Productivity of morphological patterns and social domain analysis in Papuan Malay

Angela Kluge
SIL International
SEALS25
Payap University, Chiang Mai, 27-29 May 2015
Outline

- Productivity of morphological patterns
  - Multifaceted approach
  - Diglossic situations with closely related LOW and HIGH varieties
  - Papuan Malay

- Examined language factors
  - Language internal factors
  - Language external factors

- Prefix BER- ‘VBLZ’
  - Language internal factors
  - Language external factors
  - Summary & conclusions
Introduction

- Papuan Malay
  - Diglossic distribution (see Ferguson 1972, Weinreich 1953)
    - LOW: Papuan Malay
    - HIGH: Indonesian
  - 16-hour corpus

(based on Donohue to appear, and Kim et al. 2007)
Introduction

- Attested morphologically complex lexical items
  - TER- ACCIDENTAL
  - PE(N)- AGENT
  - BER- VERBALIZER
  - -ang PATIENT
  - -nya POSSESSOR
  - ke-/ang NOMINALIZER
Introduction

- Attested morphologically complex lexical items
  - TER- ACCIDENTAL
  - PE(N)- AGENT
  - BER- VERBALIZER
  - -ang PATIENT
  - -nya POSSESSOR
  - ke/-ang NOMINALIZER
Introduction

- Investigating the productivity of BER- ‘VBLZ’
  - Productivity testing
    - ‘Productivity experiment’ (Aronoff 1978)
Introduction

- Investigations of productivity in Papuan Malay
  - Sociolinguistic profile of Papuan Malay
  - High degree of linguistic relatedness between Papuan Malay and Indonesian
  - Formal setting of a test situation

⇒ Unfeasibility of productivity testing
  - Interference from Indonesian (see Weinreich 1953)
  - Skewing of testees’ naïve judgments (see Grosjean 1992:59)

⇒ Multifaceted approach to investigate productivity
Introduction

- Language internal factors
  1. Syntax
  2. Type and token frequencies
  3. Relative token frequencies of derivations and base words
  4. Form-function relationship between the derivation and its base
  5. Alternative strategies
  6. Formally complex words
  7. Code-switches with Indonesian

- Language external factors
  1. Speaker education levels
  2. Topics
  3. Relationships between interlocutors
Language internal factors

1. Syntactic properties
   - Polyfunctionality (Booij 2002: 89-91; see also Zwanenburg 2000)

2. Type and token frequencies
   - High type frequency plus low token frequency (Plag 2006:542)

3. Relative token frequencies of derivations and bases
   - Token frequencies for derivations lower than for their bases (Hay and Baayen 2002)

4. Form-function relationship between derivations and bases
   - Transparent form-function relationship (Booij 2007: 240, 323)

⇒ Affixation process is more productive
Language internal factors

5. Alternative strategies
   - Use of unaffixed lexemes which express the same meaning

6. Formally complex words
   - Large numbers of formally complex words with non-compositional semantics (Booij 2007:17)

7. Code-switches with Indonesian
   - Large number of derivations that are code-switches with Indonesian rather than part of the Papuan Malay lexicon

⇒ Affixation process is less productive
Language external factors

- Domain analysis (Fishman 1965:86)
  - Investigated factors
    - Speaker education levels
    - Topics
    - Relationships between interlocutors
Language external factors

1. Speaker education levels
   - Derivations are used more often by better-educated speakers than by less-educated speakers

2. Topics
   - Derivations are used more often for HIGH topics (politics, education, religion, etc.) than for LOW topics (casual daily-life issues) (see Fishman 1965:71)

3. Relationships between interlocutors
   - Derivations are used more often with higher-status interlocutors than with equal/lower-status interlocutors (see Fishman 1965:75)

⇒ Affixed lexemes constitute code-switches with Indonesian
Language internal and external factors

- Language internal factors
  1. Syntax
  2. Type and token frequencies
  3. Relative token frequencies of derivations and base words
  4. Form-function relationship between the derivation and its base
  5. Alternative strategies
  6. Formally complex words
  7. Code-switches with Indonesian

- Language external factors
  1. Speaker education levels
  2. Topics
  3. Relationships between interlocutors
BER- ‘VBLZ’ (78 items with 695 tokens)

- **Verbal bases:** 29 derivations (227 tokens)
  
  (1) waktu saya ber-buru saya perlu makang pinang
  
  time 1SG VBLZ-hunt 1SG need eat betel.nut
  
  ‘when I hunt I need to chew betel nuts’

- **Nominal bases:** 30 derivations (362 tokens)
  
  (2) bapa pergi ber-kebung saya ikut
  
  father go VBLZ-garden 1SG follow
  
  ‘(whenever my) father went to do farming I went with (him)’

- **Numeral and quantifier bases:** 3 derivations (13 tokens)

- **Formally complex words lacking a base:** 16 items (93 tokens)
BER-: Language internal factors

1. Syntactic properties: Verbal bases
   □ Derivations with monovalent bases: 11 items (32 tokens)

<table>
<thead>
<tr>
<th>Base</th>
<th>Item</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>bingung</td>
<td>berbingung</td>
<td>‘be confused’</td>
</tr>
<tr>
<td>diam</td>
<td>badiam</td>
<td>‘be quiet’</td>
</tr>
<tr>
<td>ibada</td>
<td>beribada</td>
<td>‘worship’</td>
</tr>
<tr>
<td>jalang</td>
<td>berjalan</td>
<td>‘walk’</td>
</tr>
<tr>
<td>sandar</td>
<td>bersandar</td>
<td>‘lean’</td>
</tr>
<tr>
<td>tobat</td>
<td>bertobat</td>
<td>‘repent’</td>
</tr>
</tbody>
</table>
BER-: Language internal factors

1. Syntactic properties: Verbal bases
   - Derivations with bivalent bases: 18 items (195 tokens)

<table>
<thead>
<tr>
<th>Base</th>
<th>Item</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derivations: Intransitive uses (13 items)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bicara</td>
<td>berbicara</td>
<td>‘speak’</td>
</tr>
<tr>
<td>kerja</td>
<td>bekerja</td>
<td>‘work’</td>
</tr>
<tr>
<td>pikir</td>
<td>berpikir</td>
<td>‘think’</td>
</tr>
<tr>
<td>Derivations: Mono- &amp; intransitive uses (5 items)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>buru</td>
<td>berburu</td>
<td>‘hunt’</td>
</tr>
<tr>
<td>harap</td>
<td>berharap</td>
<td>‘hope’</td>
</tr>
<tr>
<td>rebut</td>
<td>bribut</td>
<td>‘trouble’</td>
</tr>
</tbody>
</table>
BER-: Language internal factors

1. Syntactic properties: Nominal bases
   - Attested: 30 items (362 tokens)

<table>
<thead>
<tr>
<th>Base</th>
<th>Item</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>arti</td>
<td>brarti</td>
<td>‘mean’</td>
</tr>
<tr>
<td>bua</td>
<td>berbua</td>
<td>‘have fruit’</td>
</tr>
<tr>
<td>doa</td>
<td>berdoa</td>
<td>‘pray’</td>
</tr>
<tr>
<td>hasil</td>
<td>berhasil</td>
<td>‘succeed’</td>
</tr>
<tr>
<td>kebung</td>
<td>berkebung</td>
<td>‘do farming’</td>
</tr>
<tr>
<td>sykur</td>
<td>bersykur</td>
<td>‘give thanks’</td>
</tr>
</tbody>
</table>
BER-: Language internal factors

1. Syntactic properties: Numeral and quantifier bases
   - Attested: 3 items (13 tokens)

<table>
<thead>
<tr>
<th>Base</th>
<th>Item</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeral bases (2 items with 7 tokens)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>empat</td>
<td>berempat</td>
<td>‘be four’</td>
</tr>
<tr>
<td>satu</td>
<td>bersatu</td>
<td>‘be one’</td>
</tr>
<tr>
<td>Quantifier bases (1 item with 6 tokens)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>brapa</td>
<td>bebrapa</td>
<td>‘be several’</td>
</tr>
</tbody>
</table>
BER-: Language internal factors

2. Type and token frequencies: Derivations with ≤10 tokens (56/62 items – 90%)

- 27/29 derivations with verbal bases
  - berburu ‘hunt’ (10 tokens)
  - bertobat ‘repent’ (8 tokens)
  - berpikir ‘think’ (8 tokens)
  - berbicara ‘speak’ (7 tokens)

- 29/33 derivations with nominal, numeral, or quantifier bases
  - berdosa ‘sin’ (6 tokens)
  - bersaksi ‘testify’ (6 tokens)
  - berhasil ‘succeed’ (6 tokens)
  - berkwasa ‘be powerful’ (4 tokens)
BER-: Language internal factors

3. Relative token frequencies of derivations and bases (43/62 items – 70%)
   - 22/29 derivations with verbal bases
     - berjalang ‘walk’ (1 vs. 480 tokens)
     - berbicara ‘speak’ (7 vs. 333 tokens)
     - bekerja ‘work’ (5 vs. 191 tokens)
     - berbuat ‘make’ (7 vs. 100 tokens)
   - 21/33 derivations with nominal, numeral, or quantifier bases
     - bermalam ‘overnight’ (2 vs. 191 tokens)
     - berbahasa ‘speak’ (2 vs. 136 tokens)
     - bebrapa ‘be one’ (6 vs. 109 tokens)
     - berempat ‘be four’ (1 vs. 66 tokens)
BER-: Language internal factors

4. Form-function relationship
   - 27/29 derivations with verbal bases: Same semantics
     - Bivalent bases (16/18)

   (3) ... kitorang bingung pikir itu penjaga kuburang
       1PL be.confused think D.DIST guard grave
       ‘[so] we’re confused to think (about), what’s-its-name, a guard (for) the grave’

   (4) ... dong tida taw ber-pikir itu
       3PL NEG KNOW VBLZ-think D.DIST
       ‘[but these kids] they don’t know (how) to think (about) those (feelings of mine)’
4. Form-function relationship

Derivations with nominal, numeral, or quantifier bases

- General meaning: ‘be/have/do BASE’

(5) … jadi saya dengang kaka Nofita masi br-ade-kaka
    so 1SG with eSb Nofita still VBLZ-siblings
    ‘so I and older sister Nofita are still siblings’

(6) … ada yang berduri ada yang tida …
    exist REL VLBZ-thorn exist REL NEG
    ‘[there are … two kinds (of sago palms),] ones that have thorns and ones that don’t (have thorns) …’
5. Alternative strategies

- Derivations with nominal, numeral, or quantifier bases
  - Preference for analytical constructions
    - ‘do BASE’: Use of alternative verbs

(7) bapa pergi ber-kebung saya ikut
    father go VBLZ-garden 1SG follow
    ‘(whenever my) father went to do farming I went with (him)’

(8) kalo di Arbais prempuang bisa biking kebung
    if at Arbais woman be.able make garden
    ‘as for Arbais, (there) the women can work a garden’
BER-: Language internal factors

6. Formally complex words with non-compositional semantics
   - High-frequency derivations: e.g. berdoa ‘pray’ (136 tokens)
     (9) bebang masala de punya dia perlu …
         burden problem 3SG POSS 3SG need
         harus di-ber-doa
         have.to UV-VBLZ-pray
     [Conversation about problems of a church congregation:]
     ‘(all) burdens (and) problems (that) it has, (the congregation) needs … has to be prayed for’
6. Formally complex words with non-compositional semantics
   - Derivations lacking a base (16 items with 93 tokens)
     (16/78 items – 20%)
     - bakalai  ‘fight’
     - bergaul  ‘associate’
     - berhenti  ‘stop’
     - berjuang  ‘struggle’
     - berlabu  ‘anchor’
     - bertriak  ‘scream’
BER-: Language internal factors

7. Code-switches with Indonesian (tentatively identified)
   (49/62 items – 79%)
   - 24/29 derivations with verbal bases
     - bekerja ‘work’
     - berbicara ‘speak’
     - berbuat ‘make’
     - berjalang ‘walk’
   - 25/33 derivations with nominal, numeral, or quantifier bases
     - bebrapa ‘be serval’
     - berbua ‘have fruit’
     - berkebung ‘do farming’
     - bersaksi ‘testify’
BER-: Language external factors

- Domain analysis
  - Speaker education levels
  - Topics
  - Relationships between interlocutors

- Investigated items: 56 derivations with 164 tokens
  - 27 lexemes with verbal bases (94 tokens)
  - 29 lexemes with nominal, numeral, quantifier bases (70 tokens)
BER-: Language external factors

- Excluded from domain analysis
  - Five derivations with >50 tokens each
    - 2 lexemes with verbal bases
      - blajar ‘study’, brangkat ‘leave’
    - 3 lexemes with nominal bases
      - berdoa ‘pray’, berdiri ‘stand’, brarti ‘mean’
  - One derivation with unreliable data
    - berusaha ‘attempt’
BER-: Language external factors

- Domain analysis
  - 148/164 tokens (90%)
    - Conditioned by variables of the communicative event
      - Speaker education levels
      - Topics
      - Relationships between interlocutors
  - 16/164 tokens (10%)
    - Unaccounted for
      - -EDC speakers conversed with -STAT fellow-Papuans about LOW topics
BER-: Language external factors

- Domain analysis
  - 148/164 tokens (90%) accounted for
  - 16/164 tokens (10%) unaccounted for

<table>
<thead>
<tr>
<th></th>
<th>+EDC-SPK</th>
<th>-EDC-SPK</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH topic</td>
<td>71</td>
<td>24</td>
<td>95</td>
</tr>
<tr>
<td>Outsider</td>
<td>14</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>LOW topic</td>
<td>18</td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>+STAT-SPK</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>-STAT-SPK</td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>61</td>
<td>164</td>
</tr>
</tbody>
</table>
BER-: Language external factors

- Domain analysis
  - 148/164 tokens (90%) accounted for
  - 16/164 tokens (10%) unaccounted for

<table>
<thead>
<tr>
<th></th>
<th>+EDC-SPK</th>
<th>-EDC-SPK</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH topic</td>
<td>71</td>
<td>24</td>
<td>95</td>
</tr>
<tr>
<td>Outsider</td>
<td>14</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>LOW topic</td>
<td>18</td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>+STAT-SPK</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>-STAT-SPK</td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>61</td>
<td>164</td>
</tr>
</tbody>
</table>
BER-: Language external factors

- Domain analysis
  - 148/164 tokens (90%) accounted for
  - 16/164 tokens (10%) unaccounted for

<table>
<thead>
<tr>
<th></th>
<th>+EDC-SPK</th>
<th>-EDC-SPK</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH topic</td>
<td>71</td>
<td>24</td>
<td>95</td>
</tr>
<tr>
<td>Outsider</td>
<td>14</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>LOW topic</td>
<td>18</td>
<td>11</td>
<td>45</td>
</tr>
<tr>
<td>+STAT-SPK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-STAT-SPK</td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>61</td>
<td>164</td>
</tr>
</tbody>
</table>
**BER-: Language external factors**

- **Domain analysis**
  - 148/164 tokens (90%) accounted for
  - 16/164 tokens (10%) unaccounted for

<table>
<thead>
<tr>
<th></th>
<th>+EDC-SPK</th>
<th>-EDC-SPK</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIGH topic</td>
<td>71</td>
<td>24</td>
<td>95</td>
</tr>
<tr>
<td>Outsider</td>
<td>14</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>LOW topic</td>
<td>18</td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>+STAT-SPK</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>-STAT-SPK</td>
<td></td>
<td>16</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>103</td>
<td>61</td>
<td>164</td>
</tr>
</tbody>
</table>
BER-: Language external factors

- Domain analysis
  - 148/164 tokens (90%) accounted for
  - 16/164 tokens (10%) unaccounted for
    - 13/16 items: Classified as code-switches with Indonesian
      - berbeda ‘be different’
      - bersodara ‘be sibling’
      - bebrapa ‘be several’
    - 3/16 items: Classified as Papuan Malay
      - berhasil ‘be successful’
      - bruba ‘change’
      - bagaya ‘put on airs’
BER-: Summary & conclusions

- Language internal factors
  - Polyfunctionality
    - Verbal, nominal, numeral, quantifier bases
  - Form-function relationship
    - Transparent for derivations with nominal, numeral, and quantifier bases
  - Type and token frequencies
    - Large number of low frequency words
  - Relative token frequencies of derivations and their bases
    - Most derivations have lower frequencies than their bases

⇒ BER- affixation is rather productive
BER-: Summary & conclusions

Language internal factors
- Form-function relationship
  - Same semantics as their bases (for verbal bases)
- Alternative strategies
  - Analytical constructions are preferred (for nominal bases)
- Formally complex words
  - Substantial number of derivations lacking a base
  - High frequency derivations with non-compositional semantics
- Code-switches
  - Most derivations classified as code-switches with Indonesian

⇒ BER- affixation is rather unproductive
BER-: Summary & conclusions

- Language internal factors
  - Productivity of BER- affixation is questionable
BER-: Summary & conclusions

- Language external factors
  - 90% of tokens (148/164) accounted for in terms of variables of the communicative event
    - Speaker education levels
    - Topics
    - Relationships between interlocutors
  - 10% of tokens (16/164) unaccounted for
  - **BER- affixation is rather unproductive**
Conclusion

Ber- affixation is rather unproductive

(10) satu kali tong pergi berdoa
    one time 1PL go pray
    ‘one time we went pray’

(11) kalo saya berburu tida pake anjing ...
    if 1SG hunt NEG use dog
    ‘if I hunt (and) don’t take dogs, …’
References