The effect of syntactic frequency on sentence comprehension in standard Indonesian Broca's aphasia

Bernard A. Jap, Silvia Martinez-Ferreiro & Roelien Bastiaanse

Routledge
Taylor & Francis Group

The effect of syntactic frequency on sentence comprehension in standard Indonesian Broca's aphasia

Bernard A. Jap [a,c,d], Silvia Martinez-Ferreiro[b,c] and Roelien Bastiaanse[c]

[a]International Doctorate for Experimental Approaches to Language and Brain (IDEALAB), University of Groningen, the Netherlands; University of Potsdam, Germany; University of Newcastle, UK; University of Trento, Italy; and Macquarie University, Sydney, Australia; [b]Biolinguistic Initiative Barcelona (BIB), Universitat de Barcelona, Spain; [c]Center for Language and Cognition Groningen (CLCG), University of Groningen, The Netherlands; [d]Science, Technology, and Society (STS) Research Group, Tarumanagara University, Jakarta, Indonesia

ABSTRACT

Background: Comprehension of reversible sentences that have derived word order has often been reported as impaired in agrammatic aphasia. Most accounts of this phenomenon refer to the syntactic differences between derived and base word order of the arguments. However, it has been demonstrated that in agrammatic spontaneous speech in standard Indonesian (SI) passives are produced at a rate that is proportional to that of healthy speakers. The main difference between passives in SI and in other languages is the frequency with which passives are used: passives in SI are highly frequent. The high frequency can be explained by the fact that passives are used for politeness reasons, saliency of the passive morphology, earlier acquisition, and formal simplicity of the passive structure.

Aims: The purpose of the current study is to investigate comprehension of the passive as a derived structure in SI and the influence of frequency.

Methods & Procedures: A sentence-to-picture matching task was developed to test four reversible sentence types (active, passive, subject cleft and object cleft). There are three variables that are of interest, that is, word order, embedding and relative frequency of structures. Eleven agrammatic speakers classified as suffering from Broca's aphasia were tested.

Outcomes & Results: The passive sentences were comprehended equally well as the active sentences. Embedding had limited effects: subject clefts were understood as well as actives and passives. Object clefts, however, were understood poorly and significantly worse than the three other sentence types.

Conclusions: The sentence comprehension deficit pattern shown in SI individuals with Broca's aphasia introduces frequency of a syntactic structure as an additional factor to consider. Whether frequency or pragmatic constraints protects against erosion of the passive in Broca's aphasia in SI remains an open question.

CONTACT Bernard A. Jap ✉ b.a.j.jap@rug.nl

Introduction

One of the defining characteristics of aphasia is that in almost all its types, patients show deficits in sentence comprehension (Goodglass, Kaplan, & Barresi, 2001). For agrammatic Broca's aphasia[1] in particular, there is substantial evidence to indicate that reversible sentences with derived word order are more vulnerable to breakdown (Bastiaanse & Edwards, 2004; Burchert, de Bleser, & Sonntag, 2003; Caramazza & Zurif, 1976; Grodzinsky, 2000, among others). Although different theories have been formulated to account for these findings (Bastiaanse & Van Zonneveld, 2005, 2006; Caplan & Futter, 1986; Grodzinsky, 1995, 2000; Schwartz, Linebarger, Saffran, & Pate, 1987), overall these studies suggest that the order of the arguments influences performance on sentence comprehension tasks in Broca's aphasia. The Derived Order Problem Hypothesis (DOP-H; Bastiaanse & Van Zonneveld, 2005, 2006) assumes that every language has a base order (e.g., SVO for English, SOV for Dutch). All other orders are derived by linguistic operations. Sentences with derived order are harder to produce than sentences with base word order for individuals with Broca's aphasia. For comprehension, this implies that all sentences with derived order require more processing capacity (see Bastiaanse & Van Zonneveld, 2006). This is hard to measure with offline tasks for most constructions, but for semantically reversible sentences, the order of the arguments is crucial: when the agent and theme are in base order (i.e., agent precedes the theme) sentences are relatively easy; when the agent and theme are in derived order (i.e., theme precedes the agent, such as in the passive and object relative sentences in English), comprehension will be impaired. Notice that the DOP-H is a processing account: it assumes that derived order is more difficult than base order, but not that derived order is impossible, such as representational accounts like the Trace Deletion Hypothesis (e.g., Grodzinsky, 2000) suggests. The DOP-H focuses on word order in particular and is based on empirical findings across languages (for a review of the hypothesis, Abuom, Shah, & Bastiaanse, 2013).

Data on standard Indonesian (SI) aphasia are relatively scarce compared to those of Indo-European languages. The available data suggest that SI speakers with Broca's aphasia not only comprehend passive sentences correctly, but also produce them at a rate that is proportionate to healthy speakers (Anjarningsih, Haryadi-Soebadi, Gofir, & Bastiaanse, 2012; Postman, 2004). A marked difference between passive structures in SI and other languages is the frequency in which these structures are used (Sneddon, 1996). The current study aims to present an additional factor to be considered when examining comprehension performance: the role of relative frequency of structures. We examine comprehension patterns in SI speaking individuals with Broca's aphasia speakers, focusing on three variables, that is, word order, embedding and relative frequency of structures. We start by providing the relevant background on SI word order and passive structures, which is followed by an overview of theories in Broca's aphasia that aim at describing sentence comprehension deficits. Then, previous studies in agrammatic Broca's aphasia in SI will be reviewed. We conclude the introductory section by stating the predictions of the current study.

Word order and passives in standard Indonesian

SI is a member of the Austronesian language family under the Western Malayo-Polynesian subdivision. It has 23 million native speakers and over 140 million L2

speakers (Lewis, Simons, & Fennig, 2013). SI is the language used in education, governmental activities and other formal settings. Though initially SI is acquired as a second language with regional dialects spoken as L1, native speakers of SI continue to grow in number (Postman, 2011).

SI is a zero-marking language (Nichols & Bickel, 2013). Zero marking is defined as the absence of overt morphological markers that are usually present in the core arguments of a predicate (Sinnemäki, 2010); in this context, SI has neither case nor gender markings. Transitive verbs are usually only inflected for voice (active or passive). With the exception of particular reduplicated verb constructions that signify an iterative aspect (Mistica, Andrews, Arka, & Baldwin, 2009), typical use of SI usually involves no verb inflection for tense, aspect or agreement. For example, the verb *memasak* in (1b) provides a lexical entry as well as information regarding voice, which is active and indicates transitivity. The base clause of SI has two obligatory components: the subject and the predicate. The subject of a clause is, in general, what is being discussed (the topic). It is usually produced in the form of a noun or a pronoun phrase (though nominal clauses can also appear in subject position). Clauses are either non-verbal (1a) or verbal, with a verb as the predicate centre (1b),

(1a) Andi   di   rumah
      Andi   at   home
      "Andi is at home"
(1b) Andi   **me**masak   nasi
      Andi   **ACT-**cook   nasi
      "Andi is cooking rice"

SI morphology (or its lack thereof) suggests a rigid word order, though in certain constructions such as WH-questions (Stack, 2005) and predicate nominalisation (McCune, 1979), the ordering of constituents can be flexible. The base word order in an SI clause is subject–predicate. An object follows the verb in transitive verbal clauses indicating that SI has an SVO base word order. The examples below (2a, b) clarify the types of sentences that will be discussed in the present study.

(2a) Base order, simple active (agent-theme)
      Perempuan   itu   **me**manggil laki-laki itu
      girl   the   **ACT**-call   boy   the
      "the girl is calling the boy"
(2b) Derived order, simple passive (theme-agent)
      *Laki-laki itu **di**panggil*   (oleh)   perempuan   itu
      boy   the **PAS**-call   (by)   girl   the
      "the boy is called by the girl"

The base word order with an active voice marking (*meN-* reduced to "me" due to assimilation) on the verb is shown in (2a). The examples above, as with the materials we used for testing, are semantically reversible. The passive[2] is expressed by the *di*-prefix on the verb in (2b) where "the boy" is the theme. As in English, the by-phrase in the passive (2b) is optional. Additionally, the preposition *oleh*: "by" may be omitted in

both spoken and written form when nothing is placed in between the passivised verb and the agent phrase. The NP–V–NP structure of the passive when the agentive preposition is omitted appears to be the mirror image of the active structure.

The passive construction plays a more vital role in SI than in other languages documented so far. It is acquired appropriately at a very early age, sometimes under 2 years (Gil, 2006). In comparison, Gil (2006) mentions that English-speaking children do not use the passive voice until age 4 or 5. Moreover, the input frequency of passive structures is estimated to be between 28% and 35% in SI, compared to 4–5% in English. Among adults, SI passives become increasingly more frequent as they age, in both spoken and written context. Another contrast with English is shown by a corpus study (Roland, Dick, & Elman, 2007). Passives appear four to five times more often in written corpora compared to spoken ones in English, demonstrating it is primarily a written structure in this language. In SI, however, Kaswanti-Purwo (1991) found that 30–40% of the verbs in written SI have the passive *di-* prefix, compared to approximately 9% of English verbs displaying passive morphology (Givón, 1979). The highly frequent SI passive is also found in other Malay languages such as classic Malay, a language from which SI is derived (Suwarso, 1994). The saliency of the passive in SI is attributed to the unambiguous voice morphology that provides a straightforward schema of the *meN-* prefix for active and the *di-*prefix for passive. This salience can also be observed in the fact that, unlike the *di-* prefix, the active prefix *meN-* contains a schwa and is often reduced in spoken Indonesian to stem-initial assimilation (e.g., *menyapu -> nyapu*: to sweep).

The passive in Indonesian has a functional use: it makes sentences more polite. Randriamasimanana (1999) observed the usage of verbs in passive form in letters sent by native SI speakers (parents) to their sons and daughters studying in the United States in the late 1970s. In an example, one of the individuals was reported to write three letters: one to a civil servant, one to the individual's elder son and another to the individual's younger son. The letter directed to the civil servant had a large proportion of passive verbs (57.1%; 32 out of 56), compared to that of the elder son (at 29.5%; 18 out of 61) and that of the younger son (16.3%; 8 out of 49). This example illustrates that passives are considered to be more polite, and thus deemed more appropriate in certain contexts. The frequent use of passives in SI can also be motivated by specific discourse functions that are distinct from those of active verb forms. Using a discourse analysis, Kaswanti-Purwo (1988) described the *di-* verbs' functions as foregrounding, describing punctual and/or factual events, as well as introducing actions that come in sequences. On the other hand, *men-* verbs function as beginning of discourse (background), describing habitual and/or nonfactual events, providing parenthetical information, and also breaking or closing narrative flows. Verhaar (1978) also noted some contexts where the *di-* passives are more "compatible", for instance when the verb form is not reduplicated or when a sentence does not provide information on duration.

Subject and object clefts (3a, b) are relatively infrequent compared to actives and passives. In other languages such as English, the use of cleft constructions is highly restricted and rarely used in both spoken and written English, occurring at a rate of 0.8 cleft construction per 2000 words in the British component of the International Corpus of English (Nelson, 1997). While there were no formal corpora analyses comparing clefts with other structures in SI, using a corpus of colloquial spoken data, Englebretson (2008) found

83 occurrences of cleft constructions within 8744 intonation units (IU) and 24,074 words by tagging the word *yang* (who/that), an obligatory relativiser. Another study of colloquial Jakarta Indonesian (Ewing & Cumming, 1998) reported that the corpus, consisting of 1360 IUs, had 74 relative clauses of which 25 were identified as "clefts", and they added that clefts in their observed corpus were restricted by transitivity; only about 20% of intransitive relative clauses were classified as clefts. The sentence referred here as object cleft has also been called a "bare" passive or passive type-two (Nomoto, 2010; Sneddon, 1996) while others refer to it as an object-preposed construction (Postman, 2004). The structure (3b) shows the embedded, derived order condition. In this sentence type, the verb occurs in bare form, as the restriction on extracting objects of verbs prefixed with *meN-* can be explained by seeing the active voice marker as lacking an Extended Projection Principle feature (Cole & Hermon, 2008). However, Postman (2002) remarked that the SI grammar may evolve to a point where object extraction with *meN-* active verbs is allowed. While instances of these are relatively uncommon and still generally regarded as ungrammatical, they are recorded to have occurred in formal registers of Indonesian (Hassall, 2005).

(3a)   Base order, subject cleft (agent-theme)

Perempuan	itulah	yang	**me**manggil	laki-laki	itu
girl	that is	who	**ACT**-call	boy	the

"that is the girl who is calling the boy"

(3b)   Derived order, object cleft (theme-agent)

Laki-laki	*itulah*	*yang*	*perempuan*	*itu*	*panggil.*
boy	that is	who	girl	the	0-call.

"that is the boy who the girl is calling"

Derived order problem hypothesis: agrammatic sentence comprehension

Individuals with Broca's aphasia have problems assigning the correct thematic roles in a reversible sentence that has derived word order. When a sentence such as *the cat is scratched by the dog* is read aloud and the participants have to select the correct picture, they may have problems identifying who does the action (the agent) and who under-goes it (the patient/theme). Consequently, they may fail to identify the matching picture in a sentence-to-picture matching task. Representational accounts of the comprehension impairment suggest that individuals with Broca's aphasia cannot understand derived structures such as passives and object-clefts, and thus resort to a guessing strategy. An example of a representational account is the Trace Deletion Hypothesis (Grodzinsky, 2000). From the point of view of minimalist linguistic theory, sentences with base and derived word order have the same underlying structure, but for derived word order an extra operation is needed. Consider the following example. In a simple active sentence (in English), the base word order is Agent–Verb–Theme. In a passive sentence, word order is derived, that is, the order is Theme–Verb–Agent. In order to get this derived structure, a linguistic operation is needed. (We refrain from a discussion whether this happens through movement, merge or any other operation; the idea is that the word order is not the base order.) It has been shown repeatedly and in many languages that comprehension of sentence with Theme–Agent order is difficult for agrammatic speak-ers. This has been shown for both production and comprehension in several structures

in different languages: passives in English (e.g., Grodzinsky, 1995); object clefts in Swahili (Abuom et al., 2013); sentences with scrambled objects in Turkish (Yarbay Duman, Ozgirgin, Altinok, & Bastiaanse, 2011). The DOP-H captures these word-order problems and also accounts for the pattern of performance observed with embeddings: embedding requires an extra operation and thus adds to the problems with derived order (Abuom et al., 2013).

Sentence comprehension in SI aphasia

As previously mentioned, data on aphasia in SI speakers are very scarce compared to those of Indo-European languages. There are, however, two studies on complex constructions in SI aphasia: one by Anjarningsih et al. (2012) and one by Postman (2004).

Anjarningsih et al. (2012) analysed the agrammatic spontaneous-speech of SI individuals with Broca's aphasia and found that, in addition to characteristics of agrammatism also present in other languages (such as reduced use of functional elements, lower speech rate and lower mean length of utterances), there are two phenomena that seem to be unique to SI: normal verb production and normal production rate of non-canonical structures such as passives. The occurrence of passive sentences was found to be proportionate to normal speakers.

Postman (2004) conducted a case study on a non-fluent SI aphasic speaker using puppets in an acting-out task. He successfully comprehended all single-clause sentences; therefore, a condition including complex sentences (two verbal clauses) was added to the design (4). The results of the study show that he comprehended and enacted the first clause of all trials including verbs with passive markers perfectly, but interpreted sentences inaccurately when the first clause was either a passive (Theme–Verb–Agent) or an object-topicalised structure (Theme–Agent–Verb).

(4) Susan   **di**cium   Nando   dan   **di**peluk   Allen
     Susan   **PASS**-kiss   Nando   and   **PASS**-hug   Allen
     "Susan is kissed by Nando and hugged by Allen"

While he comprehended that "Nando kissed Susan", he incorrectly interpreted *Nando* (instead of *Susan*) as the theme of the second action. This outcome draws further questions to whether aphasic SI speakers can consistently and accurately parse sentences with derived order, and whether word order interacts with the increasing complexity of the sentence. In Postman's (2004) study, complexity is defined in terms of number of clauses and total length of the sentence, where in the current study, the focus is shifted to word order and embedding as sentence complexity factors.

Structural frequency in sentence processing

Models of language processing, for example, the constraint-based model (Trueswell & Tanenhaus, 1994) and the competition model (MacWhinney, 1987), have considered sensitivity towards statistical and probabilistic aspects of language, such as the frequency details of lexical items used and how it affects processing of sentences. It has been shown that the lexical frequency of a word (MacDonald, Pearlmutter, &

Seidenberg, 1994) and its occurrence in different syntactic constructions (Juliano & Tanenhaus, 1994) affect processing decisions in non-brain-damaged speakers (NBDs).

MacDonald et al. (1994) proposed that, at the sentence level, the relative frequency of constructions affect parsing decisions. This is more clearly seen in sentences containing syntactic ambiguities. The fact that information from linguistic regularity is conveyed in the initial interpretation of sentences suggests that "exposure-based strategies" are indeed prevalent in sentence processing (Mitchell, 1994). Moreover, Mitchell, Cuetos, Corley, and Brysbaert (1995) presented evidence against models that rely exclusively on "fine-grained" lexical details. They argued that lexical constraints can be weak, and a more viable account of language processing requires frequency records on statistical regularities beyond the word level.

Eye movement patterns while processing unambiguous English sentences have been taken as evidence that passives (both reversible and irreversible) are more complex to process than actives—both in terms of accuracy and time required to respond to the comprehension question in NBDs (Ferreira, 2003). Difficulty in the atypical assignment of thematic roles is not attributed to frequency, however: actives, which are frequent, and subject clefts, which are not frequent, are comprehended equally well. To illustrate, if one were to predict processing difficulty solely based on frequency, (5a) would be best comprehended followed by the three structures, (5b–d) in no particular order since all three are infrequent in English. However, we know that in Broca's aphasia, the comprehension of subject clefts (5c) is relatively unimpaired in comparison to other derived structures (Grodzinsky, 2000; Abuom et al., 2013). This makes a frequency-based account not valid, at least when it is used as a sole factor.

(5a)	The horse bites the cow.	(active)
(5b)	The cow is bitten by the horse.	(passive)
(5c)	It is the horse that bites the cow.	(subject cleft)
(5d)	It is the cow that the horse bites.	(object cleft)

In aphasia, competing theories have suggested that various factors make certain syntactic structures more difficult to comprehend and/or produce than others. According to St John and Gernsbacher (1998), sentences with more noun phrases and/or verbs are more difficult than the ones with fewer noun phrases and/or verbs. This adds up to the fact that derived word order is more difficult than the canonical word order. Despite the fact that lexical frequency has been traditionally taken to be important in the fields of human memory, language acquisition and language processing, frequency of grammatical constructions has not been incorporated into theories that explain aphasic comprehension of different syntactic structures. Gibson, Sandberg, Fedorenko, Bergen, and Kiran (2015) found that individuals with aphasia (IWAs) rely more on plausibility than on syntax in interpreting sentences.

A comparison to a computational model based on frequency and a collection of aphasic comprehension data showed a matching performance for the frequent active and less frequent passive structures (St John & Gernsbacher, 1998). However, one criticism of this study is that it only investigates actives and passives, and this is insufficient to draw a conclusion on the basis of frequency, when the latter is both more complex from a syntactic and morphological point of view and less frequent.

Though more frequent, passives are more difficult for individuals with Broca's aphasia than subject clefts and subject relatives, in both production and comprehension (Abuom et al., 2013). Also, linguistic complexity (i.e., derived word order) rather than frequency of the grammatical construction has been shown to be the predictive factor for agrammatic sentence production in Dutch (Bastiaanse, Bouma, & Post, 2009).

A review of usage-based effects at the sentence level (Gahl & Menn, 2016) suggests that the influence of frequency extends to other syntactic contrasts (such as sentences with unaccusative verbs and sentences with other types of verbs as well as subject and object relatives). Additionally, Gahl and Menn (2016) argued that frequency at the sentence level interacts with other factors such as the frequency in which a certain verb is used in certain constructions. Gahl et al. (2003) found that there is a modulation effect of lexical bias where, for example, passive sentences with passive-bias verbs are comprehended better than passive sentences with active-bias verbs.

To recap, frequency of sentence structure plays a role in normal-language processing and has an effect on sentence processing in aphasia. However, the nature of the relationship between relative syntactic frequency and word order in aphasic sentence comprehension has yet to be established.

If the frequency of the passive construction with its derived order does indeed play a role in SI sentence comprehension of individuals with Broca's aphasia, as expected on the basis of the data of Postman (2004) and Anjarningsih et al. (2012), relatively good comprehension of this sentence type is to be expected, even though the SI passive has derived word order, the frequency and pragmatic constraints thus overruling the DOP-H. However, this can only be concluded when influence of derived word order is shown by poor comprehension of SI object cleft sentences. Noting that the object cleft in SI are both infrequent and use a derived order of agent and theme, this structure is predicted to be impaired.

Methods

Participants

Twenty-three participants were included in the sentence comprehension study: 11 IWAs who were classified as having Broca's aphasia on the TADIR/SI aphasia battery and spoke agrammatically, and 12 NBD standard Indonesian speakers. Aphasic participants were recruited from six nursing homes in several cities of Central Java, Indonesia (Surakarta, Brebes, Semarang, Ungaran, Bantul and Sleman). Their demographic profiles were acquired from the caretaker of the nursing home and individual interviews. The NBD group was comprised of healthy individuals from two nursing homes in Surakarta and Brebes. While the participants' first languages were not SI, the experimenter as well as nursing staff and interns (for the nurse profession) communicated to the participants using SI. Since these were state-owned nursing homes, many of the civil servants and interns working in the nursing homes come from other areas of Java, where they may speak a different dialect of Javanese or perhaps an entirely different language (e.g., Sundanese from West Java). Additionally, the experimenter has little to no proficiency in Javanese and therefore looks for

Table 1. Demographics of the participants.

n	Type	Education (years)	Age (sd)	Gender	Handedness	TPO (years)	Token Test
12	NBD	11.5 (1.1)	66.9 (5.3)	Male = 2	Right	–	49.2 (0.9)*
11	Broca	12.1 (1.5)	68.7 (8.3)	Male = 5	Right	5.9 (4.3)	31.5 (3.8)

Notes: Mean numbers are given with (sd).
TPO = Time post onset.
*The Token Test norm is from a non-age-matched group (n = 26, mean age = 28.3).

participants who are highly proficient in SI. Thus, high proficiency in SI is part of the inclusion criteria of the current study.

Due to the lack of aphasia diagnosis in nursing homes, first a screening test was administered to all stroke victims who had problems communicating or suffered from a right hemiparesis. An SI translation of the Token Test from the Dutch *Aachener Aphasie Test* (AAT: Graetz, de Bleser, & Wilmes, 1992) was used. An error rate of more than 15 on the Token Test indicates a high probability of aphasia when hearing and vision are intact. Those participants who had over 15/50 errors were then tested with the *Tes Afasia untuk Diagnosis, Informasi, dan Rehabilitasi* (TADIR: Indonesian Aphasia Test for Diagnosis, Information, and Rehabilitation; Dharmaperwira-Prins, 1996). A total of 24 post-stroke individuals were interviewed and tested with the Token Test, and 11 were tested with TADIR and participated in the sentence-comprehension study. These 11 aphasic individuals produced agrammatic speech and had relatively good comprehension and were classified as suffering from Broca's aphasia.[3] Table 1 provides an overview of the NBD participants and IWAs involved. Detailed individual information on aphasic participants can be found in Appendix 1.

Individual characteristics (such as related motor disorders and corrected hearing and vision) were noted. A written informed consent form for the interview and testing was either signed or finger-stamped after being read to every participant.

Materials and procedure

The sentence comprehension test was adapted from the subtask for sentence comprehension of the Verb and Sentence Test (VAST; Bastiaanse, Edwards, Mass, & Rispens, 2003). It contains 40 semantically reversible sentences distributed equally over four conditions representing four sentence types (10 actives, 10 passives, 10 subject clefts and 10 object clefts). In addition, a set of practice items is used to introduce the four conditions. Each item was presented as a set of four pictures: one target and three distractors (see Figure 1).

In Figure 1, for the target sentence *Sapi ditendang kuda:* "The cow is kicked by the horse", the target action is (A). This picture is contrasted with a distractor that has reversed assignment of thematic roles labelled as "reversed role distractor" (D). In (C), a semantically related verb with the same order of thematic roles as the target is used to form a sentence, referred to as "lexical distractor". Finally, the lexical distractor is presented with these roles reversed, forming the "reversed role lexical distractor" (B). The position of the target and distractor pictures was balanced.

The procedure was as follows: while looking at the practice item, the participant was asked whether he/she could see each picture clearly, and whether he/she could see all four
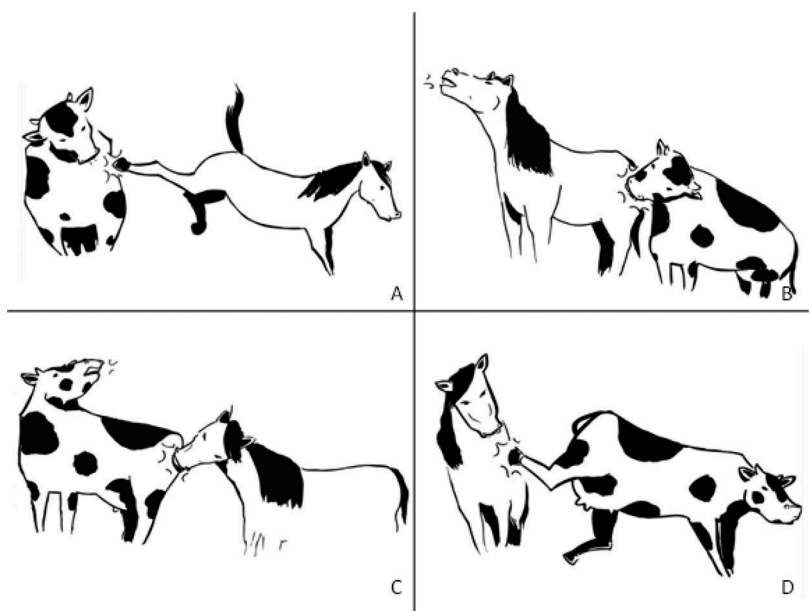
Figure 1. An example of an item set from the sentence comprehension test.

pictures. For the practice items prior to starting the task, the participant heard all four-sentence types, one by one. When a practice item was answered incorrectly, the participant was corrected. After the practice items, no feedback was given. The auditory stimulus was repeated one time on request. If the participant asked for a second repetition, this was given, but the answer was counted as incorrect. Self-corrections were counted as correct.

Scoring

Quantitative and qualitative error analyses were performed. One point was given for each correct response. In the event of an incorrect response, the error was noted and classified per type.

Results

Table 2 summarises the group results for all participants. Scores for each aphasic individual can be found in Appendix 2.

The control group of 12 NBDs performed close to ceiling level (mean = 0.97, range = 38–40 out of 40). The NBD group scored significantly higher in comparison to

Table 2. Mean numbers correct and standard deviations (sd) on the sentence comprehension test.

n	Type	Mean (sd)	Active	Subject cleft	Passive	Object-cleft
12	NBD	38.5 (1.0)	10 (0)	9.8 (0.4)	9.9 (0.3)	8.8 (1.0)
11	Broca	24.2 (3.9)	6.8 (1.8)	6.9 (2.1)	6.7 (1.4)	3.7 (1.1)
	Max.	40	10	10	10	10

Note: NBD = Non-brain-damaged participants.

Table 3. Mean number of sentence comprehension errors (sd) of the individuals with Broca's aphasia.

N	mean (sd)	RR	LD	RRLD
11	15.7 (4.0)	11.7 (2.8)	3.1 (2.1)	0.9 (1.2)

Notes: Maximum number of errors is 40.
RR = Reversed role distractor; LD = lexical distractor; RRLD = reversed role lexical distractors.

the IWAs (Mann-Whitney U test: $U = 0$, $p = .001$). We use a logistic generalised linear model with correctness as the dependent variable while including the maximum random-effect structure shown by the data. The random effects in this model include participants, items and position of the target picture (whether the target picture is on the right, to exclude possible visual field deficits).

In our model, the main effect of condition (sentence type) remained significant when all random effects are included. A multiple comparisons test (Tukey contrasts) was conducted to reveal significant difference between the conditions. The scores for object clefts were significantly lower than those for actives ($Z = -3.44$, SE = 0.36, $p < 0.01$), passives ($Z = 3.45$, SE = 0.36, $p < 0.01$) and subject clefts ($Z = 3.65$, SE = 0.36, $p < 0.01$). Actives did not differ significantly from passives ($Z = 0.02$, SE = 0.36, $p = 1$). Finally, subject clefts did not differ from actives ($Z = 0.13$, SE = 0.37, $p = 0.99$) and passives ($Z = 0.12$, SE = 0.37, $p = 0.99$).

Mismatches were classified into three categories according to the three distractor types: reversed role distractors, lexical distractors, and reversed role lexical distractors. Table 3 shows mean numbers for the different types of errors. The individual error rates by type can be found in Appendix 3.

Role reversal errors were made significantly more often than both the lexical errors (t(1) = 10.9; $p < 0.001$) and the lexical errors with role reversals (t(10)10.32, $p < 0.001$).

Discussion

The current study investigated the comprehension of sentences with base and derived word order and with and without clefting in individuals with Broca's aphasia and NBDs. The NBD group scored at ceiling and significantly higher than the aphasia group. For the whole group with Broca's aphasia, there was no difference between comprehending active and passive sentences. However, this group had lower scores on object-clefts than on any other structure, including subject clefts. Thus, the effect of word order is only apparent in the cleft conditions. With regard to embedding, there is no significant difference between actives and subject clefts. While passives are comprehended better than object clefts, it can be argued that the cause of this cannot be embedding, or at least the increased complexity entailed by embedding alone (since the subject clefts are relatively intact). This implies that the influence of embedding (via the use of subject and object relatives) that was reported for sentence comprehension of individuals with Broca's aphasia in English and Swahili (Abuom et al., 2013) is not found in SI.

The DOP-H predicts passives and object clefts to be comprehended less well than actives and subject clefts, respectively. Thus, the data on the SI passive do not support the DOP-H. In line with previous studies in SI, the current data show that the frequent passive constructions in SI are relatively well spared in individuals with Broca's aphasia.

Relative syntactic frequency in sentence comprehension

As previously mentioned, the four conditions tested in the sentence comprehension task differ not only in word order and embedding, but also in frequency. Gil (2006) has shown that the frequency of passives in both adult and child speech is much higher in SI than it is in other languages such as English. For that reason, both the passive and the active are marked as "frequent". Though no formal spoken corpora were used for the exact figures, Gil (2006) used a corpus of child and adult speech to confirm that the frequency of passive structures is higher in SI than in English and that the passive construction is acquired earlier in SI. Additionally, with the optional omission of the preposition *oleh* or "by", passive structures in SI have an NP–V–NP structure similar to actives. However, the reason that the passive construction is relatively spared may also be related to its pragmatic and discourse functions: (1) using an active construction when addressing someone who is higher in hierarchy is very impolite (Randriamasimanana, 1999); (2) the SI (*di*-) passives are associated with foregrounding, stating punctuality and/or factuality, as well as expressing sequential events (Kaswanti-Purwo, 1988). These functions explain the high frequency of the passive in SI. It would be interesting to disentangle the effect of word order and frequency-related effects spurred by topicalisation and pragmatic constraints. The data on the passive construction in SI so far show that frequency and/or pragmatic constraints protect the passives from being affected in SI Broca's aphasia, both in comprehension and in production (Anjarningsih et al., 2012; Postman, 2004; current study). Thus, the data also support the idea of Gahl and Menn (2016) that the conditions under which grammatical constructions are used influences aphasic performance.

The present results open a new door for future studies: comprehension (and production; Anjarningsih et al., 2012) of sentence structures in aphasia should no longer be defined solely in terms of syntactic structure. Frequency of construction (current study) and lexical items used in particular constructions (Gahl & Menn, 2016) should also be considered as a factor.

Notes

1. We do not go into the discussion on the differences and/or similarities between agrammatism and Broca's aphasia. Here, we follow Goodglass et al.'s (2001) definition of Broca's aphasia: agrammatic speech and relatively good (word) comprehension. Since in the literature on sentence comprehension the term 'Broca's aphasia' is most common, this is the term we use throughout the paper.
2. The use of the term "passive" to refer to all un-affixed and *di*- verb clauses has been widely debated. Several researchers (Hopper, 1983; Rafferty, 1982) claim that some of the traditionally considered passive constructions are passives, while others should be classified as ergative. However, a review on passive constructions being ergative concluded that none of the arguments successfully demonstrated that Indonesian/Malay is an ergative language (Cumming & Wouk, 1987).
3. Two agrammatic speakers were classified as suffering from transcortical motor aphasia, because their repetition was relatively good. However, since they spoke in prototypical telegraphic speech, we included them in the Broca group.

Acknowledgements

Disclosure statement

Funding

ORCID

Bernard A. Jap http://orcid.org/0000-0001-5276-9805

References

Abuom, T. O., Shah, E., & Bastiaanse, R. (2013). Sentence comprehension in Swahili-English bilingual agrammatic speakers. *Clinical Linguistics & Phonetics*, *27*, 355–370. doi:10.3109/02699206.2013.775346

Anjarningsih, H. Y., Haryadi-Soebadi, R. D., Gofir, A., & Bastiaanse, R. (2012). Characterising agrammatism in standard Indonesian. *Aphasiology*, *26*, 757–784. doi:10.1080/02687038.2011.648370

Bastiaanse, R., Bouma, G., & Post, W. (2009). Linguistic complexity and frequency in agrammatic speech production. *Brain and Language*, *109*, 18–28. doi:10.1016/j.bandl.2008.12.004

Bastiaanse, R., & Edwards, S. (2004). Word order and finiteness in Dutch and English Broca's and Wernicke's aphasia. *Brain and Language*, *89*, 91–107. doi:10.1016/S0093-934X(03)00306-7

Bastiaanse, R., Edwards, S., Mass, E., & Rispens, J. (2003). Assessing comprehension and production of verbs and sentences: The Verb and Sentence Test (VAST). *Aphasiology*, *17*, 49–73. doi:10.1080/729254890

Bastiaanse, R., & Van Zonneveld, R. (2005). Sentence production with verbs of alternating transitivity in agrammatic Broca's aphasia. *Journal of Neurolinguistics*, *18*, 57–66. doi:10.1016/j.jneuroling.2004.11.006

Bastiaanse, R., & Van Zonneveld, R. (2006). Comprehension of passives in Broca's aphasia. *Brain and Language*, *96*, 135–142. doi:10.1016/j.bandl.2005.06.012

Burchert, F., de Bleser, R., & Sonntag, K. (2003). Does morphology make the difference? Agrammatic sentence comprehension in German. *Brain and Language*, *87*, 323–342. doi:10.1016/S0093-934X(03)00132-9

Caplan, D., & Futter, C. (1986). Assignment of thematic roles to nouns in sentence comprehension by an agrammatic patient. *Brain and Language*, *27*, 117–134. doi:10.1016/0093-934X(86)90008-8

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, *3*, 572–582. doi:10.1016/0093-934X(76)90048-1

Cole, P., & Hermon, G. (2008). Malay/Indonesian syntax from an Austronesian perspective: An introduction. *Lingua*, *118*, 1431–1439. doi:10.1016/j.lingua.2007.08.003

Cumming, S., & Wouk, F. (1987). Is there 'discourse ergativity' in Austronesian languages?. *Lingua*, *71*, 271–296. doi:10.1016/0024-3841(87)90075-1

Dharmaperwira-Prins, R. I. (1996). *Tes Afasia untuk Diagnosis Informasi Rehabilitasi*. Jakarta: Balai Penerbit FK UI.

Englebretson, R. (2008). From subordinate clause to noun-phrase: Yang constructions in colloquial Indonesian. In Laury, R. (Ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions* (pp. 1–33). Amsterdam: John Benjamins.

Ewing, M. C., & Cumming, S. (1998). Relative clauses in Indonesian discourse: Face to face and cyberspace interaction. In S. L. Chelliah & W. J. De Reuse (Eds.), *Papers from the fifth annual meeting of the Southeast Asian Linguistics Society* (pp. 79–96). Tempe, AZ: Arizona State University.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203. doi:10.1016/S0010-0285(03)00005-7

Gahl, S., & Menn, L. (2016). Usage-based approaches to aphasia. *Aphasiology*, 1–17. Advance online publication. doi:10.1080/02687038.2016.1140120

Gahl, S., Menn, L., Ramsberger, G., Jurafsky, D. S., Elder, E., Rewega, M., & Audrey, L. H. (2003). Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. *Brain and Cognition*, *53*, 223–228. doi:10.1016/S0278-2626(03)00114-3

Gibson, E., Sandberg, C., Fedorenko, E., Bergen, L., & Kiran, S. (2015). A rational inference approach to aphasic language comprehension. *Aphasiology*, 1–20. Advance online publication. doi:10.1080/02687038.2015.1111994

Gil, D. (2006). The acquisition of voice morphology in Jakarta Indonesian. In N. Gagarina & I. Gülzowm (Eds.), *The acquisition of verbs and their grammar: The effect of particular languages* (pp. pp. 201–227). Dordrecht: Springer Netherlands.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston diagnostic aphasia examination* (3rd ed.). Philadelphia: Lippincott, Williams and Wilkins.

Graetz, P., de Bleser, R., & Wilmes, K. (1992). *De Akense Afasietest* [The Aachen Aphasia test]. Lisse: Swetz & Zeitlinger.

Grodzinsky, Y. (1995). A restrictive theory of agrammatic comprehension. *Brain and Language*, *50*, 27–51. doi:10.1006/brln.1995.1039

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences*, *23*, 1–21. doi:10.1017/S0140525X00002399

Hassall, T. (2005). Taboo object relative clauses in Indonesian. SEALS XV, 1

Hopper, P. (1983). Ergative, passive, and active in Malay narrative. In F. Klein-Andreu (Ed.), *Discourse perspectives on syntax* (pp. 67–88). New York: Academic Press.

Juliano, C., & Tanenhaus, M. K. (1994). A constraint-based lexicalist account of the subject/object attachment preference. *Journal of Psycholinguistic Research*, *23*, 459–471. doi:10.1007/BF02146685

Kaswanti-Purwo, B. (1988). Voice in Indonesian: A discourse study. *Passive and Voice*, *16*, 195.

Kaswanti-Purwo, B. (1991). Perkembangan bahasa anak: Pragmatik dan tata bahasa. In S. Dardjowidjojo (Ed.), *PELLBA 4* (pp. 157–202). Jakarta: Penerbit Kanisius.

Lewis, M. P., Simons, G. F., & Fennig, C. D. Eds. (2013). *Ethnologue: Languages of the world: Seventeenth edition*. Dallas, TX: SIL International. Retrieved from http://www.ethnologue.com

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703. doi:10.1037/0033-295X.101.4.676

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. pp. 249-308). Hillsdale, NJ: Lawrence Erlbaum.

McCune, K. (1979). Passive function and the Indonesian passive. *Oceanic linguistics*, *18*, 119–169.

Mistica, M., Andrews, A., Arka, I. W., & Baldwin, T. (2009, December). Double double, morphology and trouble: Looking into reduplication in Indonesian. In *Australasian Language Technology Association Workshop 2009* (p. 44).

Mitchell, D. C. (1994). Sentence parsing. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 375–409)). New York: Academic Press.

Mitchell, D. C., Cuetos, F., Corley, M. M., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, 24, 469–488. doi:10.1007/BF02143162

Nelson, G. (1997). Cleft constructions in spoken and written English. *Journal of English Linguistics*, 25, 340–348. doi:10.1177/007542429702500408

Nichols, J., & Bickel, B. (2013). Locus of marking: Whole-language typology. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures*. Munich: Max Planck Digital Library.

Nomoto, H. (2010). Making sense of the optionality of voice marking in Malay/Indonesian. In *Proceedings of workshop on Indonesian-type voice system* (pp. 37–44).

Postman, W. A. (2002). *Thematic role assignment in Indonesian: A case study of agrammatic aphasia* (Doctoral Dissertation). Cornell University.

Postman, W. A. (2004). Processing of complex sentences in a case of aphasia in Indonesian: Thematic vs. linear strategies. *Journal of Neurolinguistics*, 17, 455–489. doi:10.1016/j.jneuroling.2004.09.001

Postman, W. A. (2011). Some critical concerns for adapting the bilingual aphasia test to Bahasa Indonesia. *Clinical Linguistics & Phonetics*, 25, 619–627. doi:10.3109/02699206.2011.566665

Rafferty, E. (1982). *Discourse structures of the Chinese Indonesian of Malang*. Jakarta: University Atma Jaya.

Randriamasimanana, C. (1999). The Malayo-Polynesian origins of Malagasy. In N. A. Kirk & P. J. Sidwell (Ed.), *From Neanderthal to Easter Island: A tribute to, and a celebration of the work of W. Wilfried Shuhmacher presented on the occasion of his 60th birthday* AHL Studies in the Science & History of Language 2 (Vol. 2, pp. 26–43). Melbourne: Association for the History of Language.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, 57, 348–379. doi:10.1016/j.jml.2007.03.002

Schwartz, M. F., Linebarger, M. C., Saffran, E. M., & Pate, D. S. (1987). Syntactic transparency and sentence interpretation in aphasia. *Language and Cognitive Processes*, 2, 85–113. doi:10.1080/01690968708406352

Sinnemäki, K. (2010). Word order in zero-marking languages. *Studies in Language*, 34, 869–912. doi:10.1075/sl.34.4

Sneddon, J. N. (1996). *Indonesian: A comprehensive grammar*. London: Routledge.

St John, M. F., & Gernsbacher, M. A. (1998). Learning and losing syntax: Practice makes perfect and frequency builds fortitude. In A. F. Healy & L. E. Bourne Jr (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 231–255). Mahwah, NJ: Lawrence Erlbaum Associates, Incorporated.

Stack, M. (2005). Word order and intonation in Indonesian. *LSO Working Papers in Linguistics 5: Proceedings of WIGL*. 168–182.

Suwarso, S. (1994). *Struktur kalimat pasif dalam Bahasa Melayu Klasik*. Unpublished research report, University of Indonesia, Depok, Indonesia.

Trueswell, J., & Tanenhaus, M. (1994). Toward a lexical framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–179). Hillsdale, NJ: Erlbaum.

Verhaar, J. W. (1978). Some notes on the verbal passive in Indonesian. *NUSA: Linguistic Studies in Indonesian and Languages in Indonesia*, 6, 11–19.

Yarbay Duman, T., Ozgirgin, N., Altinok, N., & Bastiaanse, R. (2011). Sentence comprehension in Turkish Broca's aphasia: An integration problem. *Aphasiology*, 25, 908–926. doi:10.1080/02687038.2010.550629

Appendix 1. Background information of the participants with aphasia

No	Education (years)	Age	Origin	Gender	Handed-ness	Time post-onset (years)	First language	Token Test
1	12	70	Brebes	M	Right	1.5	Javanese	28/50
2	12	71	Brebes	M	Right	10	Betawi	33/50
3	9	53	Brebes	F	Right	13	Javanese	30/50
4	12	66	Surakarta	M	Right	3	Javanese	37/50
5	12	65	Surakarta	F	Right	3–4	Javanese	37/50
6	16	57	Semarang	M	Right	5	Javanese	37/50
7	12	73	Semarang	F	Right	0.5	Javanese	29/50
8	12	82	Semarang	F	Right	10	Javanese	26/50
9	12	68	Ungaran	F	Right	7	Javanese	28/50
10	12	81	Sleman	F	Right	0.5	Javanese	30/50
11	12	70	Sleman	M	Right	11	Javanese	31/50

Appendix 2. Individual sentence comprehension scores for participants with aphasia

	Total	Active	s-cleft	Passive	o-cleft
1	25	8	7	5	5
2	23	9	5	7	2
3	31	9	9	10	3
4	29	9	9	5	6
5	23	6	7	5	5
6	28	6	10	8	4
7	18	3	5	7	3
8	18	5	4	6	3
9	24	7	7	7	3
10	22	7	4	7	4
11	25	6	9	7	3
Maximum	40	10	10	10	10
Mean	24.18	6.82	6.91	6.73	3.73
sd	3.93	1.80	2.07	1.42	1.14
%	60.45	28.20	28.57	27.82	15.41

Appendix 3. Individual errors for participants with aphasia (maximum number of errors 440)

	Total	Reversed role	Lexical distractor	Reversed role lexical distractor
1	15	12	2	1
2	17	16	1	0
3	9	8	1	0
4	11	8	3	0
5	17	12	3	2
6	11	9	1	1
7	22	15	6	1
8	22	14	4	4
9	16	15	1	0
10	18	10	7	1
11	15	10	5	0
Total	173	129	34	10
Mean	15.73	11.73	3.09	0.91
sd	4.02	2.80	2.07	1.16
%		74.57	19.65	5.78