Repetition reduction in Papuan Malay prosody

Constantijn Kaland, Christoph Bracks, Nikolaus P. Himmelmann

Institute of Linguistics, University of Cologne, Germany
{ckaland, c.bracks, sprachwissenschaft}@uni-koeln.de

Abstract

It has frequently been shown that speakers prosodically reduce repeated words in discourse. This phenomenon has been claimed to facilitate speech recognition and to be language universal. However, virtually all evidence for repetition reduction comes from English. The current study investigates to what extent repetition reduction in prosody is found in Papuan Malay. The prosody of Papuan Malay is under-researched and not yet well understood. In the current study, we hypothesize that Papuan Malay speakers show prosodic reduction of repeated words. In order to investigate this, an acoustic analysis is carried out on repeated words in short stories produced by native Papuan Malay speakers. The results show that for repeated mentions duration is reduced and mean pitch is higher. It is concluded that these findings are partially compatible with current theories on repetition and prominence.

Index Terms: prosody, reduction, repetition, prominence, Papuan Malay.

1. Introduction

Research has shown that speakers tend to prosodically reduce repeated words in discourse (e.g. [1]). For example, when a speaker starts mentioning the new car he or she bought, the first production of the word "car" is generally more carefully articulated compared to subsequent productions of the same word. That is, the word's prosodic characteristics such as pitch, duration and intensity are generally reduced for repeated mentions. Repetition reduction has also been found in other domains than prosody. For example, upon repetitively referring to the same entity, speakers tend to use lexical alternatives that are shorter; i.e. "it" instead of "the car" (e.g. [2]) and tend to reduce their co-speech gestures [3]. Most research on the prosodic effects of repetition has been done on well-known languages, in particular on English. As repetition effects overlap with (de)accentuation in this language, more research needs be done on languages that have fundamentally different prosody. In this study, we investigate Papuan Malay, a language spoken in the western part of Papua (Indonesia), which presumably does not make use of pitch accents. The remainder of the introduction is organized as follows. First, literature on repetition reduction in prosody is discussed. Second, existing literature on the prosody of Papuan Malay and related languages is presented.

1.1. Repetition reduction in prosody

Which linguistic factors exactly drive prosodic reduction has been topic of various studies. It has been shown that information redundancy [4] as well as lexical frequency and probability in a given context [5] affect the prosodic reduction of words. That is, the more a word is redundant, frequent or probable in a given context, the more likely the speaker is to reduce this word prosodically. In the case of repetition, a word is more redundant because it has been mentioned in recent discourse. The Smooth Signal Redundancy Hypothesis (SSRH; [4]) assumes a language universal inverse relationship between informational redundancy and the acoustic signal, such that for low redundant words more acoustic energy is spent, whereas for highly redundant words, the acoustic signal is attenuated. In this way, the acoustic signal, by means of prosodic prominence, has a mediating effect on language redundancy to maintain a robust communication process. The SSRH is compatible with current theories on prosodic prominence and intonation. For example, Germanic languages commonly use pitch accents to highlight new information and deaccent given information (e.g. [6]). New information, generally low in redundancy, stands out as more prominent compared to given information, which is likely to be redundant or predictable from context.

Experimental investigations have focused on the conditions under which repetition leads to prosodic reduction in speech production. For example, it has been shown that repeated words are shortened when produced in a meaningful context (i.e. a phrase or a read paragraph) instead of a list. Even more shortening occurs when a listener is physically present ([7], [8]). Generally, content words are taken into account in studies on repetition reduction, as function words show little to no reduction [9]. From these studies it appears that repetition effects are likely to occur where meaningful inferences can be made from the linguistic context (i.e. discourse context and content words). However, more recent studies have shown that repetition effects in prosody can also be found when these inferences cannot be made. That is, it has been shown that reduction occurs upon lexical repetition only, even when the discourse referent of the first and second mention is not the same [10]. Furthermore, auditory memory appears to influence whether a repeated mention is prosodically reduced. That is, second mentions that are homophonous to the first mention show reduction [11]. These recent studies show that the effects of repetition on prosody exist in roughly two categories. On the one hand, there are repetition effects that closely correlate with information status changes (new/given). On the other hand, the literature reports independent effects that can only be attributed to repetition proper (lexical/homophonous repetition).

Some studies also investigated how repeated mentions are perceived. In [1] it was found that second mentions were less intelligible than first mentions when presented in isolation. Crucially, this effect was not found when second mentions were presented in context. A similar effect was found in [12], where fluent speech and speech with errors followed by self-repairs were compared. It was found that second mentions were less intelligible for self-repairs, not for fluent speech. In [12], a second experiment was carried out to test the effect of

10.21437/SpeechProsody.2018-2

discourse status; i.e. whether second mentions were referring to discourse new or given information. Reduction was then only found for repeated words referring to given information. These studies showed that in case of perception, (the presence of) referential context appears to be crucial to understand repetition effects.

Studies on repetition reduction most often investigate duration as acoustic parameter. Pitch and intensity are often not taken into account and if they are, they do not show consistent effects. That is, in some studies, pitch showed no effect ([1], [13]) and in another study pitch was lowered for repeated mentions [11]. However, according to the SSRH [4] pitch and intensity are, like duration, expected to correlate with prosodic prominence to maintain language redundancy. Furthermore, virtually all of the studies on repetition reduction have investigated English speaking participants. In the current study we argue that more research is needed on lesser known languages. This will not only test the universality of the SSRH, it will also advance our knowledge of the prosody of lesser known languages. In particular, it has to be investigated to what extent repetition reduction occurs in languages that do not make use of pitch accents like English. In the current study we therefore focus on Papuan Malay.

1.2. Trade Malay prosody

Papuan Malay is a variety of Trade Malay spoken in the Indonesian provinces Papua and Papua Barat [14]. Other than what is mentioned in descriptive grammars of Trade Malay varieties, little is known about their prosody. To date, there are two empirical studies that investigated the prosody of Ambonese Malay [15] and Papuan Malay [16] respectively. For Ambonese Malay, no consistent evidence for the use of pitch accents was found [15]. It was concluded that the common rising intonation pattern observed on the (pen)ultimate syllables in a phrase could be interpreted as a 'floating boundary tone'. Similarly, in Papuan Malay it was found that native annotators reached considerably more agreement on indicating prosodic boundaries compared to indicating prosodic prominences [16]. From these recent empirical studies, it appears that the main function of prosody in Papuan Malay is rather demarcative. Such a conclusion breaks with earlier claims on the commonly observed prominences at the right edge of a phrase, as either being lexical stresses or pitch accents (e.g. [17], [18]).

No study to date investigated prosodic reduction in discourse in Trade Malay varieties and most of the work on repetition reduction is based on English speaking participants. In the current study we focus on prosodic effects of repetition only, as no corpus is available to assess word frequency or probability. On the basis of the universality of the SSRH [4], we hypothesize that speakers of Papuan Malay prosodically reduce repeated mentions in similar ways compared to known languages. In other words, the cognitive mechanism that smoothens the redundancy of the speech signal in order to maintain robust communication is assumed to be equally important for any spoken language. Alternatively, it is possible that that repetition reduction does not occur in Papuan Malay, or occurs in a different way. Previous work suggested that prosody in Trade Malay varieties functions to mark boundaries rather than prominences ([15],[16]). Following these conclusions, it is plausible that the redundancy of repeated information is does not affect the prosody of these languages.

2. Methodology

To investigate the extent to which speakers of Papuan Malay prosodically reduce repeated mentions in discourse, a story retelling task was carried out. In this task speakers were instructed to watch a short video clip and retell what they had seen to an interlocutor who did not see the video. The video clip showed a short story about a man picking pears. The actors in the video clip did not use any speech. The video clip has been previously used in cross-linguistic studies on narrative production (Pear Film; [20]).

2.1. Design and procedure

Recordings were made at the Center for Endangered Languages Documentation (CELD) in Manokwari, West Papua ([21]). Participants received instructions about the experimental procedure before the start of the task. They were verbally instructed to watch the video clip in order to retell the story displayed in the clip to an interlocutor who did not see it. The video clip was shown to the participants on a laptop. Thereafter, participants were introduced to their interlocutor, to which they were then instructed to retell the story from the video clip. The participants and interlocutor were seated next to each other during the retelling. The interlocutor was allowed to ask clarification questions during the participant's retelling. This happened up to three times per telling.

No soundproof or silent rooms were available at the recording location. Recordings were therefore made outside, behind a building where background noise was minimal. The recordings were made using a Sony ECM-MS957 microphone connected to a Sony HDR-SR11 portable video camera. The microphone was placed in front of the participant and interlocutor and recorded the speech of both. The experimenter supervised the entire recording procedure. The duration of the collected recordings ranged between two and five minutes.

2.2. Participants

All participants were students at the University of Papua. There were 10 male and 9 female participants (M_{age} = 22, age range = 20-28). All were native speakers of Papuan Malay without speech problems.

2.3. Data selection and acoustic analysis

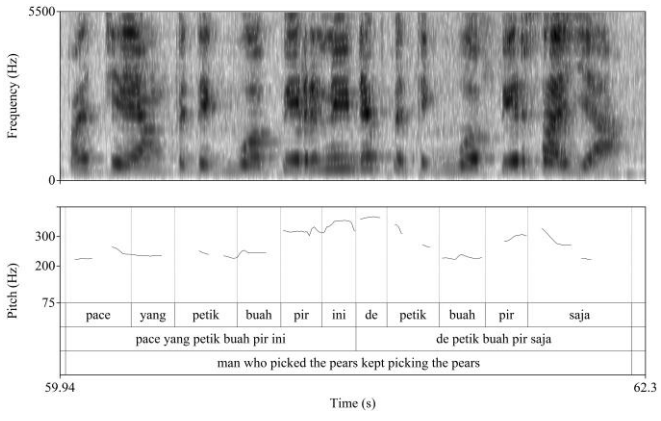

Figure 1. Example of annotated speech represented by a spectrogram and F0 contour with a Papuan Malay word tier (top), Papuan Malay intonation unit tier (middle), and English gloss tier (bottom).

Audio-tracks (48 kHz, 16 bit, stereo) extracted from the recordings of the portable video camera were translated, transcribed and segmented into intonation units [22] by native speakers of Papuan Malay. A group of six labelers annotated all words produced by the participants using Praat text grids ([23]; see Figure 1). All labelers received phonetic training to set label boundaries by auditory and visual inspection of the wave-form and its spectrogram.

A subset of the labelled words was selected on the basis of the following criteria. Words that were part of an utterance that was interrupted or cut off were omitted. Reduplicated words (e.g. *tiba-tiba*) were omitted, whereas single occurrences (e.g. *tiba*) were taken into account. To avoid prosodic effects of final-lengthening, words that occurred in the final position of the intonation unit were omitted. Words that were lengthened due to hesitations were also omitted. Words with loud background noise (e.g. motorcycle, bird or clicks) were omitted as well. Selected words were repeated at least one time. Any later repetitions of words were not taken into account for two reasons. First, later repetitions have increased distance from the first mention, which could be confounded with topic or referential shifts. Second, a considerable number of words were repeated only once (23%), and these words would not allow for analysis of later repetitions. The subset was divided into content words (nouns, verbs, adverbs) and function words (demonstratives, conjunctions, question words, prepositions, negations, pronouns), based on word lists in [14]. Note that words that translate to adjectives in English are generally expressed by stative verbs in Papuan Malay [14]. The selected subset including both first and second mentions contained 856 content words and 436 function words.

Using a script, two acoustic measures were taken from the selected words in Praat [23]. First, raw word duration was measured, as this appeared the most reliable cue to repetition reduction (i.e.[1], [8], [13]). Second, pitch mean was measured by taking the average F0 value of the word, as this measure is predicted to show reduced values for repeated mentions [4].

2.4. Statistical analysis

Statistical analyses were done using R [24] and the lme4 package [25]. Linear mixed model analyses fit by maximum likelihood (using Satterthwaite approximations to degrees of freedom to calculate p-values) were carried out on the acoustic measures of duration (ms.) and pitch mean (Hz) as dependent variables, with mention (2 levels: first, second) and word type (2 levels: content, function) as fixed factors, and with participants (speakers) and items (words) as random factors.

3. Results

Table 1. Means (*SDs*) of duration (ms) and mean pitch (Hz) of repeated content and function words.

Measure	Word type	1st mention	2nd mention
Duration	Content	273.63 (*112.97*)	259.81 (*102.37*)
	Function	218.54 (*125.77*)	199.49 (*94.40*)
Pitch (*M*)	Content	195.82 (*58.31*)	203.64 (*64.94*)
	Function	194.39 (*60.68*)	205.23 (*63.82*)

Results (Table 1 and 2) show a significant effect of mention, in that second mentions (M = 239.45) were overall shorter compared to first mentions (M = 255.04). The effect of word type indicates that content words (M = 266.72) are significantly longer compared to function words (M = 209.02). The lack of interaction effect between mention and word type indicates that there is no significant difference in repetition shortening between content words and function words.

As for pitch, second mentions show significantly higher values compared to first mentions, both for content words (M_1 = 195.82, M_2 = 203.64) and for function words (M_1 = 194.39, M_2 = 205.23). The factor word type did not show any significant (interaction) effects for pitch.

Table 2. Results of the linear mixed effect model analysis on acoustic measures of duration and mean pitch. None of the interactions were significant.

Measure	Factor	Estimate	*SE*	*t*	*p*
Duration	(Intercept)	260.28	17.82	14.61	< .001
	Mention	-19.06	7.41	-2.57	< .05
	Word type	48.78	19.64	2.48	< .05
Pitch (*M*)	(Intercept)	178.70	11.53	15.49	< .001
	Mention	10.63	4.74	2.24	< .05
	Word type	1.24	9.83	.12	n.s.

4. Discussion

In this paper we have shown that upon repetition, Papuan Malay words are shortened and their mean pitch is higher. This finding is partially confirming earlier findings on repetition reduction in English and partially contradicting predictions of the SSRH [4].

4.1. Duration and pitch

Previous work using mainly English speakers recurrently showed that repeated mentions have shorter durations. This has been traditionally explained according to the high informational redundancy of repeated words. The Papuan Malay data analyzed in the current study are in line with these earlier findings, as second mentions were indeed shorter compared to first mention.

We furthermore hypothesized that pitch would show reduced values for repeated mentions, following previous experiments [11] and the predictions of the SSRH [4]. The results, however, showed that the mean pitch of repeated words is higher in Papuan Malay, thus contradicting earlier predictions and findings. Although all acoustic parameters are commonly taken as correlates of prosodic prominence, the current results show that duration and pitch have opposite effects. It is not a priori clear why this is the case in the current study. One possibility is that different prosodic correlates fulfill different (linguistic) functions. Such an explanation is in line with the multiple source account to prosodic prominence [26], which claims that listeners do not process all acoustic correlates of prominence in an equal way. In fact, in [10] lexical repetition affected both duration and intensity, whereas referential repetition only affected intensity. In the current study, duration could be the main correlate of repetition reduction, whereas pitch could be more related to phrasal intonation phenomena. Although more empirical research is needed to confirm such an explanation, existing work usually

investigates duration as the primary correlate of repetition reduction (see section 1). Furthermore, phrase intonation is generally described by F0 contours [6], which may be cross-linguistically variable. Given the limited research on Papuan Malay phrase prosody, it remains unclear to what extent the pitch results in the current study are affected by it.

Another hypothetical explanation for the diverging acoustic effects may be found in the narrative style of Papuan languages. That is, it is common for speakers to fully or partially repeat the last clause of the previous phrase in the next, also called tail-head linkage (i.e. [27]). Therefore, the first mention is likely to occur late in the phrase, whereas the second mention is likely to occur early in the next phrase. Given the declination of pitch over the course of a phrase, it is plausible that second mentions had a higher mean F0 compared to first mentions. A post-hoc chi-square test on the distribution of first and second mentions in the head (217_{first}, 221_{second}) or tail (192_{first}, 196_{second}) was carried out to test this explanation. In this analysis only content words were taken into account as these refer to topics in discourse. Words occurring in the first half of the phrase were labelled "head", whereas word occurring in the second half of the phrase were labelled "tail". Single word phrases were ignored. The test did not reveal significant differences: $\chi^2(1, N = 826) = .0003$, n.s. Thus, it was not the case that significantly more first mentions were found in the tail compared to the head. Such a distribution would be needed to cause a significant F0 difference between first and second mentions. Therefore, tail-head linkage is ruled out as an explanation of the effect on pitch mean.

In addition, it is possible that other phrase intonation phenomena play a role of which to date there is no knowledge. In this respect, it deserves to be mentioned that the repetitions analyzed in this study occurred both within and across intonation units. A possible declination reset between intonation units was not controlled for, which could have affected the pitch results.

4.2. Word type

Concerning word type, the current results do not support the hypothesis that repetition reduction mainly affects content words (i.e. [9]). Overall, content words are longer than function words. However, both word types are subject to durational reduction when mentioned a second time. And, crucially, both word types show an increase in pitch for second mentions. These results, then, question to what extent the informational value of a word in terms of redundancy [4] actually plays a role in repetition in Papuan Malay. That is, it has been shown that homophony between first and second mention can lead to reduction [11]. If homophony is "enough" to prosodically reduce a second mention, the actual informational value of the word might not be the most important factor determining reduction. The absence of word type differences in the current results seem to better fit a model where ease of articulation of a string of speech sounds determines the shortening of later produced sequences, regardless of their content.

4.3. Outlook

Several issues are left open for future investigation. Foremost, a planned follow-up study will investigate to what extent the intelligibility of repeated mentions is reduced. This will shed more light on the perception of second mentions with decreased duration and increased pitch. Furthermore, future work should take into account more acoustic parameters such as intensity or more complex pitch measures such as pitch range. Pitch range is particularly interesting to investigate as repetition reduction could be signaled by pitch range compression rather than the absolute pitch values that were measured in this study. More in general, phrase level prosody of Papuan Malay is in need of further investigation to assess the role of pitch in this language.

5. Acknowledgements

This research is funded by the Deutsche Forschungsgemeinschaft (DFG), as part of the collaborative research centre SFB-1252 "Prominence in Language". The authors are thankful to Jean Lekeneny, Sonja Riesberg, Yusuf Sawaki, Emanuel Tutorop, Volker Unterladstetter and Boas Wabia for support in one or more stages of the research, being: recording, transcribing and translating the collected data (funded by a documentation grant to SR and NPH from the Volkswagen Foundation within its DoBeS program), and for facilitating the perception experiment. They also thank Lisa Barz, Pascal Coenen, Jan-Niklas Linnemeier and Katja Wiesner for help with the segmentation and two anonymous reviewers for valuable comments on an earlier version of this paper.

6. References

[1] Fowler, C. A., and Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *JML* 26, 489-504. doi: 10.1016/0749-596X(87)90136-7

[2] Clark, H. H., and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition* 22, 1-39. doi: 10.1016/0010-0277(86)90010-7

[3] Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E, and Swerts, M. (2015). Reduction in gesture during the production of repeated references. *JML* 79–80, 1-17. doi: 10.1016/j.jml.2014.10.004

[4] Aylett, M., and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *L&S* 47, 31-56. doi: 10.1177/00238309040470010201

[5] Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. D. (2001). Probabilistic relations between words; Evidence from reduction in lexical production. In J. Bybee and P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*, 229-254. Amsterdam: Benjamins. doi: 10.1075/tsl.45.13jur

[6] Ladd, D. R. (2008). *Intonational Phonology* (2nd ed.). Cambridge: Cambridge University Press.

[7] Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *L&S* 31, 307-319. doi: 10.1177/002383098803100401

[8] Mcallister, J., Potts, A., Mason, K., and Marchant, G. (1994). Word Duration in Monologue and Dialogue Speech. *L&S* 37(4), 393-405. doi: 10.1177/002383099403700404

[9] Bell, A., Brenier, J. M., Gregory, M. L., Girand, C., and Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *JML* 60, 92-111. doi: 10.1016/j.jml.2008.06.003

[10] Lam, T. Q., and Watson. D. G. (2014). Repetition Reduction: Lexical Repetition in the Absence of Referent Repetition. *J. of Exp. Psychology* 40(3), 829-843. doi: 10.1037/a0035780

[11] Jacobs, C. L., Yiu, L. K., Watson, D. G., Dell, G. S. (2015). Why are repeated words produced with reduced durations? Evidence from inner speech and homophone production. *JML* 84, 37-48. doi: 10.1016/j.jml.2015.05.004

[12] Bard, E. G., Lowe, A., and Altmann, G. (1989). The effects of repetition on words in recorded dictations. In J. Tubach and J. Mariani (Eds.), *Proc. of Eurospeech*, Paris, France, 2573–2576.
[13] Lam, T. Q., & Watson. D. (2010). Repetition is easy; Why repeated referents have reduced prominence. *Memory and Cognition* 38, 1137-1146. doi: I0.3758/MC.38.8.1137
[14] Kluge, A. (2017). *A grammar of Papuan Malay*. Berlin: LSP. doi: 10.5281/zenodo.376415
[15] Maskikit-Essed, R., and Gussenhoven, C. (2016). No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay. *Phonology 33*, 353-389. doi: 10.1017/S0952675716000154
[16] Riesberg, S., Kalbertodt, J., Baumann, S., and Himmelmann, N. P. (accepted). On the perception of prosodic prominences and boundaries in Papuan Malay. In Riesberg, S., Shiohara, A., and Utsumi, A. (Eds.), *A crosslinguistic perspective on information structure in Austronesian languages*, Berlin: LSP.
[17] Minde, D. van (1997). *Malayu Ambong: phonology, morphology, syntax*. PhD dissertation, Leiden University.
[18] Stoel, R. B. (2007). The intonation of Manado Malay. In Van Heuven, V. J. and Van Zanten, E. (Eds.), *Prosody in Indonesian Languages*, 117-150. Utrecht: LOT.
[19] Paauw, S. H. (2009). The Malay contact varieties of Eastern Indonesia: A typological comparison. PhD dissertation. State University of New York.
[20] Chafe, W. L. (1980). The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production. Norwood, NJ: Ablex.
[21] Riesberg, S. and Himmelmann, N.P. (2012-2014). Papuan Malay. Summits-Page Collection, DoBeS Archive MPI Nijmegen, http://www.mpi.nl/DOBES/
[22] Chafe, W. L. (1994). *Discourse, Consciousness, and Time*. Chicago: The University of Chicago Press.
[23] Boersma, P., and Weenink, D. (2017). *Praat: doing phonetics by computer*. Software, v.6.0.28. www.praat.org.
[24] R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Software, v.3.4.0, www.r-project.org.
[25] Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. of Statistical Software* 67(1), 1-48. doi: 10.18637/jss.v067.i01
[26] Watson, D. G. (2010). The many roads to prominence: Understanding emphasis in conversation. *Psych. of Leaming and Motivation* 52, 163-183. doi: 10.1016/S0079-7421(10)52004-8
[27] Vries, L. de (2005). Towards a typology of tail–head linkage in Papuan languages. *Studies in Language* 29, 363-384. doi: 10.1075/sl.29.2.04vri