Kalamang fillers and placeholders

Kalamang is an endangered Greater West Bomberai language spoken in eastern Indonesia. I will present a detailed analysis of fillers and placeholders in seven sample texts from the publicly available Kalamang corpus (Visser 2020a). All texts are conversations and they involve six different speakers. Totalling 5200 words, the texts contain 105 fillers and 77 placeholders. This study complements and elaborates on the first description of Kalamang fillers and placeholders in Visser (2020b). It reveals more forms for both fillers and placeholders than previously described for Kalamang, and a much higher ratio of placeholders than previously calculated.

Placeholders
The main Kalamang placeholder *neba* is originally the question word 'what'. In the sample texts, it is used as a replacement for nouns, verbs, adverbials and attributes, and is always inflected accordingly. Hence, *neba* occurs in the sample with possessive suffixes, postpositional clitics, attributive markers, focus markers, predicate linkers, aspectual markers, and reduplicated. (1) and (2) illustrate placeholder *neba* in nominal and verbal function, respectively.

(1) *in*   *bo*   *yakop*   **neba-un=ko**   *pos-un=ko*
     1PL.EX go   cockatoo   PH-3POSS=LOC   hole-3POSS-=LOC
     'We go to the cockatoo's whatsit, hole.'
(2) *pen*   *koi*   **neba~neba**   *gosomin~gosomin*
     tastiness   then   PH~ATT   disappear~ATT
     'Then the tastiness whatsits a little, disappears a little.'

Placeholder *neba* was previously reported to occur with a frequency of 6 per 1000 words (Visser 2020b:475), but in the current sample the frequency of *neba* in the function of placeholder is 13.5 per 1000 words. If we also take into account the other placeholders *don* 'thing', *apa* 'what' (< Indonesian) and the question word root *taman*, this ratio rises to 14.7 per 1000 words. The discrepancy can be explained by the more careful coding that was undertaken for this study. Kalamang has a much higher placeholder ratio than Russian, which has 5 per 1000 words (Podlesskaya & Kibrik 2006), or Chinese, which has 6.68 per 1000 words (Zhao & Jurafsky 2005).

The raw frequency of the different placeholders can be found in Table 1.

Table 1. Frequency of placeholders.

Fillers
The Kalamang sample analysed here reveals a great number of fillers. The most common one is *nain* 'like', followed by *a* 'ehm'. Thereafter, lengthening and repetition are commonly used strategies. Other words used as fillers include *eba* 'then', *jadi* 'so' (< Indonesian) and *koi* 'then'. For all fillers, the exact phonetic form was coded as well. Nine different pronunciations for *a* (Table 2) and eight different pronunciations for *nain* (Table 3) were counted. When taking into account the speakers, it turns out that most speakers use [a:] for *a* and [nain] or [nain:] for *nain*. This data can help choosing a standardised spelling for fillers.
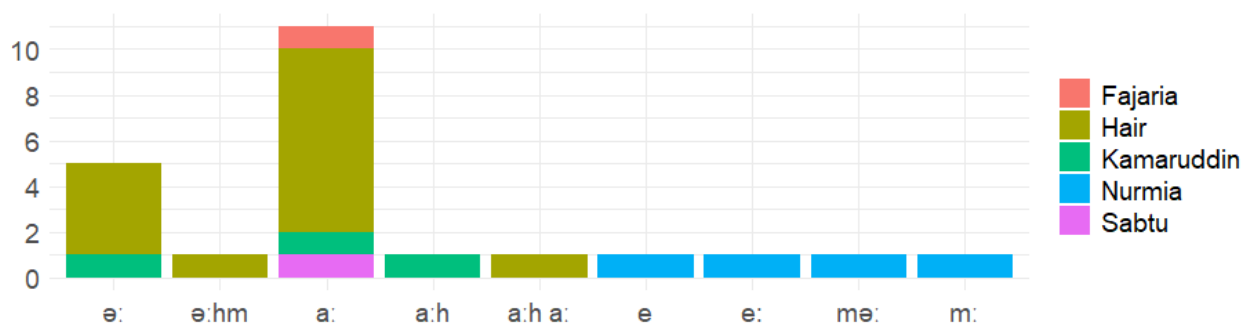
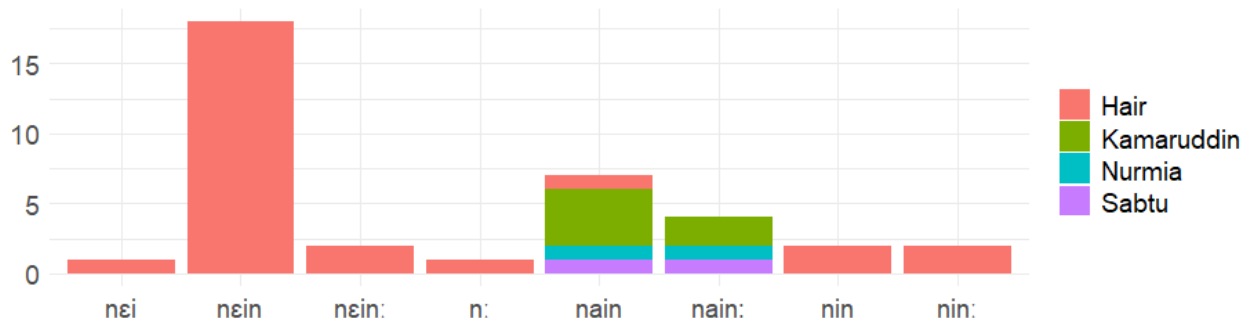Table 2. Pronunciation variants for the filler *a* 'ehm'.

Table 3. Pronunciation variants for the filler *nain* 'like'.

The frequency of fillers is 20 per 1000 words. Notable is the large range in frequencies for individual speakers: in one text, one speaker has 103 fillers per 1000 words, while his interlocutor only uses 13 fillers per 1000 words.

This study shows that careful coding of fillers and placeholders can:
- reveal higher frequencies than is commonly assumed
- help choosing a standardised spelling for fillers
- reveal big individual differences in frequency of use

Abbreviations not in Leipzig Glossing Rules
ATT = attenuative

References

Podlesskaya, Vera I. & Kibrik, Andrej A. 2006. Korrekcija v ustnoj monologičeskoj reči po dannym korpusnogo issledovanija (A corpus study of repairs in spoken monologues). *Russkij jazyk v naučnom osveščenii 2*(12): 7–55.

Visser, Eline. 2020a. *The Kalamang collection: An archive of linguistic and cultural material from Karas*. Lund: Lund University Humanities Lab. http://hdl.handle.net/10050/00-0000-0000-0003-C3E8-1

Visser, Eline. 2020b. *A grammar of Kalamang, the Papuan language of Karas Islands*. Lund University PhD thesis.

Zhao, Yuan & Jurafsky, Dan. 2005. A preliminary study of Mandarin filled pauses. In *Proceedings of DiSS'05, Disfluency in Spontaneous Speech Workshop*. Aix-en-Provence, France, 10–12 September 2005, 17–182.