

Crosslinguistic Studies of Clause Combining

The multifunctionality
of conjunctions

edited by Ritva Laury

John Benjamins Publishing Company

Crosslinguistic Studies of Clause Combining

Typological Studies in Language (TSL)

A companion series to the journal *Studies in Language*. Volumes in this series are functionally and typologically oriented, covering specific topics in language by collecting together data from a wide variety of languages and language typologies.

General Editor

Michael Noonan
University of Wisconsin-Milwaukee

Assistant Editors

Spike Gildea
University of Oregon

Suzanne Kemmer
Rice University

Editorial Board

Wallace Chafe
Santa Barbara

Matthew S. Dryer
Buffalo

Paul J. Hopper
Pittsburgh

Ronald W. Langacker
San Diego

Doris L. Payne
Oregon

Sandra A. Thompson
Santa Barbara

Bernard Comrie
Leipzig / Santa Barbara

John Haiman
St Paul

Andrej A. Kibrik
Moscow

Charles N. Li
Santa Barbara

Frans Plank
Konstanz

Dan I. Slobin
Berkeley

R.M.W. Dixon
Melbourne

Jerrold M. Sadock
Chicago

Edith Moravcsik
Milwaukee

Andrew Pawley
Canberra

Bernd Heine
Köln

Volume 80

Crosslinguistic Studies of Clause Combining. The multifunctionality of conjunctions

Edited by Ritva Laury

Crosslinguistic Studies of Clause Combining

The multifunctionality of conjunctions

Edited by

Ritva Laury

University of Helsinki

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Studies of clause combining : the multifunctionality of conjunctions / edited by Ritva Laury.

p. cm. (Typological Studies in Language, ISSN 0167-7373 ; v. 80)

Includes bibliographical references and index.

1. Grammar, Comparative and general--Clauses. 2. Grammar, Comparative and general--Conjunctions. I. Laury, Ritva.

P297.S78 2008

415'.7--dc22

2008021013

ISBN 978 90 272 2993 9 (Hb; alk. paper)

© 2008 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

List of contributors	VII
Introduction	IX
<i>Ritva Laury & Sandra A. Thompson</i>	
From subordinate clause to noun-phrase: <i>Yang</i> constructions in colloquial Indonesian	1
<i>Robert Englebretson</i>	
On quotative constructions in Iberian Spanish	35
<i>Ricardo Etxepare</i>	
Bulgarian adversative connectives: Conjunctions or discourse markers?	79
<i>Grace E. Fielder</i>	
Projectability and clause combining in interaction	99
<i>Paul J. Hopper & Sandra A. Thompson</i>	
Conjunction and sequenced actions: The Estonian complementizer and evidential particle <i>et</i>	125
<i>Leelo Keevallik</i>	
Clause combining, interaction, evidentiality, participation structure, and the conjunction-particle continuum: The Finnish <i>että</i>	153
<i>Ritva Laury & Eeva-Leena Seppänen</i>	
The grammaticization of <i>but</i> as a final particle in English conversation	179
<i>Jean Mulder & Sandra A. Thompson</i>	
Quotative <i>tte</i> in Japanese: Its multifaceted functions and degrees of “subordination”	205
<i>Shigeko Okamoto & Tsuyoshi Ono</i>	
Quoting and topic-marking: Some observations on the quotative <i>tte</i> construction in Japanese	231
<i>Ryoko Suzuki</i>	
Index of names	247
Index of subjects	251

List of contributors

Ritva Laury
Department of Finnish Language
and Literature
PL 3, Fabianinkatu 33 (4th floor)
00014 University of Helsinki
Finland
ritva.laury@helsinki.fi

Robert Englebretson
Dept. of Linguistics, MS23
Rice University
6100 Main St.
Houston, TX 77005-1892
USA
reng@rice.edu

Ricardo Etxepare
Antzaran Plaza 16-2, 4
20301 (Gipuzkoa)
Spain
retxepare@gmail.com

Grace E. Fielder
Department of Russian
and Slavic Studies
PO Box 210105
Tucson, AZ 85721-0105
USA
gfielder@email.arizona.edu

Paul J. Hopper
Carnegie Mellon University
Department of English
Pittsburgh, PA 15213
USA
hopper@cmu.edu

Leelo Keevallik
Department of Modern Languages
Box 636
Uppsala University
751 26 Uppsala
Sweden
Leelo.Keevallik@moderna.uu.se

Jean Mulder
School of Languages and Linguistics
University of Melbourne, Victoria 3010
Australia
jmulder@unimelb.edu.au

Shigeko Okamoto
Language Program
1156 High Street
Cowell Academic Services
University of California
Santa Cruz, CA 95064
USA
sokamoto@ucsc.edu

Tsyoshi Ono
Department of East Asian Studies
University of Alberta
Edmonton, AB T6G 2E6
Canada
tono@ualberta.ca

Eeva-Leena Seppänen
Department of Finnish language
and literature
PL 3, Fabianinkatu 33 (4th floor)
00014 University of Helsinki
Finland
Eeva-Leena.Seppanen@helsinki.fi

Ryoko Suzuki
Faculty of Economics
Keio University
4-1-1- Hiyoshi
Kohokuku Yokohama 223-8521
Japan
ryoko@hc.cc.keio.ac.jp

Sandra A. Thompson
Department of Linguistics
University of California, Santa Barbara
Santa Barbara, CA 93106-3100
USA
sathomps@linguistics.ucsb.edu

Introduction

Ritva Laury and Sandra A. Thompson

University of Helsinki, Finland/University of California, Santa Barbara

1. Overview

This volume is a collection of articles based on papers presented at the panel on clause combining at the 6th International Pragmatics Conference at Lake Garda in July, 2005. The papers by Suzuki and Hopper & Thompson were not part of the panel. Suzuki was a discussant in the panel, and Thompson was an original participant with her paper coauthored with Jean Mulder. Although the authors of that paper were not able to present it at Garda, fortunately it is included in the volume. The papers represent eight languages from four different language families: Bulgarian, English, German, and Spanish; Estonian and Finnish; Indonesian; and Japanese.

Several major themes important to the functional-typological school of linguistics are represented in the papers. Many of the papers address the issue of linguistic categoriality. The papers by Englebretson, Fielder, Keevallik, Laury & Seppänen, and Okamoto & Ono demonstrate the polyfunctionality and fuzzy edges of linguistic categories, and their findings support the point that linguistic items achieve categoriality through use in discourse, originally made by Hopper and Thompson (1984). Englebretson's paper makes an additional important typological point based on his examination of *yang* constructions in Indonesian. Following Dryer (1997), who argues for a conception of grammar as language- and construction-specific, Englebretson points out the fallacy of assigning cross-linguistically valid labels to constructions found in very different languages; if we take seriously the claim that categoriality is a result of use patterns in discourse, then we need to consider that similarity in the behaviour of linguistic items in different languages is a result of certain degrees of isomorphism between structure and function.

Quite a different look at polyfunctionality is offered in the paper by Etxepare, who examines the use of the Peninsular Spanish *que* as a root complementizer, that is, as a clause-initial item which is not associated any (overt) main clause. Etxepare's paper adds theoretical breadth to the volume. It focuses on the same

issue, the use of complementizers without any associated main clause, as several other papers in the volume, but from a very different theoretical framework.

Closely linked to the issue of categoriality and polyfunctionality is the process through which linguistic items develop new uses. The papers by Mulder & Thompson and by Suzuki focus on the grammaticalization of items used as conjunctions into additional uses as particles. The papers by Keevallik and Laury & Seppänen also touch upon the historical development of the complementizers whose wider uses their papers concern.

A further step in the development of novel uses of linguistic items and constructions is that with increased use, they often become lexically specific and formulaic, allowing little variation in their internal structure. In such a process, their internal structure gradually becomes opaque and they become set phrases strongly associated with particular functions. The paper which most clearly illustrates this process is the one by Hopper & Thompson, in which they examine the use of certain clausal constructions which participants in conversation use to project upcoming actions.

2. Categoriality

As noted, one of the main themes of the volume has to do with categoriality of linguistic forms, in particular forms which function as conjunctions (such as relativizers, complementizers, quotatives and adversatives), but which also have other uses in syntactic environments that do not involve the joining of one clause to another. This is the main topic of papers by Englebretson, Fielder, Keevallik, Laury & Seppänen, and Okamoto & Ono. In addition, the papers by Mulder & Thompson and by Suzuki consider the grammaticization processes through which items that function as conjunctions develop other uses.

Robert Englebretson's paper discusses the use of the Indonesian *yang*, which has been classified as a relativizer in many grammars. In his highly innovative paper, Englebretson shows that the constructions introduced by *yang* range from traditional relative clauses to headless referring expressions serving as presupposed information in clefts, to referring expressions functioning as nominal arguments of predicates. Moreover, Englebretson shows that the *yang*-constructions in his conversational data form a continuum, so that there is no clear distinction between the various types of constructions in which *yang* is used in spoken Indonesian. The uses of *yang* thus illustrate the blurring of the distinction between relative clause and noun phrase. On the basis of the use of *yang* in spoken interactions, Englebretson suggests that it does not easily fit into any of the grammatical categories provided by existing theories. Following Dryer (1997), Englebretson argues that

the positing of crosslinguistically valid labels may be misleading, and that instead, if grammatical forms can be shown to emerge from their use in discourse, the resultant forms may in fact be language-specific and unique to the language in which they appear, instead of fitting neatly into a well-defined crosslinguistically valid category.

Grace Fielder's paper discusses the multiple uses of the Bulgarian adversative connectives *no*, *ama*, and *ami*, which can all be translated into English as "but", and which function as conjunctions, particles and/or discourse markers. Fielder's paper is based on a qualitative and quantitative analysis of spoken and written corpora; like Englebretson, she demonstrates that the connectives she examines are not neatly divisible into conjunctions and discourse markers, but that instead, the different uses form a continuum. Through a careful analysis of the polyfunctional uses of these items, Fielder shows that there are systematic correlations between a connective function and (a) the position of the item in initial versus medial position, as well as (b) its occurrence in a particular genre and register.

Several of the papers in the volume deal with the polyfunctionality of quotatives or linguistic elements which are used as complementizers. Leelo Keevallik's paper belongs in this group. Like the linguistic items studied by Englebretson and Fielder, the uses of the item Keevallik studies, the Estonian *et*, also form a continuum and are not neatly separable into complementizer and particle uses. Keevallik shows that the use of *et* as a clause-initial element in both oral and Internet communication serves to attribute the upcoming content to a prior speaker or writer, and thus introduces a new voice into the current discourse. Keevallik argues that when the principal of the upcoming contribution is clear from the context, no quotative "main" clause is used. She also shows that the use of *et* is connected with and implements certain types of embedded conversational actions, and argues that the use of *et* in conversational sequences is not related to any monological pattern in a straightforward way. Instead, sequences containing *et* are shown to be the mutual achievement of the participants in the conversation, and a side-product of the implemented actions as well as an integrated part of them. Keevallik shows that the meaning of *et* in each of its uses is interactionally contingent and jointly constructed in real time as the conversation proceeds. Keevallik also presents a highly interesting hypothesis regarding the prosodic reflections of the grammaticizing pathways taken by *et*, showing how it can come to be latched either to the preceding speech act verb or to the upcoming complement clause.

The paper by Laury & Seppänen concerns a cognate of the form examined by Keevallik, the Finnish *et(tä)*. Like its Estonian counterpart, *et(tä)* also functions both as a complementizer and a particle in ordinary conversation. Through an analysis of the uses of *et(tä)* in a multi-party conversation, Laury & Seppänen also show that a division into particle and complementizer uses is not consistent

with the data. They also suggest that the description of *että*-clauses as complements is questionable, and that, interactionally, in both its complementizer and particle uses, *et(tä)* functions to regulate the participation framework (Goffman 1981) and to effect changes in footing. That is, like the Estonian *et*, the Finnish *et(tä)* also functions to introduce a new voice to the interaction. Like Keevallik, Laury & Seppänen also discuss the historical development of *et(tä)*. Although it is commonly assumed that particle uses developed from the clause-linking use of complementizers, the authors argue that it is quite possible that the particle use of *et(tä)* developed alongside of, or even before, the complementizer use.

The paper by Etxepare also concerns a complementizer, the Peninsular Spanish *que*. Using constructed examples, and approaching the topic from the perspective of formal syntax, Etxepare focuses on the use of *que* in which it heads a main clause and argues that such structures in Spanish involve an extra speech eventuality, which is syntactically represented as a complex verbal predicate, composed of a light verb and an aspectual projection which takes as a complement an utterance. Etxepare compares this use of *que* with elements which have been called Quotative Verbs in other languages and which are used to frame direct and semi-direct speech. In this collection, Etxepare's paper shows that even when approached from a different theoretical perspective, the uses of complementizers in positions where they do not combine two clauses, such elements are connected with the quoting of earlier utterances; that is, changes in footing, to use a term invoked by other authors in this volume.

Okamoto & Ono's contribution to the volume focuses on the use of the Japanese quotative particle *tte* in naturally occurring conversations. The authors argue that the five different usages of *tte*, which they identify based on syntactic, semantic and pragmatic properties, constitute subcategories of the particle ranging from an object complement marker (OCM) to a sentence-final particle (SFP), that they form a continuum, and that the SFP usage derives from the OCM usage. The study sees grammatical categories as inseparable from usage, in the sense that usage shapes grammar, and like many others in the volume, the authors also demonstrate the non-discrete nature of grammatical categories such as complementizer vs. particle and main clause vs. subordinate clause.

3. Grammaticalization

The papers by Suzuki and Mulder & Thompson both consider particles which also function to combine clauses. Suzuki's paper concerns the historical development of the Japanese quotative *tte*, also discussed by Okamoto and Ono. The paper is based on diachronic data taken from the conversational portions of novels from the

1830s to the 1980s, in addition to transcripts of informal conversations in modern Japanese. Suzuki focuses on the development of *tte* from a marker of quotation to an element which serves to introduce or track a referent, arguing that, once again, the uses are arranged on a continuum, with the referent-introducing and tracking uses having developed from the quotative function. Crucial to this development, according to Suzuki, is the nature of dialogic interaction, where speakers take parts of another's utterance and repeat them, thus building their contributions on preceding talk.

Mulder & Thompson's paper deals with the grammaticization process currently affecting the use of the English *but*, in which it is developing from a prosodic-unit-initial conjunction to a prosodic-unit-final discourse particle. The authors examine both Australian and American English conversational data, finding that as *but* moves along the continuum from conjunction to particle, its conversational function changes in a way that is consistent with what has been described in the grammaticization literature. They note that in the earlier stages of development, based on both prosodic and sequential evidence, a turn-final *but* indicates that an implication has been left "hanging", so that the material contrasting with the clause preceding *but* is only implied rather than overtly expressed. However, in the Australian data, the use of a turn-final *but* has progressed further; it is uttered with turn-final prosody, and instead of a construction with an implication "hanging", a construction emerges in which the semantically contrastive material is supplied in the prosodic unit ending with the final *but*. Thus, in Australian English, final *but* has developed into a final particle which marks contrastive content. The authors show that in a diaspora situation where speakers of the same language are widely separated, speech communities may be positioned at different points in this grammaticization process, suggesting that the development of *but* is only one example of many such processes which can be seen in English and other languages. They claim that the evidence argues strongly for an emergentist view of language structure, whereby large-scale patterns are created from repeated small-scale interactions, as suggested by other authors in the volume, for example Englebretson, Keevallik, and Suzuki.

4. The nature of biclausal constructions

The final paper in the volume, by Hopper and Thompson, widens the perspective by considering the nature of clause combining in general. The authors consider several 'multi-part' construction types in English and German, questioning the description of such constructions as 'biclausal', since the second part may not resemble a syntactic clause at all, but may rather be an indeterminate stretch of

discourse without a consistent syntactic structure. Instead, the first clause functions to project this upcoming segment. Indeed, in some such constructions, such as Extraposition and Pseudocleft, the first part is also not necessarily best viewed as a “clause”, but rather a fixed sequence allowing few lexical choices. The authors suggest that such apparently “biclausal” constructions can be more realistically understood in terms of real-time social activities as single, partially formulaic formats which are among the resources speakers use to manage interaction.

Based on a typologically broad sample of languages and a variety of theoretical approaches, these papers represent a significant rethinking of the nature and richness of the morphology, prosody, and syntax of clause combining.

References

- Dryer, M.S. 1997. Are grammatical relations universal? In *Essays on Language Function and Language Type: Dedicated to T. Givón*, J. Bybee, J. Haiman & S.A. Thompson (Eds), 115–143. Amsterdam: John Benjamins.
- Goffman, E. 1981. *Forms of Talk*. Oxford: Basil Blackwell.
- Hopper, P.J. & Thompson, S.A. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60: 703–752.

From subordinate clause to noun-phrase

Yang constructions in colloquial Indonesian

Robert Englebretson
Rice University

This paper addresses the blurring of a distinction between subordinate clause and nominal phrase in colloquial Indonesian conversation. Specifically, it presents a discourse-based analysis of the forms and functions of constructions introduced by the morpheme *yang*. Previous literature has generally described these constructions as relative clauses; however, this paper shows that they form a continuum around three general foci. They range from traditional relative clauses serving as modifiers of head nouns, to headless referring expressions serving as presupposed information in clefts, to referring expressions functioning as direct nominal arguments of predicates. The findings call into question the label “relative clause” for these Indonesian constructions, and support a view of grammatical categories as emerging from their use in discourse.

1. Introduction

Recent work on the grammar of clause combining in natural discourse, including many of the contributions to the present volume, has examined the putative distinction between main and subordinate clauses.¹ In some cases the distinction has been shown to be a continuum, and in other cases the distinction has been shown not to exist at all. Thompson and Mulac (1991a, 1991b) demonstrate the continuous, noncategorical nature in English of constructions ranging from complement

1. I would like to thank Ritva Laury, Sandy Thompson, Mike Ewing, Scott Paauw, and Susanna Cumming for fruitful discussion and comments at various stages of this project. I also appreciate the input from my fellow panel members at the Ninth International Pragmatics Conference in Riva del Garda, Italy (July 2005). As with all of my work on colloquial Indonesian, I am forever grateful to the numerous individuals who consented to participating in recordings for the Indonesian conversational corpus in Yogyakarta, and to the three Research Assistants who helped in data transcription and glossing. I alone of course bear full responsibility for any errors or misunderstandings presented in this work.

main clauses to epistemic adverbial phrases. Laury and Seppänen (this volume) show similar findings for constructions involving the Finnish complementizer *että*. Specific cases where the distinction between main and subordinate clause has been argued not to exist at all include object complements in conversational English (Thompson 2002), and complementation in colloquial Indonesian (Englebretson 2003). These works all highlight the blurring of a distinction between main and subordinate clause. The present paper shows the converse situation, namely the blurring of the distinction between subordinate clause and phrasal constituent. This paper addresses a particular grammatical construction in colloquial Indonesian, hereafter referred to as the *yang* construction, which ranges along a cline from subordinate clause to phrasal constituent. On one end of the continuum are relative clauses serving as modifiers of head nouns, and on the other end of the continuum are referring expressions serving as direct arguments of predicates. The following two examples illustrate *yang* constructions at both ends of this continuum.²

- (1) Restrictive Relative Clause (“Pencuri” IU 2832–2833)
2832

Bagi orang yang *nggak* bawa duit,
for people yang NEG bring money

2833

tidak terpengaruh.
NEG NV-influence
‘As for people who aren’t carrying money, they aren’t influenced.’

- (2) Referring Expression as Argument of a Predicate (“Pencuri” IU 3819)

Enak yang itu.
delicious yang that
‘That one is delicious.’

In (1) the *yang* construction *yang nggak bawa duit* is a subordinate clause which modifies the head noun *orang* “people”, forming a restrictive relative clause “people who don’t bring money”. It has a head and is itself a clause. In contrast, the *yang* construction in (2) *yang itu* “that one” is not a relative clause; it does not have a head noun, and it in fact is not even a clause at all. It is a referring

2. For ease of identification, in these and all subsequent examples, *yang* constructions are italicized, and relative clause heads (when present) are underlined. Since the purpose of this paper is to discuss the *yang* morpheme itself, I will gloss it eponymously as *yang*, rather than attempting a standard grammatical gloss. See Appendix A for a summary of glossing abbreviations.

expression consisting of the *yang* morpheme followed by the demonstrative/determiner *itu* “that”. It functions like an NP, refers anaphorically to a referent previously mentioned in the conversation, and serves as the single argument of the intransitive verb *enak* “pleasant/delicious”.

These two examples show that colloquial Indonesian *yang* constructions do not form a homogeneous group in terms of grammatical form or discourse function. Throughout this paper, I will argue that this construction in colloquial Indonesian must be understood to be continuous, nondiscrete, and polyfunctional, and that conventional labels such as “relative clause” are inadequate to characterize its full range of uses. The present paper follows researchers such as Dryer (1997) and Croft (2001) who argue that grammatical structures need to be analyzed on a case-by-case, language-specific and construction-specific basis. It also draws from the insights of Hopper and Thompson (1984) regarding acategoriality in grammar. Hopper and Thompson suggest that the lexical categories “noun” and “verb” arise from specific uses in local discourse contexts; the lexical category of a particular form is not given a priori, but “is imposed on the form by discourse” (Hopper & Thompson 1984: 747). Similarly, in this paper I suggest that Hopper and Thompson’s insights likewise apply to larger levels of phrasal and clausal forms as well; in the colloquial Indonesian data, categories like “relative clause” or “nominal phrase” are not a priori structures, but rather these categories emerge from specific discourse uses. Whether the colloquial Indonesian *yang* construction is realized as a “relative clause” or a “nominal phrase” depends on how it is being used in a particular discourse context.

This paper is organized as follows. I will begin with a brief summary and overview of the corpus data and methodology. I will then illustrate the formal continuum of Indonesian *yang* constructions, showing that there are three primary foci of this construction: full-fledged restrictive relative clauses modifying a head noun, headless clauses in equational sentences, and *yang* constructions occurring as referring expressions. I will outline the range of pragmatic functions of these constructions: restriction and attribution, grounding, framing, identification, and ad hoc reference formulation. I will conclude by showing how these construction types form a continuum; while there are three central foci of *yang* constructions, there is also overlap among them in terms of both form and function, and there are instances of constructions on the borders between these three foci. Thus, Indonesian *yang* constructions illustrate a blurring of the distinction between relative clause and nominal phrase. I propose that the continuous and polyfunctional nature of *yang* constructions makes it necessary to eschew more general typological labels such as “relative clause” for the colloquial Indonesian data, since ultimately it is only through their particular uses in discourse that the specific functions of this construction come into being.

2. Data and methodology

While a thorough discussion of the language variety designated here as colloquial Indonesian lies outside the scope of this paper, some background information may be helpful, especially for readers unfamiliar with the social and linguistic situation of present-day Indonesia. The Republic of Indonesia is a linguistically diverse country. The official language is Indonesian, the standard literary variety of which is promoted by the government and used in official capacities such as education, business, and the media. In addition to standard Indonesian, several hundred local languages are widely spoken (including Javanese, Balinese, Sundanese, etc.), as well as numerous regional varieties of Malay. Indonesians tend to be highly diglossic with a wide stylistic repertoire, as speakers' language choices are sensitive to a wide range of social situations.

By colloquial Indonesian, I mean the language variety that tends to be spoken informally among predominantly young, urban Indonesians who generally do not otherwise share a local language in common. There are significant differences between this variety and standard written Indonesian.³ This has at least two consequences for the current paper. First, findings of prior research on this topic in standard Indonesian do not necessarily apply to the language variety of the corpus used in the current study. And, secondly, findings of the current study are likewise not necessarily generalizable to other varieties of Indonesian. Discussion of *yang* constructions in descriptive grammars (Alwi et al. 1993; Sneddon 1996, inter alia), as well as published research specifically about *yang* constructions (Ewing 1991; Steinhauer 1992) have focused almost exclusively on the standard written variety of Indonesian. One exception to this is Ewing and Cumming (1998) who compare *yang* constructions in a corpus of internet NewsNet postings with a corpus of relatively informal conversation. One of the key findings of Ewing and Cumming's research is that the form and function of *yang* constructions is highly sensitive to and constrained by genre and register factors. For this reason, readers should be aware that the findings of the current paper may not be relevant to other genres, registers, or varieties of Indonesian. Whether, and to what extent, the *yang* constructions in colloquial Indonesian can be compared with those found in other varieties such as standard written Indonesian promises to be a fruitful area for future research.

The present study focuses on informal conversation – the most-ordinary and most-prevalent type of spoken language, and the locus in which the interplay

3. For a thorough discussion of colloquial Indonesian, see Ewing 2005a. For specific details about the colloquial features found in my corpus, see chapter 1 of Englebretson 2003.

between grammar and social interaction is most readily observable (cf. papers in Ochs et al. 1996, *inter alia*). The database for this study consists of four transcribed audio recordings of naturally-occurring, informal, colloquial Indonesian conversation. These segments comprise face-to-face interaction among college-age students who are friends or roommates and know each other well. Conversations are unprompted and spontaneous. These segments are a subset of the corpus fully described in Englebretson (2003), based on nearly 25 hours of data recorded during 1996 fieldwork in the city of Yogyakarta in Central Java. Because of the current study's explicit focus on informal interaction and everyday language as the locus for observing grammar in use, and in order to maintain genre consistency across speech events, the database for the present study includes only those four speech events out of the larger corpus characterized as informal, face-to-face, casual conversation among friends. Speech events in the corpus were transcribed using a slightly-modified version of the Du Bois et al. (1993) Discourse Transcription system. (See Appendix B for a brief overview of transcription conventions.) Table 1 lists the title, number of words, number of Intonation Units (IU's), and length of time for each segment in the database for the current study.

Table 1. Overview of Indonesian “informal” sub-corpus

Title	IU's	Words	Time
Pencuri	3,902	10,974	55:10
Blewah	1,331	3,823	24:23
Dingdong	1,685	4,053	30:14
Wisuda	1,826	5,224	41:40
	8,744	24,074	2:31:07

Each token of the *yang* morpheme in the corpus was tagged, and its associated construction was then coded in a relational database for features of both form and function. The resulting database yielded 410 *yang* constructions. Of these, 32 were truncated mid-IU and therefore could not be analyzed further and were subsequently excluded from the study. This left a total of 378 *yang* constructions to serve as the database for this paper. The following sections outline the form and function of the 378 *yang* constructions occurring in the corpus.

3. Formal and functional characteristics of *yang* constructions

Grammars of standard Indonesian have typically characterized *yang* constructions as relative clauses (Alwi et al. 1993; Sneddon 1996, *inter alia*), and descriptions of colloquial Indonesian likewise use this term (Ewing 2005a). Specific studies of