

## Base Selection in Analogical Change in Yiddish

ADAM ALBRIGHT

*University of California, Los Angeles*

### 0. Introduction

A notable difference between Yiddish and German verb paradigms is that Yiddish has no vowel alternations in the present tense.<sup>1</sup> Whereas Middle High German (MHG) and Modern German (NHG) often have alternations among the singular forms (1a), or between the singular and plural (1b), Yiddish never does (2).

(1) MHG present tense vowel alternations

a. 'dig'	sg.	pl.	b. 'know'	sg.	pl.
1st	<i>grabe</i>	<i>graben</i>	1st	<i>weiȝ</i>	<i>wiȝzen</i>
2nd	<i>grebest</i>	<i>grabet</i>	2nd	<i>weist</i>	<i>wiȝzet</i>
3rd	<i>grebet</i>	<i>graben</i>	3rd	<i>weiȝ</i>	<i>wiȝzen</i>

(2) Yiddish paradigms have no vowel alternations<sup>2</sup>

a. 'dig'	sg.	pl.	b. 'know'	sg.	pl.
1st	<i>grɔb</i>	<i>grɔbən</i>	1st	<i>veys</i>	<i>veysən</i>
2nd	<i>grɔbst</i>	<i>grɔbt</i>	2nd	<i>veyst</i>	<i>veyst</i>
3rd	<i>grɔbt</i>	<i>grɔbən</i>	3rd	<i>veyst</i>	<i>veysən</i>

As I will show in Section 1, the form that has been extended in Yiddish is always the expected 1sg form. Interestingly, although this change is across the board in Yiddish, it is apparently unattested in any other German dialect.

Paradigmatic levelings of this sort, in which some members of the paradigm are rebuilt based on other forms, pose a well-known problem in historical linguistics. On the one hand, they occur frequently, and seem natural and unsurprising. In

<sup>1</sup>The Yiddish data in this paper concerns the eastern dialects of Central and Eastern Europe; I do not know if the same holds true of the western dialects.

<sup>2</sup>For Yiddish examples, I will use YIVO transliteration (<http://www.yivoinstitute.org/yiddish/alefbeysh.htm>), with two minor modifications: I will use the IPA symbol ɔ instead of YIVO *o* for *komets-aleph*, and I will use *-ən* instead of YIVO *-en/-n* for the infinitive/1pl/3pl suffix. For MHG forms, I will use the standardized orthography of Paul, Wiehl, and Grosse (1989, §§18–20), in which *^* marks long vowels, *ɛ* is a short open [e], and *ȝ* is a coronal sibilant fricative, possibly fortis, possibly postalveolar (Paul et al, §151). The change of MHG short [a] > Yiddish [ɔ] reflects a regular sound change; the correspondences between MHG ⟨w⟩ and Y ⟨v⟩, MHG ⟨ei⟩ and Y ⟨ey⟩, MHG ⟨s⟩ and Y ⟨z⟩, and MHG ⟨ȝ(ȝ)⟩ and Y ⟨s⟩ are also completely regular.

fact, it is often suggested that the desire for nonalternating paradigms is simply a primitive of language, sometimes referred to as “Humboldt’s Universal”, or, more recently, as Uniform Exponence (Kenstowicz 1995) or Paradigm Uniformity (Steriade 2000). On the other hand, a blanket preference for uniform paradigms can only go so far in explaining levelings: it can tell us an alternation is likely to be leveled, but not necessarily in which direction. Why was the 1sg extended, and not the 3sg, yielding paradigms like \**greb*, \**grebst*, \**grebt*? And what distinguished Yiddish from other German dialects, otherwise very similar morphologically?

Many proposals over the years have attempted to explain the direction of leveling. The usual approach, pioneered by Kuryłowicz (1947) and Mańczak (1958) and continued by Bybee (1985) and others, has been to focus on tendencies, or groups of factors that compete to determine the direction of a change. Under such an approach, it is possible to derive typological predictions – leveling is often to the isolation form, the most frequent form, the 3sg, and so on – but it is impossible to make predictions about a given language at a given time, because we do not know which factors will win in that particular case. In this paper, I will pursue a different approach, in the tradition of Paul (1920) and Kiparsky (1965), which focuses on the role of the learner in language change. In particular, I will pursue the hypothesis that language learners impose a structure on paradigms that helps them to construct phonological and morphological grammars that generate unknown forms as accurately or as confidently as possible. The way that they do this, I will claim, is by seeking a base form within the paradigm that is “maximally informative” – that is, that suffers the least serious phonological and morphological neutralizations – and then deriving the remaining forms in the paradigm from the base form. Under this approach, we can use the direction of the grammar (base form → derived forms) to predict the direction of possible analogical changes. In section 2, I will show that the 1sg form preserves the most contrasts in Yiddish, and thus would be selected as the base form in the proposed model. Finally, in section 3, I will argue that the advantages of the 1sg are unique to Yiddish, due to small but crucial differences between Yiddish and other German dialects. Thus, the proposed analysis gives us insight not only into the question of why verbs were leveled to the 1sg in Yiddish, but also into why this did not occur elsewhere.

### **1. Leveling to the 1sg in Yiddish present tense paradigms**

As shown in (2), Modern Yiddish has no vowel alternations in the present tense (Rockowitz 1979). In this section, I will show that in virtually all cases, the vowel of the 1sg has been extended to the rest of the paradigm. To show this, we will consider the following candidates as sources for the modern present tense stem, eliminating all except the 1sg<sup>3</sup>: 1sg, 2sg, 3sg, 1pl, 2pl, 3pl, infinitive, and UR. I

---

<sup>3</sup>This list includes almost all of the verb forms that occur in Yiddish; the only other forms are the present participle, stem (*shtam*), past participle, and imperative. The present participle and stem are always based on the infinitive, so any conclusions regarding the infinitive hold of them as well. The past participle is also demonstrably not the source of the modern present stem. The singular imperative form is always identical with the 1sg, and could equally well have served as the base for the leveling discussed here. For expository ease, I will refer to the 1sg, but I cannot preclude the possibility that it was the singular imperative instead.

## Base Selection in Analogical Change

will start with the assumption that the origin of Yiddish was some form of MHG, so it is useful to begin by considering the possible types of present tense paradigms that occurred in MHG (Paul, Wiehl, and Grosse 1989, chap. 7).

### 1.1. MHG present tense patterns

Most MHG verbs had the same vowel throughout the entire present tense, with no alternations, as in (3); verbs of this type included the “strong” classes I, IIIa, and some of VII, as well as all of the “weak” verbs.

- (3) No alternations (Strong I, Strong IIIa, some Strong VII, all weak)
- |           |               |              |          |               |              |
|-----------|---------------|--------------|----------|---------------|--------------|
| a. ‘live’ | <i>lēbe</i>   | <i>lēben</i> | b. ‘say’ | <i>sage</i>   | <i>sagen</i> |
|           | <i>lēbest</i> | <i>lēbet</i> |          | <i>sagest</i> | <i>saget</i> |
|           | <i>lēbet</i>  | <i>lēben</i> |          | <i>saget</i>  | <i>sagen</i> |

In another set of verbs, an *a* in the root surfaced as an *e* in the 2sg and 3sg due to a process known as *umlaut*, as in (4). This occurred in strong class VI and the remainder of strong class VII.

- (4) 2sg, 3sg different due to Umlaut (*a* ~ *e*): Strong VI, some Strong VII
- |       |                |               |   |
|-------|----------------|---------------|---|
| ‘dig’ | <i>grabe</i>   | <i>graben</i> | (Also <i>varn</i> ‘travel’, <i>laden</i> ‘invite’,              |
|       | <i>grebest</i> | <i>grabet</i> | <i>slahen</i> ‘beat’, <i>halten</i> ‘hold’, <i>lâzen</i> ‘let’) |
|       | <i>grebet</i>  | <i>graben</i> |   |

Finally, a third set of verbs showed vowel alternations between the entire singular and the entire plural. This pattern occurred in two types of verbs. In some, the present tense derived from a Proto-Indo-European perfect, and the singular/plural alternation reflected a PIE alternation in the perfect tense (*ablaut*); these verbs are known as *preterite presents* (5a). In others, the alternation was due to a phonological process in Old High German that raised mid vowels before a following high vowel, causing the singular to diverge from the plural (5b); this pattern is sometimes referred to as *Wechselflexion* (“alternating inflection”), and occurred in strong class II (*ie* ~ *iu*), as well as IIIb, IV, and V (*ë* ~ *i*).

- (5) Singular ~ plural alternations
- |  |              |                          |   |               |              |
|--|--------------|--------------------------|---|---------------|--------------|
| a. Preterite presents                    |              | b. <i>Wechselflexion</i> |   |               |              |
| ‘know’                                   | <i>weiȝ</i>  | <i>wiȝzen</i>            | ‘give’                                  | <i>gibe</i>   | <i>gēben</i> |
|  | <i>weist</i> | <i>wiȝzet</i>            |   | <i>gibest</i> | <i>gēbet</i> |
|  | <i>weiȝ</i>  | <i>wiȝzen</i>            |   | <i>gibet</i>  | <i>gēben</i> |
| (Also <i>kunnen</i> ‘can’, <i>durfen</i> |              |                          | (Also <i>nēmen</i> ‘take’, <i>ēȝzen</i> |               |              |
| ‘need’, <i>suln</i> ‘should’, etc.)      |              |                          | ‘eat’, <i>gieȝen</i> ‘pour’, etc.)      |               |              |

### 1.2. Yiddish present tense patterns

Let us now consider the fate of these patterns in Yiddish. Unsurprisingly, verbs with no alternations in MHG continue to have no alternations in Yiddish, as seen in (6).

- (6) Non-alternating verbs remain non-alternating in Yiddish
- |           |              |              |          |              |              |
|-----------|--------------|--------------|----------|--------------|--------------|
| a. ‘live’ | <i>leb</i>   | <i>lebən</i> | b. ‘say’ | <i>zɔg</i>   | <i>zɔgən</i> |
|           | <i>lebst</i> | <i>lebt</i>  |          | <i>zɔgst</i> | <i>zɔgt</i>  |
|           | <i>lebt</i>  | <i>lebən</i> |          | <i>zɔgt</i>  | <i>zɔgən</i> |

## Adam Albright

Turning to umlaut alternations (1sg *grabe*, 2sg *grebst*), these were leveled to the non-umlaut (*a*) alternant (7). Thus, it appears that the base, or pivot, of the leveling was not the 2sg or 3sg, or else the modern Yiddish paradigm would have *e*.

(7) Umlaut verbs leveled to non-umlaut (*a*) alternant: *grəbən* ‘dig’

1st	<i>grəb</i>	<i>grəbən</i>	infin.	<i>grəbən</i>
2nd	<i>*grebst &gt; grəbst</i>	<i>grəbt</i>		
3rd	<i>*grebt &gt; grəbt</i>	<i>grəbən</i>		

Considering next the preterite present verbs, we find that for these, the Yiddish present tense forms come from MHG singular forms. This is shown in (8) for the verbs *darfən* ‘need’ and *visən* ‘know’, whose present tense forms are derived from the MHG singular forms *darf-* and *weiz-*, and not the plural *dürf-/durf-* and *wiz-*. Other examples include *muzən* ‘must’ (< MHG sg. *muoz*, not pl. *müezzen*), *tərən* ‘must’ (< MHG sg. *tar*, not pl. *türren*), and *zolən* (< MHG sg. *sol*, not pl. *süln*)

(8) Preterite Present verbs leveled to singular

a. ‘need’	sg.	pl.	b. ‘know’	sg.	pl.
1st	<i>darf</i>	<i>*dūrfən &gt; darfən</i>	1st	<i>veys</i>	<i>*visən &gt; veysən</i>
2nd	<i>darfst</i>	<i>*dūrft &gt; darft</i>	2nd	<i>veyst</i>	<i>*vist &gt; veyst</i>
3rd	<i>darf</i>	<i>*dūrfən &gt; darfən</i>	3rd	<i>veys(t)</i>	<i>*visən &gt; veysən</i>
infin.	<i>*dūrfən &gt; darfən</i>		infin.	<i>visən</i>	
UR	<i>/dūrf-/</i> , <i>/darf-/ &gt; /darf-/</i>		UR	<i>/vis/</i> , <i>/veys/</i>	

We can conclude that the generalized form was not a plural form or the infinitive – in fact, most infinitives of preterite presents were also rebuilt based on singular forms. Furthermore, the two MHG stem alternants (*darf-*, *dūrf-*) cannot easily be reduced to a single UR, since they involve an idiosyncratic alternation that occurs in only one other verb, and it is not clear how to derive *ü* from *a* or vice versa. The most promising analysis is to list two URs for these verbs (e.g., */darf/*, */dūrf/*), in which case generalized form does match one of the MHG UR’s (*/darf/*). However, simply saying the UR has been generalized does not explain why one UR was chosen over the other. Putting this conclusion together with that from the umlaut verbs, we have now eliminated the 2sg, 3sg, all of the plural, the infinitive, and the UR as sources of Yiddish present tense forms. Thus, it appears that the 1sg is only possible source.

The data so far converge on the 1sg as the source for Yiddish present tenses. Unfortunately, when we turn to the *Wechselflexion* verbs, the situation is more complicated. From what we have seen thus far, we would expect these verbs to generalize the *i* of the singular, and indeed this is what we find with *gebən* ‘give’ (9a). For most MHG *Wechselflexion* verbs, however, Yiddish seems to have generalized the *e* of the plural/infinitive, as in *nemən* ‘take’ (9b).

(9) Fate of *Wechselflexion* verbs

a. Generalized <i>i</i> from sg: <i>gebən</i> ‘give’	sg.	pl.	b. Generalized <i>e</i> from pl: <i>nemən</i> ‘take’	sg.	pl.
1st	<i>gib</i>	<i>*gebən &gt; gibən</i>	1st	<i>*nim &gt; nem</i>	<i>nemən</i>
2nd	<i>gibst</i>	<i>*gebt &gt; gibt</i>	2nd	<i>*nimst &gt; nemst</i>	<i>nemt</i>
3rd	<i>gibt</i>	<i>*gebən &gt; gibən</i>	3rd	<i>*nimt &gt; nemt</i>	<i>nemən</i>
infin.	<i>gebən</i>		infin.	<i>nemən</i>	

### *Base Selection in Analogical Change*

The pattern of generalized *e* is found not only in *nemən*, but also in *esən* ‘eat’ (*es*, \**is*), *fargesən* ‘forget’ (*farges*, \**fargis*), *zeyn* ‘see’ (*zey*, \**zi*), *vern* ‘become’ (*ver*, \**vir*), *helfən* ‘help’ (*helf*, \**hilf*) etc. Why do these verbs show a different pattern from all other verbs? Is this an exception to generalization of the 1sg form?

I would like to argue that verbs like *nemən* are not exceptions, but rather that they already contained *e* in the 1sg at the time that Yiddish diverged from other German dialects.<sup>4</sup> I began this section with the assumption that Yiddish began as some form of MHG, exemplified by the literary MHG forms in (3)-(5). However, the history of the *Wechselflexion* in German is somewhat complicated, and it is not clear that the paradigm in (5) is the correct starting point for Yiddish. According to the standard account (Paul et al. 1989 §§31-35), *Wechselflexion* was due to a phonological process in OHG raising /e/ to [i] when a high vowel (*u*, *i*) was in the following syllable. Since singular suffixes had high vowels and plural suffixes had mid vowels, this led to an alternation between *i* in the singular and *e* in the plural (*issu* ~ *ëssê* ‘eat-1sg/1pl’). In MHG, all suffix vowels were reduced to schwa, making the *i*~*e* alternation a purely morphological difference between the singular and the plural (*isse* ~ *ëssen*). This pattern is found in all MHG texts until the mid-15th C (Dammers, Hoffmann, and Solms 1988, §148.4). Finally, during late MHG or early NHG times, the vowel of the 1sg lowered back to *e* (*isse* > *esse*), probably under the influence of the umlaut pattern (1sg vs. 2,3sg, as in (4) above). 1sg forms with *e* began to occur regularly in “middle German” (Fränkisch, Thüringisch, Böhmisch, Schlesisch) during the fifteenth century (Paul et al. 1989, §242, note 1; Philipp 1980, p. 66), appearing earlier in the west than in the east (Dammers, et al. 1988, §148.4). The change proceeded verb-by-verb, with considerable variation even between occurrences of the same verb in the same text (Kern 1903, pp. 47-60; Geyer 1912, §31-§32), but eventually all *Wechselflexion* verbs were affected.

What we see, then, is that 1sg forms with *e* in German predate the Middle Yiddish period (16th-17th C). Thus, I hypothesize that Yiddish already had *e* in the 1sg of *Wechselflexion* verbs prior to leveling. If this is the case, then the *e* of *nemen* is not an exception to the generalization that leveling was always to the 1sg.

I have found only three exceptions to generalization of the 1sg in Yiddish: (1) *zayn(ən)* ‘to be’ retains a suppletive paradigm (2) the future auxiliary *velən* derives from a conditional form, not the 1sg present indicative, and (3) *gefelən* ‘be pleasing’ is used predominantly in the 3rd person, and derives from a 3sg form (*gefelt* ‘it is pleasing’).<sup>5</sup> These exceptions are not all that surprising – two are extremely high frequency, and the third has semantic restrictions. In sum, for every type of MHG verb, it appears that the 1sg form has been extended to the remainder of

<sup>4</sup>I am not making any commitment as to when Yiddish ceased to be a sociolect of German, except to suppose that the two probably continued to co-evolve at least until the beginning of the Middle Yiddish period (c. 16th C), when Yiddish literature began to flourish in the east, eastward migrations trickled off, and significant east-west dialect differences emerged (Weinreich 1980, p.724-726).

<sup>5</sup>This effect, in which lexical semantics influences the direction of analogy, is discussed by Tiersma (1982) under the rubric of *local markedness*. However, this is the only example in Yiddish, so it seems extravagant to invoke local markedness to explain just one case. It may also derive from a MHG variant of *gefallen*; another example is Yiddish *fregən* ‘ask’, derived from MHG *vrëgen*, a variant of *vragen* (Paul et al. 1989, §30).

the paradigm in Yiddish. This leveling has been remarkably complete, affecting virtually all verbs. In the next section, I will consider the question of why Yiddish paradigms were rebuilt on the basis of this, and not some other form.

## 2. The 1sg as the optimal base in Yiddish

### 2.1. Identifying the optimal base

Why did the 1sg have a privileged status in Yiddish? In this section, I will argue that it was the “maximally informative” form, suffering from the fewest phonological neutralizations, and maintaining the most contrasts. In order to show this, I will examine a pre-leveling version of Yiddish, considering which parts of the paradigm suffered from neutralizations, and how many verbs were affected in each case.

Yiddish, like German and English, disallows coda clusters of obstruents with voicing disagreement ( $*bs]_{\sigma}$ ,  $*pd]_{\sigma}$ , etc.). When a suffix containing voiceless obstruents (2sg *-st*, 3sg/2pl *-t*) is added to a root ending in a voiced obstruent, the root-final obstruent is devoiced. The result is that in the 2sg, 3sg, and 2pl, the contrast between root-final voiced and voiceless obstruents is neutralized.

(10) Neutralization in the 2sg/3sg/2pl: voicing assimilation to suffix

	<i>libən</i> ‘to love’	<i>zipən</i> ‘to sift’
1sg	<i>lib</i>	<i>zip</i>
2sg	<b><i>lipst</i></b>	<b><i>zipst</i></b>
3sg	<b><i>lipt</i></b>	<b><i>zipt</i></b>
1pl	<i>libən</i>	<i>zipən</i>
2pl	<b><i>lipt</i></b>	<b><i>zipt</i></b>
3pl	<i>libən</i>	<i>zipən</i>
infinitive	<i>libən</i>	<i>zipən</i>

This neutralization affects all obstruent pairs with a voicing contrast, of which Yiddish has seven ( $p/b$ ,  $t/d$ ,  $k/g$ ,  $f/v$ ,  $s/z$ ,  $\int/\zeta$ ,  $t\int/d\zeta$ ). A hypothesis of the current approach is that the seriousness of a neutralization depends not only on the number of phonemes involved, but also on the number of lexical items whose underlying form cannot be recovered due to the neutralization. In order to estimate of the number of verbs whose final segment would be ambiguous due to voicing assimilation, I counted the number of verbs ending in these 14 obstruents in the German portion of CELEX.<sup>6</sup> For CELEX counts, I considered only verb lemmas that had a token frequency of 1 or greater and did not contain a separable initial element (separable prefix, incorporated object, adverb); this left a total of 4877 verbs. As it turns out, 1988 of these end in obstruents with voicing contrasts, meaning approximately 41% of all verbs are ambiguous in the 2sg, 3sg, and 2pl.

Another set of neutralizations in Yiddish comes from a ban on word-internal geminates. For example, adding the 2sg suffix *-st* to a verb ending in *s* or *z* should

<sup>6</sup>Ideally, we count a lexicon of Middle Yiddish, but this does not exist in searchable form, and counts from German form a reasonable approximation. There are certainly numerous lexical differences between Yiddish and German, and even some phonological ones – e.g., Yiddish has verb roots ending in [v] and [dʒ], which are rare or absent in German. However, most common Yiddish verbs are shared with German, and there is no reason to believe that the lexical differences would significantly alter the proportion of major classes like obstruent-final verbs, strident-final verbs, etc.

*Base Selection in Analogical Change*

yield the sequence *-sst* (with devoicing of *z* to satisfy voicing agreement). This sequence actually surfaces as degeminated *-st*: /veys-st/ → [veyst], not \*[veysst]. The result is that *s-* and *z-*final verbs are neutralized with vowel-final verbs in the 2sg, as seen in (11a). For the 3sg and 2pl, the suffix is *-t*, and an equivalent degemination of *tt* (fed by /d/ → [t] devoicing) applies (11b).

(11) Neutralizations caused by degemination

a. Neutralization in the 2sg:

	devoice <i>z</i> , degeminate of <i>ss</i>		
	<i>geyn</i> 'go'	<i>visən</i> 'know'	<i>vayzn</i> 'show'
1sg	<i>gey</i>	<i>veys</i>	<i>vayz</i>
2sg	<b><i>geyst</i></b>	<b><i>veyst</i></b>	<b><i>vayst</i></b>
3sg	<i>geyt</i>	<i>veys(t)</i>	<i>vayst</i>
1pl	<i>geyən</i>	<i>veysən</i>	<i>vayzən</i>
2pl	<i>geyt</i>	<i>veyst</i>	<i>vayst</i>
3pl	<i>geyən</i>	<i>veysən</i>	<i>vayzən</i>
infin	<i>geyən</i>	<i>visən</i>	<i>vayzən</i>

b. Neutralization in the 3sg/2pl:

	devoice <i>d</i> , degeminate <i>tt</i>		
	<i>faltən</i> 'fold'	<i>falən</i> 'fall'	<i>redən</i> 'talk'
1sg	<i>falt</i>	<i>fal</i>	<i>red</i>
2sg	<i>fal(t)st</i>	<i>falst</i>	<i>retst</i>
3sg	<b><i>falt</i></b>	<b><i>falt</i></b>	<b><i>ret</i></b>
1pl	<i>faltən</i>	<i>falən</i>	<i>redən</i>
2pl	<b><i>falt</i></b>	<b><i>falt</i></b>	<b><i>ret</i></b>
3pl	<i>faltən</i>	<i>falən</i>	<i>redən</i>
infin	<i>faltən</i>	<i>falən</i>	<i>redən</i>

How many lexical items would be affected by these neutralizations? The voicing neutralization of *s/z* and *t/d* was already included in the count for voicing assimilation above, but degemination means that vowel-final roots are also ambiguous in these forms – an additional 227 words in CELEX, or 5% of the verbal vocabulary.

So far, we have examined neutralizations in forms with obstruent suffixes – the 2sg, 3sg, and 2pl. Turning to the 1pl, 3pl, and infinitive forms, the suffix for all of these forms is *-ən*. Since this suffix is vowel-initial, and Yiddish allows vowels to occur in hiatus, it does not give rise to illegal sequences to trigger assimilation or deletion, with one exception: if the verb root ends in a schwa (e.g., *pərə* 'fiddle with'), then the 1pl/3pl/infinitive form ends simply in *-ən*, not *\*-əən*. The reduction of /əə/ to [ə], motivated by a ban on long schwa (\*[ə:]), means that in these forms, schwa-final and non-schwa-final verbs are neutralized.

(12) Neutralizations in the 1pl/3pl/infinitive: stem-final /ə/

	<i>pərən</i> 'to match'	<i>pərən</i> 'to fiddle with'
1sg	<i>pər</i>	<i>pərə</i>
2sg	<i>pərst</i>	<i>pərəst</i>
3sg	<i>pərt</i>	<i>pərət</i>
1pl	<b><i>pərən</i></b>	<b><i>pərən</i></b>
2pl	<i>pərt</i>	<i>pərət</i>
3pl	<b><i>pərən</i></b>	<b><i>pərən</i></b>
infinitive	<b><i>pərən</i></b>	<b><i>pərən</i></b>

How serious is this neutralization? German does not have schwa-final verbs, so we cannot use CELEX to estimate the number of lexical items that would be affected by it. Instead, I took a sample from Weinreich (1990), counting all of the verbs beginning with [l]. (This segment was chosen to avoid skewing the sample by including uniquely Slavic onsets like *shtsh-* or *tl-*, or characterically Hebrew onsets like *mə-*; [l]-initial words [l] seem to come from Germanic, Slavic, and Hebrew in

representative proportions.) Of the 90 [l]-initial verbs, 9 of them (10%) have stem-final ə. Thus, a contrast that is seen in a significant portion of the Yiddish verbal vocabulary is neutralized in the 1pl/3pl/infinitive forms.

As with other neutralizations, it is worth considering whether the presence of stem-final [ə] is truly neutralized in the 1pl/3pl/infinitive forms, or whether it could be predicted using secondary cues. In casual speech in many dialects, the [ə] of the *-ən* suffix may be lost, resulting in a syllabic nasal agreeing in place with a preceding consonant: [libən] ~ [libm̩] ‘love-1pl/3pl/inf.’. This process affects suffix [ə], but not stem-final [ə] – meaning that verbs with stem-final [ə] might be distinguished by lack of a [ə]-less variant (*pərən/pər̩n* ‘match’ vs. *pərən/\*pər̩n* ‘fiddle with’). This difference would be rather poor evidence about the status of final [ə], however. First, it requires distinguishing a syllabic nasal from a schwa-nasal sequence, which is not always easy to do, especially after continuants. Furthermore, this form is only informative if it is determined to end in a syllabic nasal; if it ends in *-ən*, no conclusion can be drawn. Finally, reduction of *-ən* to syllabic *-n* does not seem to occur in all environments. This is reflected in the YIVO orthography, which uses *-en* after *m*, *n*, *ng*, *nk*, and syllabic *l*, but *-n* elsewhere. In practice, reduction is probably not as categorically restricted as the orthography implies, but occurs most often after stops, least often after vowels, and so on. Therefore, we would be able to use the 1pl/3pl/infinitive form to infer a lack of final [ə] for at best only a subset of verbs in the language.

Another potentially relevant fact is that virtually all [ə]-final verbs come from Slavic or Hebrew. If a verb can be identified as non-Germanic, perhaps because it contains a sequence that is illegal in German (e.g., *pyesh̩tshən* ‘caress’, *tl̩ən* ‘smolder’, *str̩ashən* ‘threaten’), it is much more likely to have a stem-final [ə]. In addition, there are two derivational suffixes with final [ə]: the verbal suffix *-eve* (e.g., *ratev̩ən* ‘rescue’, *bush̩ev̩ən* ‘rage’), and the mimetic suffix *-ke* (e.g., *shush̩kən* ‘whisper’, *haf̩kən* ‘bark’). Therefore, verbs ending in *-ken* and *-even* are extremely likely to have final [ə]. These two facts make it somewhat easier to guess whether a new word should have final [ə], but it is still far from predictable. In fact, there are a number of other minimal or near-minimal pairs, including *br̩ayən* ‘brew’ vs. *br̩ayən* ‘talk endlessly’, *blankən* ‘gleam’ vs. *bl̩nkən* ‘stray’, and *kv̩etshən* ‘squeeze’ vs. *kv̩itshən* ‘squeak’. The upshot is that although it may be possible to guess about the status of a final [ə] in some cases, it is still easier to choose a form that shows it unambiguously (a singular form or the 2pl).

The neutralizations discussed so far affected forms with overt suffixes – that is, all forms except the 1sg. The 1sg did not suffer such severe neutralizations, because no phonological processes affected segments in stem-final position.<sup>7</sup> It would not have been completely free from neutralizations, however; in fact, two properties of verbs could not have been predicted from the 1sg form alone. Umlaut verbs like *f̩ərən* would have had the same vowel (ɔ) as non-umlaut verbs like *pərən* in the 1sg, and preterite present and *Wechselflexion* verbs would likewise have been indistinguishable from non-alternating verbs in this form. A crucial difference from

<sup>7</sup>It appears that an earlier stage of Yiddish did have final devoicing, but this was lost early on in most dialects; see King (1980) for discussion.

### Base Selection in Analogical Change

the neutralizations discussed above, though, is that umlaut, preterite present, and *Wechselflexion* verbs would have been ambiguous with non-alternating verbs not only in the 1sg, but in *all* parts of the paradigm. (Recall that we are considering here a version of Yiddish prior to paradigm leveling; in actual Modern Yiddish, all of these verbs have uniform paradigms (7-9).) The shading in (13) shows that in some parts of the paradigm, these verbs were neutralized with the vowel in the middle column, while in other parts of the paradigm, they were neutralized with the vowel in the final column.

(13) Neutralizations that include the 1sg (pre-leveling forms)

a. Umlaut and non-umlaut verbs

	<i>fərən</i> 'travel'	<i>pərən</i> 'match'	<i>hern</i> 'hear'
1sg	<i>fər</i>	<i>pər</i>	<i>her</i>
2sg	<i>ferst</i>	<i>pərst</i>	<i>herst</i>
3sg	<i>fert</i>	<i>pərt</i>	<i>hert</i>
1pl	<i>fərən</i>	<i>pərən</i>	<i>hern</i>
2pl	<i>fərt</i>	<i>pərt</i>	<i>hert</i>
3pl	<i>fərən</i>	<i>pərən</i>	<i>hern</i>
infin.	<i>fərən</i>	<i>pərən</i>	<i>hern</i>

b. Preterite presents and *Wechselflexion*

	<i>visən</i> 'know'	<i>heysən</i> 'order'	<i>vishən</i> 'wipe'
1sg	<i>veys</i>	<i>heys</i>	<i>vish</i>
2sg	<i>veyst</i>	<i>heyst</i>	<i>vishst</i>
3sg	<i>veys(t)</i>	<i>heyst</i>	<i>visht</i>
1pl	<i>visən</i>	<i>heysən</i>	<i>vishən</i>
2pl	<i>vist</i>	<i>heyst</i>	<i>visht</i>
3pl	<i>visən</i>	<i>heysən</i>	<i>vishən</i>
infin.	<i>visən</i>	<i>heysən</i>	<i>vishən</i>

Since these neutralizations affect all parts of the paradigm, they do not favor any particular choice of base, and it is perhaps unnecessary to count the number of lexical items involved. It may be noted, however, that compared with the neutralizations discussed above, these affected a very small number of words. In MHG, umlaut occurred in a handful of verbs, mostly in strong classes VI and VII – perhaps less than two dozen altogether (Paul et al. 1989, §§ 251-253). Added to these were about a dozen preterite present verbs (§§ 269-275) and around 70 verbs in the *Wechselflexion* classes (IIIa, IV, V; §§ 247-250), totaling about 2% of verbs.

The combined effect of these neutralizations is summarized in Figure 1, which shows the number of lexical items whose underlying form could not be unambiguously recovered from each part of the paradigm. In sum, the 1sg form preserves the greatest number of phonemic distinctions, including the voicing of stem-final obstruents, the presence of stem-final *t*, *d*, *s*, and *z*, and the presence of stem-final *ə*. Thus, given a 1sg form, it would be possible to predict every form of every word with absolute certainty, with the exception of the 2sg/3sg of umlaut verbs and the plurals of preterite present and *Wechselflexion* verbs.

#### 2.2. Using the 1sg as the base to derive Yiddish verb paradigms

Suppose that you are acquiring a version of Yiddish prior to paradigm leveling. Your goal is to be able to produce and comprehend all forms of all verbs, and in order to do this, you need to learn their distinctive properties. I have shown that the 1sg provides almost all of them, and would thus be the optimal choice of base form to predict other forms. In the model proposed here, once the learner has identified the base, she goes on to develop a grammar to derive the rest of the paradigm from that form. For Yiddish, the grammar would include morphological rules like suffixing *-st* to form a 2sg, *-t* to form a 3sg or 2pl, *-ən* to form a 1pl/3pl/infinite,

## Adam Albright

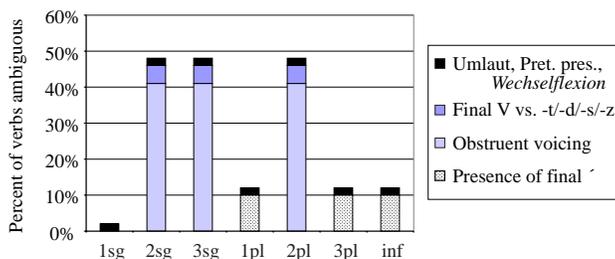


Figure 1: Summary of neutralizations affecting Yiddish verb forms

as well as phonological processes like obstruent voicing assimilation, degemination of /tt/ and /ss/, and elision of schwas in /ə/ sequences. If we use the 1sg as an input for these rules, they will yield the correct result for almost all forms of all words. The only exceptions are the 2,3sg of umlaut verbs, for which we predict incorrect forms like \**fɔrst* and \**fɔrt*, and the plurals of preterite present and *Wechselflexion* verbs, for which we predict incorrect forms like \**veysən* and \**gibən*. Under this approach, then, forms with umlaut (*ferst*) and with sg.~pl. alternations (*visən*) must be learned as exceptions.<sup>8</sup> If a speaker forgets or is unable to access the correct exceptional form, she will use the grammar to produce an “overregularized” form (*fɔrst*, *veysən*). If these mistakes are accepted and adopted by the speech community, they will eventually replace the old, exceptional forms. There are clearly many factors that determine how willingly a community adopts new forms; the thoroughness of the change in Yiddish may have been facilitated by the lack of a standard language or widespread literacy, and perhaps even by a conscious desire to differentiate Yiddish from German. The model that I am presenting here simply attempts to predict what the *potential* overregularizations would have been.

### 3. Comparison with other German dialects

The Yiddish leveling seems like a very natural change, even if its completeness is a bit striking. If it is really so natural, however, we would expect that it might also have occurred in some other related dialects. An informal survey of dialect descriptions revealed several candidates for dialects that superficially resemble to Yiddish in their present tense forms, but all turned out to have different explanations. In Dutch and some northern German dialects, the present singular paradigm is always uniform (*graaf*, *graaft*, *graaft* ‘dig’), as in Yiddish. However, these dialects never had umlaut to begin with, so this uniformity is not due to leveling. Some southern German dialects (Schwabian, Frankish, Bavarian, etc.) did historically have umlaut, and also have uniform present tense paradigms (e.g., Bavarian *grab*, *grabsd*, *grabd*) (Schirmunski 1962; Zehetner 1989). However, these dialects show leveling only of the singular forms, while maintaining singular ~

<sup>8</sup>It does not matter for present purposes whether these are stored as whole-word exceptions, or whether we posit rules that apply only to words that are lexically specified for them. All that matters is that *ferst* and *visən* cannot be derived productively, and require an overriding word-specific mechanism that may fail.

### *Base Selection in Analogical Change*

plural alternations in preterite present verbs. It appears that these dialects have lost the umlaut rule for the 2,3sg, rather than undergoing true paradigm leveling. Finally, Early NHG texts occasionally have *e* in the 2,3sg of *Wechselflexion* verbs, such as Fischart's *schmelzt* 'melts-3sg' (Standard German *schmilzt*) (Geyer 1912, §23.6, reminiscent of the generalization of *e* in Yiddish (9b)). This seems to happen mainly with verbs that are also sometimes given regular (weak) pasts, however. Weak verbs never have *Wechselflexion*, and its loss was probably part of a larger trend to create weak counterparts of strong verbs in Early NHG. In none of these cases do we find compelling evidence of paradigm leveling of the kind seen in Yiddish.

In contrast, there have been numerous changes in German that have introduced new alternations. The change from *i* to *e* in the 1sg of *Wechselflexion* verbs, for example, is usually seen as an extension of the umlaut pattern (raising/fronting in the 2,3sg), and umlaut has been extended to other verbs as well. In addition, many verbs have been rebuilt on the basis of 3sg forms, such as *ziemen* from Strong IV *zëmen*, and *wiegen* 'rock' from Strong V *wëgen* 'move'.

It would be difficult to prove that an equivalent leveling has never occurred in any other form of German, but my tentative conclusion is that German has generally gone in a different direction. So why would Yiddish have departed so radically in this respect? Considering the differences between Yiddish and German, we find that two of the neutralizations discussed above do not occur in German. First, the degemination of /dt/ and /tt/ to [t] ((11b) above) is found in only a few dialects (Schirmunski 1962), meaning the 3sg and 2pl forms preserve the voicing contrast between stem-final *t* and *d*, and keep both distinct from stem-final vowels. This is significant, because 562, or 12% of the verbs in the CELEX corpus end in coronal stops. Furthermore, German has no stem-final [ə], eliminating a major source of ambiguity in the 1pl/3pl/inf. forms. For these reasons, the 1sg form is not uniquely informative in Standard NHG; the plural and infinitive forms are just as good, and even the 2sg/3sg/2pl forms are not as ambiguous as in Yiddish.

#### **4. Discussion and conclusion**

In this paper, I have shown that the 1sg served as the base of a paradigm leveling in Yiddish, affecting almost every verb of the language. A comparison of neutralizations showed that before the change, the 1sg would have been uniquely revealing about the underlying form of the verb root. This is because the 1sg suffered from the fewest phonological neutralizations, involving the fewest lexical items. Comparing the neutralizations of Yiddish and German also provided some insight into why the 1sg may not be so privileged in related dialects.

The strategy of comparing neutralizations is rooted in a general model of paradigm acquisition, developed in Albright (in progress). The premise is that learners must be able to produce and understand forms they have never heard before, and they do this by focusing on the part of the paradigm that reveals properties of the word as unambiguously as possible. A hypothesis of this model is that learners must select a single surface form as the base or UR, even if it does not preserve every single contrast. This is similar in spirit to a proposal by Lahiri and Dresher (1984), who suggested that learners pay more attention to nominatives when learning the morphological class of nouns; the current approach is an attempt to generalize this,

and explain how learners might discover which forms to pay more attention to.

The calculations that I have been using in this paper are rather crude, but may serve as a conceptual example for a more rigorously defined, computationally implemented algorithm described in more detail in Albright (in progress). This algorithm considers each member of the paradigm as a potential base, and constructs stochastic grammars of morphological and phonological rules to derive the remaining forms. It then compares how “effective” these grammars are by calculating their accuracy, the number of exceptions needed, the reliability of the stochastic rules, and so on. The algorithm is shown to select the right base in not only the Yiddish case, but in others as well, including in Latin noun paradigms (Albright, to appear), Lakhota paradigm innovations, and Spanish verbs.

It is useful to compare this model against one without the single surface form restriction. Under a traditional approach, learners could notice that some contrasts (like obstruent voicing and final schwas) are seen in some forms, while other contrasts (like umlaut) are seen in others, combining multiple surface forms to create a lexical entry that captures *all* unpredictable information. By comparing 1sg and 3sg forms, for example, learners could set up an underlying distinction between non-alternating forms (/pɔr-/ ‘match’) and alternating forms (/fɔr-/~/fer-/, or /fɔr-/[+umlaut] ‘travel’). Under this model, there are various possible sources of error. A learner could have incomplete information about a word, failing to learn or recall that it has both [ɔ] and [e] allomorphs (or that it undergoes the *umlaut* rule), and incorrectly produce 2sg \*fɔrst without *umlaut*. A speaker could forget or not know that the singular of a particular *Wechselflexion* verb uses a different root allomorph, and incorrectly extend the plural vowel, producing 1sg \*vis instead of *veys*. There are few formal models of how learners learn URs and reason about only partially known words, and it is difficult to make exact predictions without one. However, the general point is this: if learners can construct URs from multiple parts of the paradigm, then we expect different verbs could potentially level to different parts of the paradigm. A traditional model does not explain why contrasts preserved in a particular surface form (the 1sg) were consistently maintained, while contrasts neutralized in that form were systematically lost.

The single surface base restriction, on the other hand, prevents learners from storing absolutely all unpredictable information in the UR, and requires more forms to be listed as exceptional. The learning procedure mitigates this problem by finding the URs and rules that minimize the number of stored exceptions, by selecting the base form that preserves the most contrasts. The prediction is that contrasts preserved in the base will be maintained, while contrasts neutralized in the base will be open for leveling – which, in the case of Yiddish, appears to be correct.

## References

- Albright, A. (in progress). *The identification of bases in paradigms: A computational approach*. Ph. D. thesis, UCLA.
- Albright, A. (to appear). The morphological basis of paradigm leveling. In *Proceedings of the ZAS Conference on Paradigm Uniformity, March 2001*.

*Base Selection in Analogical Change*

- Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing Company.
- Dammers, U., W. Hoffmann, and H.-J. Solms (1988). *Grammatik des Frühneuhochdeutschen, Vol. 4: Flexion der starken und schwachen Verben*. Heidelberg: Carl Winter Universitätsverlag.
- Geyer, A. (1912). *Die starke Konjugation bei Johann Fischart*. Hofbuchdruckerei C. A. Kaemmerer & Co.
- Kenstowicz, M. (1995). Base-identity and uniform exponence: Alternatives to cyclicity. In J. Durand and B. Laks (Eds.), *Current Trends in Phonology: Models and Methods*. CNRS and University of Salford Publications.
- Kern, P. O. (1903). *Das starke Verb bei Grimmelshausen: Ein Beitrag zur Grammatik des Frühneuhochdeutschen*. Chicago: University of Chicago.
- King, R. D. (1980). The history of final devoicing in Yiddish. In M. I. Herzog, B. Kirshenblatt-Gimblett, D. Miron, and R. Wisse (Eds.), *The Field of Yiddish: Studies in Language, Folklore, and Literature, Fourth Collection*, pp. 371–430. Philadelphia: Institute for the Study of Human Issues.
- Kiparsky, P. (1965). *Phonological Change*. Ph. D. thesis, MIT.
- Kuryłowicz, J. (1947). The nature of the so-called analogical processes. *Diachronica* 12(1), 113–145. (trans. Margaret Winters).
- Lahiri, A. and B. E. Drescher (1984). Diachronic and synchronic implications of declension shifts. *The Linguistic Review* 3, 141–163.
- Mańczak, W. (1958). Tendences générales des changements analogiques. *Lingua* 7, 298–325 and 387–420.
- Paul, H. (1920). *Prinzipien der Sprachgeschichte* (5th ed.). Halle: Niemeyer.
- Paul, H., P. Wiehl, and S. Grosse (1989). *Mittelhochdeutsche Grammatik* (23 ed.). Tübingen: Max Niemeyer Verlag.
- Philipp, G. (1980). *Einführung ins Frühneuhochdeutsche: Sprachgeschichte, Grammatik, Texte*. Heidelberg: Quelle und Meyer.
- Rockowitz, A. (1979). *201 Yiddish Verbs, fully conjugated in all the tenses*. Woodbury, NY: Barron's Education Series, Inc.
- Schirmunski, V. (1962). *Deutsche Mundartkunde*. Berlin: Akademie-Verlag.
- Steriade, D. (2000). Paradigm Uniformity and the Phonetics/Phonology boundary. In J. Pierrehumbert and M. Broe (Eds.), *Papers in Laboratory Phonology, vol. 6*. Cambridge University Press.
- Tiersma, P. (1982). Local and general markedness. *Language* 58(4), 832–849.
- Weinreich, U. (1990). *Modern English-Yiddish, Yiddish-English Dictionary*. New York: Yivo Institute for Jewish Research.
- Zehetner, L. (1989). *Bairisches Deutsch*. Hugendubel.

Adam Albright  
Dept. of Linguistics, UCLA  
405 Hilgard Ave.  
Los Angeles, CA 90095-1543

aalbrigh@ucla.edu