Network on Code-Switching and Language Contact

SUMMER SCHOOL
CODE-SWITCHING
and LANGUAGE CONTACT

Ljouwert/Leeuwarden, 14–17 September 1994

FRYSKE AKADEMY

1995
Mijn tetangga bilang erreg mooi! Dutch - Indonesian codeswitching and bilingual speech production models

Herman Giesbers
Ottersum, The Netherlands

1. Introduction

In the 1980's we studied codeswitching (CS) between a dialect and the Dutch standard language with the objective to relate 'functions' of CS to its structural-linguistic aspects and to some characteristics of the speech production process.

By 'function' here we do not mean the societal or social-psychological function, but the function of a particular switch on the discourse level. Actually, it is a matter of: "Why does this CS appear on this point in this stretch of discourse?" Hence we prefer the term 'extra-linguistic characterization' of a switch.

Based on an extensive study of literature we made an inventory of all kinds of 'functions' as mentioned by other researchers. This led to a large number of characterizations, but for the moment it suffices to mention just the five major categorizations we formulated for our analyses. These are:

1) Situationally motivated CS, i.e. CS in co-occurrence with changes in the interactional situation; actually, this is switching to another unmarked code in correspondence with [as a reaction to] a new situation.

2) Intentional CS, among which, first of all adaptation to the language proficiency and/or preference of the addressee(s); furthermore, intentional CS includes pragmatically motivated switches which are intended to lead to a specific interpretation of an utterance (e.g. mitigating, showing emotional involvement, etc.), or to some effect for the course of the conversation like closing a topic. Intentional CS includes also stylistic and textual switches like emphasizing, contrasting, quoting, etc.

3) CS as a mode of speaking or 'unmarked CS'. The speaker has apparently chosen for an alternating use of two languages, but without a clear specific intention as in the CS described under 2/. S/he can do this for a shorter or longer period, and/or only in some circumstances.

4) Contextual CS is not intended CS, but switching rather triggered by elements of the situational and/or linguistic context, like the conversational topic, preformulated speech patterns or idioms and "lexical need".

5) Lastly, we distinguished performance CS. These switches are clearly not intended nor triggered by contextual cues, but are caused by the fact that the speaker has command of more than one language and is using them both. All such cases are unambiguously instances of CS as opposed to language contact phenomena as e.g. interference.

Examples of these CS cases are given in section 3.

Note that categories 4/ and 5/ are not distinguished in CS models concerning the social factors underlying CS like e.g. the Markedness-Model (Myers-Scotton 1993b a.o.). Yet, these forms of CS appear to be very relevant to the discussion on the structural factors in
CS, particularly if one wants to avoid rather ad hoc notions like e.g. 'nonce borrowing'. Cf. section 4.

2. Some results from a dialect – standard language study

As could be expected, CS between two highly interdependent varieties leads to a mishmash, a "chaos". We found 189 different switch positions and violations of all the "constraints" which were formulated at that time.

On the other hand, relating extra-linguistic characterizations of switches to their structural-linguistic properties shows surprising regularities. It appeared that especially intentional CS behaviour leads to CS in larger and/or complete units on turn, sentence or constituent boundaries while contextual and performance switches have predominantly an insertional nature and are realized intrasententially.

Admittedly, despite these regularities we still found a large number of different switch positions within every extra-linguistic characterization, but the probability of violation of a constraint is apparently larger within the unintentional group (4/ and 5/) than within the other categories.

In discussing these findings we proposed to account for CS in one (psycholinguistic) model for speech production, i.e. the model of Kempen et al. (cf. Kempen 1987, De Smedt & Kempen 1987 a.o.).

CONCEPTUAL MODULE

L conceptual structure --------------- M

E ---- LEXICO-SYNTACTIC MODULE O

X syntactic structure --------------- N

I ---- MORPHO-PHONOLOGICAL MODULE I

C phonological structure --------------- T

O ARTICULATORY MODULE O

N phonetic structure --------------- R

ARTICULATORY APPARATUS

Figure 1. A model for sentence production. (De Smedt & Kempen 1987:366)

From figure 1 it becomes clear that Kempen's model and Levelt's speaking model (Levelt 1989) are not that different and, what is relevant here, they share essential assumptions which can be summarized as follows:
- Conceptualization, formulation and articulation partially run in parallel
- Sentence construction is decentralized because of specialized procedures
- Syntactic rules are lexically driven (lemma – lexeme distinction)
- Syntax generates continual (open) structures (incremental speech production).
Kempen assumes three specialized procedures within his lexico-syntactic module: functional (e.g. Subject), categorical (e.g. NP) and lexical procedures. Lexical procedures produce as their output a 'pointer' which has to look for the right lexeme.

As a result we hypothesized three basic forms of CS, intentional, contextual and performance CS (the latter including CS as a mode of speaking) and that each of these can be related to a different level of the Incremental Grammar (IG) model in figure 1. Intentional CS should be accounted for in the conceptual module, contextual CS in the lexico-syntactic module and performance CS in the morpho-phonological module.

A conversion of these ideas to Levelt's model can be found in De Bot (1992).

In sections 3 and 4 we aim at a refinement of these ideas, a.o. by confronting Dutch/Indonesian CS data with Myers-Scotton's MLF-model (Myers-Scotton 1993a, b).

3. Dutch – Indonesian codeswitching

The Indonesian data come from interviews with Indonesians from Jakarta and its suburbs who still have a good, and sometimes even a near-native command of Dutch. Naturally, the background of their proficiency in Dutch lies in the Dutch colonial past in Indonesia which lasted until the late 1940's. Due to this particular background, the informants are 50 years or more in age, with the exception of two of them, aged 28, who learned Dutch at home and subsequently studied Dutch at the Universitas Indonesia (B and W in, a.o., (4) hereafter).

In conducting the interviews we tried to stimulate informal conversational patterns, preferably between the informants themselves.

We have replicated some analyses from the dialect – standard language study to verify the main conclusions as formulated in section 2 in a situation with typologically different languages.

As far as the 'constraints discussion' is concerned we limit ourselves to confront the data with Myers–Scotton's MLF-model, firstly because we think that such a model leads to a more general level of discussion and, secondly, because she also tries to link CS with a speech production model.

The analyses have been done essentially in the same way as in the 1989 study, i.e. utterances with (intra-, intersentential or single word) switches have been transcribed and hence coded for extra-linguistic characterization, constituent structure, switch position, etc. The extra-linguistic characterization is hierarchical, i.e. when more than one interpretation is possible, the 'most intentional' one is chosen. As a consequence, when e.g. 'adaptation to addressee' has contextual features as well, it is coded as 'adaptation to addressee'. Admittedly, in particular the differences between contextual and performance CS can be somewhat fuzzy sometimes. Words for Indonesian food, drinks, institutions etc. ('cultural loans') as well as words which are being discussed ('self-referring position', e.g. colok in (4) iv) are not included in the data. Lastly, all "doubtful" transcriptions have been checked (and corrected when needed) by a native speaker of (Jakarta) Indonesian.

Based on frequency of occurrence, the extra-linguistic characterization has been ultimately comprised into five cover categories:

a) Contextual switches as for instance in (1) – (3). See Appendix for transcription conventions. Jaksa in (1) is an example of 'preformulated speech', petani and buruh
in (2) have been triggered by the topic, and dagang in (3) can be seen as an example of 'lexical need'. Note that all examples in this paper are Dutch/Indonesian CS, unless otherwise stated.

(1) Ik heet (-) naar mijn grootvader. Die was dus 'n jaksa.
I name ((verb form)) (-) after my grandfather. That was thus a public prosecutor.
I'm named after my grandfather. He was a public prosecutor then.

(2) Vijfenzestig, dat is de Chinees. Dat is dus de petani. Dus niet de buruh, niet de werkers maar, en die zijn wred-er.
Sixty five, that is the Chinese ((sing. form)). That is so the peasant. So not the labourer, not the workers but, and those are cruel-COMPARATIVE.
Sixty five [= the year 1965], those are the Chinese. Those are peasants then, not labourers, not the workers but, and those [= peasants] are more cruel.

(3) De Padang-er die denkt ik zal maar gauw vrienden maken dan kan ik weer gaan dagang weet je wel.
The Padang-PERSONIFICATION that thinks I shall but soon friends make then can I again go trade know you well.
People from Padang think let's soon make friends then I can get back into business again, you know.

b) Adaptation to addressee. Instances are (4) iv, v and x, and (17) – (18) in section 4.

(4) ((B and W are helping to connect the recorder, but they do not know that recording has already been started using the batteries. They are talking to each other in Indonesian))

(i) B > W: Uda di-colok?
Already PASS-PREFIX-connect?
Already connected?

(ii) W: Uda:h, heb ik ge-colok, Ik heb ge-colok.
Already, have I PAST PART PREFIX-connect. I have PAST PART PREFIX-connect.
Sure, I have connected. I have connected.

I nearly disconnected it, it's already connected. It remains connected.

(iv) B > He: 'Colok' weet je? Aan.
'Colok' know you? On.
You know 'colok'? On.
(v) W > He: Die moet je eerst colokk-en.  
That have-to you first connect-Vinf SUFFIX.  
You have to connect it first.

(vi) He.: ‘Colokken?’  
‘Colok-Vinf SUFFIX’?  
‘Colokken’?

(vii) W: Aansluit-in. (-)  
Connect-VERB SUFFIX ((here: referring to 'putting'; -in is a Jakarta variant of 'kan'))  
Plug in.

(viii) W: Ja: // O die is nog niet  
Ya: Oh that is yet not  
Ya: Oh that one is not yet

(ix) B > W: Vanwege wat? ((W says something from a distance))  
Because of what?  
Because of what?

(x) B > W: negak ada apa apa.  
Not be what what.  
There's nothing/It's doesn't do anything.  
((Indonesian/Dutch))

c) Other intentional CS, e.g. (5) – (7). The intersentential switch from Dutch to Indonesian in (5) is pragmatically motivated, expressing some personal touch and/or mitigating the otherwise maybe somewhat intrusive directive. (6) and (7) are stylistic switches which focus and emphasize the switched elements, respectively.

(5) ((S invites her guests to move to another room for a meal))  
Pindah? Nah.  
Move? PHATIC PARTICLE ((henceforth: 'phatic'))  
Change rooms? Let's go.

(6) ((S is telling about popcorn))  
Op een dag ik, ik probeer om zoiets te maken, ANGus. ((laugh))  
PREP a day I, I try to something-like-that to make, burnt.  
One day I, I try to make something like that, it burnt.

(7) RAjin ya zij is.  
Industrious PHATIC she is.  
She really works hard, doesn't she.
d) Performance CS. Consider:

(8) Hoe noem je dat eh als eh (-) zakkerolle of Nee nee nee, nee juga dieven ja? How call you that uh if uh (-) pick-pockets or No no no, no also thieves PHATIC?

What do you call that uh if uh (-) pick pockets or No no, no 'thieves' also, isn't it?

e) CS as a mode of speaking, e.g. (4), especially W in (ii) – (vii), or (11) – (12) in section 4.

In this paper the structural analysis has been limited as well, i.e. to a distinction 'intersentential' vs. 'intra-sentential' CS. Intersentential switches are switches with its switch position on turn or sentence boundaries, the latter including switching before a conjunction, discourse marker, tag, etc., and which at the same time form a complete linguistic unit, structurally as well as prosodically. Clear cases are (4) i or (5). All other cases are intrasentential switches, including a switch on a turn boundary as in (7).

Table 1 and 2 (next pages) show the five extra-linguistic characterizations broken down by their intra- or intersentential nature for the Indonesian as well as the dialect/standard language data.

From the tables it can be concluded that the main tendency is the same. Intentional CS gives rise to more complete switched units, unintentional CS is predominantly of an intrasentential nature. However, dialectal CS leads to more intrasentential switches for all characterizations, possibly due to the grammatical interdependence of dialect and standard language. On the other hand, the Dutch – Indonesian data show also a high amount of different switch positions (92), although only a few with a high frequency, namely those on turn- or sentence boundaries and to nouns.

Apparently, the same mechanisms prevail and this finding brings us to the following discussion.

4. Discussion

To our opinion, the MLF-model is a robust model for CS, particularly when counter-examples are evaluated only quantitatively. In addition, the model has interesting psycholinguistic implications. This raises the question as to what extent the foregoing can still be relevant.

To answer this question let us take a look to some cases which, in one way or the other, may be problematic for the MLF-model.

To start we would like to give the following assumptions.
- CS can be realized in different phases of language production (cf. also De Bot & Schreuder 1994), e.g. during syntactization or lexicalization, and this process is, among other things, linked to the extra-linguistic function of the switch. Moreover, we assume that a bilingual can retrieve language elements from at least two languages at a time, the "selected language", which is maximally activated for controlling the speech output, and the "active language", which works parallel to the selected language, but has no access to the outgoing speech channel (De Bot 1992).
- The matrix language (ML) changes much more often and in a more flexible way than Myers-Scotton seems to suggest.
Table 1 Extra-linguistic characterizations related to intra- or intersentential CS. Indonesian data.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Intrasentential CS</th>
<th>Intersentential CS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual CS</td>
<td>229</td>
<td>41</td>
<td>270</td>
</tr>
<tr>
<td>Adaptation to addressee</td>
<td>12</td>
<td>116</td>
<td>128</td>
</tr>
<tr>
<td>(Other) intentional CS</td>
<td>32</td>
<td>93</td>
<td>125</td>
</tr>
<tr>
<td>Performance CS</td>
<td>90</td>
<td>19</td>
<td>109</td>
</tr>
<tr>
<td>CS as a mode of speaking</td>
<td>36</td>
<td>36</td>
<td>72</td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td>305</td>
<td>704</td>
</tr>
</tbody>
</table>

Table 2 shows the dialect – standard language data.

Table 2 Extra-linguistic characterizations related to intra- or intersentential CS. Dialect – standard language data (from Giesbers 1989:263).

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Intrasentential CS</th>
<th>Intersentential CS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual CS</td>
<td>1210</td>
<td>86</td>
<td>1296</td>
</tr>
<tr>
<td>Adaptation to addressee (incl. 'reactive accommodation')</td>
<td>525</td>
<td>1273</td>
<td>1798</td>
</tr>
<tr>
<td>(Other) intentional CS</td>
<td>387</td>
<td>810</td>
<td>1197</td>
</tr>
<tr>
<td>Performance CS</td>
<td>990</td>
<td>35</td>
<td>1025</td>
</tr>
<tr>
<td>CS as a mode of speaking</td>
<td>126</td>
<td>32</td>
<td>158</td>
</tr>
<tr>
<td>Total</td>
<td>3238</td>
<td>2236</td>
<td>5474</td>
</tr>
</tbody>
</table>
Firstly, see fragment (4) in section 3. Following Myers–Scotton's work (at least the way we interpreted it) the ML here should be Indonesian because the unmarked code before and after (4) is Indonesian. Indeed, Myers–Scotton accepts the possibility of a change of ML even within a single conversation, but as a result from changing situational factors, and that is quite different from the data we are discussing (cf. Myers–Scotton 1992:12, 1993b). (4) iii and vi support the position of Indonesian as the ML with no Dutch word order nor system morphemes, but we have to assume that the ML in ii and v is Dutch, otherwise these utterances run counter to Myers–Scotton's predictions because of the Dutch word order and system morphemes. Note that ii and iii come from one continuing speech turn.

Confer also (7) which shows no Dutch word order, but comes from an almost entirely Dutch discourse and has been uttered by a woman with a near-native proficiency in Dutch.

Secondly, some questionable cases in following the MLF-model seem to be related to performance CS, as in (9) and (10) where yang is a system morpheme par excellence, because it "singles one item out of a class" (Wolff et al. 1987:158). Kalau ('if') is probably also a system morpheme.

9) ((In (9) and (10) S is discussing Chinese elements in Indonesian))

'emang, dat is memang yang Hokian-s en dat is gewoon de Hokian, memang. Indeed, that is indeed LIGATURE PARTICLE Hokian-s ((Dutch suffix)) and that is ordinary the Hokian, indeed.

Yes, that's indeed the Hokian language and that is usual Hokian, that's right.

10) Oo dat is 't beleefde, kalau eh naar eh naar iemand yang, yang (die ze moesten) respecteren.

Oh that is the polite, if uh for uh for someone LIGATURE PARTICLE ((2 x)) (that they had-to) respect.

Oh that is the polite, if uh for uh for someone they had to respect.

The sentences (11) – (14) (with performing CS in (11) and (14), and CS as a mode of speaking in (12)–(13)) seem all to be problematic because of embedded language (EL, here: Indonesian) word order and/or EL system morphemes.

11) Di rumah si ibu juga komt ook nog weer.

The ibu is also still coming in our house. (Ibu: a.o. respectful term for female servants)

12) (→) Kalo als ik loop, jadi musti 'n beetje jinjit-jinjit, anders ber-bunyi. (((laughing)))

If I'm walking, then I have to tiptoe a little, otherwise ((verb-forming prefix)) ber-noise.

If I'm walking, then I have to tiptoe a little, otherwise it makes a noise.

13) En soms eh: maak ik eh: wel gebruik van die: terrasso, kalo pakai leren zool "sju:"

And sometimes uh: make I uh: well use of that uh: terrazzo, if use leather sole "sju:"

And sometimes I make use of that terrazzo, with a leather sole "sju:"

2 In addition, we can observe the same state of affairs regarding the quote in the title. This utterance is from one of the most popular TV commercials in 1993 with Indonesian as its ML. However, word order and system morphemes are obviously Dutch.
We would like to suggest that CS of this type is realized as a lexeme, that is in the morpho-phonological module (figure 1): A syntactic frame is already available from the foregoing module, but for some reason or another the speaker takes more or less unconsciously a lexeme from the 'active' language. Perhaps this process can also account for the 'double forms' in examples like (11) juga – ook, (12) kalo – als, (15) die – yang, and the self-repair in (10): yang becomes die.

What is more, CS which in itself is predictable from the MLF-model, can also be accounted for in this way. See (15) ('CS as a mode of speaking') and the 'performance' switch in (16).

On this line of argument we can suppose that intentional CS is planned on a higher/earlier level of the language production process, i.e. the intended language is selected and the correspondent structures are built up in the lexico-syntactic module, but it is possible that lexemes from the "active" language are not always avoidable. See, once again, the intentional CS in (7), where we could assume that this sentence is meant to be Indonesian, but the words zij is are lexemes from Dutch.

Contrarily, confer (17) and (18), in which CS as 'adaptation to addressee' after a slight pause are perfectly planned and realized in the intended/selected language, and therefore are completely in accordance with the procedures started in the lexico-syntactic module.

(17) (In (iii) L speaks to N, a student of Dutch, in (iv) she speaks to He. again))

(i) L: Daar in Mangga Mangga, Mangga Besar of Mangga
There in (-) Mangga Besar or Mangga

(ii) N > L: Mangga Duo = Mangga Duo

(iii) L > N: = Bukan Mangga Dua Mangga itu apa Prinsenpark dulu toh? (eh) Mangga Besar,
Not Mangga Dua Mangga that QUESTION PART. Prinsenpark formerly PHATIC
Mangga Besar,
Not Mangga Dua, that Mangga wasn't it 'Prinsenpark' in former times? Mangga Besar,
(iv) L > He: hebben ze ook restaurant die zo'n dinges verkopen, slangen.
   They also have a restaurant that sells such things, snakes.
   - in Mangga Besar, there are also restaurants which sell such things eh, snakes.
   *((Indonesian/Dutch))*

(18) ((In (i) D has switched from Dutch to Indonesian when explaining something to her
   friend The. about some governmental measures in the fifties. In (ii) she switches back
   to Dutch, because He. is addressed again))

(i) D. > The.: Itu, PP sepuluh limapuluhlima (limapuluh enam) ((speaking at the same
   time, in Indonesian))
   That, PP ten fifty five (fifty six)
   That is, PP -ten in fifty five (fifty six)

(ii) D continuing: Nah itu! Masih, ik geef lessen in het Chinees. Wan zij spreken
   allemaal Chinees.
   PHATIC that! Still, I give lessons in Chinese. Because they speak all Chinese.
   That's it! Still, I give lessons in Chinese, because all of them speak Chinese.
   *((Indonesian/Dutch))*

Lastly, we would like to point to differences as can be found in (19) – (22).

(19) ((W is telling about Indonesian politics in the early 1960s, a.o. about Soekarno's
   'confrontation politics', the policy of political and armed confrontation with
   Malaysia.))

(i) W: (-) maar, de: verdediging Maleisië was niet de Maleiser (-). Dus als wij moesten
   gaan vechten daar, dan, vecht je tegen de, (-) Commonwealth, troepen, (-) maar niet
   de Maleur, (-) van de honderd zijn er alleen maar tien Maleur-s in die tijd. (-)
   but, the: defense Malaysia was not the Malayan (-) So when we had to go fight
   there, then, fight you against the, (-) Commonwealth, troops, (-) but not the
   Malayan, (-) of the hundred are there only but ten Malayan-PLUR (Dutch suffix)
   in that time. (-)
   but the Malaysian defense were not Malaysans. So, if we had to go to fight there,
   you're fighting against the Commonwealth troops, not the Malaysans, (-) out of
   hundred there are only ten Malaysans at that time.

(ii) W: (-) wanneer 't kom dan gaan de godong-s open en dan kunnen wij, de, de de, de
   senjata-s halen hè.
   when it come then go the storehouse-PLUR (Dutch suffix) open then can we, the,
   the the, the weapon-PLUR (Dutch suffix) take eh.
   when it starts the depots are going to be opened and we can get the weapons eh.

(20) ((S asks He whether he has a problem with kopitubruk, the typical Indonesian
   coffee))
   Mag 't tubruk?
   May it collision? (Lit. translation of tubruk)
   Can it be 'tubruk'? / Do you mind if it's 'tubruk'?

(21) ((One is talking about former street names in Jakarta))
(i) M: En, Jalan Sabang was vroeger -NAME- ((p)) (-) Is er nog, Jalan Jalan Jawa?
And, Jalan Sabang was formerly -NAME- ((p)) (-) Is there still, street Jawa?
And, the Sabang street was formerly (-NAME-) ((p)) (-) Is the Jawa street still there?

(ii) W: En al die: namen van die jalan-s. ((p))
And all those uh: names of those street-PLUR ((Dutch suffix)) ((p))
And all those street names. ((p))

(22) Dat ligt aan de Jalan Pemuda.
That's situated in the Pemuda street.

On the one hand, we see Indonesian nouns without a plural suffix (ML + EL constituents) but with a Dutch article, on the other hand, there are those nouns with Dutch articles and plural suffixes; compare e.g. (19) i, ii where it happens to be the same speaker in the same conversation. See also (21) and (22).

Perhaps, these phenomena show different ways to realize (mostly contextual) CS. Forms without morphological marking can be the product of a lemma procedure in the lexico-syntactical module that leads to a so called bare form, while the switches with the complete morphological marking from the ML (or: 'selected language') are lexemes from the EL which are inserted in the right slots. The latter implies that actually a lemma from the ML has been chosen including its consequences on all levels of grammatical structure, although the speaker eventually went back to his 'active language' to get the right word (lexeme) at that very moment. Reasons could be pressure of time, a strong activation because of the topic, etc.

We realize that these suggestions raise many questions, certainly when we add an example like (20), where we can ask what triggered the omission of the Dutch copula zijn. According to the above way of reasoning this example could imply that tubruk is also a 'bare form' which means that, at least to Indonesian speaking persons, a copula is grammatically comparable to morphological endings rather than to verb forms. One explanation of this could be due to the optional character of the copula in Indonesian as opposed to Dutch.

In view of the above we have come to the following conclusions.  
- The MLF-model is a robust model, but it could be modified by relating it to extra-linguistic characterizations of CS.  
- A further refinement could be to integrate it with psycholinguistic theories concerning language production models and the bilingual lexicon. Within such a broad psycholinguistic framework a "blocking filter" would be redundant.  
- ML assignment appears to be much more flexible than Myers-Scotton suggests. A 'turnover' in ML assignment is not only possible within a single conversation but even within a single speech turn.  
- More detailed studies of changes in ML assignment are needed. In addition, these studies should particularly focus on pauses, hesitations, etc. in speech production.
Acknowledgments

Fieldwork in Indonesia has been made possible by the Ministry of Education & Science of the Netherlands.

I am indebted to prof. dr. A.M. Moeliono, Universitas Indonesia, for his helpful comments with regard to grammatical characteristics of Indonesian. Naturally, flaws and inconsistencies remain my responsibility.

Finally, many thanks are due to Juwita Samosir who transcribed parts of the interviews, conscientiously checked my own transcriptions, and explained to me many aspects of the Jakarta variety of Indonesian. "Thanks berat!"

Appendix

Transcription conventions.

The present Dutch and Indonesian spelling is used unless the informant's pronunciation is heavily deviating from everyday language use. Violations of tense rules are not corrected; this phenomenon is very common in the use of European languages by Indonesians.

Every example is explained by 2 translations, a morpheme-by-morpheme translation and a paraphrase in italics. In all other cases Indonesian words are underlined including in the plain text.

. sentence final falling intonation
, clause final intonation/micro pause
! exclamation
? question
: elongation of the preceding sound
- cut off sounds
((p)) relatively long pause/silence
CAPITALS emphasized syllable(s)
= quick start after previous speech
(utterance) uncertain transcription
(-) omission by the transcriber
((words)) comment or description by the transcriber/analyst
// overlapping speech (including back channels) but S continues
[word] additional information by the transcriber/analyst
References


