetically-minimal vowel? Had *e* been the unmarked vowel at the level of morphological formation, it would have been the conscious choice for acronyms as well. Below is description of the major classes of acronyms found in Israeli Hebrew.

Ravid presents a general classification of Israeli Hebrew acronyms, as well as a model that will more or less predict what acronym-formation device will be chosen.²¹ Nir classifies acronyms by the canonical discontinuous word formation pattern (*miškal*) to which they seem to belong, and points to the compound-blend-acronym continuum.²² Two types described by both are irrelevant to this discussion: orthographic acronyms, such as WWII in English, that are never pronounced as such, and are read always

²¹ D. Ravid, "Internal structure constraints on new word formation devices in Modern Hebrew," *Folia Linguistica* 24:34 (1990) 289–347.

²² R. Nir, *Darxey ha-yecira ha-milonit be-ivrit bat yameynu* [Lexical Formation Strategies in Modern Hebrew] (Tel Aviv: The Open University, 1993).

“in full”;²³ and letter acronyms, such as MP in English, in which the full name of each letter is pronounced.²⁴

A third type, existing word acronyms, deliberately seeks to form sequences that correspond to existing “catchy” or “relevant” words (cf. English SALT):

(28) Existing word acronymy

(upper case = the acronym's orthographic representation)

Acronym	Existing Word	Gloss	Source	Acronym	Gloss
<i>XeN</i>	grace, beauty		<i>Xel Našim</i>		women's corps
<i>YéŠa'</i>	salvation		<i>Yehuda Šomron 'áza</i>		Judea Samaria Gaza

Being quite sophisticated and clearly intentional, this last subclass is relatively marginal, and does not reflect natural or spontaneous neologizing processes. At the same time, it does illustrate the conscious nature of acronym formation.

The major, most productive, acronym-forming device is what Ravid refers to as “root acronyms,” where word-initial consonants or consonant sequences are treated as root radicals, and realized as pronounceable acronyms in *miškalim* (discontinuous canonical patterns), some of them in the segolate patterns *CéCeC* or *CóCeC*:

(29) *CéCeC* acronyms

Form	Gloss	Source	Source	Gloss
<i>éŠeL</i> ²⁵	per diem	<i>'axila Štiya Lina</i>		eating, drinking, hotel
<i>ŠéKem</i>	Israel's PX	<i>Šerut Kantnot Miznomim</i>		canteen and eateries service
<i>'óMeC</i> ²⁶	IDF veterans Assoc.	<i>'irgun Mašuxrerey Cáhal</i>		organization of those (honorably) discharged from IDF

²³ As, for instance, in the following (upper case = the acronym's orthographic representation):

Form	Gloss	Source = Pronunciation
<i>'XHC</i>	afternoon	<i>aXar Ha-Cohoráim</i>
<i>XOX</i>	God forbid!	<i>Xas Ve-Xalíla</i> (<i>v</i> and <i>o</i> are orthographically identical)

²⁴ As, for instance, in the following:

Orth. Form	Pronunciation	Gloss	Source	Source Gloss
<i>ŠB</i>	<i>Šin Bet</i>	security police	<i>Šerut Bitaxon</i>	security service
<i>MK</i>	<i>Mem Kaf</i>	section commander	<i>Mefaked Kita</i>	

²⁵ This is a “word acronym” as well, meaning ‘tamarisk.’

²⁶ This is actually a “word acronym,” independently meaning ‘courage.’ It is introduced here to show that the segolate type may be *CóCeC* as well as *CéCeC*.

This pattern has two “guttural” variants *CáCaC* or *CéCaC* when the second or third consonant is “guttural”:

(30) *CáCaC* acronyms

Form	Gloss	Source	Source	Gloss
<i>BáHaD</i>	training base	<i>Bsis HaDraxa</i>		instruction base
<i>TáHaL</i>	Israel water authority	<i>Tixnun Ha-mayim Le-yisra'el</i>		planning water for Israel
<i>MéLaX²⁷</i>	emergency storage	<i>Maxsanim Liš'at Xerum</i>		warehouse for emergency time

Ravid also posits a separate category of pronounceable acronyms which she calls “stem acronyms,” which involve vocalic components:

(31) “Stem” acronyms

Form	Gloss	Source
<i>'UM</i>	The United Nations	<i>Umot Me'uxadot</i>
<i>SaTIL</i>	missile boat	<i>Sfinat TILim</i>
<i>SaKUM</i>	silverware	<i>Sakin Kaf UMazleg</i> “knife, spoon, and fork”

Most “root acronyms,” however, are realized in sequences such as *CaCaC* and *CaCCaC*, that is, they contain the vowel *a* only:

(32) Acronyms with *a* (only)

Form	Gloss	Source	Source	Gloss
<i>MaKaM</i>	radar	<i>Megale Kivun Makom</i>		finder of direction and location
<i>TaLaG</i>	gross national product	<i>Tocar Le'umi Golmi</i>		
<i>ŠaBaK</i>	internal security police	<i>Šerut Bitaxon Klali</i>		general security service
<i>MaQaK</i>	heavy machine gun	<i>MaQlea' Kaved</i>		
<i>BaBLaT</i>	nonsense, garble	<i>Bilbul Beycim Le-lo Taxlit</i>		mixing up eggs/balls without purpose
<i>MaTKaL</i>	general staff	<i>MaTe KLali</i>		
<i>MaNKaL</i>	general manager	<i>MeNahel KLali</i>		
<i>RaMaTKaL</i>	chief of general staff	<i>Roš MaTe KLali</i>		
<i>SaMaNKaL</i>	associate general manager	<i>Sgan MeNahel KLali</i>		

²⁷ This is also a “word acronym,” independently meaning ‘salt.’ It is introduced here to demonstrate the *CéCaC* segolate type.

“Root” acronyms may or may not be derived from root+*miškal* configurations as suggested by Ravid. *Miškalim* may well be involved in generating segolate acronyms (*CéCeC*, *CóCeC*) or segolate derivatives (*CáCaC*, *CéCaC*); *CaCiC* may be occasionally required for forms like *SaTiL* ‘missile ship,’ and *CaCuC* for *SaKUM* ‘silverware (= knife, spoon, fork)’. Bat-El suggests that the majority acronym group, which contain only *a*, may be derived from an (*a*, *a*) vocalic pattern (i.e., a sort of *miškal*), through a mechanism based on syllable structure constraints and other principles, regardless of how many consonants are involved, in the same way that similar reduplicated forms are generated.²⁸ *a*-acronyms, however, do not *have* to involve a *miškal*. They may be generated by some semi-linear derivation process, be it the mechanism proposed by Bat-El or some variant of it, that does not require reference to a *miškal*.²⁹ After the consonants have been selected, based on a variety of considerations, and the minor *miškalim* (*segolate* and others) have been identified, default *a*-insertion applies to all remaining cases, an overwhelming majority.³⁰ The insertion of *a* applies in a manner that makes the sequence pronounceable, while optimizing the syllabic structure of the output along principles suggested by Bat-El. Her system will generate all of these *a*-acronyms, from trilateral to five-consonant ones.

Regardless of the mechanism involved, it is clear that *a* is the unmarked vowel in acronym formation. A comprehensive survey is required to confirm this observation statistically. Merely counting acronyms listed in dictionaries will not suffice because of the unavailability of current frequency ranking, and because dictionaries tend to mostly contain unpronounceable orthographic acronyms, many of them going back to rabbinic literature. Some tentative data, however, may be derived from one list that is readily available. The 1983 supplement to Even-Shoshan’s (1970) dictionary contains an appendix with “current acronyms.”³¹ Of those listed, 436 cases are pronounceable (the number of orthographic, non-pronounceable acronyms far exceeds the number of pronounceable ones). The table below presents the distribution of vowels among them. The numbers at the headings (1, 2...) represent the number of tokens of a particular vowel in a pronounceable acronym (i.e., acronyms containing

²⁸ O. Bat-El, “Phonology and word structure in Modern Hebrew,” UCLA Ph.D. dissertation (1989), p. 79.

²⁹ Ora Schwarzwald (personal communication).

³⁰ Note that unlike regular extraction of “root” consonants (or consonant clusters) from existing lexical items and their application onto a *miškal*, there is considerable flexibility in the choice of consonants for the acronym. Existing words and other suggested similarities play a role, but in the case of *MaQaK* and *MaTKaL* above there are additional considerations: orthographic *MK* is already taken, realized as the letter acronym *mem kaf* ‘section commander.’ So *MaQaK*, *MaTKaL* and *MaNKaL* evolve.

³¹ A. Even-Shoshan, *Ha-milon he-xadaš* [The New Dictionary] (Jerusalem: Kiryat Sefer, 1970, 1983).

one *a*, acronyms containing two *a*'s, etc.); the numbers underneath them are the totals for each particular vowel in the relevant category. The percentages are out of the total number of vowel tokens in all pronounceable acronyms ("occ." = occurrence(s), V = vowel, "prcnt" = percent).

(33) Vowels in pronounceable acronyms listed in the 1983 supplement to Even-Shoshan

#	occ./V	1	prcnt	2	prcnt	3	prcnt	4	prcnt	5	prcnt	total	prcnt
<i>a</i>	129		14.5%	464	52.3%	51	5.7%	8	.9%	5	.56%	657	74%
<i>e</i>	58		6.5%	40	4.5%	3	.34%	4	.45%			105	11.8%
<i>i</i>	65		7.3%	6	.67%	3	.34%					74	8.3%
<i>u</i>	27		3.1%									27	3.1%
<i>o</i>	25		2.8%									25	2.8%
Grand Total Vowels												888	100%

Among these 436 pronounceable acronyms, 353 belong to the most productive category which encompasses all pronounceable acronyms other than what Ravid refers to as "letter acronyms" or "word acronyms." The following table represents their distribution.

#	occ./V	1	prcnt	2	prcnt	3	prcnt	4	prcnt	5	prcnt	total	prcnt
<i>a</i>	81		11.6%	448	64%	42	6%	8	1.1%	5	.7%	584	83.4%
<i>i</i>	39		5.6%	----		3	.43%					42	6%
<i>e</i>	20		2.9%	16	2.3%							36	5.2%
<i>u</i>	24		3.4%									24	3.4%
<i>o</i>	14		2%									14	2%
Grand Total Vowels												700	100%

To the extent that a dictionary listing is reliable, the 1983 supplement to Even-Shoshan (1970) demonstrates clearly the predominance of *a* in pronounceable acronyms.

There is little doubt, then, that when speakers form acronyms, they consciously select the *a* vowel as their "default" realization. This points clearly to its being the unmarked vowel for conscious, psychologically based vowel choice. *a* is selected because of its prominence, as manifested by its unmarked status in the five-vowel system, in its sonority, and in its statistical prominence throughout the language owing to the conflation of what used to be *patax*, *kamac*, and *xatax-patax*.

V. Possible Implication to the Teaching of Hebrew

As in the writing systems of other Semitic languages, Hebrew orthography represents only some of its vowels. Naturally, this presents no problem for native speakers, but at the initial stage, learners of Hebrew as a foreign language find it difficult. Instructors may, of course, use the vowel marks developed originally by the Masoretic scholars to control correct reading of the Bible, a system still used in poetry and in children's books. Since, however, the instructional goal is to bring learners to a reasonable proficiency in reading normal, unvocalized Hebrew texts, excessive dependence on vowel marks might create a dependency, making it very difficult to cross from vocalized to unvocalized texts. Thus, most instructors tend to avoid this dependency to start with and accustom their students to unvocalized text from the very beginning. The question is whether or not *partial* marking of unrepresented vowels can be selectively employed at the initial instructional phase. Even if that option is ruled out, for the pedagogical reason stated above, partial marking can still be used for glossaries and cumulative dictionaries.³² Clearly, the vowels *i*, *o* and *u* do not constitute a serious problem, since the consonantal symbols for *y* and *w* have traditionally been used for vowel representation as well, albeit not consistently. Even in the Hebrew Bible, *yod*, the symbol for *y*, may be also used for *i*, and *waw*, the symbol for *w*, for either *u* or *o*, although one cannot predict whether the symbol would or would not be used for any particular occurrence. So maximizing the use of the *matres lectionis*, with a dot distinguishing *u* from *o*, could be helpful at the initial instructional stage (or at least in glossaries/dictionaries). Since the Hebrew Language Academy sanctions *u* and *o* representation by *waw* essentially everywhere in *plene* writing, and *i* by *yod* in open syllables (as well as in some closed ones where *i* is not predictable owing to low pattern frequency), this practice is not likely to be contradicted by what students find in authentic texts they see outside of the classroom. But what about *a* and *e*?

One way of dealing with the representation of these two vowels is proposed by Boložky, whose suggestion is based on the unmarked status of *a*.³³ If instructors wished to introduce beginning students of Hebrew to the idea of a syllabary so that they understand that a non-final consonant is likely to be a CV sequence, s/he could mark texts for *e* and for a quiescent *šwa* and instruct students to assume that unless a non-final consonant is marked for *e* or for zero (by a *šwa*), it should be read as *Ca*. If a (zero) *šwa* is hard to

³² As proficient as students may become, glossaries and cumulative dictionaries would always require at least minimal vowel marking.

³³ S. Boložky, "‘Al *simun ha-tnu'ot*..."

pronounce, it should be read as *e*, which would occur naturally in a two-*šwa* sequence, as well as in a syllabic onset or coda, when the sonority hierarchy is violated (*yeladim*, *levana*, *filem*, *hér cel*). The instruction to “pronounce an unpronounceable zero *šwa* as *e*” is linear, following the progress of reading. Thus, in the case of תִּלְמִד /*tilmdu*/, for instance, the first *šwa* is read as zero, resulting in *til*; since the one under the *mem* is unpronounceable as zero, *m* is realized as *me*, just as the *y* in /*yladim*/ becomes *ye*. The only “generalizations” the student needs to remember, then, are that the *šwa* is zero, and that any unmarked, non-final consonant is pronounced *Ca*. Thus, one captures the high likelihood of a non-final consonant being followed by *a*, as well as the dual role of the *šwa* as a zero vowel as well as *e*. Note that if the statistics in the Boložky (1990) vowel count noted above are adjusted to exclude all occurrences of *o* and *u*, and those of *i* in an open syllable—that is, all vowels which are (or may be) represented by separate *matres lectionis* letters as approved by the Academy—the percentages for the remaining symbols are:

(35) Vowel count excluding *o*, *u*, and *i* in an open syllable

Vowel/Zero	% in Boložky (1990)
<i>a</i>	49.3
<i>e</i> from <i>segol/cere</i>	19.3
<i>e</i> from <i>šwa</i>	9.7
<i>i</i> in a closed syllable	9.2
zero (quiescent <i>šwa</i>)	<u>12.5</u>

Total 100.00

Since statistically about 50% of all consonant occurrences that are not followed by *matres lectionis* make an *Ca* syllable, automatic assignment of *a* as the “default” vowel will provide for the most economical representation.

VI. CONCLUSION

It has been shown that while *e* is the unmarked vowel in Israeli Hebrew from an automatic-phonetic point of view, the high sonority of *a*, its unmarked position in the five-vowel system of Hebrew, and its statistical dominance, make it prominent enough to constitute the unmarked vowel at the conscious psychological level, and thus the default choice in acronym formation. Its unmarked status can also serve a practical purpose in teaching beginning students how to read, or at least in making vowel representation in glossaries and dictionaries more efficient.

HEBREW STUDIES

A Journal Devoted to
Hebrew Language and Literature

Volume XL, 1999

ZIONY ZEVIT, Editor
FREDERICK E. GREENSPAHN, Editor
EDWARD GOLDMAN, Associate Editor
RACHEL BRENNER, Associate Editor
DAVID MARCUS, Book Review Editor

Published at the University of Wisconsin-Madison by
The National Association of Professors of Hebrew
in Institutions of Higher Learning

איגוד הפרופסורים לעברית בארצות-הנרית