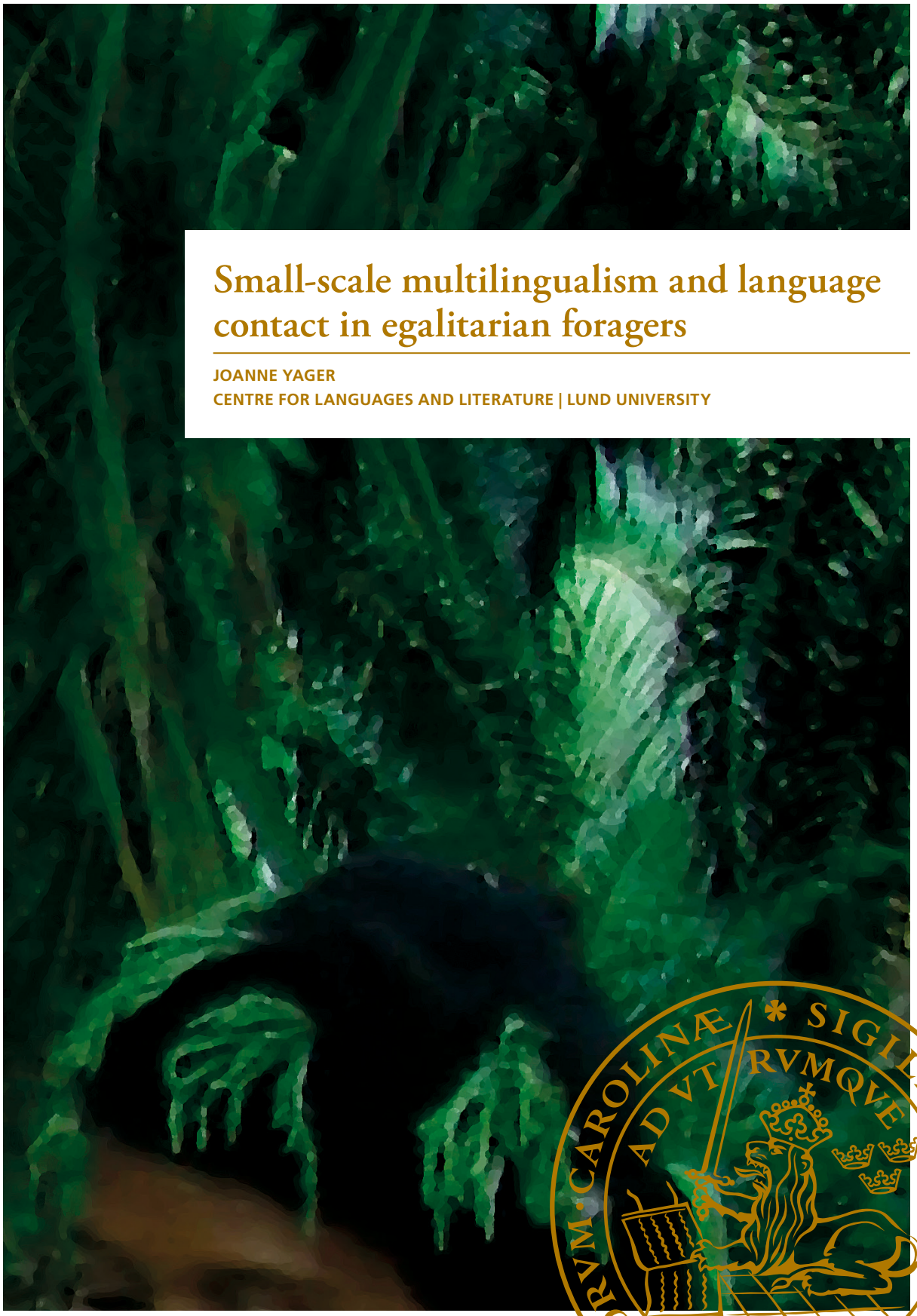
Small-scale multilingualism and language contact in egalitarian foragers

Yager, Joanne

2020

Document Version:
Publisher's PDF, also known as Version of record



Small-scale multilingualism and language contact in egalitarian foragers

JOANNE YAGER
CENTRE FOR LANGUAGES AND LITERATURE | LUND UNIVERSITY

Small-scale multilingualism and language contact in egalitarian foragers

Joanne Yager

LUND
UNIVERSITY

Cover photo by Joanne Yager

Faculties of Humanities and Theology
Centre for Languages and Literature

ISBN 978-91-88899-73-6 (print)
ISBN 978-91-88899-74-3 (digital)

Printed in Sweden by Media-Tryck, Lund University
Lund 2019

To all speakers of lesser-known languages

Table of Contents

Acknowledgements

This thesis grew out of a desire to conduct fieldwork in lesser-known linguistic contexts, and out of a curiosity about a mysterious language variety spoken in the village of Rual in northern Peninsular Malaysia. The thesis contained in these pages is the product of the unique opportunities I've been given over the past few years to follow these interests. There are a number of people without whom the completion of this thesis would not have been possible. First, I extend my most important and profound thanks to all the people at Rual who have helped me in countless ways both big and small, who spent hours patiently teaching me, answering my questions, and sharing stories, meals, knowledge and a great sense of humor. I would especially like to thank Naʔ Ya, Naʔ Bi, Taʔ Jlmɔl, Siti and Salleh for your help and for your good company. My time at Rual has taught me so much, far beyond what is written in these pages.

For the unique meeting of research interests that birthed this thesis project and for your expertise, inspiring depth of knowledge, and grand vision, I owe a thousand thanks to my supervisors Niclas Burenhult, Marianne Gullberg and Nicole Kruspe. Thank you for your careful reading of my writing, for helping me find the answers to the many questions that arose over the course of the project, and for your insightful feedback that has helped me to develop my ideas. I thank Niclas and Nicole for introducing me to the world of Aslian linguistics and helping me navigate fieldwork in a Malaysian context. I'm extremely lucky to have happened upon the rare chance to work so closely with two Aslianists with such a broad knowledge of Aslian and Austroasiatic linguistics, and to have had the chance to meet several other Aslianist and Austroasiaticist colleagues throughout the course of the project. I've had the good fortune of having my office in Villa Blix, thanks to my fellow blixlings for being such good company over the years. To the top-quality office mates I've been blessed with – Eline Visser, Sandra Cronhamn and Felix Ahlner, thanks for always making it fun to come to work. Thanks to my colleagues and fellow PhD students at the department and to my friends and family for listening to my ranting about various things related to the project. I look forward to introducing you to my post-PhD persona.

I'm extremely honored to have Pieter Muysken as faculty opponent for the defense of my thesis, and to have Brigitte Pakendorf, Michael Dunn and Jonas Granfeldt as my academic committee. Thanks to Henrik Gyllstad for acting as reserve, to Sven Strömqvist for acting as chair, and to Petra Bernardini for her suggestions, questions and comments during my mock defense. Thanks also to Asifa Majid, one of my methodological heroes, for help and inspiration at different points during the project. I've also benefited from discussions with other researchers working in small-scale multilingual contexts at the inaugural Typology of small-scale multilingualism conference in Lyon and the Uppladoc workshop on Language documentation in Uppsala.

Last but not least, I would like to thank the people and organizations that have made my fieldwork in Malaysia possible. Anthropologist Kamal Solhaimi Fadzil from the University of Malaya was endlessly generous and took me on a whirlwind tour of Rual and some other Aslian communities in northern Malaysia in June 2013. Anthropologist Alberto Gomes kindly met with me in Melbourne and put me in touch with Kamal. Zanisah Man has been my Kuala Lumpur family, I am immeasurably grateful to her for helping me with anything and everything, always offering me a homely place to stay in Kajang to regroup on my way to or from Rual, and showing me the best places to eat around KL. I also extend warm thanks to Prof. Dr. Kamila Ghazali from the University of Malaya for acting as my academic counterpart in Malaysia. I acknowledge the Economic Planning Unit and the Department of Orang Asli Development for granting permission to conduct research in Malaysia. Finally, several funding bodies have made my fieldwork possible. For a travel and equipment grant from the Lars Hierta Memorial Foundation, and travel grants from The Birgit Rausing Language Programme, Fil. Dr. Uno Otterstedt Foundation and The Knut and Alice Wallenberg Foundation, I am extremely grateful. Funding bodies who see the importance of primary data collection in lesser-known contexts are vital to the continuation of this work.

Abstract

Situations of multilingualism and language contact in which language varieties are small in scale and relatively equal in social status are important phenomena affecting processes of language change throughout human history. Despite this, our knowledge about the outcomes of multilingualism and language contact in this kind of setting remains limited. The current thesis provides insight into the linguistic consequences of interaction between closely-related, recently-described, small-scale language varieties in the community and in the minds of bilinguals, and works to overcome some of the methodological challenges associated with the study of language contact and multilingualism in this type of setting. The studies of the thesis investigate lexical and semantic outcomes of multilingualism and language contact in egalitarian foragers speaking the closely-related Northern Aslian (Austroasiatic) language varieties Jedek and Jahai in northern Peninsular Malaysia. Study I provides grammatical description of the newly discovered Northern Aslian language variety Jedek. Study II finds a high degree of lexical convergence in the language production of Jedek and Jahai speakers in contact, and presents a novel methodology for investigation of the linguistic consequences of language contact. Studies III and IV highlight the role of both social and linguistic factors in influencing bilingual outcomes, and provide evidence of symmetric (Study IV) and asymmetric (Study III) semantic interaction in two groups of Jedek-Jahai bilinguals in two different semantic domains. By combining perspectives from the fields of multilingualism, language contact and primary linguistic documentation and description, the thesis points to the potential of research in lesser-known linguistic settings to advance our theories of multilingualism, language contact and language change.

List of papers

I. *Jedek: A newly discovered Aslian variety of Malaysia*
Yager, J. & Burenhult, N. (2017). *Linguistic Typology* 21(3). 493–545.

II. *Outcomes of small-scale egalitarian contact in closely-related languages: Evidence from Southeast Asian foragers*
Yager, J. (submitted)

III. *Asymmetric semantic interaction in Jedek-Jahai bilinguals: Spatial language in a small-scale, non-standardized, egalitarian, long-term multilingual setting in Malaysia*
Yager, J. & Gullberg, M. (2019). *International Journal of Bilingualism*, doi:10.1177/1367006918814378.
(Published online by SAGE Journals, May 6, 2019)

IV. *Highly similar semantic systems become more similar in egalitarian bilingual foragers*
Yager, J. (submitted)

The contribution of the authors of the two co-authored papers was as follows.

Study I

Coauthor Yager conducted the fieldwork and data collection, conducted the analysis of the grammatical data, and wrote the sections of the paper in which Jedek phonology, morphology and syntax are described and in which sociolinguistic and multilingual aspects of Jedek speakers' situation are presented. Coauthor Burenhult wrote the majority of the introductory and concluding sections of the paper. Both authors were active in the editing and revision process.

Study III

Coauthor Yager conducted the fieldwork, data collection and analyses, and wrote the paper. Coauthor Gullberg contributed to the conceptualization of the study and both authors were active in the editing and revision process.

1 Introduction

Languages do not and have never existed in isolation. Over time, languages come into contact with one another and this contact leaves traces in the features and structures of languages. The nature of these traces and the dynamics governing the form they take are the central questions occupying the fields of language contact and multilingualism. These questions are of key importance for our understanding of fundamental issues in linguistics including the mechanisms behind language change, the nature of the human language processing system, and the histories and prehistories of languages and their speakers. Since the beginning of human history, situations in which speakers of small-scale language varieties come into contact with one another in egalitarian relations have been common the world over. Indeed, language contact of this kind is proposed to have played an integral role in human language evolution (Evans, 2018). Despite the ubiquity of small-scale, egalitarian multilingual contexts and their importance for our understanding of processes of language change, language processing and language evolution, our understanding of the dynamics of multilingualism and language contact in such settings is limited. How can studies in small-scale, egalitarian multilingual contexts inform our theories about multilingualism and language contact? And how can a combination of study of interaction between languages in the community (language contact) and in the minds of bilinguals (multilingualism) help us to better understand the outcomes of this interaction?

The current thesis explores these questions through investigation of multilingualism and language contact in egalitarian foragers speaking the closely-related Northern Aslian (Austroasiatic) language varieties Jedek and Jahai. The studies of the thesis seek to identify 1) the lexical outcomes of contact between Jedek and Jahai speakers in a multilingual speech community in northern Peninsular Malaysia, and 2) the semantic outcomes of interaction between Jedek and Jahai in the minds of Jedek-Jahai bilinguals. Importantly, the thesis also provides the first linguistic description of the newly discovered Aslian language variety Jedek. Methodological approaches from typology, psycholinguistics, sociolinguistics and historical linguistics are used to shed light on the linguistic consequences of language contact and multilingualism in this small-scale egalitarian setting. By combining perspectives from the fields of language contact, multilingualism and primary linguistic documentation and description, the thesis probes the potential of lesser-known contexts to advance our understanding of multilingual and language contact phenomena.

2 Background

Situations of contact between speakers of small-scale language varieties are thought to be the closest available analogues to the contact settings characterizing the human experience for the bulk of human history, and may provide an important window onto human language evolution (Evans, 2018). Yet this kind of contact setting is poorly represented in the literature on multilingualism and language contact. In recent years, scholars have begun calling for expansion of the scope of the empirical basis on which our theories about multilingualism and language contact are based through investigation of multilingual and contact phenomena in small-scale contexts (Gullberg, 2012; Lüpke, 2016; Vaughan & Singer, 2018). Developing an understanding of how multilingualism and language contact play out in this type of setting represents an important opportunity for advancing our knowledge of the full range of possible multilingual and contact scenarios, and for informing our theories about multilingualism, language contact and language change.

The current thesis is motivated by these research goals. The thesis investigates the outcomes of multilingualism and language contact in egalitarian foragers speaking the closely-related Northern Aslian language varieties Jedek and Jahai. The studies of the thesis investigate the linguistic consequences of interaction between languages in the community (language contact), and of interaction between languages in the minds of bilinguals (multilingualism), in the context of primary description and documentation of lesser-known language varieties. The thesis operates at the intersection of several fields, including the multilingualism, language contact, and language description and documentation fields. This section presents issues of relevance to the studies of the thesis from these three areas of research, as well as issues of relevance to the lesser-known type of contact setting in focus in the thesis (small-scale egalitarian multilingualism and language contact, contact and multilingualism in closely-related languages, and language contact and multilingualism in foragers). Finally, after discussion of the fields of research on semantic typology and bilingual semantic interaction (the focus of Studies III and IV of the thesis), this section presents the cultural and linguistic setting of the thesis and gives a brief overview of the studies of the thesis.

2.1 Multilingualism

It is often said that there are more multilingual individuals in the world than monolinguals, and thus that multilingualism represents the standard situation worldwide – this despite the tendency in the linguistics literature to treat monolinguals as the norm (Grosjean, 1989). Indeed, it has been suggested that humans have been predominantly multilingual since the beginning of the evolution of human language, and that this multilingualism may have played a key role in the process of language evolution (Evans, 2018). The field of multilingualism research is a broad field of study encompassing diverse lines of investigation including research on the mental lexicon, borrowing, code switching, effects on general cognitive abilities, education and literacy, and social aspects of multilingualism.

One of the central questions of the multilingualism field concerns the ways in which the languages of bilinguals interact with one another. A number of studies provide evidence that the languages of bilinguals interact at a range of levels of language (see e.g. Van Hell & Dijkstra, 2002; Hartsuiker, Pickering & Veltkamp, 2004; Bullock & Gerfen, 2004). This interaction means that bilinguals do not behave like "two monolinguals in one person" (Grosjean, 1989) – bilinguals tend to perform differently to monolinguals in a range of linguistic domains (see e.g. Bullock & Gerfen, 2004; Ameel, Malt, Storms & Van Assche, 2009). Studies involving second language learners find influence from the first language of learners in their developing second language, as well as influence from the second language on the first (see e.g. Cook, 2003; Brown & Gullberg, 2008, 2011). These kinds of effects in bilinguals and second language learners have been seen in both production and comprehension and are referred to in the literature using the terms *transfer*, *crosslinguistic influence*, *convergence* or *interaction* (see Jarvis & Pavlenko, 2008 for an overview). The effects of interaction between the languages of bilinguals can be bidirectional, with effects seen in both languages, or unidirectional, with effects seen in only one of the languages. The directionality of bilingual effects is generally conceived of in the multilingualism literature as resulting from the linguistic features of languages, or characteristics of bilingual speakers such as proficiency, language dominance and usage patterns.

The terms bilingual and multilingual are used interchangeably in the thesis to denote the use of two or more languages or language varieties. Where 'multilingual' is used in contrast to 'bilingual' it refers specifically to the use of more than two languages. The terms bilingual and multilingual refer in the thesis to *functional* bilingualism (Baetens Beardsmore, 1982) – that is, individuals who are able to function in two or more languages are considered bilingual. This definition of bilingualism follows e.g. Weinreich (1953), Appel and Muysken (1987) and Grosjean (1989).

2.2 Language contact

Language contact research is concerned with phenomena resulting from interaction between speakers of different languages, and typically focuses on contact-induced change in language systems. Understanding the consequences of language contact for the languages involved is crucial to our understanding of processes of language change and informs our understanding of the historical relationships between languages and peoples. The language contact field includes research on borrowing of words and structural features, convergence, pidgins and creoles, mixed languages, and language death (see e.g. Hickey, 2010 for an overview). Language contact may result in language shift, whereby a community of speakers adopt the use of another, usually more socially dominant, language and cease to use the language they used prior to the contact (see e.g. Appel & Muysken, 1987 for an overview). In some cases, language contact results in the emergence of a new language variety, as in the case of pidgins and creoles, language varieties that arise in situations of extensive contact between languages in the absence of bilingualism (see e.g. Holm, 2000). In other cases, mixed languages may emerge as a result of contact. Mixed languages are language varieties whose features combine components from two typologically distinct languages, and may emerge in situations of widespread bilingualism (see Matras & Bakker, 2003).

One of the issues that have occupied researchers in the language contact field is the relative importance of linguistic and social factors in influencing the outcomes of language contact. The effect of linguistic factors such as the structural similarity of languages in contact received much focus in early work (see e.g. Weinreich, 1953), with suggestions that greater structural similarity of languages leads to greater contact effects. Social aspects have received greater attention since the 1980s (Thomason and Kaufman, 1988; Thomason, 2001; Muysken, 2013), with factors such as the relative prestige levels of languages and numbers of speakers said to influence contact outcomes (see Muysken, 2013). Where languages differ in prestige levels, this generally results in asymmetric influence of the more dominant language on the less dominant language (Aikhenvald & Dixon, 2001). Some authors argue that social aspects play a much more important role in affecting the outcomes of contact (e.g., Thomason & Kaufman, 1988; Bowern, 2013), while others argue that contact outcomes are shaped by an interplay of linguistic and social factors (e.g., Sankoff, 2001; Muysken, 2013).

2.3 Primary documentation and description

A large number of the world's languages lack primary documentation and description, and thus our knowledge of "the full diversity of human linguistic potential" (Quinn, 2013: 3) is at present limited. Primary linguistic description and documentation of

lesser-known language varieties is essential to linguists whose work involves making general claims about the nature of human language. Since language provides an important window onto cognition, this kind of work is also of great value to researchers interested in a range of phenomena in human cognition. It is often said that around half of the world's languages are at risk of disappearing within the present century, and that the loss of a language means the loss of a unique worldview as encoded in the forms and functions of that language (Quinn, 2013). Recent decades have seen recognition of the language documentation and description endeavors as independent fields of research (Himmelmann, 1998), as well as increased funding opportunities for primary linguistic documentation and description as a result of a growing awareness of the value of lesser-known linguistic contexts for our linguistic theories.

Language documentation is concerned with the collection and preservation of linguistic data such that these data can be used by other researchers or by members of the communities who speak the languages (Himmelmann, 1998, 2012). Language description is the task of investigating and describing the linguistic structures of a language using "primary language data gathered through interaction with native-speaking consultants" (Chelliah & de Reuse, 2010: 7). Description of a language usually includes descriptions of the phonology, morphology, and syntax of a language. Ideally, a collection of texts and a dictionary will also be included (Evans & Dench, 2006: 10–16). Descriptive work is often carried out by researchers who also engage in documentary work.

2.4 The intersection of the fields

The current thesis combines perspectives from the multilingualism, language contact and documentation and description fields – elements of which are rarely brought to bear in the context of a single research project. While language contact and multilingualism research are both concerned with interaction between languages and the outcomes of this interaction, perspectives from the two fields are rarely accounted for in combination. Since the early days of research on multilingualism and language contact, it has been acknowledged that issues related to contact at the level of linguistic systems and issues related to contact between languages in the minds of bilinguals are fundamentally linked (see e.g. Haugen, 1950; Weinreich, 1953; Appel & Muysken, 1987). In practice, however, multilingualism and language contact phenomena are usually studied in relative isolation from one another, and are treated in separate academic journals and conferences (Muysken, 2013). Indeed, the fields are in turn often further split into work focusing on historical linguistic, sociolinguistic, and psycholinguistic aspects, and research on the psycholinguistic aspects of multilingualism is further divided into work focusing on bilinguals on the one hand, and second language learners on the other. The separate treatment of phenomena

resulting from interaction between languages means that opportunities are missed for gaining an understanding of the ways in which languages interact and the consequences of this interaction for the languages involved (Muysken, 2013).

Further, multilingualism research rarely includes insights from lesser-known linguistic settings, which are often rich in multilingual phenomena (Gullberg, 2012). Research in the language contact field has more commonly included findings from such settings, but most of this work focuses on the influence of larger-scale languages on small-scale languages rather than the consequences of contact between small-scale languages (Vaughan & Singer, 2018). This means that while language contact is ubiquitous in all parts of the world, and most people in the world speak more than one language, a large proportion of these speakers remain underrepresented in our theories about multilingualism and language contact (Gullberg, 2012; Lüpke, 2016). Conversely, the output of primary documentation and description work rarely considers multilingual phenomena, and tends to conceive of languages as if they were largely monolithic entities with clearly-defined boundaries. The current thesis argues that combination of methodologies and research questions from the above fields has the potential to greatly serve the goals of each of the fields. Just as linguistic description of languages often leads to the discovery of previously unknown linguistic phenomena (Evans & Dench, 2006: 4), the description of unexplored multilingual and language contact settings promises to uncover previously unrecognized multilingual and language contact phenomena.

2.5 Small-scale egalitarian multilingualism and language contact

There is a growing awareness of the need for research into the consequences of language contact and multilingualism in small-scale language communities in the absence of substantial prestige differences between languages. A number of terms have been used in characterizations of this kind of contact setting. The terms egalitarian bilingualism or multilingualism (François, 2012; Haudricourt, 1961; Vaughan, 2018), small-scale multilingualism (Lüpke, 2016), non-polyglossic multilingualism (Vaughan & Singer, 2018) and indigenous multilingualisms (Vaughan & Singer, 2018) have all been used to describe situations of this kind. Where the phrase small-scale multilingualism is used in the literature, the designation "small-scale" is intended to refer not only to the size of speech communities, but also to the social organization of language use: to situations in which "multilingual interaction is not governed by domain specialization and hierarchical relationships" (Lüpke, 2016: 35). The term egalitarian bi-/multilingualism is similarly used to refer to multilingual situations in which the social status of languages is relatively equal (see François, 2012 and Haudricourt, 1961). The terms small-scale multilingualism and small-scale language contact as used in the current thesis refer to

situations of interaction between languages whose speaker communities are small in scale and whose social status is relatively equal. The terms egalitarian multilingualism and egalitarian language contact are used to refer to interaction between languages in the absence of major differences in the social status of languages.

The impact of the small scale of language communities on the dynamics of multilingualism, language contact and language change is not yet well understood. One claim about small-scale language communities is that they may tend toward linguistic differentiation to a greater extent than communities speaking large-scale languages, due to a cultural bias toward the "constructive fostering of variegation" (Evans, 2010: 14). Some support for this idea is found in François' (2011) comparison of the Oceanic languages of the Torres-Banks linkage in Vanuatu, which have widely divergent lexica despite relatively recent descent from a common ancestor. Many small-scale languages have not been subject to the standardizing forces of media or political institutions, and many are undescribed or only recently described by linguists (Lüpke, 2016). This means that study of multilingual and language contact phenomena is often challenging in small-scale contexts, since it can be difficult to define the boundaries between ways of speaking where language varieties are non-standardized or where little information is available about language varieties. Difficulties in defining language boundaries can in turn lead to challenges in identifying features that are present in languages as a result of multilingualism or language contact.

Egalitarian prestige relations between languages are thought to result in very different outcomes compared to situations of uneven prestige. Egalitarian relations are predicted to result in bidirectional influences in multilingualism and language contact, with both (or all) languages undergoing change as a result (e.g. Dixon, 1997; Muysken, 2013). In contrast, prestige differences between languages tend to lead to asymmetric outcomes, with unidirectional influence from the language with higher prestige on the language with lower prestige (Aikhenvald & Dixon, 2001; Muysken, 2013). The outcomes of egalitarian multilingualism and language contact thus differ fundamentally from those of processes of language shift involving the spread of large-scale languages at the expense of smaller languages.

2.6 Contact and multilingualism in closely-related languages

Interaction between speakers of languages that have relatively recently diverged from a common ancestor is common in the world, since genealogically-proximal languages are also likely to be found in geographical proximity to one another (Epps, Huehnergard, & Pat-El, 2013). Related languages tend to be more similar structurally, since less time has passed since their split from a common ancestor. The extent to which linguistic

factors such as the typological and structural similarity of languages in contact play a role in influencing the outcomes of contact is a key question in language contact research. An early assumption was that transfer of linguistic features across language boundaries is only possible in the context of structural similarity (Weinreich, 1953; Moravcsik, 1978). This view is not widely held today, and there is much evidence from both language contact and multilingualism research of contact effects in the presence of large structural differences (see e.g. Thomason & Kaufman, 1988; Jarvis & Pavlenko, 2008). There are however suggestions in the literature that the degree of similarity of language varieties in contact may play a role in influencing contact outcomes. For example, it is sometimes suggested that lexical borrowing may be more frequent between typologically similar languages than between typologically very different languages (see e.g. Edwards, 2004: 18). In addition, similarity of languages may lead to an increase in the amount of contact between speakers, since this similarity may facilitate communication. This may in turn affect the outcomes of contact, since the degree of contact between speakers is said to be the best predictor for the extent of contact effects (Thomason & Kaufman, 1988).

Situations of small-scale multilingualism and language contact (see Section 2.5 above) often involve closely-related language varieties. In light of the suggested tendencies toward differentiation in small-scale language communities (Evans, 2010: 14; discussed in Section 2.5 above), combined with predictions of greater contact effects between typologically similar languages, such situations represent a meeting point between two seemingly opposing forces of language change. For this reason, situations of small-scale multilingualism and language contact between closely-related language varieties are potentially a valuable testing ground for our theories about the dynamics of language change. Currently, our understanding of the linguistic consequences of language contact and multilingualism in this kind of setting is not yet well developed (Epps et al., 2013). Relatively little is known about the outcomes of contact between speakers of closely-related languages, since research in the language contact field tends to focus on contact between unrelated or only distantly related languages (Epps et al., 2013). One prediction is that typological and/or lexical similarity of languages in contact will tend to lead to symmetric contact influences, in which both languages change as a consequence of the contact (Muysken, 2013). A factor often discussed in relation to situations of language contact in related language varieties is the methodological difficulty in this kind of contact situation in separating features that are shared as a result of contact and features that are shared due to common inheritance. Typological similarity of languages tends to represent less of a methodological challenge to research in the multilingualism field, where effects of the genetic proximity of languages and factors such as the cognate status of words have been seen (see e.g., Clyne, 2003; de Groot, 1993).

2.7 Language contact and multilingualism in foragers

The current thesis investigates the outcomes of small-scale multilingualism and language contact in egalitarian foragers. A number of claims have been made about the nature of the languages and contact patterns of forager groups. The forager category is generally defined as groups who gain their livelihood without the exertion of control over the reproductive cycle of their caloric resources (Güldemann, McConvell & Rhodes, in press). The forager subsistence mode is often associated with small group size, low population density, egalitarian social structure, exogamous marriage traditions, and flux in band composition. Many forager groups are also characterized by high levels of mobility or nomadism, although most fall on a cline of mobility, mixing sedentary and nomadic periods. The thesis primarily uses the term *forager* as this is emerging as the preferred term in the literature; this term includes not only hunting and gathering but also activities such as fishing that are of importance for many non-agricultural groups (Güldemann et al., in press). Where the term *hunter-gatherer* is used in the thesis, this is used with the same meaning as *forager*. While generalizations are often made about foragers as a category, the internal diversity within the forager category means that these generalizations are to be considered tendencies rather than universal characteristics (Kelly, 1995; Güldemann et al., in press; Epps, Bowern, Hansen, Hill & Zentz, 2012).

The relevance of the forager category for linguistic phenomena is debated by scholars. The complex of social traits often linked to forager groups has led to ideas that there may be systematic differences between languages spoken by foragers and languages spoken by food producers. One generalization is that the languages of foragers tend to have less complex numeral systems than those of food producers (see e.g. Dixon, 1980: 107–108 for such a characterization of Australian languages). Epps et al. (2012) however find no correlation between subsistence pattern and numeral complexity in a comparison of a large number of languages – any correlations were at a regional level only. Similarly, in a large-scale comparison of a number of structural features in languages spoken by foragers and food producers, Bickel and Nichols (in press) find no evidence of fundamental grammatical differences between the languages of forager and agricultural groups. There are suggestions in the literature that the vocabularies of languages spoken by foragers may differ systematically from those of food producers, due to social or cultural differences (see e.g. Bickel & Nichols, in press). There is evidence of culturally-determined semantic patterns in Aslian languages spoken by foragers and food producers in domains such as verbs of eating and drinking (Burenhult & Kruspe, 2016). Psycholinguistic differences between Aslian-speaking forager and food producer groups have also been found, for example in differential codability of domains such as color and olfaction (Majid & Kruspe, 2018).

Investigation of potential differences between foragers and food producers in the dynamics and outcomes of multilingualism and language contact may be key to understanding patterns of multilingualism, language contact and language change in prehistory. While contemporary forager groups should not be seen as direct equivalents to forager groups in human prehistory, they are the closest analogues available for study today. Thus if we were to find fundamental differences between patterns of language contact and multilingualism in present-day forager societies and food-producing societies, this could have profound implications for our understanding of prehistoric patterns of language change. In addition, since human language evolved in forager communities who were most likely exogamous and multilingual (Evans, 2018), an understanding of contact and multilingual phenomena in present-day exogamous, multilingual foragers has potential implications for our understanding of human language evolution. Claims have been made about patterns of language contact and change in foragers based on assumptions about the movement patterns of forager groups. It was long thought that foragers do not expand into new territories, and as a result, that diffusion rather than language splitting is the main source of language change in forager groups (Bowern et al., 2011). It was previously assumed that borrowing rates in the languages of foragers are higher than in languages spoken by food producers (e.g. Dixon, 1997). This assumption is shown by Bowern et al. (2011) to be inaccurate – borrowing rates were not found to differ systematically between languages spoken by foragers and languages spoken by food producers, and borrowing rates differed greatly between groups within the forager category. Some evidence of differences in the contact patterns of foragers and food producers has however been seen within certain regions. Research from the Aslian context reveals increased rates of lexical change in the languages spoken by Aslian forager groups as compared to Aslian food producing populations (Dunn, Kruspe & Burenhult, 2013). This difference was seen to coincide with subsistence mode divisions and cross-cut the genealogical subgrouping of the languages.

Multilingualism is widely reported among forager groups, and is argued to result from patterns of mobility and exogamy (out-marriage). Exogamous partnerships between individuals speaking different languages mean that exogamous speech communities are often composed of individuals with different language backgrounds, and that children grow up with several languages in their environment. Mobility may also contribute to multilingualism, since mobile forager groups may travel through different language territories. Multilingualism is thought to be a socially valuable resource in forager communities in that it allows access to resources in a wider area (Güldemann et al., in press). Despite the widespread multilingualism reported for forager groups, there is little representation of foragers in multilingualism and language contact research. This is likely in part due to the fact that foragers represent a small proportion of the contemporary human population, and in part due to the fact that the languages of foragers are often poorly documented (Epps et al., 2012). The patterns and outcomes

of multilingualism and language contact likely differ among forager groups (see e.g. the variation in borrowing rates in languages spoken by foragers in Bowern et al., 2011), and this may depend on a number of factors. For example, postmarital residence patterns are likely to affect language change processes in exogamous communities, since they influence language use and language contact patterns.

2.8 Semantic typology and bilingual semantic interaction

Studies in the field of semantic typology reveal that there is a great deal of diversity in how semantic space is divided up in the lexicon of different languages. Evidence comes from a number of semantic domains including odor, taste, color, kinship, objects, spatial relations, and actions (see Malt & Majid, 2013 for an overview). Semantic typological research has developed into a fruitful area of research over the past few decades and scholars in the field have developed a number of tools for comparison of semantic categorization in different semantic domains. The current thesis makes use of some of these tools in investigation of bilingual semantic interaction in the domains of topological relations and placement events. Semantic typological research into the expression of topological relations within the spatial domain began in the 1990s (Bowerman & Pederson, 1992; Bowerman, 1996; see also Levinson, 2003; Levinson & Meira, 2003; Levinson & Wilkins, 2006). Later, the extension of semantic typological research to events saw investigation of a number of event domains, including events of cutting and breaking (Majid, Bowerman, van Staden & Boster, 2007), reciprocal events (Majid, Evans, Gaby & Levinson, 2011; Evans, Gaby, Levinson & Majid, 2011) and caused motion events (Kopecka & Narasimhan, 2012). These studies find a large amount of variation in how events and spatial relations in these domains are categorized in different languages.

Typically, multilingualism and language contact research focuses on morphosyntactic or lexical aspects of language, and semantic aspects are rarely taken into account. In recent years, researchers in the multilingualism field have begun to investigate how bilinguals deal with differences in semantic categorization in the languages that they speak. This program of research provides a valuable window onto the interaction of languages in the minds of bilinguals. Research on bilingual semantic interaction reveals influences across the semantic systems of languages in bilinguals and second language learners (see e.g. Gathercole & Moawad, 2010; Alferink & Gullberg, 2014; Malt & Lebkuecher, 2017). This influence is typically seen in increased similarity in the form-meaning mappings of the languages of bilinguals and differences between bilinguals and monolinguals. While semantic typological work has maintained a strong focus on diversity and includes a great deal of evidence from lesser-known linguistic contexts, research on bilingual semantic interaction has maintained a Western focus. Two of the studies of the thesis focus on semantic aspects of the interaction between languages in

the minds of bilinguals in a non-Western, untutored setting. The thesis thus represents a broadening of the empirical base of our knowledge about bilingual semantic interaction. Studies III and IV focus on semantic interaction in bilinguals of the Northern Aslian language varieties Jedek and Jahai. Semantic typological work on the Aslian languages has found a tendency to encode detailed meanings in monolexemic verbs in a range of semantic domains, including caused motion verbs, verbs of eating and drinking, perception verbs and motion verbs (see e.g. Burenhult, 2012; Kruspe, Burenhult & Wnuk, 2014: 466–467; Burenhult & Kruspe, 2016; Wnuk, 2016; Burenhult & Purves, in press). Thus the Aslian setting in focus in the thesis provides a valuable context for study of bilingual semantic interaction in the context of rich verb semantics.

2.9 The cultural and linguistic setting of the thesis

The studies of the current thesis speak to the issues introduced in the above sections through investigation of the outcomes of language contact and multilingualism in closely-related language varieties in groups of foragers in northern Peninsular Malaysia known ethnographically as the Semang. The Semang are traditionally nomadic rainforest foragers inhabiting inland areas of the Malay Peninsula between central Pahang in the south and the southern regions of Thailand in the north. The Semang are generally associated with the "negrito" phenotype[1], and are characterized by a "highly egalitarian ethos, linked with a high degree of personal autonomy-cum-communality" (Benjamin, 2011: 170). Semang social organization is characterized as non-competitive egalitarian, emphasizing the right of individuals to equality of wealth, power and prestige (Endicott & Endicott, 2008; see Woodburn, 1982 for a general discussion of egalitarian societies).

Benjamin (1985) argues that the Semang must maintain social relations with a broad network of others, since they are dependent on adaptation to changing conditions and fluctuating availability of resources. Semang social patterns are characterized by movement and flux, with "constant dispersal in time, place, and consociation" (Benjamin, 1985: 228). The membership of Semang bands (small groups of around 15–50 individuals generally related by kinship) is impermanent and changes throughout time (Gomes, 2007). The tradition of band exogamy (out-marriage) associated with Semang groups means that most Semang bands include speakers of several language varieties (Benjamin, 1985). It is suggested that Semang populations have long been connected by "a continuous mesh of communication with each other extending from Isthmian Thailand right down to central Pahang" resulting from

[1] This contested term is used in the literature to refer to Southeast Asian populations of short stature, dark skin and frizzy hair (Benjamin, 2013)

"small-group nomadism coupled with wide-ranging intermarriage" (Benjamin, 1985: 234).

The Semang speak varieties of Aslian, the branch of Austroasiatic languages spoken in the Malay Peninsula. Aslian languages are divided into the Northern, Central and Southern subbranches, which roughly correspond to the ethnographic subgroupings Semang, Senoi, and Aboriginal Malay (Benjamin, 1985). The Semang bands that form the focus of the current thesis speak the Northern Aslian language varieties Jahai and Jedek. Dunn, Burenhult, Kruspe, Tufvesson and Becker (2011) propose a first-level split within the Northern Aslian language varieties between a group of varieties they term Maniq/Menraq-Batek, and Ceq Wong (the only Northern Aslian language variety not spoken by Semang groups), and a secondary subgrouping within Maniq/Menraq-Batek into the Maniq (Kensiw, Kintaq and Ten'en) and Menraq-Batek (Batek, Jahai, Menriq) varieties.

The current thesis focuses on language contact and multilingualism in the Jedek- and Jahai-speaking Semang bands inhabiting the Sungai Rual resettlement area, located on the Rual river approximately 10 km southwest of the town of Jeli in northwest Kelantan, in northern Peninsular Malaysia. With government-sponsored resettlement in the 1970s, six bands of Semang foragers who were living along the middle reaches of the Pergau watershed at the time settled in the resettlement area (Gomes, 2007). Four of these bands identify with the ethnographic labels Menriq or Batek, and two identify with the label Jahai. Together, these bands established three hamlets within the resettlement area – Rual Tengah, Kalok and Manok[2]. Three of the 'Menriq/Batek' bands settled at Rual Tengah, one 'Menriq/Batek' and one Jahai band settled at Kalok, and the remaining Jahai band settled at Manok. In the 1980s, an additional Jahai band joined the Kalok hamlet from the neighboring state of Perak (Gomes, 2007). Anthropologist Alberto Gomes' (2007) social demographic study based on fieldwork conducted between 1975 and 2006 outlines some of the demographic implications of resettlement for the bands living in the resettlement area.

Little was known about the language varieties of the bands inhabiting the Sungai Rual resettlement area until 2006 when a linguistic survey of Semang groups including collection of Swadesh-based lists of basic vocabulary was carried out by Niclas Burenhult. Phylogenetic analysis of the lexical survey data (Dunn et al., 2011) suggested that the 'Menriq/Batek' and 'Jahai' bands of Sungai Rual speak distinct language varieties, both placed within the Menraq-Batek subbranch of Northern Aslian (see the phylogenetic network in Figure 1 reprinted with permission from Dunn et al., 2011). While one of the wordlists collected at Sungai Rual appeared to represent a variant of Jahai, the other appeared to represent a previously unrecognized Northern

[2] The hamlet names are Malay renderings of the indigenous names of local watersheds, *Rwɨl*, *Kalɔʔ* and *Manɔk*.

Aslian language variety. This is the variety that is referred to in the current thesis using the label Jedek. The current thesis represents the first investigation into the ways of speaking of the Semang bands of the Sungai Rual resettlement area.

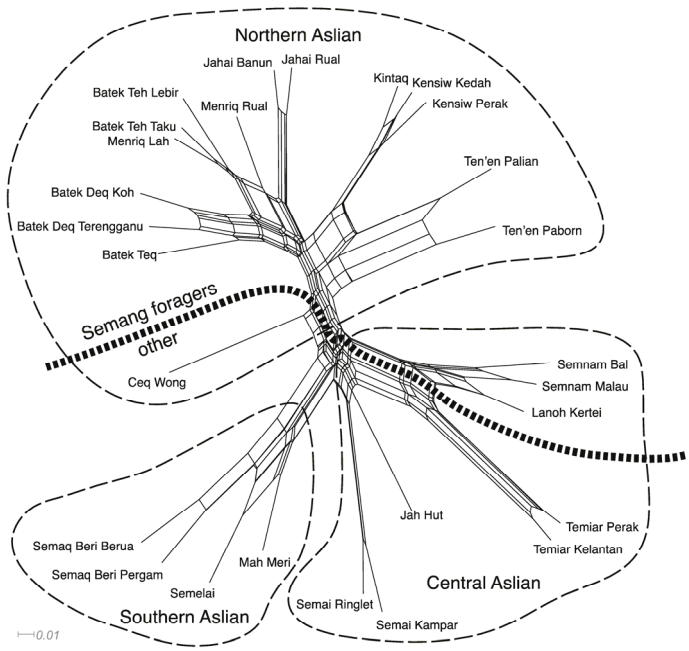


Figure 1. NeighborNet of the distances between Aslian languages from Dunn et al., 2011

2.10 Overview of the thesis

In four studies, the current thesis provides the first description of the newly discovered Northern Aslian language variety Jedek and explores the outcomes of language contact and multilingualism in the Jedek- and Jahai-speaking Semang foragers who reside at Rual in northern Peninsular Malaysia. The studies probe the lexical and semantic consequences of interaction between Jedek and Jahai in the community and in the minds of bilingual speakers, and explore the impact of social and linguistic factors on the outcomes of this interaction. The studies of the thesis explore the potential of underdescribed linguistic settings to inform our theories of multilingualism and language contact, and propose novel methodological approaches in response to some of the challenges inherent in investigating language contact and multilingual phenomena in populations speaking non-standardized, closely-related language varieties, in the context of primary linguistic documentation and description.

3 The setting

The current thesis deals with multilingualism and language contact in the Jedek- and Jahai-speaking bands of the Rual Tengah and Kalok hamlets in the Sungai Rual resettlement area. The two hamlets are referred to collectively in the thesis as 'Rual' (see the maps in Figures 2 and 3 below; see also Map 1 of Study I and Figure 1 of Study II). A survey of the language identities and origins of the approximately 170 current adult residents of Rual, conducted as part of the fieldwork for the current thesis (see Appendix B for the survey data) suggests that just over 60% of Rual residents identify as Jedek speakers, around 30% identify as Jahai speakers, and 6% identify as speakers of other Aslian languages. Rual Tengah is today inhabited by just under 70 Jedek-identifying residents and around 25 Jahai-identifying residents, while Kalok is inhabited by roughly 15 Jedek-identifying residents and 25 Jahai-identifying residents.

Figure 2. The Sungai Rual resettlement area with its three hamlets: Manok, Rual Tengah and Kalok.

Figure 3. The part of the resettlement area in focus in the thesis, 'Rual', encompassing the hamlets Rual Tengah and Kalok.

The built environment of Rual primarily consists of concrete block and wooden houses, along with some thatched bamboo houses. The concrete and wooden houses, along with a Malay-built and -run preschool, primary school and mosque were built as part of a government-sponsored development and assimilation program. During the period of the fieldwork, the livelihood of most Rual residents was based around subsistence on foraged and cultivated products such as vegetables and fish, and the sale of products such as foraged medicinal plants and cultivated rubber or bananas. Very few Rual residents were engaged in paid employment at the time of the fieldwork. The majority of the children of Rual attend the Malay-medium primary school which services the children of the resettlement area, including the children of Manok who travel the short distance from Manok in a school minibus. Many of the children aged between 13 and 17 years attend the Malay-majority high school in Jeli, living at a boarding house nearby the school and returning home to Rual during school holidays.

3.1 Exogamy, movement, multilingualism

Although nowadays settled, the Rual bands continue to maintain the majority of the cultural practices associated with the Semang pattern. While traditional nomadism (which involved moving every few days, weeks or months) is no longer practiced by the bands since resettlement, life at Rual continues to be characterized by a high degree of movement, flux and multilingualism. The following sections outline these patterns.

3.1.1 Exogamy and movement

The Rual Semang practice band exogamy, according to which individuals must marry outside their own band. This often results in marriages between speakers of different Aslian language varieties, indeed over half of all current partnerships at Rual (ca. 40 couples or 60%) are between individuals with different language identities. Around two-thirds of these mixed marriages are between a Jedek and a Jahai speaker (28 couples), and over two-thirds of these mixed Jedek-Jahai couples consist of two individuals with origins at Rual (20 couples). The remainder of the Jedek-Jahai couples at Rual consist of a Jedek speaker from Rual and a Jahai speaker with origins in another geographical area (8 couples). While Jedek is only spoken at Rual, Jahai has approximately 1000 speakers, most of whom live in the neighboring state of Perak. Jahai speakers who have married into Rual from other areas most commonly have origins in Perak (5 individuals), or in Manok, the Jahai-majority hamlet located two kilometers upstream of Rual (6 individuals).

Marriages between Rual residents and speakers of Aslian languages other than Jedek and Jahai are also relatively common, with 17% of current partnerships at Rual

consisting of a Jedek or Jahai speaker and a speaker of another Aslian language. At the time of fieldwork, in-married Rual residents from other Aslian groups consisted of three Menriq (Northern Aslian) speakers, five Temiar (Central Aslian) speakers, one Lanoh (Central Aslian) speaker, and two Semaq Beri (Southern Aslian) speakers. Although there do not appear to be any strict prescriptions concerning the postmarital residence of couples, the bias at Rual is towards matrilocality. This was reported by speakers and is reflected in the numbers of the language identities survey (see Appendix B) – at the time of fieldwork there were 18 in-married men and six in-married women living at Rual. Interestingly, this deviates from the "patri-bias" and "relative patrilocality of domestic groupings" reported by Geoffrey Benjamin on the basis of enquiry among other Semang groups in the late 1960s and early 1970s (Benjamin, 2011: 178; although see Endicott, 1974: 219, who reports that there is no postmarital residence rule among the Batek).

Band exogamy does not always equate to linguistic exogamy, since each language variety is spoken by several bands. Just under one-third of the couples at Rual consist of two Jedek speakers with origins in the different Jedek-speaking bands at Rual (20 couples or 32%). A smaller number of Rual marriages consist of two Jahai speakers (3 couples or 5%). Since resettlement, exogamy does not always result in relocation, since Rual is composed of several bands. In total, over half of the couples currently living at Rual (40 couples or 63%) consist of two individuals with origins at Rual, of which half (20 couples) are Jedek-Jedek marriages, and half (20 couples) are Jedek-Jahai marriages. Where exogamy does involve relocation, this relocation is not necessarily permanent. It is not uncommon for individuals or families to reside for periods in different areas, moving between the places of origin of partners, or to other Aslian settlements. Individuals or small groups may spend periods of time in different locations as seasonal foraging or employment opportunities arise, returning to more permanent settlements in between. It is also not uncommon for an individual to marry several times in their lifetime, often resulting in relocation of individuals several times throughout the lifespan. Thus, movement and flux remain strong themes in the way of life of Rual residents. This movement tends to remain primarily within an Aslian sphere – very rarely do Rual residents choose to relocate to locations outside of Aslian contexts. Rual residents are part of a larger network of Aslian communities, with sometimes wide-reaching networks of contacts and relatives in other communities, and the practice of band exogamy helps to maintain this pattern.

3.1.2 Multilingualism

A high degree of multilingualism characterizes the Rual setting as a result of these patterns of exogamy, movement and flux, and as a result of the cohabitation of Jedek- and Jahai-speaking bands at Rual. Multilingualism in individuals is the result of acquisition in childhood and in adulthood; individuals commonly report having

knowledge of a language because this was the language of one of their parents or of their spouse. It is however not the case that all individuals speak the language of their spouse or of both their parents – not uncommonly, individuals report that they do not speak or understand the identity variety of their spouse or one of their parents. For some individuals, acquisition of a language variety takes place in both childhood and adulthood, for example where an individual is exposed to a language variety in childhood and then comes into renewed contact with the variety in adulthood upon marriage with a speaker of that variety. Individuals may also have knowledge of a language due to periods of residence in other Aslian communities. All Rual residents have active knowledge of at least two language varieties (that is, at least one Aslian variety and Malay, the majority language of the areas surrounding Rual). Thus, multilingualism is the norm at Rual. Every individual at Rual is in some way affected by multilingualism; each household contains people with different language identities and/or who have several languages in their repertoires.

Knowledge of Malay, the unrelated Austronesian majority language, is necessary for communication with outsiders and is the medium of schooling and broadcast entertainment (television and radio). All communication with outsiders (that is, outside the Aslian sphere) is conducted in Malay, and thus knowledge of Malay is essential for economic activities such as trading foraged products and purchasing food. The Malay population of the areas surrounding Rual has grown rapidly in recent decades, with the construction of roads and Malay settlements. However, contact between the Semang bands of Rual and Malays is not a new phenomenon; the Semang contact pattern has long been characterized by opportunistic trading with outsiders (Benjamin, 1985).

3.2 Language use in the community

In-depth study of language use patterns in the Rual community is beyond the scope of the current project; however some observations can be made. Communication between Jedek and Jahai speakers at Rual appears to primarily take place through receptive bilingualism (see e.g. ten Thije & Zeevaert, 2007; see also Lüpke, 2016). That is, each individual speaks their own language variety while understanding the language variety used by their interlocutor. In multilingual households, Rual residents report that each adult speaks their own language to their spouse and to their children. However, there is some evidence that actual language practices may not follow this reported pattern. For example, some individuals report that one of their parents spoke a particular language but that they themselves do not understand or speak this language.

While Jedek and Jahai speakers at Rual tend to communicate through receptive bilingualism, speakers of other Aslian languages must learn Jedek or Jahai in order to participate in daily life in the Rual community. While most Jahai speakers who have

married into Rual from other areas claim not to speak Jedek, some have acquired Jedek and report primarily using Jedek on a daily basis. One Jahai speaker who married into Rual from Manok reported having forgotten Jahai to some extent, as she has acquired and primarily uses Jedek since moving to Rual. This pattern does not however apply to all Jahai speakers who relocate to Rual for marriage – indeed the brother of this speaker, who relocated to Rual *before* his sister, claims not to speak Jedek at all. Such differences may in part reflect potential differences in the socialization of in-marrying individuals at Rual. For example, the immediate social group of this pair of siblings is primarily composed of Jedek-speaking women and Jahai-speaking men. Social factors of this kind are likely to play a role in influencing the language use of in-marrying Jahai speakers.

3.3 Language ideology and identity

3.3.1 Language names and ideologies

To outsiders, Rual residents tend to characterize Rual as a mix of speakers of Jahai and 'Menriq' or 'Batek' (used interchangeably or expressed as a mix, 'Menriq/Batek mixed'). While Jahai speakers at Rual consistently refer to their way of speaking using the label Jahai, the use of language names to refer to the variety labelled Jedek in the current thesis is not systematic. During the fieldwork, the names Menriq, Batek and Jedek were all used by speakers to refer to the variety. The label Jedek was chosen for use in the current thesis, as it is used by many Rual residents, and is used by Aslian groups in other areas to refer to the Rual variety. It is however important to note that the label is not used systematically by speakers, and is not used by all speakers. Jedek is proposed as the scientific label for the variety; the label refers to a linguistic entity and is not intended as an ethnographic designation. The kind of ambiguity in the use of language labels encountered at Rual is not unheard of among other groups speaking Northern Aslian language varieties. For example, the label Menriq is the name used by government bodies for the language of the Semang bands today residing in the resettlement village of Kuala Lah some 75 km south of Rual. While these groups accept Menriq as the official label for their language, their own label (also used by Semang in other areas) is *L?pa?*. Another example is the language variety previously referred to in the literature using the label Mintil (see e.g. Benjamin, 1976). While this label is used for the variety by the Semaq Beri (Nicole Kruspe, p.c.) and by some Batek groups in Pahang (Rudge, 2017), it is not used by the speakers themselves, who consider it derogatory (Lye, 2013). The variety is referred to using different names in the literature, e.g., Lye Tuck-Po uses the label Batek Tanum (see e.g. Lye, 2013), and Ivan Tacey uses the label Manya' (Tacey, 2018).

Rual residents often expressed confusion around what the labels Menriq, Batek, Jahai and Jedek refer to. Some simply said that they did not know what their language is called. During the course of the fieldwork at Rual it became increasingly clear that the official language labels were not generally used by speakers in communication amongst themselves. Jedek speakers tend to refer to their way of speaking as *basaʔ hiʔ*, 'our language', or may refer more generally to the ways of speaking at Rual as *basaʔ mnraʔ* or *kliŋ mnraʔ*, 'Aslian language', labels that can be used to refer to Jedek or Jahai, to both varieties collectively, or to refer to all Aslian languages. When making a distinction between Jedek and Jahai, speakers tend to refer to Jedek as *basaʔ ʔiŋ* 'the *ʔiŋ* language' and to Jahai as *basaʔ yɛʔ* 'the *yɛʔ* language' (or simply *ʔiŋ* and *yɛʔ*, the first-person singular pronoun of Jedek and Jahai respectively). As an illustration of the ways in which Rual residents relate to language labels, consider the following conversation between two Rual residents, one of whom (Speaker A) identifies as a Jedek speaker and the other (Speaker B) as a Jahai speaker. The exchange took place in the context of an interview conducted as part of the language identities survey (see Appendix B), and followed discussion of the language affiliation of a Rual resident who the interviewees characterized as a speaker of Batek (the label used by these speakers to refer to Jedek at this time).

A:	*Batɛk*	*way*	*ʔayaŋ*	*ma=Batɛk*	*d=ʔũh.*	*Mɛnriʔ*
	Batek	EMPH	NEG	IRR=Batek	CONTR=here	Menriq
	'Wait, *Batek*, we're not Batek here.					

	ʔiŋ	*rasaʔ*	*d=hiʔ*	*ʔũh,*	*manton*	*lɛh.*
	1SG	to.feel	CONTR=1PL.INCL	here	like.that	EMPH
	Menriq I think is us here, yes.					

	Batɛk	*gin*	*sah*	*Lbi.*
	Batek	3PL	group	Lebir
	The Batek are those people from Lebir.'			

B:	*Gin*	*kdih*	*Jdɛk*	*Jdɛk*	*ton*	*teʔ ʔoʔ teʔ?*
	3PL	to.say	Jedek	Jedek	that	which
	'Those Jedek, which ones are they?'					

A:	*ʔeh*	*haʔiʔ,*	*Jdɛk*	*ʔacoh*	*lɛh.*	*Hiʔ*	*ʔũh*	*bhaʔ*
	EXCLAM	yes	Jedek	don't.know	EMPH	1PL.INCL	here	DUB
	'Oh yes, Jedek – I don't know. Is that us here, Jedek?'							

d=Jdɛk.
CONTR=Jedek

B: *ʔũh ʔũh d=Jdɛk, mey. Yɛʔ kjeŋ ʔnsɔʔ*
here here CONTR=Jedek other 1SG to.hear to.die
'The ones that say *ʔũh* aren't Jedek. I heard the late

Yaʔkdɨh wɛy, YaʔcanLew
grandmother to.say past grandmother SOURCE=Kuala Lah
Grandma what was her name, the late Grandma … from Kuala Lah,

wɛy, YaʔCãr wɛy ton –
past grandmother Cār past that
the late Grandma Cār – '

A: *Batɛk lɛh dah d=gin sah kdɨh wɛy*
Batek EMPH EMPH CONTR=3PL group to.say past
'The Batek are those people from Aring in the past.'

ʔariŋ wɛy.
Aring past

B: *Hã lɛh Batɛk lɛh gin sah ʔariŋ wɛy. ʔacoh*
yes EMPH Batek EMPH 3PL group Aring past don't.know
'Yes, the Batek are those people from Aring in the past. I don't know

lɛh Jdɛk.
EMPH Jedek
about the Jedek.'

A: *Jdɛk lɛh d=hiʔ ʔũh.*
Jedek EMPH CONTR=1PL.INCL here
'Jedek, that's us here.'

B: *Bhaʔ* *lɛh* *d=ʔoʔ* *ʔũh* *ʔũh* *ton* *lɛh* *Jdɛk*
DUB EMPH CONTR=3SG here here that EMPH Jedek
'Ah yes, those ones who say *ʔũh* are Jedek, huh?'

k=bət.
SUB=whatever

A: *Jdɛk* *btul.* *Hɛ̃n* *Batɛk* *gin* *sah* *ʔariŋ.*
Jedek to.be.true yes Batek 3PL group Aring
'Jedek is right. Yes, the Batek are those people from Aring.'

B: *Haʔiʔ* *Batɛk* *lɛh* *d=gin* *sah* *ʔariŋ* *wɛy,* *ʔoʔ*
Yes Batek EMPH CONTR=3PL group Aring past 3SG
'Yes, the Batek are those people from Aring in the past,

d=nɛŋ *nɛŋ*[3].
CONTR=*nɛŋ* *nɛŋ*
those ones who say *nɛŋ nɛŋ*.'

Jedek and Jahai share many lexical and structural features due to their genetic proximity (see Section 2.9 above), and likely in part also due to a long history of contact (see Benjamin, 1985 for claims about contact patterns between speakers of Northern Aslian language varieties). To what extent are Jedek and Jahai seen as distinct language varieties by their speakers? Jedek and Jahai are not considered mutually intelligible by their speakers, and Rual residents frequently give examples of lexical items and features that differ between the varieties. At times, however, speakers refer to Jedek and Jahai as 'the same, mixed'. Thus, there appears to be some ambiguity in the ways in which the separation of the varieties is conceived of by speakers. Ideologies about Jahai appear to be particularly complex at Rual, since ways of speaking Jahai appear to differ to some extent in different geographical locations (see the results of Study II for evidence of this). Some individuals report that Jahai speakers at Rual mix Jahai and Jedek in their way of speaking. There is some evidence of this in the data; impressionistically, and as seen in the results of Study II, the language production of individuals with origins at Rual who identify as Jahai speakers does not appear to differ greatly from that of Jedek speakers. For example, Jahai speakers who grew up at Rual tend to use Jedek variants of Jedek-Jahai cognate sets differentiated by regular sound changes (such as the Jedek

[3] *Nɛŋ* is the negative marker used by Batek Deq speakers in Aring. This lexeme is often given as an example of the speech of this group by Rual residents.

form *buŋaʔ* 'flower', rather than *buŋɛʔ*, the form used by Jahai speakers outside Rual). In addition, it is not uncommon for Jahai speakers with origins at Rual to claim that they do not understand the speech of Jahai speakers from other areas. Several times during the fieldwork, Jedek or Jahai speakers from Rual referred to Jahai speakers in other areas as 'real Jahai'. Meanwhile, Jahai speakers at Rual systematically use certain lexemes that set their speech apart from that of Jedek speakers. These forms include the first-person singular pronoun *yɛʔ* (cf. Jedek 1SG *ʔiɲ*), the second-person singular pronoun *mɔh* (cf. Jedek 2SG *bɛʔ*), and the goal-marking preposition *ba=* (cf. Jedek goal-marking *da=*). Thus while the speech of Jedek and Jahai speakers at Rual appears to share many features, certain linguistic markers differentiate speakers of the two varieties.

3.3.2 Language identities

While some individuals refer to themselves as 'mixed Jahai/Jedek' and may label their way of speaking as Jedek or Jahai, most individuals identify as either Jedek or Jahai speakers. The pattern that emerges from the language identities survey data (see Appendix B) suggests that individuals tend to identify as Jedek speakers if one or both of their parents identify as Jedek-speaking. Where an individual has grown up at Rual but neither of their parents identify as Jedek speakers, they tend to identify as Jahai-speaking if at least one of their parents identifies as a Jahai speaker. Unlike in some other small-scale multilingual contexts, in which strong preferences are found for identification with a language of male ancestors and avoidance of the use of the mother's language (Vaughan & Singer, 2018), at Rual, language identity and use do not appear to be determined by the gender of the parent that speaks a language. There do not appear to be any systematic differences in language attitudes toward Jedek and Jahai at Rual.

A methodological decision has been made in the current thesis to use speakers' own language identity in defining which individuals should be considered speakers of a variety. Thus, where 'Jedek speakers' and 'Jahai speakers' are referred to in the thesis, this is based on the language identity of individuals themselves. Similarly, rather than using terms such as 'native language' or 'first' and 'second' language, the term *identity variety* is used in the thesis to refer to the language variety with which speakers identify. Where the label 'Jedek' is used in the thesis, this refers to the linguistic features used in the speech of individuals who identify as speakers of the variety referred to in the current thesis as Jedek. 'Jahai' refers to the features of the speech of individuals who identify as speakers of Jahai. In contexts where it is necessary to differentiate the ways of speaking of individuals from different Jahai-speaking communities, location names are used, as in 'Banun Jahai', 'Manok Jahai' and 'Rual Jahai'.

In adopting this approach to the linguistic affiliation of participants and to the definition of language varieties, the thesis does not impose categories on speakers to

which they do not themselves adhere, and does not attempt to use outside measures to define the linguistic affiliation of individuals. Indeed, it is unclear what an appropriate measure for this purpose might look like in the Rual context, where the language varieties in focus are non-standardized and not subject to prescriptive norms, and where individual variation is reportedly "as much idiolectal as dialectal" (Benjamin, 2009: 20, in characterization of Semang patterns of linguistic variation). In such a context, to whose language production should a standardized measure of language affiliation conform? While self-identification is potentially influenced by biases of different kinds, a number of studies find that speakers' self-reported knowledge of a language tends to match their abilities as measured with a range of tests (see e.g. Marian, Blumenfeld & Kaushanskaya, 2007; Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012). For these reasons self-identification was thought to be appropriate for the purposes of the studies of the thesis.

4 Methods

4.1 Fieldwork and data collection

The materials that form the empirical basis of the thesis were collected during four field trips of two to six weeks' duration between 2013 and 2017. A total of three and a half months was spent in the field. After an initial daytrip to Rual in June 2013 facilitated and accompanied by anthropologist Kamal Solhaimi Fadzil, all fieldwork was carried out unaccompanied. The bulk of the field data were collected within the village limits of Rual (in both Rual Tengah and Kalok), but the corpus also includes a small number of recordings made at Manok and in the rainforest areas surrounding Rual. Communication was initially in Malay but transitioned to Jedek within a matter of weeks as the Jedek skills of the researcher surpassed her Malay skills. Data collection consisted of video and audio recordings made with the prior consent of those recorded, as well as lexical, grammatical and sociolinguistic notes made outside of recording sessions.

The corpus of materials collected during fieldwork includes video recordings of natural speech including traditional stories told by a number of Jedek and Jahai speakers, video and audio recordings of elicited speech from a large number of Jedek and Jahai speakers, recorded interviews surveying kinship, language backgrounds and language identities of Rual residents, and recorded grammatical discussions between the researcher and one or two Jedek grammatical consultants. The corpus also includes a Jedek lexicon containing lexical items collected during the fieldwork. The different types of data were collected in parallel. The recordings of elicited speech were collected using several stimulus kits produced by the Max Planck Institute of Psycholinguistics as part of the Language & Cognition Field Manuals (http://fieldmanuals.mpi.nl/), including the Topological Relations Picture Series (Bowerman & Pederson, 1992), the PUT project task (Bowerman, Gullberg, Majid & Narasimhan, 2004), the Reciprocal Constructions and Situation Type task (Evans, Levinson, Enfield, Gaby & Majid, 2004), the Man and Tree Space Games (Levinson et al., 1992), and the Cut and Break Clips (Bohnemeyer, Bowerman & Brown, 2001). Elicited basic vocabulary data were collected using an adapted version of the Swadesh (1952) list (see Section 4.3 below, and see Appendix A for the full list with Malay translations). Ethnobiological information about animal species names was collected using color books containing

photographs of mammal, bird and reptile species (Francis, 2001; Davison & Chew, 2003; Cox, van Dijk, Nabhitabhata & Thirakhupt, 1998).

Access to the corpus of collected materials, held in the Repository and Workspace for Austroasiatic Intangible Heritage (RWAAI) is available upon request at http://hdl.handle.net/10050/00-0000-0000-0003-E6ED-9@view.

4.2 Participants

The grammatical description of Jedek in Study I is informed by speech examples from a number of Jedek speakers, and by discussions in which grammatical hypotheses were checked with Jedek speakers. The study of the lexical consequences of Jedek-Jahai contact in Study II is based on elicited language production data from a total of 71 speakers, including Jedek and Jahai speakers from Rual, and Jahai speakers from the Jahai-majority settlements of Banun and Manok. For the investigation of semantic interaction in Jedek-Jahai bilinguals in Studies III and IV a subset of the data analyzed in Study II is used, including data from Jedek monolinguals and bilinguals, Jahai bilinguals at Rual, and Banun Jahai speakers. Studies III and IV include additional bilingual data from Jedek and Jahai bilinguals at Rual (allowing for within-speaker comparison across varieties). As discussed above in Section 3.3.2, 'Jedek speakers' are participants who self-identify as speakers of Jedek, and 'Jahai speakers' are participants who self-identify as speakers of Jahai. Similarly, in Studies III and IV, 'bilinguals' are participants who self-identify as speakers of both Jedek and Jahai, and 'monolinguals' are participants who self-identify as speakers of Jedek or Jahai, but not both. 'Jedek-identifying bilinguals' are participants whose identity variety is Jedek, but who also self-identify as having knowledge of Jahai. Likewise, 'Jahai-identifying bilinguals' are participants whose identity variety is Jahai, but who also self-identify as having knowledge of Jedek. The specific numbers of speakers of each group included in the datasets of each study can be found in the relevant sections of Studies II, III and IV.

4.3 Materials and procedure

The grammatical description in Study I is based primarily on the recordings of Jedek natural speech and grammatical discussions contained in the corpus, and is also informed by the recordings of elicited speech from Jedek speakers. Studies II, III and IV are based on elicited speech recordings collected in director-matcher tasks using the Topological Relations Picture Series (TRPS; Bowerman & Pederson, 1992), the PUT project (Bowerman et al., 2004), and Reciprocal Constructions and Situation Type (Evans et al., 2004) tasks, and on recorded basic vocabulary data elicited in

questionnaire format using a wordlist based on Swadesh's (1952) list, with added items appropriate to Aslian contexts[4] (see Appendix A for the full wordlist).

The director-matcher task used for elicitation of the topological relations, caused motion event and reciprocal event descriptions was set up such that two Jedek or Jahai speakers interacted with one another to achieve a communicative goal. The 'director' described each scene to the 'matcher', who located a picture of that scene on the pages of a folder containing pictures of each scene. While the TRPS, PUT project and Reciprocal Constructions and Situation Type stimulus kits were not designed to be run as director-matcher tasks, this procedure was chosen for two reasons. First, since director-matcher tasks involve communication between two participants, it was thought that this would reduce the risk of Foreigner Talk (Ferguson, 1975) and lead to a greater amount of more natural speech in the recordings. Second, the director-matcher procedure allowed for more control over the language mode (Grosjean, 1998) of participants – sessions in Jedek were performed with a Jedek matcher, and sessions in Jahai were performed with a Jahai matcher. In the questionnaire-style basic vocabulary elicitations, participants were asked to give the Jedek or Jahai (depending on the identity variety of the participant) translation for each wordlist item, which were presented orally in Malay by the researcher. Where the Malay categories were inadequate for elicitation of Jedek or Jahai forms (e.g., for pronouns or kinship terms where the Jedek and Jahai categories are more fine-grained than the Malay system), elicitation was through oral description in Jedek, and for body-part terms for which pointing was a more effective method of elicitation, elicitation was through pointing.

4.4 Data treatment

For the grammatical description of Jedek in Study I, utterances from the transcribed materials of the corpus (including grammatical discussions, natural speech and traditional stories) were entered into a spreadsheet along with utterances from natural speech examples noted down during the fieldwork. This document formed a searchable corpus of speech examples on which the grammatical analysis was based, supplemented with speech examples from the corpus of elicited speech data. For the analyses in Studies II, III and IV, the responses of each participant to each scene of the tasks were transcribed and the basic vocabulary items (in the Swadesh list data), verbs (in the caused motion event and reciprocal event data) and topological relations markers (TRMs; in the topological relations data) were entered into a dataframe.

[4] The added items were based on Benjamin's (1976) adapted Aslian wordlist. The Malay translation of the adapted Swadesh list was compiled prior to commencing fieldwork; list items were ordered semantically rather than alphabetically.

For the analysis in Study II, the responses to each stimulus were coded as identical or non-identical across participants[5], and the sequences of coded responses for each dataset were used to construct phylogenetic networks using the NeighborNet algorithm (Bryant & Moulton, 2004) in the SplitsTree4 package (Huson & Bryant, 2006). The use of phylogenetic networks was chosen in order to allow for comparison of the lexical features of the ways of speaking of a large number of individuals. The NeighborNet algorithm is capable of representing conflicting splits, and was thus thought to be a suitable approach to representing patterns of micro-variation between individual speakers.

For the analyses in Study III, the topological relations description data were aggregated at the group level (in four groups: Jedek monolinguals, Jahai monolinguals, Jedek-identifying bilinguals, and Jahai-identifying bilinguals). For the analyses in Study IV, the placement event description data were aggregated at group level in the same way. In Studies III and IV, analysis was in the form of extension maps and congruence analyses. The extension maps show the semantic extensions of each TRM or verb over the scenes of the task, for each group of participants. The extension of a TRM or verb includes scenes for which more than 33% of the speakers of a group used that TRM or verb (following Bowerman, 1996; Levinson & Meira, 2003; cf. Indefrey, Şahin & Gullberg, 2017). The congruence analyses quantify the degree of overlap of the extensions of Jedek and Jahai TRMs or verbs across the monolingual groups and across varieties within the bilingual groups. Congruence is defined as the number of scenes for which a pair of Jedek-Jahai forms were most frequently used in Jedek *and* Jahai divided by the number of scenes for which at least one of the two forms was most frequently used in Jedek *or* Jahai. Extension maps and congruence analyses are useful tools in investigation of bilingual semantic interaction as they afford comparison of the semantic extensions of wordforms across languages and across participant groups. Since they allow for comparison of ways of speaking in the absence of well-defined monolingual norms, these methods are particularly suited to study of multilingual outcomes in small-scale multilingual contexts. The analyses of Study III focus on the TRM *klɛŋ* and related TRMs, while the analyses of Study IV focus on the full set of placement verbs used in participants' responses.

As a result of the small size of the Rual speech community, the sample sizes of the studies are relatively small (the pool of individuals who are Jahai-identifying Jedek-Jahai bilinguals is particularly small). For this reason the studies present descriptive statistics rather than inferential statistics.

[5] For more detailed description of the coding principles see the methods section of Study II; for the full datasets and python script used to code the data see Appendices A–E of Study II.

5 The empirical studies

The four empirical studies contained in the thesis explore the potential of lesser-known linguistic contexts for enriching our understanding of multilingualism and language contact phenomena. Study I presents Jedek and establishes its status as a distinct Northern Aslian language variety. Studies II, III and IV explore the linguistic consequences of interaction between Jedek and the closely-related language variety Jahai in small bands of egalitarian foragers. Study II assesses the lexical consequences of contact between Jedek and Jahai speakers in the Rual community, providing an indication of the extent of lexical contact effects in this site of long-term, intense language contact. Next, Studies III and IV explore the semantic consequences of interaction between Jedek and Jahai in the minds of Jedek-Jahai bilinguals. Studies III and IV investigate bilingual semantic interaction in the domains of topological relations (Study III) and placement event descriptions (Study IV).

5.1 Study I

Study I presents Jedek, a previously unrecognized Northern Aslian (Austroasiatic) language variety spoken by four bands of Semang foragers residing in the Sungai Rual resettlement area in northern Peninsular Malaysia. Previous to the commencement of the fieldwork conducted in connection with the current thesis, little was known about Jedek. The study provides the first description of the phonological, morphological and syntactic features of the variety. On the basis of the linguistic analysis presented, Study I establishes the status of Jedek as a Northern Aslian variety that is as distinct from its closest Northern Aslian relatives as these varieties are from one another. The analysis reveals a number of features that are shared by Jedek and neighboring Aslian varieties, and a number of features that set Jedek apart from its neighbors, at all levels of language. Study I reveals pronominal distinctions in Jedek that are not shared by its closest relatives, and Jedek allows open final syllables, not allowed in neighboring Aslian varieties. Jedek's demonstrative paradigm is less elaborate than those of its closest relatives, and the paradigms of Jedek interrogative and irrealis forms lack direct parallels in other Aslian languages. Jedek negation and argument-marking strategies also appear to be distinct, and preliminary lexical comparison points to a lexical conservatism in Jedek not found in its close relatives.

Given the results of Study I, Jedek and Jahai are considered to be two distinct but closely-related, lexically and typologically similar Northern Aslian varieties. Speakers of the two varieties are co-settled at Rual since the 1970s, and have likely been in regular contact since before resettlement. Rual is thus a site of small-scale multilingualism and language contact in egalitarian foragers speaking closely-related language varieties, a highly relevant context for exploration of the issues in focus in the thesis. Study I points to the very apparent situation of underdocumentation in the Rual context – indeed Jedek is not only previously undescribed, but previously unknown to linguists. In order to investigate the consequences of multilingualism and language contact in settings like Rual, basic linguistic description is a necessary part of the workflow, highlighting the need for communication between the multilingualism, language contact and primary linguistic documentation and description fields.

5.2 Study II

Study II investigates the lexical consequences of contact between Jedek and Jahai speakers at Rual. In response to the methodological challenges associated with study of egalitarian, small-scale language contact between closely-related, recently-described language varieties (see Sections 2.5 and 2.6 above), Study II adopts a novel approach focusing on the language production of individual speakers. Lexical language production data from individual speakers are compared in phylogenetic networks, allowing for analysis of contact outcomes without the need to operate with abstract definitions of language systems and predefined groups. Data from individual Jedek and Jahai speakers at Rual and from Jahai speakers living in two other geographical locations are compared. These data are from four domains: basic vocabulary, topological relations descriptions, caused motion event descriptions and reciprocal event descriptions.

The results of Study II suggest that the cohabitation of and contact between Jedek and Jahai speakers at Rual has resulted in a high degree of lexical convergence in the ways of speaking of Rual residents. This convergence is such that the lexicon of Jedek and Jahai speakers is not differentiated for three of the four domains studied. Meanwhile, the lexicon of Jahai speakers in other geographical areas is seen to be distinct from that of both Jedek and Jahai speakers at Rual. Differentiation between the lexicon of Jedek and Jahai speakers at Rual is seen in the basic vocabulary data, suggesting that the ways of speaking of Jedek- and Jahai-identifying individuals at Rual have not converged completely. The patterns seen in the phylogenetic networks of Study II reflect the contact histories of speakers. The lexical features of Jahai speakers in the geographically-distant Jahai settlement Banun are most distinct from those of Rual residents, while the differentiation of Rual residents and Jahai speakers in the geographically more proximal Jahai settlement Manok is slightly less clear-cut across the datasets. Finally, the least

differentiation is seen in the lexical features of Jedek and Jahai speakers who cohabitate at Rual.

Study II further reveals a pattern wherein the lexicon of the youngest (both Jedek- and Jahai-speaking) individuals of the Rual sample tends to be the most divergent from that of Jahai speakers in other areas. This suggests a change in progress in which ways of speaking at Rual are diverging from ways of speaking in other areas. In this way, the study reveals patterns of lexical convergence within Rual, as well as lexical divergence between geographically-defined groups. The results of Study II demonstrate how comparison of language production data from individual speakers in phylogenetic networks can reveal contact effects without the need to define abstract language systems for comparison. In addition, due to its focus on the language production of individual speakers, this method can also reveal patterns of language change within communities. The methodology presented in Study II represents a novel approach to investigation of the linguistic consequences of language contact, which allows for the assessment of the extent of contact effects in situations of small-scale egalitarian contact between closely-related, recently-described language varieties. The use of phylogenetic networks to compare language production data from individual speakers is not only useful in studies of language contact in this type of setting but has a range of potential applications in a range of contexts. Study II points to the variation in ways of speaking within and across language varieties in the Rual setting and highlights the complexity with which the ways of speaking of individuals may relate to named language codes in contexts of intense language contact.

5.3 Study III

Study III explores semantic interaction in Jedek-Jahai bilinguals at Rual in the domain of topological relations. Study III aims to broaden the empirical base of our understanding of bilingual semantic interaction, extending the scope of analysis to a non-Western, non-standardized, small-scale egalitarian multilingual setting. The study tests the predictions of Muysken's (2013) framework of *bilingual optimization strategies* for the impact of the social and linguistic features of multilingual settings on the outcomes of multilingualism. The typological and lexical similarity of Jedek and Jahai, their relatively equal levels of prestige, and the low levels of normativity at Rual are predicted to lead to bidirectional influence between the varieties in bilinguals. Meanwhile, the greater number of Jedek speakers than Jahai speakers at Rual is predicted to favor unidirectional influence of Jedek on Jahai in bilinguals.

Two groups of Jedek-Jahai bilinguals (one group of Jedek-identifying bilinguals and one group of Jahai-identifying bilinguals) and two groups of self-defined Jedek and Jahai monolinguals described the Topological Relations Picture Series (Bowerman &

Pederson, 1992), a set of 71 line drawings depicting a range of topological relations, in a director-matcher task (see Section 4.3 above). The monolinguals completed the task once, in their identity variety, and the bilinguals completed the task once in Jedek and once in Jahai. In extension maps and congruence analyses (see Section 4.4 above), the semantic extensions of Jedek and Jahai topological relation markers (TRMs) are compared in monolinguals and in the two groups of bilinguals. The analyses focus on the TRM *klɛŋ* whose form is identical but whose semantic extension differs in Jedek and Jahai. In Jahai, *klɛŋ* is a relational noun meaning 'inside', while in Jedek, *klɛŋ* is a semantically general preposition. The results of Study III reveal evidence of semantic interaction in the topological relations descriptions of both groups of bilinguals, but with different patterns of directionality in the two groups. Symmetric interaction is seen in the extensions of the Jahai-identifying bilinguals; this group of bilinguals use *klɛŋ* with general meaning in both Jedek and Jahai, akin to the Jedek monolingual use of *klɛŋ*. At the same time, their use of *klɛŋ* for containment scenes is expanded in both varieties as compared to the Jedek monolinguals, revealing influence from the Jahai pattern. Meanwhile, asymmetric interaction in the direction of Jedek is seen in the Jedek-identifying bilinguals. This group of bilinguals use *klɛŋ* as a general location marker in both varieties, as in the Jedek monolingual pattern, and no influence from Jahai is seen in their use of *klɛŋ* in Jedek.

The finding of symmetric semantic interaction in Jahai-identifying bilinguals in Study III is in line with Muysken's (2013) predictions based on the similarity of Jedek and Jahai, the egalitarian nature of Jedek-Jahai contact and the low normativity levels in the Rual speech community. At the same time, the asymmetric pattern seen in Jedek-identifying bilinguals is in line with the prediction that the more limited access to Jahai speakers at Rual should favor asymmetric interaction in the direction of Jedek. The results of Study III suggest that the social features of multilingual settings may play a role in affecting the outcomes of multilingualism. At the same time, the findings point to a complexity in the relationship between the different factors affecting multilingual outcomes. The study points to the need for research in a broader range of multilingual scenarios in order to allow for the formulation of more precise predictions about the effects of social and linguistic factors on the outcomes of multilingualism and language contact. In the psycholinguistic multilingualism literature, the effects of social factors are rarely accounted for. The results of Study III suggest that more emphasis needs to be placed on the social features of multilingual settings if we are to develop a fuller understanding of the factors that lie behind multilingual outcomes.

5.4 Study IV

Study IV extends the investigation of bilingual semantic interaction in Jedek-Jahai bilinguals to the verb lexicon, in the semantic domain of placement events. The analyses of Study IV treat the placement event domain as a whole, rather than focusing on one particular form as in Study III. The Jedek and Jahai placement verb lexica are highly similar both in terms of wordforms and in terms of the semantic content of verbs. Thus, while most studies in the literature investigate semantic interaction in language pairs with key differences in semantic granularity, study of the Jedek and Jahai placement verb lexicon offers us the chance to investigate interaction in highly similar semantic systems. Study IV adopts the same methodological approach as that used in Study III, providing a base for comparison between results from the two different domains. The study tests the predictions of Muysken's (2013) framework of *bilingual optimization strategies* in the placement event domain. Two groups of Jedek-Jahai bilinguals (Jedek-identifying and Jahai-identifying bilinguals) and two groups of self-identified Jedek and Jahai monolinguals responded to the PUT project task (Bowerman et al., 2004), a set of 63 short video clips depicting a range of caused motion events, in a director-matcher task. The semantic extensions of the Jedek and Jahai verbs used to describe the placement event scenes of the task are analyzed in extension maps and congruence analyses.

Study IV reveals increased congruence in the semantic extensions of identical Jedek-Jahai verbs in bilinguals as compared to monolinguals. Thus, the study provides evidence of bilingual semantic interaction in the absence of differences in semantic granularity, suggesting that even where the languages of bilinguals are highly similar, bilinguals are still able to find ways to make them more similar. In contrast to the results from the topological relations domain in Study III, the interaction seen in the placement event domain is symmetric in both groups of bilinguals. The effects of semantic interaction are seen in both the Jedek and Jahai of bilinguals, in parts of the domain in which there exist minor differences between the semantic extensions of Jedek and Jahai monolinguals. The symmetric pattern of interaction seen in the results is in line with Muysken's (2013) prediction that egalitarian relations between varieties, low levels of normativity, and similarity of language varieties will favor symmetry in multilingual outcomes. However, the prediction that uneven numbers of speakers in a community may lead to asymmetry in multilingual outcomes (as seen in the results for Jedek-identifying bilinguals in Study III) is not borne out.

The differences and points of similarity in the findings of Studies III and IV suggest a complexity in the interaction of social and linguistic factors in influencing multilingual outcomes. In the topological relations domain, in which there are greater differences between the semantic systems of Jedek and Jahai monolinguals, we saw above that Study III finds symmetric semantic interaction in Jahai-identifying bilinguals, and

asymmetric semantic interaction in favor of Jedek in Jedek-identifying bilinguals. In Study IV, a subset of Jedek and Jahai placement verbs reveal this same pattern of asymmetry. Greater differences between Jedek and Jahai monolinguals are seen in semantically-specific placement verbs, used for one scene of the task each. The pattern of semantic interaction seen in these verbs is symmetric in Jahai-identifying bilinguals and asymmetric in favor of Jedek in Jedek-identifying bilinguals. The combined findings of Studies III and IV may be interpreted as an indication that domains or parts of domains in which there is more similarity between languages are more likely to undergo symmetric patterns of interaction. The results of the two studies highlight the importance of both social and linguistic factors in influencing multilingual outcomes.

6 Conclusions and future work

The thesis is concerned with the ways in which small-scale, egalitarian multilingual speech communities can inform our theories of multilingualism and language contact. The findings reported suggest that they may do this in a number of ways. We have seen how a broadening of the scope of multilingualism research to a wider range of multilingual settings can enrich our understanding of the ways in which multilingual outcomes are affected by social and linguistic factors. Further, we have seen how expansion of the empirical base of language contact research in contexts of egalitarian contact between small-scale, closely-related language varieties can provide important insights into processes of language contact and change. Importantly, we have also seen that an approach that combines perspectives from these fields serves the interests of both fields, allowing us to better understand different aspects of the consequences of interaction between languages in the community and in the minds of bilinguals. The thesis highlights a number of gaps in current knowledge and demonstrates some ways in which the task of filling these gaps may be approached in the presence of methodological challenges.

The studies of the thesis introduce the newly discovered Aslian variety Jedek, and reveal a high degree of lexical and semantic convergence in the language production of multilingual Semang foragers speaking Jedek and Jahai. Study I presents a description of the grammatical features of Jedek and finds that Jedek and Jahai represent two closely-related but distinct Aslian language varieties. Next, Study II reveals that while Jedek speakers are lexically distinct from Jahai speakers in other geographical areas, intense and long-term contact between Jedek and Jahai speakers living in the village of Rual has resulted in a high degree of lexical convergence in their ways of speaking. This convergence is such that the lexicon of Jedek and Jahai speakers at Rual is not differentiated in the majority of the domains studied. Meanwhile, the lexicon of both Jedek and Jahai speakers at Rual differs from that of Jahai speakers in other areas. In addition to this pattern of convergence, Study II reveals a pattern in which ways of speaking at Rual are diverging from ways of speaking in other geographical areas. These results add support to a similar finding from another situation of contact between small-scale, closely-related languages (François, 2011). Importantly, however, the findings of Study II demonstrate the ways in which this kind of pattern can play out in convergence across linguistic boundaries within speech communities and divergence within linguistic boundaries between speech communities.

The results of Studies III and IV reveal semantic convergence in Jedek-Jahai bilinguals, and highlight the role of social and linguistic factors in impacting the outcomes of multilingualism. Different patterns of semantic convergence are seen in two semantic domains, and in two groups of Jedek-Jahai bilinguals. Social factors such as egalitarian prestige relations between languages, low levels of normativity in the speech community, and the number of speakers of languages in the speech community are seen to impact the outcomes of multilingualism. At the same time, linguistic factors such as the typological and lexical similarity of languages and the nature of linguistic systems in contact are also seen to play an important role, interacting with the influence of social factors in complex ways. The findings of Studies III and IV suggest that where the linguistic structures of the languages of bilinguals are more similar, this may favor bidirectional outcomes, where such an outcome is compatible with the influence of social factors. This result is seen in the patterns of symmetric semantic interaction found in bilinguals in the domain of placement events, where Jedek and Jahai semantic systems are highly similar. Meanwhile, the results suggest that social factors may come into play where differences between linguistic structures are greater. This result is seen in asymmetric bilingual semantic interaction in Jedek-identifying bilinguals in the domain of topological relations where differences between Jedek and Jahai are larger. Here, patterns of asymmetry of semantic interaction in bilinguals reflect social factors in the multilingual setting.

Support is also seen for predictions about the role of linguistic factors based on the psycholinguistic literature on bilingual semantic interaction. Study III finds that more fine-grained semantic distinctions become less fine-grained in interaction with more general semantic distinctions, in line with findings from previous research. Meanwhile, the results of Study IV demonstrate that in the absence of differences in semantic granularity, the semantic extensions of bilinguals may become more similar in areas of minor difference. Finally, the results of Studies III and IV suggest that in interaction between lexically and typologically very similar language varieties the semantic extensions of form-identical lexemes, rather than semantically-equivalent lexemes, become more similar across varieties in bilinguals. These findings demonstrate the ways in which broadening the scope of multilingualism research to a wider range of multilingual settings can advance our theories about multilingualism.

The challenges involved in the application of current methods to research involving multilingualism and language contact in small-scale and closely-related language varieties have been discussed by several authors (e.g. Lüpke, 2016; Epps et al., 2013; see also discussions in Studies II and IV below). The studies of the current thesis demonstrate methodological approaches that are of potential use in overcoming some of these challenges. Study II illustrates the advantages of an approach in which elicited language production data from a large number of speakers are compared in phylogenetic networks. The study demonstrates how this approach allows for contact outcomes to be assessed without necessitating the definition of languages as abstract,

well-defined entities. Studies III and IV illustrate the advantages of extension maps and congruence analyses for study of bilingual semantic interaction in small-scale multilingual contexts, in that they allow for identification of patterns of interaction between the languages of bilinguals in the absence of standardized language varieties. The studies also demonstrate the advantages of director-matcher tasks for use in elicitation of bilingual data in small-scale multilingual contexts, in that they allow for a degree of control over the language mode of participants. The current thesis emphasizes the need for the development of methodologies suited to research in contexts of small-scale multilingualism and language contact in recently-described, closely-related language varieties, and illustrates the advantages of methodological approaches of the kind used in the thesis.

The findings reported highlight the importance of discussions about the contribution of different factors in influencing the outcomes of interaction between languages in multilingualism and language contact. While social factors are not often accounted for in the psycholinguistic literature on multilingualism, the results of the thesis indicate that a perspective that takes into account the impact of social factors on multilingual outcomes can help us to understand the different outcomes that are seen in different multilingual settings and in different groups of multilingual speakers. We have seen for example that different patterns of directionality of semantic interaction in different groups of bilinguals speaking the same language pair can be understood through consideration of the impact of social factors. Conversely, while studies in the language contact field rarely take into account the role of individual speakers in processes of contact-induced language change, we have seen the value of perspectives focusing on the interaction of languages in the minds of bilinguals for our understanding of language contact phenomena. In this way, the thesis illustrates the potential of greater integration of findings and theoretical questions from the fields of language contact and multilingualism studies (cf. the recommendation in Muysken, 2013).

Discussions about the impact of linguistic and social factors on the outcomes of interaction between languages are found in both the multilingualism and language contact literature in different ways. But these discussions could potentially be expanded into models with which more detailed predictions could be made about the impact of specific features, allowing us to make more precise predictions about the outcomes expected in specific contexts. The formulation of this kind of a model would represent a major advance for both our language contact and multilingualism theories. Muysken's (2013) framework of bilingual optimization strategies represents an important step in this direction, bringing together findings from diverse fields of enquiry in order to form broad predictions about the impact of social and linguistic factors on multilingual outcomes. In the current thesis, we have seen that Muysken's framework is a useful tool in beginning to form predictions about the occurrence of different patterns of outcomes in different multilingual scenarios. Meanwhile, since the framework predicts outcomes based on clusters of features, and does not provide weights for the contribution of

different factors to expected outcomes, it can be difficult to interpret the ways in which specific multilingual outcomes are predicted to play out in specific multilingual settings. For example, in the Rual context where different features of the setting are predicted to favor different outcomes, we have seen that different predictions are possible on the basis of the framework. A system of weighting could potentially allow for predictions about the relative role of different factors to be made. Some suggestions about potential weighting of factors can be found in the literature, for example, in a context where the linguistic and social features of a setting result in conflicting predictions, Thomason and Kaufmann (1988) predict that social factors will tend to win out. This kind of prediction could be tested with the help of a more detailed model.

In order to be able to make more precise predictions about the impact of specific factors, our models require further development and importantly, testing in a broad range of multilingual scenarios. This testing should not only consider different multilingual settings, but different groups of individuals within these settings (see the different results for two groups of Jedek-Jahai bilinguals in Studies III and IV). Further, the development of our models should take into account the ways in which different factors interact to produce specific multilingual outcomes (see the results of Studies III and IV). Importantly, testing of our models must also take into account the potential for different outcomes in different linguistic domains, or different parts of linguistic domains, within the same language pair (see the different results for two different semantic domains in Studies III and IV; see also Alferink, 2015). All of these aspects have been seen to play a role in the studies of the thesis, highlighting the complexity involved in the development of models predicting the outcomes of language contact and multilingualism.

Evans (2018) makes the case that human language emerged in the context of small-scale multi(proto)lingualism, through a process of contact between speakers of small-scale language varieties. If this is the case, multilingualism and language contact research in small-scale multilingual settings represents an important key to advancing our understanding of processes of human language evolution. In particular, since we know that the communities in which language evolved were forager communities, studies of language contact and multilingualism in foragers are particularly valuable in this regard. Small-scale egalitarian contexts and in particular those involving foragers currently represent a conspicuous gap in our knowledge of multilingual and language contact phenomena – thus research in this kind of setting harbors great potential for shaping our understanding of multilingual and language contact phenomena, and of the mechanisms involved in human language evolution.

Given the linguistic and social diversity found within the forager category, and given the diversity of outcomes found in different contexts depending on the particular features of scenarios in which languages interact, an understanding of multilingual and language contact phenomena in foragers will need to take into account the impact of

the different features of different forager settings. Where marriage patterns are exogamous for example, language ecologies may differ greatly depending on whether in-marrying individuals typically come from one specific language group or from many different language groups. Postmarital living arrangements are likely to affect multilingual outcomes, as are norms around who should speak which language and to whom. In some cases, it may be necessary for in-marrying individuals to learn to speak one or more languages other than the language(s) they have grown up with, while this may not be necessary in other cases. Where communication is through receptive bilingualism, this will affect patterns of multilingualism and language use in the community. The similarity of language varieties and the length of the contact histories of speakers will also play a role. Thus an understanding of the impact of different social and linguistic factors on the outcomes of language contact and multilingualism will play an important role in understanding the kinds of processes that may have been involved in human language evolution.

The current thesis emphasizes the value of perspectives that take into account the role of individual speakers in multilingualism and language contact research. Studies in the language contact field tend to conceive of languages as abstract systems, removed from their use by speakers. However, such a view ignores the actual processes involved in the emergence and spread of contact features in speech communities. In reaction to this tendency in the literature, Thomason (2008: 51) argues that "any search for absolute linguistic constraints on language change, whether internally-motivated change or contact-induced change, must focus on the process of innovation by one or more individuals". The studies of the thesis illustrate the advantages of such a focus on the role of individual speakers in processes of language contact and change. Through comparison of the linguistic features of individuals' ways of speaking in Study II, processes of convergence, divergence and language change in the community are revealed. Further, in Studies III and IV comparison across varieties within bilinguals in a contact setting allows us to gain a better understanding of the outcomes of interaction between languages.

Finally, the thesis demonstrates the value of lesser-known linguistic settings for our theories about multilingualism and language contact. The studies of the thesis point to the potential of the rich multilingual settings often found in contexts in which primary linguistic documentation and description work is carried out to provide important new insights for multilingualism and language contact studies. The thesis highlights the advantages of taking the descriptive task one step further by not only describing the linguistic features of language varieties but also documenting and describing the features of the often-complex multilingual settings in which these language varieties are spoken. The thesis describes a very apparent situation of linguistic underdocumentation, with Jedek being not only previously undescribed but also previously unknown to linguists. The studies of the thesis uncover dimensions of unrecognized linguistic diversity beyond the level of named linguistic codes. To identify

Jedek as a distinct language variety, it was necessary to go beyond the use of official language labels and probe the linguistic content behind the labels. Further, by moving away from ideas about languages in contact as abstract systems removed from their use by speakers, patterns of contact outcomes and processes of language change within and between language communities are seen. Finally, a look beyond the level of the speaker, to examine the ways in which language varieties interact *within* speakers, allowed us to gain a deeper understanding of the ways in which languages may change as a result of the interaction between them. Studies of multilingualism and language contact in small-scale, egalitarian, underdocumented contexts highlight dimensions of linguistic diversity that are not often taken into account, and hold great potential for providing new insights into the human language capacity.

Abbreviations

1, 2, 3 1st, 2nd, 3rd person; BECK beckoning particle; CAUS causative; CLF classifier; COLL collective plural; CURSE cursing word; DEI deictic; DET determiner; DISTR distributive; DU dual; DUB dubitative particle; EMPH emphatic particle; EXCL exclusive; EXCLAM exclamation particle; CONTR contrast-marking preposition; GOAL goal-marking preposition; IMP imperative; INCL inclusive; INSTR instrument-marking; INV invitational particle; IPFV imperfective; IRR irrealis; ITER iterative; LOC location-marking preposition; N noun; NEG negator; NMZ nominalizer; NOM MOD nominal modifier; OBJ object-marking preposition; PL plural; PRO pronominal; PROG progressive; PROP property; PST.DIST distant past; PST.PROX recent past; Q question proclitic; QNT quantifier; REL relative; RT relational tense; SG singular; SOURCE source-marking preposition; SUB subject-marking preposition; UNIT unitization

References

Aikhenvald A. Y., & Dixon, R. M. W. (2001). Introduction. In A. Y. Aikhenvald & R. M. W. Dixon (eds.), *Areal diffusion and genetic inheritance: Problems in comparative linguistics,* 1–26. Oxford University Press.

Alferink, I. (2015). *Dimensions of convergence in bilingual speech and gesture*. Nijmegen: Radboud Universiteit Nijmegen. (Doctoral dissertation.)

Alferink, I., & Gullberg, M. (2014). French–Dutch bilinguals do not maintain obligatory semantic distinctions: Evidence from placement verbs. *Bilingualism: Language and Cognition* 17(1). 22–37.

Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language* 60(2). 270–290.

Appel, R., & Muysken, P. (1987). *Language contact and bilingualism*. Amsterdam University Press.

Baetens Beardsmore, H. (1982). *Bilingualism: Basic principles.* Clevedon: Multilingual Matters.

Benjamin, G. (1976). Austroasiatic subgroupings and prehistory in the Malay Peninsula. In P. N. Jenner, L. C. Thompson & S. Starosta (eds.), *Austroasiatic studies, Part 1,* 37–128. Honolulu: The University of Hawai'i Press.

Benjamin, G. (1985). In the long term: Three themes in Malayan cultural ecology. In K. L. Hutterer, T. Rambo & G. Lovelace (eds.), *Cultural values and human ecology in South East Asia*, 219–278. University of Michigan: Center for South and Southeast Asian Studies.

Benjamin, G. (2011). Egalitarianism and ranking in the Malay World. In K. S. Sillander & T. Gibson, (eds.), *Anarchic solidarity: Autonomy, equality and fellowship in Southeast Asia,* 170–201. New Haven, CT: Yale University Southeast Asia Studies.

Benjamin, G. (2013). Why have the peninsular "negritos" remained distinct? *Human Biology* 85(1–3). 445–484.

Bickel, B., & Nichols, J. (in press). Linguistic typology and hunter-gatherer languages. In T. Güldemann, P. McConvell & R. Rhodes (eds.), *The language of hunter-gatherers.* Cambridge: Cambridge University Press.

Bohnemeyer, J., Bowerman, M., & Brown, P. (2001). Cut and break clips. In S. C. Levinson & N. J. Enfield (eds.), *Manual for the field season 2001,* 90–96. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.874626.

Bowerman, M. (1996). Learning how to structure space for language: A cross-linguistic perspective. In P. Bloom, M. Peterson, L. Nadel & M. Garrett (eds.), *Language and space*, 385–436. Cambridge, MA: MIT Press.

Bowerman, M., Gullberg, M., Majid, A., & Narasimhan, B. (2004). Put project: The crosslinguistic encoding of placement events. In A. Majid (ed.), *Field Manual, Vol. 9,* 10–18. Nijmegen: Max Planck Institute for Psycholinguistics.

Bowerman, M., & Pederson, E. (1992). Topological relations picture series. In S. C. Levinson (ed.), *Space stimuli kit 1.2: November 1992,* 51. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.883589.

Bowern, C. (2013). Relatedness as a factor in language contact. *Journal of Language Contact* 6(2). 411–432.

Bowern, C., Epps, P., Gray, R., Hill, J., Hunley, K., McConvell, P., & Zentz, J. (2011). Does lateral transmission obscure inheritance in hunter-gatherer languages? *PloS One* 6(9). e25195.

Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition* 30(2). 225–251.

Brown, A., & Gullberg, M. (2011). Bidirectional cross-linguistic influence in event conceptualization? Expressions of Path among Japanese learners of English. *Bilingualism: Language and Cognition* 14(1). 79–94.

Bryant, D., & Moulton, V. (2004). Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21(2). 255–265.

Bullock, B. E., & Gerfen, C. (2004). Phonological convergence in a contracting language variety. *Bilingualism: Language and Cognition* 7(2). 95–104.

Burenhult, N. (2012). The linguistic encoding of placement and removal events in Jahai. In A. Kopecka & B. Narasimhan (eds.), *Events of putting and taking: A crosslinguistic perspective,* 21–36. Amsterdam: John Benjamins Publishing.

Burenhult, N., & Kruspe, N. (2016). The language of eating and drinking: A window on Orang Asli meaning-making. In K. M. Endicott (ed.), *Malaysia's original people: Past, present and future of the Orang Asli,* 175–199. Singapore: NUS Press.

Burenhult, N., & Purves, R. (in press). The spatial properties of forager motion categories: evidence from Jahai. Hunter-Gatherer Research.

Chelliah, S. L., & de Reuse, W. J. (2010). *Handbook of descriptive linguistic fieldwork.* Springer.

Clyne, M. (2003). *Dynamics of language contact: English and immigrant languages.* Cambridge University Press.

Cook, V. J. (ed.) (2003). *Effects of the second language on the first*. Clevedon: Multilingual Matters.

Cox, M. J., van Dijk, P. P., Nabhitabhata, J., & Thirakhupt, K. (1998). *A photographic guide to snakes and other reptiles of Peninsular Malaysia, Singapore and Thailand.* London: New Holland Publishers (UK) Ltd.

Davison, G. W. H., & Chew Yen Fook (2003). *Birds of Peninsular Malaysia and Singapore.* London: New Holland Publishers (UK) Ltd.

Dixon, R. M. W. (1980). *The languages of Australia.* Cambridge: Cambridge University Press.

Dixon, R. M. W. (1997). *The rise and fall of languages*. Cambridge: Cambridge University Press.

Dunn, M., Burenhult, N., Kruspe, N., Tufvesson, S., & Becker, N. (2011). Aslian linguistic prehistory: A case study in computational phylogenetics. *Diachronica* 28. 291–323.

Dunn, M., Kruspe, N., & Burenhult, N. (2013). Time and place in the prehistory of the Aslian languages. *Human Biology* 85(1/3). 383–410.

Edwards, J. (2004). Foundations of bilingualism. In T. K. Bhatia & W. C. Ritchie (eds.) *The Handbook of bilingualism,* 7–31. Blackwell Publishing.

Endicott, K. M. (1974). *Batek Negrito economy and social organization.* Cambridge, Massachusetts: Harvard University. (Doctoral dissertation.)

Endicott, K. M., & Endicott, K. L. (2008). *The headman was a woman: The gender egalitarian Batek of Malaysia.* Long Grove, Illinois: Waveland Press.

Epps, P., Bowern, C., Hansen, C. A., Hill, J. H., & Zentz, J. (2012). On numeral complexity in hunter-gatherer languages. *Linguistic Typology* 16(1). 41–109.

Epps, P., Huehnergard, J., & Pat-El, N. (2013). Introduction: Contact among genetically related languages. *Journal of language contact* 6(2). 209–219.

Evans, N. (2010). *Dying words: Endangered languages and what they have to tell us*. Chichester: Wiley-Blackwell.

Evans, N. (2018). Did language evolve in multilingual settings? *Biology & Philosophy* 32(6). 905–933.

Evans, N., & Dench, A. (2006). Introduction: Catching language. In F. Ameka, A. Dench & N. Evans (eds.), *Catching a language: The standing challenge of grammar writing,* 1–40. Berlin: Mouton de Gruyter.

Evans, N., Gaby, A., Levinson, S. C., & Majid, A. (eds.). (2011). *Reciprocals and semantic typology*. Amsterdam: Benjamins.

Evans, N., Levinson, S. C., Enfield, N. J., Gaby, A., & Majid, A. (2004). Reciprocal constructions and situation type. In A. Majid (ed.), *Field Manual Volume 9,* 25–30. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.506955.

Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics* 17. 1–14.

Francis, C. M. (2001). *Mammals of South-East Asia.* London: New Holland Publishers (UK) Ltd.

François, A. (2011). Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1. 175–246.

François, A. (2012). The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language* 214. 85–110.

Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late bilinguals: All words are not created equally. *Bilingualism: Language and Cognition* 13(4). 385–408.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition* 15(3). 594–615.

Gomes, A. G. (2007). *Modernity and Malaysia: Settling the Menraq forest nomads*. London: Routledge.

de Groot, A. M. (1993). Word-type effects in bilingual processing tasks. In R. Schreuder & B. Weltens (eds.), *The bilingual lexicon,* 27–51. Amsterdam: John Benjamins.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36(1). 3–15.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1(2). 131–149.

Güldemann, T., McConvell, P., & Rhodes, R. A. (in press). Hunter-gatherer anthropology and language. In T. Güldemann, P. McConvell & R. Rhodes (eds.), *The language of hunter-gatherers*. Cambridge: Cambridge University Press.

Gullberg, M. (2012). Bilingual multimodality in language documentation data. In F. Seifart, G. Haig, N. P. Himmelmann, D. Jung, A. Margetts & P. Trilsbeek (eds.), *Language Documentation & Conservation 6* (Special issue *"Potentials of Language Documentation: Methods, Analyses, Utilization"*), 54–63. Honolulu: University of Hawai'i Press.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological science* 15(6). 409–414.

Haudricourt, A-G. (1961). Richesse en phonèmes et richesse en locuteurs. *L'Homme* 1:1. 5–10.

Haugen, E. (1950). The analysis of linguistic borrowing. *Language* 26(2). 210–231.

Hickey, R. (2010). Language contact: Reconsideration and reassessment. In R. Hickey (ed.), *The handbook of language contact,* 1–28. Wiley-Blackwell.

Himmelmann, N. P. (1998). Documentary and descriptive linguistics. *Linguistics* 36. 161–195.

Himmelmann, N. P. (2012). Linguistic data types and the interface between language documentation and description. *Language Documentation and Conservation* 6. 187–207.

Holm, J. (2000). *An introduction to pidgins and creoles.* Cambridge University Press.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2). 254–267.

Indefrey, P., Şahin, H., & Gullberg, M. (2017). The expression of spatial relationships in Turkish/Dutch bilinguals. *Bilingualism: Language and Cognition* 20(3). 473–493. doi:10.1017/S1366728915000875 (online Jan. 2016).

Jarvis, S., & Pavlenko, A. (2008). Crosslinguistic influence in language and cognition. New York: Routledge.

Kelly, R. L. (1995). The lifeways of hunter-gatherers: the foraging spectrum. Cambridge University Press.

Kopecka, A., & Narasimhan, B. (2012). *Events of putting and taking: A crosslinguistic perspective*. Amsterdam: John Benjamins Publishing.

Kruspe, N., Burenhult, N., & Wnuk, E. (2014). Northern Aslian. In M. Jenny & P. Sidwell (eds.), *The handbook of Austroasiatic languages,* 415–474. Brill.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity.* Cambridge University Press.

Levinson, S. C., Brown, P., Danzinger, E., De León, L., Haviland, J. B., Pederson, E., & Senft, G. (1992). Man and Tree & Space Games. In S. C. Levinson (ed.), *Space stimuli kit 1.2,* 7–14. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.2458804.

Levinson, S. C., & Meira, S. (2003). 'Natural concepts' in the spatial topological domain—Adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79(3). 485–516.

Levinson, S. C., & Wilkins, D. P. (2006). *Grammars of space: Explorations in cognitive diversity* (Vol. 6). Cambridge University Press.

Lüpke, F. (2016). Uncovering small-scale multilingualism. *Critical Multilingualism Studies* 4.2. 35–74.

Lye, T-P. (2013). Making friends in the rainforest: "Negrito" adaptation to risk and uncertainty. *Human Biology* 85(1/3). 417–445.

Majid, A., Bowerman, M., van Staden, M., & Boster, J. S. (2007). The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18(2). 133–152.

Majid, A., Evans, N., Gaby, A., & Levinson, S. C. (2011). The grammar of exchange: A comparative study of reciprocal constructions across languages. *Frontiers in Psychology* 2. 1–15.

Majid, A., & Kruspe, N. (2018). Hunter-gatherer olfaction is special. *Current Biology* 28(3). 409–413.

Malt, B. C., & Lebkuecher, A. L. (2017). Representation and process in bilingual lexical interaction. *Bilingualism: Language and Cognition*, 20(5). 867–885. doi:10.1017/S1366728916000584.

Malt, B., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science* 4. 583–597.

Matras, Y., & Bakker, P. (2003). The study of mixed languages. In Y. Matras & P. Bakker (eds.), *The mixed language debate,* 1–20. Berlin: Mouton de Gruyter.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research* 50. 940–967.

Moravcsik, E. (1978). Universals of language contact. In J. H. Greenberg (ed.), *Universals of human language,* 94–122. Stanford: Stanford University Press.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16. 709–730.

Quinn, C. M. (2013). Endangered languages. In C. A. Chapelle (ed.), *The encyclopedia of applied linguistics.* Oxford, UK: Wiley-Blackwell.

Rudge, A. C. (2017). *Sound and socio-aesthetics among the Batek hunter-gatherers of Pahang State, Malaysia*. London: University College London. (Doctoral dissertation.)

Sankoff, G. (2001). Linguistic outcomes of language contact. In J. K. Chambers & N. Schilling (eds.), *The handbook of language variation and change*, 638–668. John Wiley.

Swadesh, M. (1952). Lexicostatistic dating of prehistoric ethnic contacts. *Proceedings of the American Philosophical Society* 96. 452–63.

Tacey, I. (2018). Animism and interconnectivity: Batek and Manya' life on the periphery of the Malaysian rainforest. Helsinki: University of Helsinki. (Doctoral dissertation).

ten Thije, J. D., & Zeevaert, L. (2007). *Receptive multilingualism: Linguistic analyses, language policies and didactic concepts.* Amsterdam/Philadelphia, PA: John Benjamins.

Thomason, S. G. (2001). *Language contact.* Edinburgh University Press.

Thomason, S. G. (2008). Social and linguistic factors as predictors of contact-induced change. *Journal of language contact* 2(1). 42–56.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics.* Berkeley & Los Angeles, CA: University of California Press.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review* 9(4). 780–789.

Vaughan, J. (2018). "We talk in saltwater words": Dimensionalisation of dialectal variation in multilingual Arnhem Land. *Language and Communication* 62. 119–132.

Vaughan, J., & Singer, R. (2018). Indigenous multilingualisms past and present. *Language and Communication* 62. 83–90.

Weinreich, U. (1953). *Languages in contact: Findings and problems*. Walter de Gruyter.

Wnuk, E. (2016). *Semantic specificity of perception verbs in Maniq*. Nijmegen: Max Planck Institute for Psycholinguistics. (Doctoral dissertation.)

Woodburn, J. (1982). Egalitarian societies. *Man* 17. 431–451.

Appendices

Appendix A: Wordlist for basic vocabulary elicitation

Swadesh 200 (1952) list + Benjamin's (1976) 12 added items + kinship terms
Consultant:
Identity variety:
Place of residence:
Place of origin:
Date of recording:

English	Malay	
hand	tangan	
arm	tangan	
claw/nail	kuku	
foot	kaki	
leg	kaki	
knee	lutut	
head	kepala	
ear	telinga	
eye	mata	
nose	hidung	
mouth	mulut	
tongue	lidah	
tooth (front)	gigi	
hair	rambut	
skin (of person)	kulit	
neck	leher, tengkuk	
breast	buah dada	
back (of person)	belakang	
belly	perut	
bone	tulang	
heart	jantung	
liver	hati	
guts	isi perut	
blood	darah	
I	saya	
you (PL)	awak orang	

thou/you	awak	
he	dia, ia	
we (incl.)	kita	
we (excl.)	kami	
they	dia orang	
person	orang	
child	budak (kinship: anak)	
woman	perempuan	
man (male)	laki-laki	
mother	ibu	
father	ayah	
sibling, elder	abang (m), kakak (f)	
sibling, younger	adik	
wife	bini	
husband	suami	
grandmother	nenek	
grandfather	atuk	
great-grandparent	moyang (~ ancestor)	
uncle	bapak saudara	
aunt	mak saudara	
siblings	adik beradik	
elder brother	abang	
elder sister	kakak	
cousin	sepupu	
niece/nephew	anak saudara	
child	anak	
grandchild	cucu	
great-grandchild	cicit	
spouse's parent	mentua	
spouse's sibling, sibling's spouse	kakak/abang/adik ipar	
child's spouse's parents	besan	
wife's sister's husband, husband's brother's wife	biras	
child's spouse	menantu	
animal	binatang	
bird	burung	
dog	anjing	
fish	ikan	
snake	ular	
worm	cacing	
louse	kutu, tuma	
feather	bulu burung	
tail	ekor	
wing	sayap	
egg	telur	
meat	daging	
fat (substance)	lemak	
rice	nasi, beras	

salt	garam	
knife	pisau	
blowpipe	sumpit	
quiver	tabung anak damak	
spear	lembing	
rope	tali	
road	jalan	
woods	hutan	
river	sungai	
lake	tasik	
mountain	gunung	
tree	pokok, pohon	
bark (of tree)	kulit kayu	
flower	bunga	
leaf	daun	
root	akar	
fruit	buah	
seed	biji	
rice (plant)	padi	
bamboo	buluh	
stick (of wood)	kayu, tongkat	
stone	batu	
grass	rumput	
earth (soil)	tanah	
water	air	
fire	api	
smoke	asap	
ashes	abu	
dust	debu, abu	
wind	angin	
cloud	awan	
fog	kabut	
sky	langit	
sun	matahari	
moon	bulan	
star	bintang	
day (opp. of night)	hari	
night	malam	
year	tahun	
sand	pasir	
sea (ocean)	laut	
snow	ais	
ice	ais, air batu	
name	nama	
good	baik	
bad (unsuitable)	tidak baik (evil = jahat)	
big	besar	
small	kecil	

warm (weather)	panas	
cold (weather)	sejuk	
sharp	tajam	
dull (blunt)	tumpul	
near	dekat	
far	jauh	
heavy	berat	
light (weight)	ringan	
wide	lebar	
narrow	sempit, lipis	
long	panjang	
short	pendek	
new	baru	
old	tua, lama	
wet	basah	
dry	kering	
thick	tebal	
thin	nipis (person: kurus)	
full	penuh	
straight	lurus, betul	
smooth	licin	
dirty	kotor	
rotten (of wood)	reput	
right (correct)	betul	
green	hijau	
yellow	kuning	
black	hitam	
white	putih	
red	merah	
to walk	berjalan kaki	
to fly	terbang	
to swim	berenang	
to float	timbul	
to stand	berdiri	
to sit	duduk	
to lie (on side)	baring	
to come	datang	
to turn (veer)	pusing	
to eat	makan	
to drink	minum	
to bite	gigit	
to cut	potong	
to dig	gali	
to hunt	buru	
to kill	bunuh	
to shoot (blowpipe)	sumpit	
to die	mati	
to live	hidup	

to sleep	tidur	
to wash (bathe)	mandi	
to breathe	tarik nafas	
to vomit	muntah	
to suck	kulum, isap	
to hit	pukul	
to hold	pegang	
to give	bagi	
to split	belah	
to pull	tarik	
to push	tolak	
to squeeze	perah	
to stab	tikam	
to spit	berludah	
to throw	baling	
to tie	ikat	
to rub	gosok	
to wipe	sapu	
to scratch	garu	
to sew	jahit	
to laugh	gelak tawa	
to sing	nyanyi	
to dance	menari, tari	
to play	bermain	
to fight	melawan	
to fear	rasa takut	
to say	kata, cakap	
to hear	dengar	
to smell	cium	
to know	tahu	
to count	bilang	
to think	pikir	
to fall (drop)	jatuh	
to flow	mengalir	
to rain	hujan	
to burn	bakar	
to blow (of wind)	angin bertiup	
to swell	kembang	
what?	apa?	
when?	bila?	
where?	di mana?	
who?	siapa?	
how?	macam mana?	
left (side)	kiri	
right (side)	kanan	
this	ini	
that	itu	
here	di sini	

there	di sana, di situ	
many	banyak	
few, some	sikit	
all	semua	
and	dan	
because	sebab	
if	kalau	
not	tidak	
other	lain	
at	di	
in	dalam	
with	dengan	
one	satu	
two	dua	
three	tiga	
four	empat	
five	lima	

Appendix B: Rual language identities survey data

Each pair of individuals demarcated with horizontal lines represents a married couple; partners are ordered female, male.

Language identity	Origin	Mother's language	Mother's origin	Father's language	Father's origin
Jedek	Rual	Jedek	Rual	Jahai	Perak
Jedek	Rual	Jedek	Rual	unknown	unknown
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jahai	Rual	Jahai	Perak	Jahai	Perak
Jedek	Rual	Jahai	Rual	Jahai	Rual
Jedek (deceased)	Rual				
Jedek	Rual	Jahai	Rual	Jahai	Rual
Jahai (deceased)					
Jedek	Rual	Jedek	Lebir	Temiar	Gua Musang
Jahai	Manok	Jahai	Rual	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	Manok
(separated)					
Jedek	Rual	Jedek	Rual	Jahai	Manok
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jahai	Rual	Jahai	Rual	Jahai	Rual
Jedek (deceased)	Rual	Jahai	Rual	Jedek	Rual
Jedek	Rual	Jahai	Rual	Jedek	Rual
Jahai	Manok	Jahai	Manok	Jedek	Manok
Jahai	Rual	Jahai	Rual	Temiar	
Jedek (deceased)	Rual				
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jahai	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jahai	Rual	Jedek	Rual
Jahai	Rual	Jahai	Perak	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jahai	Rual	Jahai	Rual	Jahai	Rual
Jedek	Rual	Jedek	Lebir	Jahai	Perak
Temiar	Kuala Lah	Temiar		Temiar	
Jedek	Rual	Jedek	Lebir	Jahai	Perak
Jahai	Rual	Jahai	Rual	Jahai	Rual
Jahai	Manok	Jahai	Manok	Jahai	Manok
Jedek	Rual	Jedek	Rual	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jahai	Perak
Jedek (deceased)	Rual				
Jahai	Rual	Jahai	Rual	Jahai	Rual

Jedek	Rual	Jedek	Rual	Jahai	Perak
Jahai	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jahai	
Jahai	Perak	Jahai	Klap, Perak	Jahai	Klap, Perak
Jedek (deceased)	Rual				
Jedek	Rual	Jedek	Lebir	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jahai	Banun				
Jedek	Rual	Jedek	Lebir	Jahai	Perak
Jahai	Manok	Jahai	Manok	Jahai	Manok
Jedek	Rual	Jedek	Rual	Jahai	Manok
Semaq Beri	Terengganu				
Jedek	Rual	Jedek	Rual	Jahai	Manok
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Lebir	Jahai	Perak
Jahai	Rual	Jahai	Rual	Temiar	Perak
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jahai	Perak				
Jedek	Rual	Jedek	Rual	Jahai	
Jedek	Lebir	Jahai	Rual	Jahai	Rual
Jedek	Rual	Jahai	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jahai	
Jedek	Rual	Jedek	Rual	Jahai	
Jahai	Rual	Jahai	Rual	Temiar	Perak
Menriq	Kuala Lah				
Jahai	Rual				
Temiar (deceased)	Perak				
Lanoh	Perak	Lanoh	Perak	Lanoh	
Jahai	Rual	Jedek	Rual	Jahai	
Jedek	Rual	Jahai		Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual				
Jedek (deceased)	Rual				
Jahai	Banun				
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual		
Semaq Beri	Terengganu				
Jedek	Rual	Jedek	Rual	Jedek	Rual
L?pa?	Kuala Lah				
Jedek	Rual	Jedek	Rual	Jahai	Manok
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
(separated)	(unknown)				
Jedek	Rual	Jedek	Rual	unknown	unknown
Jahai	Rual	Lanoh	Perak	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual

Jedek	Rual	Jahai	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jahai	Perak
Jedek	Rual	Jahai	Rual	Jedek	Rual
(separated)					
Jahai	Rual	Jahai	Perak	Jahai	Perak
Jedek (deceased)	Rual				
Jahai	Rual	Jahai	Perak	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jahai	Rual	Jedek	Rual
(deceased)					
Jahai	Rual				
Jedek	Rual				
Jedek	Rual	Jahai	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jahai	Rual
Temiar	Gua Musang				
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jedek	Rual	Jedek	Rual	Jahai	
Jahai	Rual	Jahai	Perak	Jahai	Perak
Jahai	Rual	Jahai	Rual	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jahai	Rual				Rual
Jedek	Rual	Jedek	Rual	Jedek	Rual
Jahai	Rual			Jahai	
Jahai	Rual			Jahai	
Jahai	Rual				
Jahai	Rual			Jahai	
Jedek	Rual	Jedek	Lebir	Jahai	Perak
Jahai	Rual				
Jedek	Rual	Jedek	Rual	Jahai	Rual
Jahai	Rual	Jahai	Rual	Jahai	Rual
Jahai	Perak			Jahai	Perak
Jahai	Manok				
Temiar	Perak				
Jahai	Rual	Jahai	Perak	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	
Jahai	Rual	Jahai	Rual	Jedek	Rual
?					
Jahai	Rual	Jahai	Perak	Jahai	Perak
Jedek	Rual	Jedek	Rual	Jahai	
Jahai	Rual	Jahai	Rual	Jedek	Rual
Jedek	Rual	Jahai	Rual	Jedek	Rual
Jahai	Rual	Jahai	Rual	Jedek	Rual
Jahai	Manok	Jahai	Manok	Jahai	Manok
Jahai	Rual				
L?pa?	Kuala Lah				
Jahai	Rual				
Jedek	Rual	Jedek	Rual	Jahai	Perak
Jahai					
(deceased)					
Jahai					
Jedek (separated)	Rual	Jedek	Rual	Jahai	Rual

Jahai Jahai (deceased)	Rual	Jahai		Jahai	
Jedek Jahai	Rual Rual	Jedek Jahai	Rual	Jahai Jahai	Perak Rual
Jahai (separated)	Rual	Jahai	Rual	Jahai	Rual
Jahai Jahai (deceased)	Rual Rual				
Jedek Jedek	Rual Rual	Jedek Jedek	Rual Rual	Jahai L?pa?	Rual Kuala Lah
Jedek Jahai (deceased)	Rual Rual	Jedek	Rual	Jahai	Rual
Jahai Temiar	Rual Perak	Jahai	Rual	Jahai	Rual
Temiar Jedek	Gua Musang Rual	Jedek	Rual	Jahai	Rual

Studies I-IV

Study I

Language Profile

Joanne Yager and Niclas Burenhult

Jedek: A newly discovered Aslian variety of Malaysia

https://doi.org/10.1515/lingty-2017-0012
Received November 22, 2016; revised October 10, 2017

Abstract: Jedek is a previously unrecognized variety of the Northern Aslian subgroup of the Aslian branch of the Austroasiatic language family. It is spoken by about 280 individuals in the resettlement area of Sungai Rual, near Jeli in Kelantan state, Peninsular Malaysia. The community originally consisted of several bands of foragers along the middle reaches of the Pergau river. Jedek's distinct status first became known during a linguistic survey carried out in the DOBES project *Tongues of the Semang* (2005–2011). This article describes the process leading up to its discovery and provides an overview of its typological characteristics.

Keywords: Aslian, Austroasiatic, grammar sketch, Jedek, undiscovered languages

1 Background

Much of the world's linguistic diversity remains undocumented and uninvestigated by science. For the majority of the world's languages there is only scant information available, and only a small proportion has been subject to in-depth grammatical and lexical description. Typically, however, languages and dialects have some degree of scientific or administrative recognition, even those which have not been targeted by systematic studies. But, as was shown by the widely publicized 2008 discovery of Koro in northeastern India (Anderson & Murmu 2010), there are languages which may have passed entirely unnoticed. For example, as in the case of Koro, their speakers may not recognize themselves as ethnically or linguistically distinct from some other community of speakers,

Joanne Yager: E-mail: joanne.yager@ling.lu.se, **Niclas Burenhult:** E-mail: niclas.burenhult@ling.lu.se, Språk- och litteraturcentrum, Lunds universitet, Box 201, Lund 221 00, Sweden

and in the absence of systematic study their distinctiveness goes unreported. In other cases, recognized dialect varieties may upon closer examination turn out to be more distinct than previously assumed and warrant more independent characterization and classification. The re-classification of Zialo, a Mande language of Guinea, is a case in point (Babaev 2010); the status of the North Germanic variety Övdalian as distinct from Swedish is another (Dahl 2008).

The Aslian group of languages, a typologically distinct branch of the Austroasiatic language family spoken in the Malay Peninsula, is no stranger to classificatory mysteries and inconsistencies. Most of these minority languages are spoken by 14 ethnolinguistic groups officially recognized in Malaysian administrative practice. But linguistic work on Aslian has tended to operate with finer distinctions. For example, an early lexicostatistical analysis and genealogical classification of Aslian by Geoffrey Benjamin included 20 varieties, sampled not only according to administrative practice but also on the basis of older sources as well as previously unreported names obtained from consultants (Benjamin 1976). All of the additional varieties included were, in a sense, concealed by the official classification.

Setting out from Benjamin's 1976 study, the *Tongues of the Semang* project – a language documentation program supported by the Volkswagen Foundation's DOBES scheme (2005–2011) – carried out targeted surveying of Aslian-speaking forager groups (known ethnographically as the Semang) in the Malaysian states of Perak and Kelantan and the southern Thai provinces of Trang and Satun. The survey, the bulk of which was carried out in March to May 2006, involved the in-situ collection of 200-item Swadesh lists and basic sociolinguistic and grammatical information from a total of 28 settlements or camps, with the purpose of providing a refined and up-to-date overview of language varieties and their endangerment status. 24 of these were located in Malaysia and inhabited by groups officially recognized as Lanoh, Kensiw, Kintaq, Jahai, Menriq, and Batek. The four locations in Thailand were inhabited by groups known linguistically and ethnographically as Ten'en or Maniq (cf. Bishop & Peterson 2003; Wnuk 2016). All of these ethnolinguistic groups speak varieties of the Northern Aslian subbranch of Aslian, except Lanoh, which is Central Aslian.

The lexical data emanating from the survey have been comprehensively explored with computational phylogenetic and phylogeographic techniques in a series of subsequent works (Dunn et al. 2011; Burenhult et al. 2011; Dunn et al. 2013; Yager 2013). These analyses refine and largely support Benjamin's 1976 sampling and classification of the relevant sectors of the Aslian family tree, showing for example that the official label Lanoh harbors several distinct language varieties. They also highlight the complex patterns of contact typical

of the highly mobile and socially flexible Aslian-speaking foragers (cf. Benjamin 1985: 234–235; see further in Section 2).

However, the survey data collected in the resettlement area of Sungai Rual, located on the Rual river near Jeli in northwest Kelantan, offered an immediate surprise. Sungai Rual is inhabited partly by people who refer to themselves as Jahai, and partly by people who, to outsiders, refer to themselves varyingly as Batek or Menriq. All three labels form part of Malaysian administrative practice, and all three ethnic groups have the bulk of their speaker populations in other locations – Jahai in the area of Lake Temenggor in Perak, Batek in southeastern Kelantan and adjacent parts of Terengganu and Pahang, and Menriq in the village of Kuala Lah in central Kelantan. But while the Jahai Swadesh list collected at Rual corresponded well with the list previously collected among the Jahai in Perak, the Batek/Menriq list from Rual diverged significantly from the lists obtained in the Batek and Menriq heartlands further south and southeast. The amount of shared cognates between the Rual variety and other Batek and Menriq varieties was between 65 and 78 percent, which is on a par with the rates observed between the recognized language varieties, e.g., between Jahai and Menriq (ca. 72%) and between Batek and Menriq (also ca. 72%). For comparison, the two Jahai lists had 89% shared cognates.[1] The Rual variety also did not show any clear signs of approaching cohabitant Jahai (ca. 68% shared cognates). The separate lexical status of the Rual list is also apparent in the later computational analyses of lexical divergence, where it is as distinct from the Jahai, Batek, and Menriq lists as these are from each other (see Figure 1). On the basis of these lexical patterns, the Rual variety together with varieties of Jahai, Batek, and Menriq are posited to form a subbranch within Northern Aslian, labeled Menraq-Batek (Dunn et al. 2011: 314).

In his ethnographic account of the Sungai Rual resettlement area, Gomes (2007: 76–77) explains that the inhabitants who called themselves either Menriq or Batek traced their origin to four different bands which prior to resettlement in the 1970s roamed the middle section of the Pergau valley. Gomes's study was not a linguistic one and thus the nature of the language variety spoken by these bands was, until the 2006 survey, unknown.[2] The unexpected lexical divergence

1 No historical reconstruction is available for the Aslian branch of Austroasiatic. In the work cited here, cognates are forms which share the same place of articulation in both the consonant onset and coda of the final syllable, with certain systematic exceptions; for details, see Dunn et al. (2011: 300–301).

2 Gomes's 2007 book remains the most significant ethnographic account of the Rual inhabitants. Further anthropological work has been carried out by Kamal Solhaimi Fadzil, Diana Riboli, and Ivan Tacey.

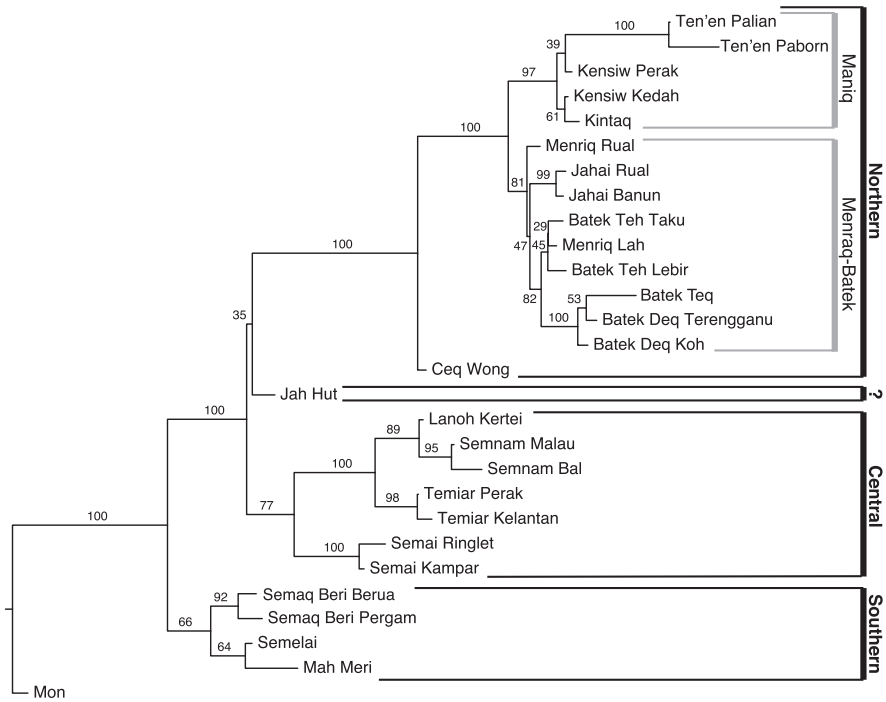


Figure 1: Aslian family tree, rooted on the Austroasiatic outlier Mon (from Dunn et al. 2011). The tree is based on basic vocabulary and is a Maximum Clade Credibility tree. Numbers on the branches indicate percentage of the tree sample supporting each bifurcation, and branch length indicates rate of lexical divergence (for details, see Dunn et al. 2011). The tree broadly reproduces the earlier proposed clades of Aslian genealogy (Benjamin 1976) but the phylogenetic aspect also reveals that the clades show very unequal rates of lexical divergence. In particular, the Maniq and Menraq-Batek varieties of Northern Aslian (corresponding to the Semang foragers) are contained within a clade which is highly divergent externally, but which has low internal diversity, suggesting a rather recent diversification. This diversification is estimated to have started around 1,500 to 2,000 years ago (Dunn et al. 2013) and hypothesized to have been boosted by the contact dynamics of the forager groups (Burenhult et al. 2011). The Dunn et al. 2011 study provided an early indication of the lexical distinctiveness of Jedek, labeled "Menriq Rual" in this chart.

identified by the survey provided a first hint that the self-designated Menriq/Batek at Rual were in fact speaking an unknown Northern Aslian variety, obscured by existing ethnonyms and therefore overlooked by previous linguistic work. Our subsequent survey data collection focused on lexical domains and grammatical classes which had already been documented in the surrounding varieties. This was done with the help of one Rual consultant in February 2008

and included animal species vocabulary, kinship terms, pronouns, and demonstratives, as well as basic sentences.

In terms of lexicon, the Rual variety harbors terms not documented among its close neighbors. Interestingly, some of these point to similarities to the distant Maniq and Kensiw languages (also Northern Aslian but from a different subgroup). Most notably: the 1st person singular *ʔiɲ* is identical to Maniq *ʔiɲ* '1SG' (cf. Jahai, Menriq, and Batek *yɛʔ* '1SG'); the variety has retained the indigenous term *pip* 'ashes, dust' just like Kensiw (Kensiw *tpip* 'ashes, dust'; cf. Jahai, Menriq, and Batek *ʔabuʔ* or *habuʔ*, from Malay *abu* 'ashes, dust'); and the term for 'tiger' (or rather 'large felid') is *ʔɔʔ*, the cognate of Maniq *taʔɔʔ* 'tiger' (cf. Wnuk 2016; cf. Jahai and Menriq *ʔap*, Batek *ʔayɔʔ*).[3] These features seem to suggest some degree of lexical conservatism on the part of the Rual variety, not shared by its closest neighbors and relatives, or possibly an historical situation of contact with a Maniq/Kensiw-type language.

While the lexical materials collected during the survey were too limited for comprehensive phonetic, phonological, and phonotactic analysis, one feature was particularly salient. In the Rual variety, /r/ is realized as a uvular or velar fricative [ʁ] or [ɣ] in syllable-initial position, and as zero in syllable-final position. The back realizations are unattested in Jahai and Menriq, which display an apical trill [r] in all positions, but do occur in some Batek varieties as well as local dialects of Malay. The Rual variety also allows open final syllables, not allowed in Jahai, Menriq, or Batek. Speakers of surrounding language varieties frequently point out that the pronunciation of /r/ is one of the features that set speakers of the Rual variety apart, along with the 1st person singular form *ʔiɲ*. However, /r/ realization is sometimes notoriously varied and unstable within and across Aslian varieties and speakers (cf. Wnuk & Burenhult 2014: 968), so no conclusions could be drawn on the basis of this limited data.

Speakers of neighboring varieties acknowledge the linguistic distinctiveness of the Rual variety, as do the speakers themselves. However, like several other Aslian ethnolinguistic groups they do not have a dedicated endonym apart from the generic *mnraʔ* ('human being' or 'indigenous person'), and when asked by outsiders for a name for themselves they typically respond with one of the officially recognized names Menriq or Batek. Speakers of surrounding varieties, the Jahai and Menriq in particular, frequently refer to speakers of the Rual variety as *Jdɛk*, a name of unknown origin. Upon further work with the group

3 Although superficially similar, the Batek term *ʔayɔʔ* cannot at this point be identified as a cognate of *ʔɔʔ*, see Footnote 1.

at Rual it was discovered that this name is also used spontaneously by the speakers themselves in addition to the labels previously recorded. Most consultants readily accept this name, and they do not consider it derogatory. The term was first recorded by Geoffrey Benjamin in Jeli in 1970 as *Jədɛk* (Benjamin, field notes). His Jahai consultant claimed it referred to an extinct Semang band that had once lived on the Jedok river, a tributary of the Thailand-Malaysia border river Golok. The headwaters of Jedok are also close to an area in which a relevant Semang band was observed in the 1930s (Rentse 1937; see further in Section 2 below).

Furthermore, the name bears a noteworthy resemblance to the ethnonym Tea-De, documented by Phaiboon (2006: 208) as referring to an enigmatic and little-known population of Northern Aslian speakers in the Waeng district of Narathiwat Province in southernmost Thailand, across the border some 20 kilometers from Rual. Phaiboon's 2006 wordlist indicates that Tea-De is part of the Maniq-Kensiw subgroup of Northern Aslian, but it is unclear if the Maniq-Kensiw-like features of the Rual lexicon are to be somehow linked to it. It is not unreasonable to assume that the two varieties were in regular contact in the past, but no such interaction can be documented at present.

Taken together, the results emerging from the survey suggested that the Rual variety was sufficiently distinct to imply a separate historical signal and merit independent description and documentation. The present article represents a first step in this descriptive endeavor, and Sections 3 and 4 provide a preliminary outline of the main structural and lexical features of the variety, based on extensive new fieldwork carried out in Rual by co-author Yager in 2013 to 2016. Given its comparatively unambiguous reference and accepted status in the community, the term *Jdɛk* (pronounced [ɟᶻəˈdɛkˀ], romanized as Jedek) is here proposed as the scientific label for the object of our linguistic inquiry. However, we do not use this term in reference to a particular ethnic group or community, and we wish to emphasize that we introduce the term solely for the purpose of disambiguation and characterization of a linguistic entity, not an ethnographic one. Jedek replaces the term Menriq Rual, used in previous reporting of the survey data (Dunn et al. 2011, 2013; Burenhult et al. 2011; Yager 2013); it is currently subsumed under the ISO code "mnq" (Menriq).

Although the Rual community is ethnographically relatively well known, Jedek remained undiscovered as a linguistic variety until the start of the present research program. We believe the reason for this is threefold:

(i) The lack of a common or standard ethnolinguistic label (exonymic or endonymic) has prevented Jedek from attracting the attention of linguists and ethnographers as a distinct entity; the speakers' habit of designating

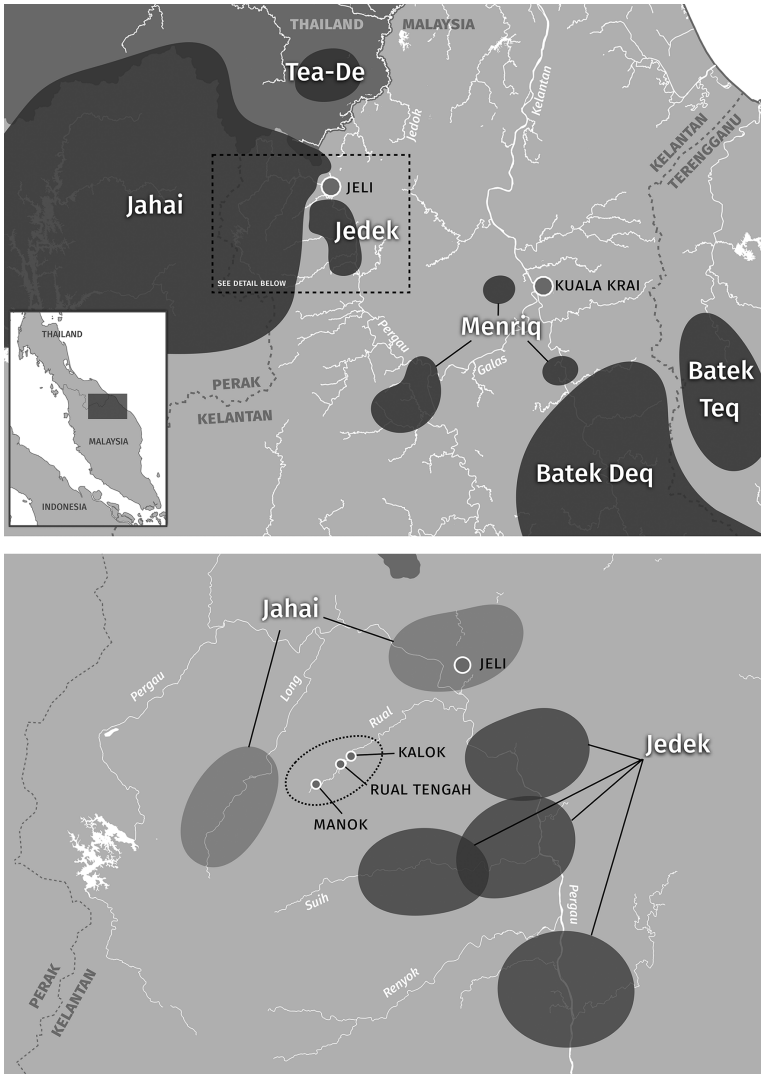
themselves as either Jahai, Batek, or Menriq to outsiders has led analysts to believe that these were also their linguistic affiliations and that the sociolinguistic situation of Rual involved a mix of Northern Aslian varieties whose existence was already known.

(ii) The high degree of multilingualism among the speakers of Jedek and their cohabitation and intermarriage with speakers with different language backgrounds (in particular Jahai), coupled with the reported Semang pattern of idiolectal variation (cf. Section 2), make for a metalinguistic fluidity in how Jedek speakers define their way of speaking. At times speakers define their way of speaking as "the same" as Jahai, or "the same as, mixed with" Jahai, and at times they are clear that "that is how it is in Jahai, in our language it's different". Thus speakers' characterizations of their own language have not been revealing as to its nature.

(iii) A general lack of knowledge about Northern Aslian varieties and a dearth of earlier surveys have prevented researchers from easily appreciating the diversity harbored by the subgroup; several of the varieties remain unexplored and it is only within the last 15 years that we have gained in-depth knowledge about some Northern Aslian varieties, and had access to materials with which we can compare new data.

2 Historical background, sociolinguistics, and endangerment

In the beginning of the 1970s, Jedek- and Jahai-speaking Semang groups in northwest Kelantan comprised at least six distinct bands scattered along the mid-section of the Pergau valley – roughly from Jeli in the north to Kampung Jebang in the south – as well as along the larger tributaries of the Pergau, especially the rivers Long, Suih, and Renyok (cf. Gomes 2007); see Map 1. Among these, the two most upstream bands (around Jeli and Long) are reported by Gomes as being primarily associated with the Jahai ethnicity, whereas the remaining four are reported as Menriq/Batek and are thus likely to have been primarily Jedek-speaking. It is possible that Jedek speakers were also present further east at an earlier stage, as suggested by the information given to Benjamin (see above) as well as an eyewitness account from the 1930s by Rentse (1937) which locates a Semang band to the Kelubi, a tributary of the Bertam river located some 20 kilometers east of Rual. This area is presently not associated with any Aslian-speaking groups. Rentse provides a hint of the linguistic identity of this band – the term *ʔabən* for 'good' (or



Map 1: The top map shows the approximate historical distribution of Northern Aslian varieties in northeastern Peninsular Malaysia and adjacent parts of southern Thailand. Speakers are currently mostly resettled in several permanent villages within these former areas of distribution. The two eastern enclaves of Menriq represent small populations who are sometimes referred to as Batek Teh; however, our current lexical data suggest their language is nearly indistinguishable from Menriq as spoken further west. The exact location and distribution of the enigmatic variety Tea-De in Thailand's Waeng district is unknown (cf. Phaiboon 2006: 208). The bottom map is a close-up of the middle and upper part of the Pergau watershed. The Sungai Rual resettlement area and its three hamlets are indicated, as is the approximate historical distribution and assumed linguistic affiliation of the six local bands settled at Sungai Rual (adapted from Gomes 2007: 77).

abøtn in his rendering; Rentse 1937: 130), later only documented in Jedek and Menriq.[4]

With government-sponsored resettlement in the 1970s, five bands from Pergau were relocated to the Rual site: three of the Jedek-speaking bands settled together to form the Rual Tengah hamlet, while one Jahai band and one Jedek-speaking band settled together to form the Kalok hamlet around 500 meters downstream. Soon after, the last remaining band of those that lived along the mid-section of the Pergau valley, a large group of Jahai, joined the resettlement area, forming a hamlet of their own, Manok, around two kilometers upstream from Rual Tengah (Gomes 2007). In the 1980s one further (smaller) Jahai band moved from the state of Perak to join the two bands already living at Kalok. Thus the Sungai Rual resettlement area was formed, consisting of three hamlets comprised of seven bands, of which three were primarily Jahai and four were primarily "Menriq/Batek" (in Gomes's terminology).

The lifestyle of Semang groups is traditionally highly mobile, manifested in small-group nomadism as well as a pattern of group breakup and regrouping into new units in response to changing subsistence conditions. The Semang also practice band exogamy, which means that intermarriage between individuals of widely dispersed bands or of different linguistic backgrounds is common. It has been suggested that these spatial and social dynamics result in particular patterns of variation and change in individual language use among the Semang (see especially Benjamin 1985: 234–235, 1987: 114; Bishop & Peterson 1993; Endicott 1997; Wnuk & Burenhult 2014). Benjamin (2001: 111) reports that Semang groups have until quite recently maintained "a continuous mesh of communication" with each other covering the entire Semang area. In our phylogenetic analyses of Aslian basic vocabulary we indeed found signals of long-term lexical exchange among those Aslian languages that are spoken by the foraging Semang populations (Burenhult et al. 2011). Due to frequent contact with surrounding communities, Aslian speakers are typically multilingual and frequently speak three or more languages fluently, including neighboring Aslian languages and at least one of the adjacent majority languages Malay and Thai. And since most Semang groups contain members of several different language backgrounds, it has been suggested that Semang speech communities display a high degree of variation in individuals' ways of speaking (Benjamin 2001). In addition, the fact that a Semang individual may

4 Schebesta (1928) makes no mention of distinct groups although he traveled through the relevant area in search of Semang communities; Benjamin (1987: 115) designates the Northern Aslian groups of the Pergau valley as Jahai.

move through several linguistic environments throughout his or her lifetime has been suggested to result in high rates of idiolectal change among Semang communities. Thus the sociolinguistic characteristics of Semang groups have been described by scholars as highly fluid, highly idiolectally varied, and highly multilingual. Our data suggest that the Jedek speech community is no exception to this pattern.

While most Rual residents continue to live off the forest in various ways, since resettlement their lifestyle is no longer characterized by small-group nomadism. Rual residents do still adhere to rules concerning band exogamy. However, the resettlement of several bands together at the Rual site means that access to eligible partners within the Rual community is increased, and thus that band exogamy no longer necessarily requires the relocation of one of the parties. Nowadays, it is common for Rual residents to marry a partner who has also grown up in Rual, in fact roughly 90% of Rual residents have origins in the original seven Jedek- and Jahai-speaking bands that were resettled at Rual. Meanwhile, with a population of around 150 adult individuals, it is not always possible to find a partner within Rual and it is not uncommon that individuals relocate to or from Rual for marriage purposes. The Rual community contains a number of residents with origins in other areas, in particular Jahai speakers from Perak, Menriq speakers from Kuala Lah, Temiar speakers from the Gua Musang area and from Perak, and Semaq Beri or Semelai from Terengganu. The majority of Rual residents are however Jedek- and Jahai-speaking, and speakers of the two varieties have been involved in a high degree of intermixing. Over 50% of Rual residents are of mixed Jedek/Jahai-speaking parentage or are themselves part of a Jedek/Jahai-speaking parental pair. Thus while the Semang are known for their high levels of multilingualism, contact, and band exogamy, as described above, the Rual context is also strongly characterized by the intense intermixing of Jedek and Jahai speakers after more than 40 years of cohabitation and intermarrying.

With a speaker population of only around 280, Jedek is an endangered language variety. However, measuring the degree of its endangerment and vulnerability is not uncomplicated. As pointed out by Benjamin (2001, 2012), the Aslian language varieties spoken by the Semang foragers probably never had large numbers of speakers. Furthermore, the Semang communities have a long-standing tradition of maintaining their ethnic and linguistic identity in an environment defined by intense contact and constant social flux. Thus, to some extent, they are culturally primed to transmit their languages to the next generation. The exact number of Jedek speakers at the time of resettlement at Rual some 40 years ago is not known, but estimates based on Gomes (2007) suggest that it was smaller than today's figure. Increases in the Jedek

speaker population are indeed likely given the demographical effects of settlement and modernized healthcare. Jedek is currently spoken by all generations and in most domains at Rual, and transmission remains unbroken. Jedek is understood and/or spoken by the majority of Rual residents. At the same time, the Jedek speakers' resettlement at Rual – which entails permanent cohabitation with speakers of other Aslian language varieties, in particular Jahai – has left and continues to leave its mark on the variety. In addition, the Rual speech community is a small one surrounded by the Austronesian majority language Malay. Paid employment, where available, invariably involves communication in Malay, and all media and schooling is in Malay. Jedek is not officially recognized by any government departments and indeed its existence is unknown to officials, who define Rual as a Jahai resettlement area. No orthography is available to Jedek (or Jahai) speakers; however, there are some local attempts at improvised orthographical representation using Malay orthography. While attitudes of Jedek speakers toward their language variety are generally positive, the attitude of Malaysia's majority population toward indigenous groups is typically ignorant or dismissive, and at times hostile. In short, Jedek is endangered and vulnerable to influence in a variety of different ways. But this does not necessarily spell imminent extinction for the variety.

3 Typological outline

This section provides a first typological overview of Jedek. Aiming to give a broad description of the main grammatical features of the language, it also highlights features that make it typologically noteworthy in the Aslian context and beyond. The description begins with Jedek's phonological features (Section 3.1) before turning to aspects of word formation including derivational morphology and cliticization (Section 3.2), followed by nominal and verbal word classes (Section 3.3), and finally phrase and clause structure (Section 3.4). The description is based on data collected during fieldwork in Rual by co-author Yager between 2013 and 2016. Data collection involved stimulus-based elicitation and conventional elicitation of grammaticality judgments as well as collection of a corpus of recordings of natural language use. Many aspects of this structural analysis are still in their initial stages and the description is to be considered preliminary. Apart from illustrating the basic structural properties of Jedek, the description also serves to identify similarities and differences between Jedek and the other Northern Aslian varieties.

In order to facilitate this comparison, the outline is structured loosely on Kruspe et al. (2015).[5]

3.1 Phonology, phonotactics, syllabic structure

3.1.1 Phoneme inventory

The Jedek phoneme inventory is given in Table 1. The system of vowel qualities is distinguished by three degrees of height in front, central, and back positions. Nine oral vowels contrast with seven nasal vowels (contrasting, e.g., /wɛʔ/ 'be quiet!' and /wɛ̃ʔ/ 'leftside'; /paw/ 'side of body' and /pãw/ 'to be different'). Phonemic nasality is a feature of all Aslian languages but is almost unheard of elsewhere in Austroasiatic.

Table 1: Jedek vowel phonemes.

Oral			Nasal		
Front	**Central**	**Back**	**Front**	**Central**	**Back**
i	ɨ	u	ĩ	ɨ̃	ũ
e	ə	o		ə̃	
ɛ	a	ɔ	ɛ̃	ã	ɔ̃

The following set of words illustrates the contrast between the nine oral vowels in Jedek:

(1) *ctis* 'long time' *gɨs* 'to climb down' *gus* 'to be together'
get 'to cut' *gəs* 'to carve' *ros* 'liver'
gɛs 'gas' *hagas* 'mosquito' *gɔs* 'to live'

The back vowels are rounded while the front and central vowels are not. Height distinctions between the three central vowels are not unproblematic and the material contains cases that on the basis of auditory impressions might

5 The practical orthography adopted in the present work for the most part follows the phonemic representation. However, in line with the orthographic conventions of Burenhult (2006 and later), Kruspe (2004), Benjamin (1976), and Diffloth (1976), this orthography departs from standard IPA representation in that the voiced palatal stop /ɟ/ is written as <j> and the palatal approximant /j/ as <y>.

suggest a fourth central vowel height. However no unequivocal instances of minimal pairs distinguishing a fourth vowel height are found in the material and thus only three are posited at this stage. Vowels in the final syllable are subject to slight palatal diphthongization if followed by a palatal consonant /c, ɲ, s/ (as in *ɓc* 'arrow', phonetically [lɔ^{i}c̚]).

The Jedek consonants (see Table 2) follow the standard Northern Aslian pattern, including bilabial, alveolar, palatal, and velar voiced (/b/, /d/, /j/, /g/) and voiceless stops (/p/, /t/, /c/, /k/) and nasals (/m/, /n/, /ɲ/, /ŋ/), voiceless glottal stop (/ʔ/), voiceless bilabial (/ɸ/), palatal (/s/) and glottal (/h/) fricatives, bilabial (/w/) and palatal (/y/) approximants, the lateral liquid /l/, and the rhotic /r/ (the realization of which is subject to individual variation, see below). The voiceless bilabial fricative /ɸ/ is a phoneme in most Northern Aslian varieties but is otherwise exceedingly rare in Southeast Asia.

Table 2: Jedek consonant phonemes.

	Bilabial	Alveolar	Palatal	Velar	Glottal
Stop	p b	t d	c j	k g	ʔ
Nasal	m	n	ɲ	ŋ	
Fricative	ɸ		s		h
Lateral		l			
Rhotic		r			
Approximant	w		y		

Where unvoiced stops /p, t, c, k/ occur in final position they are unreleased, phonetically [p̚, t̚, c̚, k̚]. Nasal consonants are prestopped in word-final position when preceded by an oral vowel, phonetically [^{b}m, ^{d}n, ^{ɟ}ɲ, ^{g}ŋ], preventing anticipatory nasalization of non-nasal preceding vowels. Nasal consonants following nasal vowels are not prestopped. Furthermore, nasal consonants cause progressive nasalization of vowels – and nasal consonants following such progressively nasalized vowels are thus not prestopped. Thus for example *tanɛm* 'to plant' is phonetically [taˈnɛ̃m] and *knɔm* 'urine' is [kəˈnɔ̃m]. The material contains a handful of forms in which progressive nasalization does not occur. Some of these are Malay loanwords in which a consonant cluster of nasal + stop has been reduced in Jedek to a simple nasal, suggesting that the stop has left a trace that prevents nasalization from spreading to following vowels. For other, indigenous forms the explanation is less clear. Such forms are analysed (and represented orthographically) as containing an underlying consonant cluster whose realization is variable and at times indiscernible, as in *ʔndaŋ* [ʔəˈna^{g}ŋ ~ ʔənˈ^{d}a^{g}ŋ] 'side' or *lmbuʔ* [ləˈmuʔ ~ ləmˈ^{b}uʔ] from Malay *lembu* 'cow'.

The realization of /r/ is subject to individual variation. Some individuals realize /r/ as an apical trill [r], some as a uvular fricative [ʁ], and some speakers varyingly produce either realization. It is likely that this can be explained by a generational shift: younger Jedek speakers (those in their 40s and below) tend to use [r] while older Jedek speakers tend to use [ʁ] or a mixture of [ʁ] and [r]. Cohabitant Jahai realizes /r/ as a distinct apical trill in all positions, so it is likely that this shift has been brought on by the presence of a larger number of Jahai-speaking individuals and larger number of children of mixed Jedek/Jahai-speaking parentage in the Rual speech community since resettlement. The uvular realization considered as a peculiarity of Jedek by speakers of nearby languages (see Section 1) is a feature encountered in some other Northern Aslian varieties (e.g., Batek Deq) as well as local dialects of Malay, but it is not present in Jedek's closest neighbors Jahai and Menriq.

3.1.2 Phonotactics

As is typical in Aslian, the full range of Jedek vowels occur in word-final syllables while only a limited range of vowels may occur in non-final syllables: none of the nasal vowels, nor /ɔ/, /ɛ/, /ɨ/, or /ə/ occur (and /e/ and /o/ are rare) in non-final syllables. Most consonant phonemes occur in both initial and final position. Exceptions are the voiced stops /b, d, j, g/ which occur only in initial position, the rare bilabial fricative /ɸ/ which occurs only in final position, as well as /r/ which does not occur word-finally but may occur as coda in non-final syllables.

3.1.3 Word and syllable structure

The analysis of word and syllable structure is based on the citation forms of Jedek words. As is typical of Aslian languages, most lexemes in Jedek are monosyllabic, sesquisyllabic, or disyllabic. Trisyllabic forms also occur, but words longer than three syllables occur only as a result of affixation. The minimal Jedek word has the form /CV/ (such as *bɛ* 'younger sibling'), while the maximal words found in the material are the tetrasyllabic (C.C.CV.CVC) *b-plagaŋ* 'to be joking' (phonetically [bəpəlaˈgaᵍŋ]) and (C.C.CC.CVC) *b-k<n>rjaʔ* 'to have a job' (phonetically [bəkənərˈjᶻaʔ]).

Phonetically, the minimal Jedek syllable consists of a consonant plus a vowel [CV]σ. The maximal syllable has the form [CVC]σ, with a simple onset, nucleus, and coda. Thus onsets are obligatory while codas are not. While

phonetically the minimal syllable in Jedek includes a vowel nucleus, this nucleus may be predictable and underspecified. In such cases vowel nuclei are treated as epenthetic. Epenthetic vowels may occur in open (/C/) or closed (/CC/) syllables, thus phonemically the minimal syllable in Jedek consists of a consonant onset only, /C/. Such syllables are termed half syllables. Half syllables are allowed only in prefinal position – word-final syllables are obligatorily full syllables (/CV(C)/). A distinction is also made in Jedek between light and heavy syllables. Light syllables are those which do not have a coda: /C/ or /CV/, and heavy syllables are those with a coda: /CC/ or /CVC/. Prefinal syllables may be heavy or light and contain either a full phonemic or epenthetic vowel nucleus. Word-final syllables may be either heavy or light but always contain a full phonemic vowel nucleus.

In most cases the epenthetic vowel of half syllables is realized as [ə]. Epenthetic vowels followed by the palatal approximant /y/ are commonly realized as [i] (e.g., *kyəm* [kiˈjəᵇm] 'lower side'). Glottal consonants (/ʔ/ and /h/) are transparent to vowel assimilation and thus epenthetic vowels followed by glottal consonants may in some cases take on an identical quality to vowels in the following syllables (e.g., *cʔay* [cᵉaˈʔaj] 'what'). Epenthetic vowels occurring in heavy non-final syllables (e.g., in /CC.CVC/ words) are most commonly realized as [ə].

In Jedek, as in other Aslian varieties, sesquisyllabic words consist of a final syllable preceded by a half penultimate syllable /C.CV(C)/ as in *ksɨy* 'husband'. The category of sesquisyllabic words is justified on morphological grounds (see the nominalizing morpheme in Section 3.2.3 below, and cf. Burenhult 2005; Kruspe 2004). Disyllabic words consist of a final syllable preceded by a full penultimate syllable: /CV.CV(C)/ as in *baboʔ* 'woman', /CC.CV(C)/ as in *tmkal* 'man', or /CVC.CV(C)/ as in *kaltoŋ* 'knee'. Trisyllabic words follow the form of disyllabic words with the addition of an initial half syllable or full, open syllable: /C(V).CV.CV(C)/ as in *klabas* 'sun bear (*Helarctos malayanus*)', /C(V).CC.CV(C)/ as in *glmhɛ̃m* (meaning unknown), /C(V).CVC.CV(C)/ as in *cmaldɨk* 'to hiccup'.

As in some other Northern Aslian varieties (such as Maniq, Kensiw, and Ceq Wong), but in contrast to Jahai, Batek, and Menriq (those varieties most closely related to Jedek), open word-final syllables are allowed in Jedek. Such syllables are partly the result of a process whereby word-final /r/ has been lost leaving final syllables formerly with an /r/ coda open. The forms *bɛ* 'younger sibling' and *ha* 'road, path' are examples of this (compare with the Jahai equivalents *bɛr* and *har*). Open final syllables also occur in many Malay loans. That is, while a handful of Malay loans with open final syllables are subject to addition of word-final consonants (e.g., *dwaʔ* from *dua* 'two'), most retain their open final

syllables, e.g., *mejə* from *meja* 'table', *buku* from *buku* 'book', *baka* from *bakar* 'to bake, burn', *pike* from *pikir* 'to think'.[6] Occasional presumably indigenous forms that may not originally have had /r/-codas also occur, including the species terms *kasɔ* 'chestnut-winged babbler (*Stachyris erythroptera*)' and *tiwɔ* 'cream-colored giant squirrel (*Ratufa affinis*)', as well as the toponyms *Kte* (a river name) and (possibly) *Swɛ* (a river which in Malay is called Suih). Vowels in word-final open syllables are phonetically lengthened, e.g., *bɛ* [bɛː] 'younger sibling', *kasɔ* [kaˈsɔː] 'chestnut-winged babbler (*Stachyris erythroptera*)'.

3.1.4 Prosodic features

Stress falls invariably on the final syllable of words. Despite suggestions that some Northern Aslian varieties display marginal tonal contrasts (Hajek 2003), we have so far found no evidence of suprasegmental contrastive strategies in Jedek, neither tone nor register.

3.2 Word formation

This section provides a brief introduction to the units, processes, and functional categories of Jedek morphology. For definitions of Jedek's word classes – nouns, verbs, prepositions, pronouns, quantifiers, classifiers, demonstratives, adverbs, negators, connectives/conjunctions, and interjections – see Section 3.3. The principles of word formation described here are similar to those described for other Aslian languages, especially those of Jedek's close relatives Jahai and Menriq (cf. Burenhult 2005, field notes; Kruspe et al. 2015).

3.2.1 Morphological units

The morphological units of Jedek involve three kinds of free forms – roots, lexemes, and bases – and two types of bound morphemes – affixes and clitics. Roots are defined as morphologically unanalysable words. Lexemes are minimal free forms but do not necessarily consist of a bare root: some Jedek lexemes are morphologically complex and analysable into morphemes that do not exist in contemporary Jedek. Bases are defined as those units to which bound

6 Note that word-final /r/ in many Malay varieties is unrealized and thus *bakar* and *pikir* in fact have open final syllables.

morphemes may be added, and may consist of either roots or lexemes. A Jedek word is defined as a morphologically free form – it may consist of a single root, base, or lexeme, or it may be a compound or a form derived through affixation or cliticization. Following Klavans (1985), affixes and clitics are analysed as phonologically bound morphemes that differ in their domain of attachment. Whereas the domain of attachment of affixes is words, the domain of attachment of clitics is phrases or clauses. Affixes may be either prefixes or infixes, not suffixes, and clitics are always proclitics. Affixes attach to bases before clitics and a derived form can in turn act as a base for further derivational processes or for the attachment of clitics. Following the convention introduced by Kruspe (2004), clitics are represented in the orthography by the equals sign (=), prefixes with a hyphen (-), and infixes with angle brackets (<>).

3.2.2 Compounding

Jedek compounds are nominal and in many cases compositional. Compounds involve the combination of two nouns of which the initial noun forms the head, and are especially frequent in species names as well as topographical and body part terms. Another compound-like construction is the associative phrase, a syntactic construction described in Section 3.4.4. Essentially, the difference between compounds and associative phrases is that the meaning of a compound is not simply the sum of the meaning of its parts, as in *tɔm naʔ* [water mother] 'big stream' and *ʔɔʔ bintaŋ* [tiger star] 'leopard', while the meaning of an associative phrase is.

3.2.3 Derivational morphology

Jedek derivational morphology primarily functions to transfer words between word classes, as in the case of the nominalization morpheme operating on verbs and the property morpheme operating on nouns, or between subclasses within a word class, as in the collective plural and unitization morphemes within nominal word classes and the aspectual/Aktionsart derivations within verbal word classes. Causative constructions are the only clear example of valency-increasing operations, while progressive and imperfective constructions are often associated with decreased valency. Derivational morphology in Jedek and other Aslian languages is rich and mostly productive, which makes it typologically unusual in the wider Austroasiatic and Mainland Southeast Asian context (cf. Matisoff 2003: 22–23).

There are two kinds of affixation process in Jedek: inner affixation whose domain of attachment is the penultimate syllable of bases, and outer affixation whose domain of attachment is the base in its entirety (the distinction was first described for Jahai, see Burenhult (2005: 46–64)). Inner affixation may involve prefixes or infixes, while outer affixation always involves prefixes. Outer affixation may result in forms that do not conform to Jedek word-structure constraints observed in citation forms (such as the constraint prohibiting half penultimate syllables in trisyllabic forms, violated in, e.g., *p-bkɨt* 'to heat'). Inner affixation is more prevalent than outer affixation.

Affixes also differ in their degree of phonological prespecification. Inner affixes may be fully prespecified, partly prespecified, or fully underspecified while outer affixes are always fully prespecified. Where affixes are phonologically underspecified the segmental content of the affix is filled through copy of segments from the final syllable of the base, specifically the consonant onset and/or coda.[7] Fully prespecified affixes do not involve any copy (as in the nominalized *bdɛl* 'to shoot' → *b<n>dɛl* 'act of shooting') while fully underspecified affixes are formed wholly through copy (as in the imperfective *ci?* 'to eat starchy food' → *c?-ci?* 'to be eating starchy food'). Affixation may result in the creation of a new syllable, as in the case of outer affixation (e.g., the causative *kjeŋ* 'to hear/listen' → *p-kjeŋ* 'to cause to hear/listen'), or inner affixation on monosyllabic words (as in imperfective *cp-cɨp* 'to be going', distributive *cɨp-cɨp* 'to go (here and there)', or iterative *lp-cɨp* 'to go (all the time)' from *cɨp* 'to go'). Or it may result in the restructuring of an existing syllable, such as creating a full syllable from a half syllable (as in imperfective *kjeŋ* 'to hear' → *k<ŋ>jeŋ* 'to be hearing') or a heavy syllable from a light one (as in imperfective *tolɛk* 'to push' → *to<k>lɛk* 'to punch (here and there)'). There are two examples in Jedek of inner affixation that does not involve copy: the allomorph of the nominalizing morpheme used for sesqui- and disyllabic bases *<n>*, and the collective plural *<ra>*, infixed without the copy of any of the segments of the base. These are also the only examples of inner affixation that may result in a trisyllabic word, as in the nominalized *h<n>aluh* from *haluh* 'to shoot with blowpipe' and the collective plural form *b<r>abo?* from *babo?* 'woman'.

3.2.3.1 Deriving nouns

Jedek derivational processes that produce nouns include the nominalizing morpheme which derives nouns from verbs and numerals, and two relatively

7 The use of the term "copy" follows Kruspe (2004).

uncommon derivations that derive nouns from nouns: the collective plural and unitization derivations.

3.2.3.1.1 Deriving nouns from verbs. Nouns may be derived from verbs by means of the nominalizing (NMZ) morpheme *n-* ~ *<n>*. Nouns derived in this way have meanings relating to the act denoted in the verb. On monosyllabic bases, nominalizing *<n>* forms a CC prefix whose onset is *n-* and whose coda is an underspecified consonant filled through copy of the coda of the base. On sesqui- and disyllabic bases, *<n>* follows the onset of the penultimate syllable, in the case of sesquisyllabic bases becoming the coda of the penultimate syllable and in the case of disyllabic bases creating a trisyllabic word where *<n>* is the onset of the penultimate syllable.

(2) a. Monosyllabic:

ciʔ 'to eat starchy food'	→	*nʔ-ciʔ* 'act of eating starchy food'
cɨp 'to go'	→	*np-cɨp* 'act of going'
cɔl 'to speak'	→	*nl-cɔl* 'act of speaking'
kap 'to bite'	→	*np-kap* 'act of biting'

b. Sesquisyllabic:

bdɛl 'to shoot'	→	*b<n>dɛl* 'act of shooting'
tbɔh 'to beat'	→	*t<n>bɔh* 'act of beating'
ʔnay 'to bathe'	→	*n-ʔnay*[8] 'act of bathing'

c. Disyllabic:

ckwɨk 'to talk'	→	*c<n>kwɨk* 'act of talking'
kijəw 'to stand'	→	*k<n>ijəw* 'act of standing'
haluh 'to shoot'	→	*h<n>aluh* 'act of shooting'

3.2.3.1.2 Deriving nouns from numerals. The nominalizing morpheme may also be used on some numerals to form a noun referring to the state of being that number, as in *tigaʔ* 'three' → *t<n>igaʔ* 'state of being three'. The most common use of this affix is on the numeral *dwaʔ* 'two', its derived form *d<n>waʔ* meaning 'both', see (3).

(3) *d<n>waʔ d<n>waʔ ton srupaʔ blakaʔ*
two<NMZ>.two<NMZ> that to.be.the.same all
'They're both the same.'

8 The root is sesquisyllabic but has irregular derivation; the pattern is found in nominalized sesquisyllabic forms with a glottal initial.

3.2.3.1.3 Deriving nouns from nouns: Collective plural (COLL) and unitization (UNIT). The infix *<ra>* may be added to human nouns to form collective nouns. The allomorph *<r>* occurs with disyllabic bases with an open penultimate syllable and the allomorph *<a>* occurs with disyllabic bases with a closed penultimate syllable and an epenthetic vowel. Another collective plural form is that of *kɛn* 'child' as in *gin gr-kɛn* 'those kids'. This is the only instance of this allomorph of the collective plural that has been recorded. The following collective plural forms have been attested to date:

(4)	*kdah* 'young woman'	→	*k<ra>dah* 'young women'
	kjih 'young man'	→	*k<ra>jih* 'young men'
	babo? 'woman'	→	*b<r>abo?* 'women'
	tmkal 'man'	→	*tm<a>kal* 'men'
	bidan 'old person'	→	*b<r>idan* 'old people'
	mnra? '(indigenous) person'	→	*mn<a>ra?* '(indigenous) people'

Another rarely-used nominal derivation is the unitization morpheme *nC-*, primarily serving to turn mass nouns into count nouns in cases of reference to discrete units of the noun, exemplified in (5).

(5) *ʔiŋ rɛŋ dah wɔŋ ʔhəy, wɔŋ*
1SG to.eat.meat already to.be.little to.be.small to.be.little
s-nc-sɛc
one-UNIT-flesh
'I've only eaten one small piece (of meat).'

Unitization is a typologically unusual phenomenon which seems largely restricted to Aslian languages. The Jedek unitization morpheme has been recorded in only a very few contexts and does not give the impression of being as fully productive as its equivalent in Jahai, for example (Burenhult 2005: 75–77). The following unitized forms have been recorded in Jedek:

(6)	*sɛc* 'flesh'	→	*nc-sɛc* 'piece of flesh'
	ʔay 'meat animal'	→	*ny-ʔay* 'unit of meat animal'
	teʔ 'ground/earth'	→	*nʔ-teʔ* 'place/location'
	tɔm 'water'	→	*nm-tɔm* 'unit of water'
	can 'foot/leg'	→	*nn-can* 'unit of foot/leg'

3.2.3.2 Deriving verbs

3.2.3.2.1 Deriving verbs from nouns. The property morpheme *b-* (most likely borrowed from Malay *ber-* also used to derive verbs from nouns) may be added to nouns through outer affixation to form property verbs. Verbs derived in this manner denote notions of being 'characterized by' the referent of the base noun and may have vaguely possessive meanings. The use of the morpheme in positive contexts is restricted to a subset of nouns: it has been recorded primarily in connection with nouns characterized by inalienability such as body parts and children or spouses as in (7a), although other nouns may also receive the morpheme (7b). Meanwhile, in negated contexts the property morpheme is productive and may be used with any noun (7c).

(7) a. *ʔoʔ b-knih*
3SG PROP-wife
'He has a wife.'
b. *brapaʔ hariʔ, b-bulan bulan lɛh dah*
how.many day PROP-month.month EMPH EMPH
'How many days (did we camp)? For months and months!'
c. *blap b-swal blap b-bajuʔ ʔoʔ wãŋ*
to.not.exist PROP-pants to.not.exist PROP-shirt 3SG to.be.naked
'He has no pants, no shirt, he's naked.'

3.2.3.2.2 Deriving verbs from verbs. A number of Jedek affixes function to derive verbs from verbs: the aspectual/Aktionsart derivations, and the causative. Examples of the Jedek verbal derivations are given in Table 3. The imperfective and distributive derivations operate according to inner affixation while the causative and progressive derivations operate according to outer affixation. The iterative derivation operates according to inner affixation on monosyllabic bases and outer affixation on sesqui- and disyllabic bases. Many of the verbal derivations of Jedek are also present in Jahai (Burenhult 2005: 94).

3.2.3.2.2.1 Causative (CAUS). The causative morpheme derives transitive verbs with meanings of 'to cause to X'. The causative involves outer affixation and has three allomorphs (*p-*, *pi-*, and *pr-*). There is no evidence for semantic differences between the allomorphs. Rather, as is common in Aslian causative morphology, their use appears to primarily be determined by the syllabic structure of bases. In addition, the allomorphs have different degrees of productivity: the allomorph *pr-* has only been attested on a handful of (mostly monosyllabic) bases

Table 3: Jedek verb-to-verb derivations.

Derivational morpheme		Monosyllabic *ciʔ* 'to eat'; *cɨp* 'to go'	Sesquisyllabic *kjeŋ* 'to listen'	Disyllabic *haluh* 'to shoot with a blowpipe'
Causative		*p-ciʔ ~ pi-ciʔ ~ pr-ciʔ*	*p-kjeŋ ~ pi-kjeŋ*	*p-haluh*
Aspect/ Aktionsart	Imperfective	*cʔ-ciʔ*	*k<ŋ>jeŋ*	*to<k>lɛk* (*tolɛk* 'to push')
	Progressive	*b-cʔ-ciʔ*	*b-kjeŋ*	*b-haluh*
	Iterative	*lʔ-ciʔ*	*l-kjeŋ*	*l-haluh*
	Distributive	*cip-cɨp*	*k<iŋ>jeŋ*	*j<ip>wap* (*jawap* 'to answer'), *p<i>ʔjiʔ* (*pʔjiʔ* 'to sleep')

(8a), while *pi-* and *p-* are more productive. *Pi-* is the preferred allomorph on monosyllabic bases (8b) and is attested on a handful of sesqui- and disyllabic bases (8c), while *p-* is preferred and used exclusively on sesqui- and disyllabic bases.

(8) a. *ʔoʔ bəw* → *ʔom=pr-bəw kbɨʔ lɛh k=karɛy*
3SG to.be.big IRR.3SG=CAUS-to.be.big fruit EMPH SUB=Karei
'He's big.' 'Karei will make the fruit big.'

b. *ʔapay tek* → *ʔapay pi-tek wɔŋ ʔapay*
1PL.EXCL to.lie.down 1PL.EXCL CAUS-to.lie.down child 1PL.EXCL
'We were lying down.' 'We put our children to bed.'

c. *ʔiŋ ʔntɨŋ* → *ʔiŋ pi-ʔntɨŋ ʔoʔ*
1SG to.be.afraid 1SG CAUS-to.be.afraid 3SG
'I'm afraid' 'I'm scaring her (causing her to be afraid).'

3.2.3.2.2.2 Imperfective (IPFV) and progressive (PROG). The imperfective morpheme derives imperfective verbs which may in turn feed derivation with the progressive morpheme. Both are used to describe situations as ongoing, in progress, or habitual. While the progressive morpheme *b-* is fully productive and may be used on any verb, imperfective-marked verbs without the progressive morpheme are relatively rare in the material. The imperfective morpheme is an example of fully underspecified inner affixation – its segments are filled entirely through copy of segments of the base. On monosyllabic bases both onset and coda of the base are copied and attached to the base as a prefix (as in *jok* 'to move around' → *jk-jok* 'to be moving around'). On sesquisyllabic bases and disyllabic bases with an open penultimate syllable, the coda of the

penultimate syllable is filled through copy of the coda of the final syllable of the base (as in *pgɛŋ* 'to hold' → *p<ŋ>gɛŋ* 'to be holding', *tolɛk* 'to push' → *to<k>lɛk* 'to be pushing'). The progressive *b-* (most likely a borrowed form of the Malay prefix *ber-* with a similar function) operates according to outer affixation. The progressive may attach to either the imperfective or bare root form of sesqui- and disyllabic verbs (with no apparent difference in meaning), but monosyllabic verbs must be marked with the imperfective or, more rarely, distributive, iterative, or causative morphemes in order to receive progressive marking.

Any subtle differences in meaning between the imperfective and progressive morphemes are difficult to discern: both may express notions of ongoingness (9a) or habituality (9b) and may be used to convey the ongoingness of a certain situation in relation to other events (9c). As a result of this as well as the paucity of imperfective forms in the material, the imperfective and progressive morphemes are considered here to be semantically equivalent. The imperfective morpheme is also present in several fossilized forms whose morphologically unanalysable root form does not exist in contemporary Jedek (such as *lclɛc* 'to be wrong', *plpɛl* 'to drip', *ʔmʔəm ~ kmʔəm* 'to hug').

(9) a. *ja=Nin lɛh d=ʔoʔ d=ʔoʔ bʔ-baʔ ton*
RT=Nin EMPH CONTR=3SG CONTR=3SG IPFV-to.carry that
'also Nin, the one that was being carried'

b. *ʔapay jk-jok lɛh d=k=sɛŋ kaduy*
1PL.EXCL IPFV-to.move.around EMPH CONTR=LOC=front PSTDIST
'We used to move around in the old days.'

c. *hiʔ pʔjiʔ dah d=hiʔ, ʔoʔ b-cl-cɔl*
1PL.INCL to.sleep already CONTR=1PL.INCL 3SG PROG-IPFV-to.speak
lagiʔ d=ʔoʔ
still CONTR=3SG
'We were already asleep, but he kept on telling (the story).'

3.2.3.2.2.3 Distributive (DISTR). The Jedek distributive morpheme operates on verbs to express non-temporal multiplicity of a situation. Such situations may involve multiple individuals as subject, as in (10a), multiple locations or directions, as in (10b, c), or other features involving multiplicity. The distributive is often used in contexts involving reciprocity among participants, as in (10a), but is also often found in non-reciprocal contexts involving multiple participants as well as contexts involving a single participant but multiple locations, as in (10b, c). Its functions are very similar to those described for Jahai (Burenhult 2011). The distributive morpheme is partially

prespecified, involving the prespecified vowel *i*. On monosyllabic bases the prespecified *i* forms the nucleus of a new penultimate syllable whose onset and coda are filled through copy of the onset and coda of the base, as in *cɨp* 'to go' → *cip-cɨp* 'to go (here and there)'. On sesquisyllabic bases the coda of the final syllable of the base is copied and becomes the coda of the penultimate syllable, with the prespecified *i* as its nucleus, as in *bdɛl* 'to shoot' → *b<il>dɛl* 'to shoot (here and there)'. On disyllabic bases with an open penultimate syllable the coda is also filled through copy of the coda of the final syllable of the base, while the nucleus is in most cases replaced with the prespecified *i*, as in *bagiʔ* 'to give' → *b<iʔ>giʔ* 'to give (here and there)'. There are no examples in the material of distributive derivation on disyllabic forms with a closed penultimate syllable.

(10) a. *wih miy-may kuy*
3DU DISTR-to.delouse head
'Those two are delousing each other.'[9]

b. *ʔoʔ b<il>dɛl pãw pãw tmpət dah lɛh*
3SG to.shoot<DISTR> different.different place EMPH EMPH
'He shot all around.'

c. *ʔoʔ ʔil-ʔɛl da=tũn da=tadeh*
3SG DISTR-to.look GOAL=there GOAL=here
'S/he looked around over there, over here.'

3.2.3.2.2.4 Iterative (ITER). The Jedek iterative morpheme signals temporal multiplicity of an action, usually involving multiple repetition of a complete action on a single occasion, as in (11a, b). It may also be used to signal multiplicity over separate occasions, as in (11c), but this use is more rare in the material. The iterative morpheme involves the prespecified onset *l* and operates according to two different morphological processes depending on the syllabic structure of the base to which it attaches. On sesqui- and disyllabic bases it operates according to the process of outer affixation, with the prespecified *l* attaching to the left edge of the base, as in *kdɨh* 'to say' → *l-kdɨh* 'to say repeatedly' and *haluh* 'to shoot with a blowpipe' → *l-haluh* 'to shoot repeatedly with a blowpipe'. On monosyllabic bases it operates according to inner

9 Description obtained during elicitation by means of the "Reciprocal constructions and situation type" task (Evans et al. 2004), a video stimulus kit designed to probe linguistic expressions of reciprocity.

affixation, the prespecified *l* forming the onset of a new penultimate syllable and the coda being filled through copy of the coda of the base, as in *ʔɛl* 'to see/look' → *ll-ʔɛl* 'to look repeatedly'.

(11) a. *ʔoʔ cirit s-ʔomo ʔomo ja=ʔoʔ lc-ʔɛc*
3SG to.have.diarrhea always.always RT=3SG ITER-to.defecate
'S/he is pooping all the time (because of diarrhea), s/he poops and poops.'
b. *bɛʔ l-ʔaŋket bɛʔ l-tulis*
2SG ITER-to.get 2SG ITER-to.write
'You pick up (your notebook) and write all the time.'
c. *ʔoʔ lp-cip ʔoʔ lk-ŋɔk s-miŋgu s-kaliʔ*
3SG ITER-to.go 3SG ITER-to.sit one-week one-time
'She always goes there and stays a week at a time.'

There are many contexts in the material in which the iterative morpheme appears to have a kind of imperative function, as in (12a, b). The semantic connection between such contexts and the more straightforwardly iterative contexts in the material is at this stage unclear.

(12) a. *pɛy pɛy lk-wek, lk-wek ʔujan*
BECK BECK ITER-to.return ITER-to.return rain
'Come, come, come home, come home, it's raining.'
b. *lp-cip l-kdih, l-kdih ll-cɔl da=Yati, ʔiɲ*
ITER-to.go ITER-to.say ITER-to.say ITER-to.speak GOAL=Yati 1SG
ma=cip da=hip
IRR=to.go GOAL=forest
'Go and tell Yati that I'm going to the forest.'

3.2.4 Clitics

Like affixes, clitics are distinct from words in that they are bound forms – they exist only attached to a base and do not appear as free forms, and they cannot receive stress. Jedek clitics have a [C], [CV], or [CVC] structure. Due to their bound status, final vowels of [CV] clitics are not lengthened and nasal codas of [CVC] clitics are not prestopped. Like outer affixes, clitics do not have different allomorphs that depend on the structure of the base to which they attach, but appear in the same form in all contexts.

Clitics differ from affixes, however, in that their domain of attachment is phrases or clauses rather than words. Clitics are invariably proclitics and attach to the left edge of a base within the phrase or clause that forms the domain of their attachment. The different clitics differ in terms of the kinds of constituents which may function as hosts to which they attach. The contrastive proclitic *d=* (described in Section 3.3.8) may be hosted by any constituent. The irrealis proclitics *ma= ~ na=* and *ʔom=* are hosted by verbs (see Section 3.3.11), and the imperative proclitics *ca=~ ka=* and *ha=* are hosted only by verbs in imperative form (see Section 3.3.11). Prepositional proclitics (see Section 3.3.8) are hosted by the first constituent of noun phrases. The relational tense proclitic *ja=* (see Section 3.3.15) and question proclitic *ha=* (see Section 3.3.13) are hosted by the first constituent of a clause.

3.2.5 Full reduplication

Full reduplication of lexemes occurs in the case of Jedek adverbs and interrogative pronouns (forming indefinite pronouns). Fully reduplicated forms do not conform to Jedek word constraints or word-level stress patterns and are thus not interpreted as the result of a morphological process forming a word-like unit. It is likely that the process is a borrowed form of a similar process in Malay.

3.3 Word classes

Jedek has the distinct open word classes of nouns (Section 3.3.1) and verbs (Section 3.3.10) and closed classes of prepositions (Section 3.3.8), pronouns (Section 3.3.2), quantifiers (Section 3.3.3), classifiers (Section 3.3.4), demonstratives (Section 3.3.6), adverbs (Section 3.3.15), negators (Section 3.3.12), connectives/conjunctions (Section 3.3.14), and interjections.

3.3.1 Nouns

Jedek nouns denote concrete or abstract concepts and occur as part of noun phrases, either as noun phrase heads or as modifiers of other nouns. Nouns may be modified by pronouns, demonstratives, quantifying expressions, another noun, or a relative clause, and may be marked with the derivational categories applied to nouns that are described in Section 3.2.3.

3.3.2 Pronouns and question words

3.3.2.1 Personal pronouns

Jedek personal pronouns distinguish singular, dual, and plural number in the 1st, 2nd, and 3rd person distinctions. The Jedek personal pronoun system is untypical among the Northern Aslian varieties in the existence of two separate forms distinguishing 2nd and 3rd person plural, a distinction which is otherwise collapsed in all Northern Aslian varieties except Ceq Wong (Kruspe et al. 2015). The system is otherwise typically Aslian in that gender is not marked, and 1st person dual and plural pronouns are marked for inclusion/exclusion.

Table 4: Jedek personal pronouns.

	Singular			Dual		Plural	
1	*ʔiɲ*			Inclusive *hɛy*	Exclusive *yɛh*	Inclusive *hiʔ*	Exclusive *ʔapay*
2	General *bɛʔ*	Familial *mɔh*	Friendship *bɔʔ*	*jɨh*		*smpay*	
3	*ʔoʔ*			*wih*		*gin*	

Three 2nd person singular forms have been recorded. Of these, *bɛʔ* is the most general and widely used, used to address interlocutors ranging from complete strangers to one's friends and acquaintances, spouse, parents, siblings, and other family members for which use of in-law avoidance pronouns (described below) is not prescribed. One's own children or the children of one's siblings may be addressed with the pronoun *mɔh*, which may also be used by children to address younger children to whom they are related. However *bɛʔ* may also be used in these contexts. The less frequent *bɔʔ* is used optionally and occasionally between close friends who are not related by kinship.

Jedek pronouns occur both as unstressed subject-markers on verbs (obligatory on dynamic verbs in realis clauses) and as noun phrase constituents (either heads or modifiers) which may receive stress. The 3rd person singular irrealis form *ʔom=* (see Section 3.3.11) may also be argued to have pronominal status, since unlike the other irrealis proclitics it it does not co-occur with the subject-marking 3SG pronoun. The form is a proclitic and is used only as an unstressed preverbal marker, not as part of a noun phrase.

3.3.2.2 In-law avoidance pronouns

As is a common feature of Aslian varieties, Jedek has a set of pronouns used specifically to make reference to and address affinal kin with whom interaction is restricted by sets of rules. Interaction with an opposite gender parent-in-law is especially restricted, but a set of rules also governs the nature of interactions between brothers- and sisters-in-law. The forms of the Jedek in-law pronouns are taken from the 2nd and 3rd person dual and plural forms of the personal pronoun system. Parents- and children-in-law address and refer to each other with the form *gin* while siblings-in-law address one another using the pronoun *jih*, and refer to one another using *wih*. Jedek and other Aslian in-law pronoun systems are a typologically unusual form of affine avoidance register, not to be confused with the more well-known honorific registers of the Southeast Asian area (Kruspe & Burenhult, submitted; cf. Fleming 2014).

3.3.2.3 Interrogative pronouns

Jedek has a number of indigenous and borrowed question words to question thing, reason, person, place, manner, time, quantity, and identity. The system is especially rich in forms, the use of which appears to be subject to within- and between-speaker variation. The forms of the Jedek question words are given below:

(13)
a.	Thing/Reason	*cʔay, cbap, cbap ʔay, baʔay*
b.	Person	*makɛn*
c.	Place	*ʔnah* (*~ pãn, ~ gɛl, ~ lɛŋ, ~ ʔirah, ~ cãn*)
d.	Manner	*mancin ~ maʔancin*
e.	Time	*ʔnah pyan, bilaʔ*
f.	Quantity	*brapaʔ*
g.	Identity	*ʔnah ʔoʔ teʔ*

The identity-questioning *ʔnah ʔoʔ teʔ* (14a) is untypical among the Northern Aslian varieties, none of which are reported as possessing a distinct term for questioning the identity of a referent. Specifically, it questions the identity of a thing, person, or place among multiple competing referents, akin to English 'which'. Person-questioning *makɛn* also questions possessor, as in *makɛn ʔasuʔ* [who dog] 'whose dog?'. The place-questioning *ʔnah* is combined with the mostly transparent *pãn, gɛl, lɛŋ, ʔirah*, and *cãn* to form place-questioning phrases. The five terms differ as to the location and direction of movement questioned. *ʔnah pãn* is the most commonly used and may question either location or direction (14b) of movement. *ʔnah lɛŋ* (14c) and *ʔnah gɛl* question

the location of a referent, *ʔnah ʔirah* questions the goal of movement, and *ʔnah cãn* questions the source of movement. Location-questioning *ʔnah lɛŋ* contains the preposition *lɛŋ* 'at/in/on', *ʔnah gɛl* contains the noun *gɛl* 'middle portion' (and cf. Menriq *gɛl* 'at/in/on'), *ʔnah cãn* contains the form also functioning as source-marking preposition *can=*, and *ʔnah ʔirah* contains the relational noun *ʔirah* 'side'. Complex WH-constructions of this type are otherwise so far only attested in Ceq Wong (cf. *han halaʔ* 'to/from where' and *han mataʔ* 'on where?') among Aslian varieties, among which location-questioning typically involves prepositional phrases (see, e.g., Kruspe 2004: 187–188 for Semelai). The thing-questioning forms (two of which, *cʔay* and *cbap*, are also used to question reason) are also somewhat analysable: as a combination of synchronically non-existing prefixes and two of the food class nouns: *bap* 'starchy food' and *ʔay* 'meat'.[10] Person-questioning *makɛn* contains the noun *kɛn* 'child' and identity-questioning *ʔnah ʔoʔ teʔ* is constructed using place-questioning *ʔnah*, 3rd person singular pronoun *ʔoʔ*, and noun *teʔ* 'earth'.

(14) a. *ʔnah ʔoʔ teʔ cnɛl ʔiŋ ma=cɔl*
which origin.being 1SG IRR=to.speak
'Which origin being story should I tell?'
b. *ʔnah pãn bɛʔ ma=cip*
where.goal 2SG IRR=to.go
'Where are you going?' [a common greeting]
c. *ʔnah lɛŋ smpay ma=gorɛŋ*
where.location 2PL IRR=to.fry
'Where are you going to fry it?'

3.3.2.4 Indefinite pronouns

Indefinite pronouns are formed through reduplication of question words: *cʔay cʔay*, 'whatever' (15a), *ʔnah ʔnah ʔoʔ teʔ* 'whichever', *ʔnah ʔnah pãn* 'wherever (location or goal)' (15b), *ʔnah ʔnah cãn* 'wherever (source)', *ʔnah ʔnah gɛl* and *ʔnah ʔnah lɛŋ* 'wherever (location)', and *brapaʔ brapaʔ* 'however many'.

(15) a. *hɛy haguʔ cʔay cʔay ja=ʔoʔ gɔ̃ŋ*
1DU.INCL to.request whatever RT=3SG to.withold
'We request whatever and s/he refuses to give it.'

10 The form *ʔay* also occurs in interrogatives of several other Aslian languages (in most of which it also means meat), cf. Ceq Wong *cɔʔ ʔay* 'what' and *biʔ ʔay* 'who', Batek *ʔay ləw* 'what', and Semaq Beri *mʔay* 'what' (Nicole Kruspe, personal communication).

b. *gin lɔy ʔnah ʔnah pãn*
3PL to.run wherever
'They ran away to wherever.'

3.3.3 Quantifiers

Quantifiers are most commonly found prenominally as modifiers of nouns or classifiers (17a). They may also form a phrasal head on their own (17b). Jedek numerals, see (16), are, with the exception of the numeral 'one', all borrowed from Malay. In addition, the borrowed Malay form *s-* from Malay *se-* 'one' may be used as a quantifying affix. Its use is especially common with Malay loans, as in (17c). The numeral *nay* 'one' is also used as a quantifier with delimiting function, meaning 'only, just' (17d), often accompanied by relativizing *d=* which attaches to the left edge of the element introduced by *nay*.

(16)

1	*nay*	6	*ʔnam*	11	*s-blas*
2	*dwaʔ*	7	*tujuh*	12	*dwaʔ blas*
3	*tigaʔ*	8	*lapan*	100	*s-ratos*
4	*ʔmpət*	9	*smilan*	1000	*s-ribuʔ*
5	*limaʔ*	10	*s-puloh*		

(17) a. *dwaʔ hariʔ da=tkɨh*
two day GOAL=behind
'two days ago'

b. *ʔapay jok ʔiɲ baʔ dwaʔ dwaʔ*
1PL.EXCL to.move.around 1SG to.carry.child two.two
'We moved around, I carried two children.'

c. *s-jam, dwaʔ jam*
one-hour two hour
'one hour, two hours'

d. *ʔoʔ ltʔet nay basaʔ hamiʔ*
3SG to.know only language Malay
'He only understands Malay.'

Jedek has three additional quantifiers: indigenous *kɔm* 'many/much' and *pãw* 'other' along with the Malay loan *blakaʔ* 'all' (from Malay *belaka* 'entire'). As described above for quantifiers in general, these quantifiers may form a phrasal head on their own (18a) or occur as prenominal modifiers of heads (18b).

(18) a. *gin k<i>mʔəm blakaʔ*
3PL to.hug<DISTR> all
'All of them are hugging each other.'
b. *ja=pãw ktɔʔ wel*
RT=other day again
'another day'

3.3.4 Classifiers

Classifiers have a relatively marginal status in Jedek. Two forms are found in the data, both borrowed from Malay: *ʔoraŋ* from Malay *orang* 'person' used for human referents, as in (19a), and *ʔekɔ* from Malay *ekor* 'tail' used for animals, as in (19b). These borrowed forms are not found as ordinary nouns in the data – indigenous Jedek equivalents *mnraʔ* 'person' and *hatɛ̃ʔ* 'tail' are used elsewhere. Classifiers are modified by numerals, forming a noun phrase with the function of specifying the number of some referent. Classifiers do not co-occur with ordinary nouns in a noun phrase but always replace them as head. Classifiers are used in combination with numerals but are not attested with the other quantifiers described above.

(19) a. *gin k<i>mʔəm gin tigaʔ ʔoraŋ*
3PL to.hug<DISTR> 3PL three CLF
'They are hugging each other, those three.'
b. *cnɛl planok, ʔoʔ [s-ʔekɔ] s-ny-ʔay*
origin.being mouse.deer 3SG one-CLF one-UNIT-meat.animal
planok nay
mouse.deer one
'The origin mouse deer, it was one mouse deer.'

3.3.5 Proper nouns: Personal names

As is common among Aslian speech communities, individual Jedek speakers typically have several names. Most children are given two names at birth: an indigenous name, and a Malay name which becomes the child's registered name and is used in interactions with outsiders. Within Rual, Jedek speakers are referred to variously by their indigenous or Malay name. After the birth of their first child, adults are most commonly known by the name of their firstborn child, as in *Naʔ Jila*, 'Jila's mum' (a practice referred to as teknonymy which also

exists in many other Aslian-speaking communities, e.g., Temiar, Semelai, Semaq Beri, Batek, and Maniq; Kruspe et al. (2015: 453); Nicole Kruspe, personal communication; Ewelina Wnuk, personal communication). Grandparents are commonly known by the name of their firstborn grandchild, as in *Yaʔ Mira* 'Mira's grandmother'. In addition to this, informal and often humorous nicknaming is common, in some cases with the knowledge of the nicknamed individual (as in *Taʔ Spɛk* 'grandpa spectacles'), in other cases not (as in *Taʔ ʔŋjulən* 'grandpa ambulance', so named after his use of a flashing head torch). Often a single individual will be known by several different nicknames, given by different people. Thus each individual is typically known by at least four different names throughout their lifetime, with the addition of any (and in some cases many) nicknames.

3.3.6 Demonstratives

3.3.6.1 Spatial demonstratives

Jedek has a single set of nine spatial demonstratives that are used both nominally and adverbially, shown in (20). Four demonstratives encode distance and accessibility distinctions (*ʔũh, ton, taniʔ*, and *tũn*), two encode elevation (*titɨh* and *tuyih*), and two encode exteriority in relation to the speech dyad (*tadeh ~ tudeh ~ tudɛʔ* and *ʈŋɨʔ*, see Burenhult (2005: 84–87, 2008, in press) for in-depth discussion of the Jahai equivalents); one demonstrative denotes referents predominantly perceived with senses other than vision (such as heard or smelled, *cɨn*).

(20)	*ʔũh*	accessible and usually proximal to speaker
	ton	accessible to addressee, attention confirmer
	taniʔ	inaccessible to and usually distant from speaker
	tũn	inaccessible to addressee, attention drawer
	tadeh ~ tudeh ~ tudɛʔ	exterior: outside speaker's side of speech dyad
	ʈŋɨʔ	exterior: outside addressee's side of speech dyad
	titɨh	superjacent to speech situation
	tuyih	subjacent to speech situation
	cɨn	perceived through its emissions (e.g., heard or smelled)

In their nominal function the spatial demonstratives are used either pronominally and then represent the heads of noun phrases, as in (21a), or adnominally and then follow a head noun, as in (21b).

(21) a. *cʔay d=tũn, ha=gulə*
what CONTR=that Q=candy
'What's that? Candy?'
b. *ma=cɨp da=taniʔ hagas kɔm*
IRR=to.go GOAL=there mosquito many
'If we go over there there's lots of mosquitoes.'

In their adverbial function they occur as part of locative prepositional phrases headed by locative prepositional proclitics (described in Section 3.3.8). The functional distinctions are akin to those documented for Jahai and Menriq, but unlike these languages Jedek does not have parallel sets of nominal and adverbial demonstrative forms.

The forms *tadeh, tudeh*, and *tudɛʔ* are all found in the material. *Tadeh* is the most commonly used of the three; *tudeh* is a relatively common variant; *tudɛʔ* occurs only rarely. All three are analysed as expressing the speaker-exterior dimension but they have yet to be explored systematically. Addressee-anchored *tũn* and *ton* additionally encode pragmatic meanings related to the addressee's attention relative to the referent. While *tũn* is used for introducing a new referent in discourse or drawing the addressee's attention to a referent *ton* denotes a referent known by the addressee (cf. Burenhult (2003, in press) for the Jahai equivalents).

3.3.6.2 Temporal demonstratives

Jedek has three demonstrative-like forms which express temporal meanings: *kaʔũn, kaduy*, and *hkɨt*. They may be used either adverbially (22a) or adnominally (22b). While *kaduy* refers to the more distant past, *kaʔũn* refers to the more recent past. *Hkɨt* refers specifically to the previous day, 'yesterday'.

(22) a. *cbap bɛʔ jim kaʔũn*
why 2SG to.cry PSTPROX
'Why were you crying just now?'
b. *buku hkɨt*
book yesterday
'the book from yesterday'

3.3.7 Relational nouns

Another set of nominal forms with locational meaning is what are treated here – following the terminology of Kruspe et al. (2015: 463) – as relational nouns. The

full set of recorded relational nouns is: *ʔatɛs* 'upside', *kyəm* 'downside', *daləm* 'inside', *ɦip* 'outside', *sɛŋ* 'front', *krəʔ* 'back', *ʔndaŋ* and *ʔirah* 'side', and *ditɛp* 'other side'. These nouns refer to spatial zones and may form part of a prepositional phrase headed by prepositional proclitics expressing goal (*da=*), source (*can=*), and location (*lɛŋ, la=*, and *k=*), as in (23a). They may also appear without these prepositional proclitics, as the head of an associative phrase, where they denote a spatial area in relation to another nominal referent, as in (23b).

(23) a. *mɔh hɔk da=sɛŋ, mɔh hɔk da=ʔatɛs, mɔh hɔk can=kyəm*
2SG to.throw GOAL=front 2SG to.throw GOAL=upside 2SG to.throw SOURCE=under
'You throw it forward, you throw it up above, you throw it under.' [on how to fish]

b. *wih ŋik-ŋɔk wih cil-cɔl, ʔatɛs mejə*
3DU DISTR-to.sit 3DU DISTR-to.speak upside table
'They are sitting and talking, on a table.'

3.3.8 Prepositions

Jedek has a number of prepositions that express location (*lɛŋ, la=*, *k=*), goal (*da=*), source (*can=*), comitative (*lɔʔ ~ ʔalɔʔ*), instrument/subject/object (*k=*), contrast (*d=*), and similarity (*laguʔ*). The prepositions combine with noun phrases to form a prepositional phrase. Many of these prepositions are proclitics which attach to the left edge of noun phrases. Others are free morphemes: the comitative preposition *lɔʔ ~ ʔalɔʔ*, the semblative *laguʔ*, and the general location-marking *lɛŋ*. The remnant of a semblative proclitic form *man=* can be found in a small set of words, most of them derived from demonstratives, but is unattested outside of these contexts.[11]

Da= marks the goal, *can=* the source, and *lɛŋ, la=*, and *k=* the location of an action, situation, or referent. The goal- and source-marking prepositions may mark concrete or abstract goals and sources, as in (24a). Location-marking *k=* is rare and is used only in combination with demonstratives and relational nouns. It may also be added to the left edge of constructions with locative *lɛŋ* and *la=*, apparently without any change in meaning, see (24b).

11 The forms attested so far are *manton* 'like that', *mantadeh* 'like that', *mancin* 'like this', and *mantūn* 'like this', derived from demonstratives, as well as *mancin* manner-questioning 'how?'. The manner-questioning *mancin* bears some resemblance to Ceq Wong manner-questioning *cin* (Nicole Kruspe, personal communication).

(24) a. *ʔoʔ cɨp da=hɨp da=cbaʔ, ja=ʔoʔ wek da=hayãʔ*
3SG to.go GOAL=forest GOAL=hill RT=3SG to.return GOAL=house
'He went to the forest, to the hills, then he returned home.'
b. *ja=d=ʔoʔ ton ʔoʔ pʔjɨʔ k=la=hayãʔ*
RT=CONTR=3SG that 3SG to.sleep LOC=LOC=house
'And that guy, he was sleeping at home.'

The instrument/subject/object preposition *k=* occurs in three kinds of context. As instrument marker, *k=* is used to mark an instrument with which the activity described in the verb is executed (25a). As subject marker, *k=* may be used to mark full subject arguments of the verb but is not obligatory (25b). Subject-marking *k=* is most commonly found on postverbal subject arguments, and its use on subject arguments to the left of the verb is rare in the material. The argument marked by subject-marking *k=* may be a stressed version of the pronoun that forms the subject marker on the verb (see Section 3.3.2 above), or may function to further specify the referent of the subject marker. *K=* may also introduce a direct object argument, turning the object into a partially-affected object at which the action is directed rather than implying that the entire object is affected (25c).[12] The construction is akin to the use of English *at* in contexts such as *eat at* and *hit at*.

(25) a. *ʔoʔ cek k=mataʔ gajah kaʔũn lɛh*
3SG to.stab INSTR=spear elephant PSTPROX EMPH
'He stabbed that elephant with a spear.'
b. *ʔoʔ ʔɔ k=ciʔgu*
3SG to.order SUB=teacher
'The teacher asked us to.'
c. *ja=bah ha=ciʔ k=nasiʔ*
RT=to.go.to.a.place Q=to.eat.starchy.food OBJ=rice
'Go and eat some rice.'

The prepositional proclitic *d=* attaches to the left edge of an argument in order to mark a focus on its referent, most often to express contrast to some other possible referent.[13] It is most commonly found on subject arguments but may also be used on object arguments. As with subject-marking *k=*, subject arguments marked with contrastive *d=* may consist of a stressed version of the

12 Jahai *k=* and Ceq Wong *kaʔ* are also used in this way.
13 The label "contrastive" (CONTR) is known from previous Aslian grammars. Despite the name, however, note that this proclitic does not always overtly mark a contrast with another referent.

pronoun that forms the subject marker on the verb, as in (26a, b), or may function to further specify its referent, as in (26c). Another function of contrastive *d=* is to mark ownership or change of ownership of a referent, as in (26d). In this context it is attached to an argument whose referent possesses or comes to possess something.

(26) a. *ʔiŋ ma=wek d=ʔiŋ*
1SG IRR=to.return CONTR=1SG
'Me, I want to go home.'
b. *d=mɔh ha=mɔh dapɛt dah?*
CONTR=2SG Q=2SG to.meet already
'Have you ever encountered (a tiger)?'
c. *ʔnah pãn ʔoʔ cip dɛn d=ʔabi kaʔũn?*
where 3SG to.go that.one CONTR=Bi PSTPROX
'Where did that Bi go?'
d. *weʔ d=ʔiŋ dah b<n>ulaŋ*
to.exist CONTR=1SG already headdress
'I've already got a headdress.'

3.3.9 Coordinating morphemes

A relative marker *d=* attaches to the left edge of the first word of a phrase or clause that is embedded within a noun phrase, creating a relative clause that modifies the head of the noun phrase, as in (27a, b). Relative clause markers of this kind are attested in other Aslian languages – see, e.g., Burenhult (2005: 122–126) for Jahai exponents – but only Jedek is recorded using the form *d=*. It is also found in two further kinds of context in the material. It may attach to the left edge of a property verb or adverb, as in (27c). It may also be used to mark direct quotes. In this context it attaches to the left edge of the phrases *ʔoʔ cɔl* or *ʔoʔ kdɨh* 's/he says' to specify that the preceding content is a direct quote, as in (27d).

(27) a. *ʔnah pãn cklaʔ d=ʔiŋ bliʔ kaʔũn*
where candy REL=1SG to.buy PSTPROX
'Where's the candy that I bought before?'
b. *ʔoʔ d=bəw liy*
3SG REL=to.be.big body
'the fat one'

c. *ciʔ d=ʔabən*
to.eat REL=good
'Eat carefully!'
d. *naʔ ʔiŋ ja=ʔiŋ rɛŋ lɛh d=ʔoʔ cɔl*
mother 1SG RT=1SG to.eat.meat EMPH REL=3SG to.speak
'I ate my mother, he said.'

3.3.10 Verbs

Jedek verbs denote actions and states, and they are defined as words that may take irrealis marking as well as the verbal derivations presented in Section 3.2.3. Jedek verbs function most commonly as predicates. The most important distinction within the class of verbs is that between dynamic and stative verbs.

Dynamic verbs denote actions or processes, such as *bay* 'to dig' and *gɔs* 'to live'. While dynamic verbs denote situations that involve some form of internal change, stative verbs do not. Stative verbs are distinct from dynamic verbs also in their morphological and syntactic characteristics. While dynamic verbs are obligatorily marked with preverbal subject- or irrealis-marking (except in the case of imperative constructions, see Section 3.3.11), this is not obligatory for stative verbs (28a). Stative verbs may also be used as adverbial modifiers of other verbs (28b), and be used attributively in their root form, as in *jkɔp bəw* [snake to.be.big] 'a big snake'.

(28) a. *taʔ ʔapɔh, weʔ makɔw ja=blap ma=lbɨt*
NEG anything to.exist tobacco RT=NEG IRR=to.be.tired
'It's ok, I've got tobacco so I won't get tired.'
b. *ʔɛl gnɔʔ, ʔɛl d=ʔabən buku ʔoʔ ʔoʔ sayɛŋ*
to.look to.be.quiet to.look REL=to.be.good book 3SG 3SG to.love
'Look quietly, look carefully at her book, it's dear to her.'

Meanwhile, dynamic verbs require nominalization (see Section 3.2.3) if they are to be used attributively, as in *baju ʔ p<n>ʔiiʔ* [shirt to.sleep<NMZ>] 'pajamas' – most likely a calque from Malay *baju tidur*. This feature of stative verbs is shared with Ceq Wong and Maniq (Kruspe et al. (2015); see Wnuk (2016) for in-depth discussion of the two verb types in Maniq). Stative verbs refer to states and include verbs of existence and non-existence (*weʔ* 'to exist' and *blap* 'to not exist', which also functions as a negator; see Section 3.3.12), and the adjective-like property verbs which denote properties of referents, such as *ʔhəy* 'to be small' and *mŋjiʔ* 'to be far'.

The citation form of dynamic verbs most often consists of the verb plus preverbal subject marking, typically the 3rd person singular, as in *ʔoʔ lwec* [3SG to.climb] 'to climb'. Irrealis-marked forms (see Section 3.3.11.1 below) may also be used as citation forms of dynamic verbs, as in *ma=cɨp* [IRR=to.go] 'to go'. For stative verbs, the unmarked form (without irrealis proclitics or preverbal subject-marking pronouns; see Section 3.3.2) is used as citation form (as in *bəw* 'to be big').

3.3.11 Markers of modality

In Jedek, two kinds of proclitics attach to verbs to express modality: the irrealis and the imperative proclitics. While their domain of attachment is the entire clause, their phonological host is always a verb.

3.3.11.1 Irrealis

Jedek has three irrealis proclitics which vaguely encode subject: *ma=* is used for all persons except 3rd person singular, for which the form *ʔom=* is used; *na=* is a less common allomorph of *ma=* variably used only in the case of 1st person singular.[14] The form *ʔom=* is likely a merging of the 3rd person singular pronoun *ʔoʔ* and a reduced form *-m*, thought to have originated in the irrealis form *ma=*.[15] It might be argued that *ʔom=* has a pronominal status: while *ma=* and *na=* may be preceded by subject-marking pronouns as in (29a, b), *ʔom=* may not. This further suggests that the form originated as a merging involving the pronoun *ʔoʔ*. The use of irrealis proclitics is obligatory in negative clauses, as in (29a). In other contexts they are used to signal the intention that an action/situation should happen, such as in (29b, c). An illustrative example of both uses of irrealis proclitics is seen in (29c).

14 *Na=* may have roots in Malay *nak*, used to express intended situations (Mintz 1994: 73).

15 The Jedek irrealis displays features documented in the diverse irrealis paradigms of some other Aslian languages. It shares with Jahai the abstract categorical distinction between 3SG and all non-3SG, although the Jahai forms are different (*wa=* vs. *ya=*, respectively; cf. Burenhult 2005: 110–112). The 3SG irrealis form also bears some resemblance to the Menriq and Batek paradigms, whose irrealis constructions involve systematic replacement of the final consonants of all pronouns with *-m*, much like the 3SG irrealis form in Jedek (Burenhult, field notes; cf. also Temiar irrealis pronouns (Benjamin 2016)). Furthermore, a proclitic *ma=* is the irrealis marker also in Semelai, a distantly related language of the Southern Aslian branch (Kruspe 2004: 161–163).

(29) a. *blap bɛʔ ma=ʔɛl da=mɛt ʔoʔ*
NEG 2SG IRR=to.look GOAL=face 3SG
'You didn't look at her face.'
b. *ʔiɲ na=wek na=pʔiiʔ*
1SG IRR.1SG=to.return IRR.1SG=to.sleep
'I'm going/want to go home and sleep.'
c. *ʔoʔ ʔɔm ʔom=makɔʔ tapiʔ blap*
3SG to.want IRR.3SG=to.be.pregnant but NEG
ʔom=makɔʔ
IRR.3SG=to.be.pregnant
'She wants to get pregnant but she's not.'

3.3.11.2 Imperative proclitics

Imperative clauses in Jedek are formed with the imperative form of verbs. The imperative form is unmarked, i.e., it is not marked by subject-marking pronouns or irrealis proclitics. Three proclitics with imperative function (*ka=*, *ca=*, and *ha=*) may optionally be added to verbs to express meanings ranging from commands (30a) to suggestions (30b). *Ka=* and *ca=* are equivalent and are placed at the left edge of a verb in imperative form, and *ha=* is attached to the second and subsequent verbs of a sequence of verbs with imperative function (30b). Invitational and beckoning particles *lah* 'come on!' and *pɛy* 'come here!' also have imperative functions. While *lah* is used in contexts where both speaker and addressee are to embark on the movement together (30c), *pɛy* is used in contexts where the interlocutor is beckoned to move toward the speaker (30d). The verb *bah* 'to go to a place' is also commonly used in imperative contexts instructing the interlocutor to move in a direction away from the speaker (as in (25c) above).

(30) a. *ka=ŋɔk ba=ʔaniʔ baliʔ*
IMP=to.sit GOAL=there CURSE
'Sit over there!' [a mother's irritated command to her daughter]
b. *ja=wek ha=ʔnay*
RT=to.return IMP=to.bathe
'Go home and bathe.' [a polite response to the greeting that one is on one's way home to bathe]
c. *lah hɛy cip hɛy bliʔ lɛh*
INV 1DU.INCL to.go 1DU.INCL to.buy EMPH
'Come on, let's go and buy (groceries)!'

d. *pɛy da=ʔũh*
BECK GOAL=here
'Come here!'

3.3.12 Negators

Jedek has a number of negators which attach to the left edge of an irrealis-marked (see Section 3.3.11.1) verb. The dominant Jedek negator is *blap*, but six additional negators have been recorded: *braʔ, ckɔʔ, makaʔ, ʔayaŋ*, and the borrowed *bukan* and *takan*,[16] as well as the prohibitive *ʔakɛt*. *ʔayaŋ* and *bukan* are typically used as nominal negators and tend to be more common in contexts where it is the identity of the negated referent that is in focus, as in (31b), but may also be used in the same way as the other negators. *Takan* tends not to be used in contexts referring to situations in the past; otherwise, no apparent semantic differences have been found between the different negators. Apart from when they are used nominally (in the case of *bukan* and *takan*), negators are placed at the left edge an irrealis-marked verb, as in (31a). Prohibitive marker *ʔakɛt* is placed at the left edge of verbs in imperative form. The form *blap* also functions as the stative verb 'to not exist'. The Malay negator *tak ~ tidak* has been borrowed as part of some common expressions such as *taʔ ʔapɔh* [NEG anything] 'it doesn't matter; it's o.k.' and *taʔ prnah* [NEG ever] 'never'.

(31) a. *blap ʔapay ma=pakɛy bajuʔ*
NEG 1PL.EXCL IRR=to.wear shirt
'We didn't wear t-shirts.'
b. *bukan pŋjɔʔ ʔiŋ blaʔ*
NEG sarong 1SG self
'It's not my own sarong (it was given to me by a friend).'

3.3.13 Interrogative

Polar questions are formed in Jedek by placing the interrogative proclitic *ha=* at the left edge of a clause, see (32).

(32) *ha=bɛʔ ma=ŋɔk can=ton d=bɛʔ*
Q=2SG IRR=to.sit SOURCE=there CONTR=2SG
'Do you intend to sit there?'

16 Cf. Malay *takkan*, a reduced form of *tidak akan* 'will not'.

3.3.14 Conjunctions/connective words

A number of conjunctions are used to connect and relate phrases or clauses to one another. One coordinating conjunction, the comitative preposition *lɔʔ ~ ʔalɔʔ* 'with/and' acts as an additive conjunction, typically used to coordinate noun phrases, as in (33a). Eight subordinating conjunctions, all borrowed from Malay, are also used: the causal *pasəl* (33b) (from Malay *pasal* 'on the subject of, about, concerning'), *malum* (from Malay *maalum* 'known, grasped, understood'), *dapun* (from Malay *ada-pun* 'now'), and *sbap* (from Malay *sebab* 'because'), conditional *ʔamun* (from Malay *amun* 'provided that'), *misaʔ* (possibly from Malay *miseh/maseh* 'yet, still, continuously'), and *kaluʔ* (from Malay *kalau*, 'if'), and the temporal *lpəs* (from Malay *lepas* 'after'). The four causal conjunctions appear to be used interchangeably, as do the three conditional conjunctions – the forms cannot be semantically differentiated at this point.

(33) a. *ʔaŋket makɔw lɔʔ haliʔ mɛʔ*
to.get tobacco and leaf mum
'Fetch mummy's tobacco and rolling-leaves.'

b. *kɔm hagas pasəl ʔoʔ hltuh k=haliʔ*
many mosquito because 3SG to.fall SUB=leaf
'There's lots of mosquitos because the leaves have fallen (from the trees).'

3.3.15 Auxiliaries and adverbs

Jedek has a number of elements that modify verbs and clauses, including a number of adverbs (Section 3.3.15.1), relational tense marker *ja=* (Section 3.3.15.2), a number of negators (see Section 3.3.12 above), interrogative proclitic *ha=* (see Section 3.3.13 above), and the emphatic particle *lɛh* (Section 3.3.15.3).

3.3.15.1 Adverbs

Jedek has a number of indigenous and borrowed adverbs. Postverbal adverbs include the indigenous *hayeʔ* 'also', *sʔoʔ* 'just, only', *wel* 'again', and *sut* 'always' and the borrowed *sʔomo* 'always' (from Malay *seumur* 'whole, entire'), *lagiʔ* 'more, still, yet' (from Malay *lagi* 'more, again, still'), *trus* 'straight away' (from Malay *terus* 'straight away'), and *dah* (from Malay *sudah* 'already'). Two

adverbs typically occur clause-initially: the borrowed *saja?* 'just' (from Malay *saja* 'just, only') and *baru?* 'newly, just now' (from Malay *baru* 'new, anew'), see (34a). The adverb *s?o?* 'just, only' is often used in polite commands or invitations, as in (34b). Property verbs may also be used to modify other verbs, thus functioning as adverbs, most commonly occurring postverbally, as in (34c).

(34) a. *baru? ?o? wek*
newly 3SG to.return
'She's just gotten home.'
b. *tutup s?o? hãŋ*
to.close only mouth
'Close your mouth.'
c. *?o? ci? nasi? lagi? ton, ?om=ci?*
3SG to.eat.starchy.food rice still that IRR.3SG=to.eat.starchy.food
s?o? hakɨy
only to.be.slow
'She's still eating rice, she's just eating.' [of someone we were waiting for]

3.3.15.2 Relational tense (RT)

The proclitic *ja=* may be attached to the left edge of a verb to express temporal or consequential meanings. In its temporal sense it is used to denote that the situation referred to in the clause is temporally either anterior (35a) or posterior (35b) to the present time or some other time referred to in the utterance. In anterior contexts *ja=* is often used interchangeably with or in combination with the adverb *dah* (35a). Used in negative contexts, *dah* denotes that a situation is no longer the case, akin to English *anymore*, as in (35c). Where *dah* is combined with irrealis proclitics (see Section 3.3.11.1) it denotes that something is or was about to happen, as in (35d).

In its consequential sense the relational tense proclitic is used to introduce a result, as in (35e). It is also used as a discourse connector, as in (35f), and may be attached to a verb in imperative form as part of commands, as seen in (25c) and (30b) above.

(35) a. *tani? ja=?o? pɨh dah Yati*
there RT=3SG to.wake.up already Yati
'There, Yati's woken up now.'
b. *pukul dwa? blas satu ?nuy ja=?iŋ p?ji?*
o'clock twelve one later RT=1SG to.sleep
'At twelve, one o'clock I'll have a nap.'

c. *blap ma=lbɨt dah ja=ʔiɲ pʔjiʔ dah kaʔũn*
NEG IRR=to.be.tired already RT=1SG to.sleep already PSTPROX
'I'm not tired anymore, I had a nap before.'
d. *ʔom=kiʔ dah*
IRR.3SG=to.vomit already
'She was close to vomiting.'
e. *kaluʔ ʔoʔ ʔɔ k=ciʔgu ʔapay ja=ʔapay*
if 3SG to.request SUB=teacher 1PL.EXCL RT=1PL.EXCL
mlawat
to.vacation
'If our teacher tells us to we'll go on a trip.'
f. *mɔh jok ja=mɔh ŋɔk la=ton, ja=mɔh*
2SG to.move.around RT=2SG to.sit LOC=there RT=3SG
b-cp-cɨp
PROG-IPFV-to.go
'You move around, then you stay there, then you walk.'

3.3.15.3 Emphasis

The particle *lɛh*, which may be placed after any element, has the function of adding emphasis. The form originates in the Malay emphatic particle *lah*. The adverb *dah* also takes on an emphatic function when used in combination with the emphatic particle *lɛh*, as in (36).

(36) *hã lɛh ma=cɔl nay ʔoʔ cɨn lɛh dah*
yes EMPH IRR=to.speak only 3SG this EMPH EMPH
'Yes, I'll just tell this one (story).'

3.4 Phrase and clause structure

3.4.1 Basic clauses

Every Jedek clause contains a verb (with the exception of non-verbal clauses, see Section 3.4.2 below). Dynamic verbs (see Section 3.3.10 above) are obligatorily marked with a preverbal subject-marking pronoun, as in (37a), which may be replaced by or combined with irrealis proclitics, as in (37b). Stative verbs (see Section 3.3.10 above) are optionally marked with subject-marking pronouns or irrealis proclitics, (37c). In imperative clauses (described in Section 3.3.11), verbs do not receive subject marking. The order of Jedek constituents is relatively flexible: full subject and object arguments may be placed to the left or the

right of the verb, however their placement to the right of the verb is more common. Considering this and the fact that preverbal subject markers and irrealis proclitics (which make up an obligatory part of dynamic clauses) are invariably preverbal, Jedek may generally be characterized as following a SVO pattern.

(37) a. *ʔoʔ cɨp*
3SG to.go
'S/he goes.'
b. *ʔiŋ ma=wek dah*
1SG IRR=to.return already
'I'm going home.'
c. *bəw liy ʔoʔ*
to.be.big body 3SG
'His body's big.'

In addition to verbs, a clause may contain noun phrases or prepositional phrases representing full subject or object arguments. Subjects may be specified by a noun phrase preceding or following the verb, or by a prepositional phrase following the verb introduced by subject-marker *k=* or contrastive *d=*, as in (38a). Noun phrases representing direct objects may follow the verb, or, less commonly, precede it. Prepositional phrases representing direct or indirect objects may follow the verb. Where a clause contains both a direct and an indirect object the direct object typically precedes the indirect object. Ellipsis is widespread: omission of either direct or indirect objects, or both, is common and no verb requires either type of object, see (38b). In addition to the verb and its arguments, a clause may contain adjuncts: elements of a clause that are not arguments of the verb (this terminology follows Burenhult (2005)). Adjuncts most commonly occur to the right of the verb and its arguments, as in (38c).

(38) a. *ʔoʔ ton blap ma=rɛŋ d=hiʔ*
3SG that NEG IRR=to.eat.meat CONTR=1PL.INCL
'We don't eat that (animal).'
b. *ʔoʔ ʔɛk*
3SG to.give
'S/he gives (something to someone).'
c. *gin bacaʔ hkit*
3PL to.read yesterday
'They read yesterday.'

3.4.2 Non-verbal clauses

Non-verbal clauses do not contain a verb, but rather consist of two noun phrases representing topic and comment. Semantically, non-verbal clauses may be identifying/presentative, as in (39), ascriptive, or locative. As in some other Northern Aslian varieties, but in contrast to Jahai (cf. Burenhult 2005: 125), Jedek does not employ an identificational phrase-coordination marker in this kind of clause.

(39) *ʔoʔ tũn hayãʔ p<n>ʔjɨʔ*
3SG that house to.sleep<NMZ>
'That one's a sleeping-house.'

3.4.3 Complex clauses

Clauses may be combined in Jedek through the use of clause chaining, relative clauses, or complement clauses. In clause chaining, two or more independent clauses are simply placed one after the other, see (40a). Clauses are relativized with relative marker *d=* which attaches to the left edge of the initial constituent of an embedded phrase or clause (see Section 3.3.9 above). A complement clause is a clause which is used as an argument of a verb. A restricted set of verbs may take a complement clause as an argument, examples of verbs recorded so far in this context are the verbs *ʔɔm* 'to want', *yeʔ* or *ʔalah* 'to refuse', *sdəp* 'to be tasty', *susah* or *payah* 'to be difficult', *ʔnsol* 'to be shy', *laŋkay* 'to be lazy', *ʔntɨŋ* 'to be afraid', *pandɛy* 'to be good at X', and *boleh* 'to be able'. Complement clauses do not receive any special marking: a clause acting as a complement clause has the same structure as when it is used elsewhere, see (40b). The complement of a negative or negated verb receives irrealis marking, as in (40c).

(40) a. *wih ŋik-ŋɔk wih cil-cɔl, ʔatɛs mejə*
3DU DISTR-to.sit 3DU DISTR-to.speak upside table
'Those two are sitting and talking, on a table.'
b. *mŋjiʔ pon ʔoʔ pandɛy ʔoʔ coŋ*
to.be.far also 3SG to.be.good.at 3SG to.follow
'Even afar, he was good at following us.'
c. *ʔiŋ yeʔ ma=ʔãm*
1SG to.refuse IRR=to.drink
'I don't want to drink.'

3.4.4 Noun phrases

Noun phrases consist of a nominal head (which may involve a simple nominal, a compound, or a nominalized verb) plus optional modifiers including pronominal determiners, quantifiers, nominal modifiers, deictic determiners (demonstratives or prepositional phrases), and relative clauses. These elements are arranged around the head noun in the following order, as in (42a, b):

(41) (DET:PRO) (QNT) **N** (NOM MOD) (REL) (DET:DEI)

(42) a. *wih* *dwaʔ* *tm<a>kal* *ton*
3DU two man<COLL> that
DET:PRO QNT N DET:DEI
'those two men there'

b. *gin* *cnɛl* *gajah* *d=bəw bəw* *ton*
3PL origin.being elephant REL=to.be.big.to.be.big that
DET:PRO N NOM MOD REL DET:DEI
'those elephant origin beings that are big'

Where a nominal head is modified by a nominal modifier, they form an "associative phrase" (cf. Kruspe 2004; Kruspe et al. 2015: 447). Such constructions express a range of functions including part/whole relationships as in *tʔaʔ hobiʔ* [green.leafy.vegetable casava] 'casava leaves', object/purpose relationships as in *bajuʔ p<n>ʔjiʔ* [shirt to.sleep<NMZ>] 'pajamas', and possession relationships as in *wɔŋ bɛ ʔiɲ* [child younger.sibling 1SG] 'my younger sibling's child', among other functions.

4 Notes on the lexicon

As mentioned in Section 1, the initial survey uncovered examples of lexicon not recorded among Jedek's immediate neighbors and relatives within the Menraq-Batek branch of Northern Aslian. Some of this vocabulary has its closest equivalents in the Maniq varieties and/or Ceq Wong and Batek Nong, geographically and genealogically more distant Northern Aslian relatives with which Jedek has no documented history of contact (recall Jedek *ʔiɲ* '1SG', *piɲ* 'ashes', *ʔɔʔ* 'tiger', which have cognates in Maniq). The ensuing data analysis has added to this list a number of further candidates, such as *tiwɔ* 'cream-colored giant squirrel (*Ratufa affinis*)' (Maniq *ciwɔ* denoting the same species; Ewelina Wnuk, personal communication), *ʔayaŋ* 'NEG' (Maniq and Ceq Wong *ʔayaŋ* 'not', also recorded in

Batek Deq; Kruspe et al. 2015: 465), and *haguʔ* 'to request' (Ceq Wong *hagoʔ* 'to ask'; Nicole Kruspe, personal communication). Furthermore, like the Maniq varieties and Ceq Wong, Jedek has retained the reflex of the proto-Aslian term for blood in the form of *bhĩm* 'blood' (Maniq and Kensiw *mhĩm*, Ceq Wong and Batek Nong *mhə̃m*; proto-Aslian form rendered as **maha:m* by Sidwell & Rau (2015: 256)). The reflex does occur elsewhere in the Menraq-Batek branch but is then limited to certain registers (e.g., myths) or has a restricted meaning, such as Jahai *bhĩm* 'menstrual discharge'.[17] Historical contact and lexical exchange with the Maniq varieties and Ceq Wong should not be ruled out, but the conservative nature of some of this vocabulary rather suggests that Jedek has retained some terms and form-meaning mappings from earlier stages of Northern Aslian which are now lost among other members of the Menraq-Batek branch.

Jedek also displays terminology or form-meaning mappings which are so far unknown elsewhere in Aslian. One conspicuous example is the word *kmɔc* 'to die'. Cognates exist in other Aslian languages but with different, typically non-verbal meanings, e.g., Jahai *kmuc* 'ghost' (Burenhult 2005: 174), Jah Hut *kmɔc* 'ghost' (Diffloth 1976: 76), Semnam *kmuuc* 'tiger' (Burenhult & Wegener 2009: 304), Semaq Beri *kmuc* 'deceased person' (Nicole Kruspe, personal communication), and Mah Meri *kəmuc* 'burial site, grave' (Kruspe 2010: 90). Other examples include the form *blap* as negator and negative existential verb 'to not exist' (corresponding in form but not function to, e.g., the Jahai adverbial *blap* 'only'), *btɛh* as a basic term for 'red', *raŋah* 'hornbill species', *karɛ̃s* 'gaur, wild ox (*Bos gaurus*)', and *put* 'porcupine species'.

Turning to basic vocabulary shared with Jedek's three closest neighbors and relatives Jahai, Menriq, and Batek Deq, we find examples of uniquely shared items with each of these languages. For example, Jedek shares *jʔɔm* 'to be dirty' and *ʔabən* 'to be good' only with Menriq; *ʔntɨŋ* 'to be afraid' and *toʔ* 'elder sibling' only with Batek Deq; and *sagup* 'cloud' and *pɨs* 'to wipe' only with Jahai. Another noteworthy term shared only with Batek Deq is *klabas* or *tlabas* 'sun bear (*Helarctos malayanus*)', cf. Jahai and Menriq *kawɨp*. These correspondences seem to suggest that the intermediate status of Jedek within Menraq-Batek which is evident in parts of the grammar (e.g., irrealis constructions, Section 3.3.11) is also manifested lexically.

Several other basic meanings are represented by identical forms in Jedek, Menriq, and Batek Deq but not in Jahai: *kawaw* 'bird' (Jahai *kawɔ̃t*), *ʔasuʔ* 'dog' (Jahai *ʔɔ̃t*), *ciʔ* 'to eat (starchy food)' (Jahai *gey*), *makɔʔ* 'egg' (Jahai *ktɨt*), *ʔikan* 'fish' (from Malay *ikan*, cf. Jahai *ʔikəʔ*), and *ʔasɛp* 'smoke' (from Malay *asap*, cf. Jahai *ʔyʔey*). This association is in agreement with a pattern of vowel

17 Our Menriq basic vocabulary data differ from Benjamin's (1976: 103) in this respect. He gives the form *bəhəm* 'blood' for the variety he collected.

correspondences (outlined in Section 5 below) which unites Jedek and Menriq and to some extent Batek, but not Jahai, and may hint at a closer historical relationship with Menriq and Batek.

5 Notes on comparative phonology

There has been no attempt at an historical reconstruction of the Northern Aslian languages apart from brief preliminary notes by Diffloth (1975: 2–6) and Phillips (2012: 199), so it is problematic to situate our current data in a comparative context. Moreover, extensive lexical exchange between the Northern Aslian varieties has led to a dizzying blend of inherited and borrowed vocabulary which complicates any quest for systematic phonological correspondences. As far as Jedek and its closest relatives are concerned, evident patterns of comparison have so far only been identified among a subset of the vowels. These suggest a closer relationship to Menriq than to either Jahai or Batek. Occasional examples suggest a specific Jedek sound change *ɔ* → *ə* where the other three languages have retained *ɔ*. Table 5 illustrates these correspondences.

Table 5: Vowel correspondences in Menraq-Batek languages.

	Jedek	Menriq	Batek Deq	Jahai
'1PL.INCL'	*hiʔ*	*hiʔ*	*heʔ*	*heʔ*
'bone'	*jʔiŋ*	*jʔiŋ*	—	*jʔeŋ*
'to make'	*diʔ*	*diʔ*	*deʔ*	*deʔ*
'thorn'	*jliʔ*	*jliʔ*	*jliʔ*	*jleʔ*
'tongue'	*lntik*	*lntik*	*lntīk*	*lntek*
'butterfly'	*tawãk*	*tawãk*	*tawãk*	*tawẽk*
'breast'	*ʔãm*	*ʔãm*	*ʔãm*	*ʔẽm*
'to eat vegetables'	*hãw*	*hãw*	*hãw*	*hẽw*
'tooth'	*hãɲ*	*hãɲ*	*hãɲ*	*hẽɲ*
'to squeeze'	*cpẽt*	*cpẽt*	*cpīt*	*cpīt*
'left side'	*wẽʔ*	*wẽʔ*	—	*wīʔ*
'to rain'	*hẽc*	*hẽc*	*hãc*	*hīc*
'tail'	*hatẽʔ*	*hatẽʔ*	*hacẽʔ*	*hatīʔ*
'to be short'	*cnhẽt*	*cnhẽt*	*cnhãt*	*cnhãt*
'other'	*pãw*	*pãw*	*pɔ̃w*	*pẽw*
'back'	*krəʔ*	*krɔʔ*	*krɔʔ*	*krɔʔ*
'underside'	*kyəm*	*kyɔm*	*kyɔm*	*kyɔm*

6 Conclusions

The above sections have introduced for the first time a previously unidentified Aslian variety, proposed here to receive the scientific label Jedek, and provided an outline of its grammatical and lexical characteristics. The study is an important one, being as it is only the second description of a variety of the Menraq-Batek group of the Northern Aslian subbranch. Previously we had detailed knowledge only of Jahai from this group and did not know to what extent its structural features were common across the group. A number of typological features of Jedek are indeed shared with Jahai as well as other described languages of the Aslian branch of Austroasiatic. For example, Jedek's phonemic distinctions, its rich and productive derivational paradigms and processes, as well as elaborate deictic classes, are shared by its Aslian relatives. However, a range of phenomena revealed by this study are not shared by Jahai, and several features of Jedek are either undocumented elsewhere in Aslian or attested only in distantly related Aslian languages. These Jedek-specific features are found at all formal levels of language, from phonetics and phonology to morphological paradigms and processes, as well as in the syntax. Unlike its closest Northern Aslian relatives, Jedek allows open final syllables. It makes pronominal distinctions which are not shared by close Northern Aslian relatives and, conversely, it has a formally less elaborate demonstrative paradigm in comparison to those same relatives. Furthermore, the Jedek paradigms of interrogatives and irrealis forms lack direct parallels in other Aslian languages, and its strategies of negation and argument-marking appear to be distinct. Jedek's lexicon is still largely unexplored, but on the basis of forms collected so far it is clear that it harbors basic terminology and principles of lexicalization which are either unrecorded elsewhere, or present only in more distantly related languages. This points to an element of lexical retention and conservatism not encountered in Jedek's close relatives. Our previous phylogenetic analyses of basic vocabulary, as well as our preliminary attempts here to map sound correspondences, similarly provide indications of a separate historical signal.

Like other unidentified languages, Jedek bears witness to the existence of not only undocumented but also entirely unrecognized linguistic diversity. It also reminds us of the existence of urgent but undiagnosed cases of endangerment. Linguistic surveying was critical to the discovery of Jedek. Although not typically a prioritized aspect of language documentation funding initiatives, surveying is clearly fundamental to the galvanization and regeneration of the documentation enterprise and to maximizing informed future coverage of the poorly charted corners of the world of languages. In the case of Jedek, brief surveying has

brought about an unexpected opportunity to deepen our understanding of the linguistic history of the Malay Peninsula, the typological diversity of Aslian, Austroasiatic, and Southeast Asia at large, as well as the dynamics of language use and maintenance in highly multilingual small-scale speech communities.

Acknowledgements: We wish to express our gratitude first and foremost to the community at Sungai Rual. We also wish to acknowledge the Economic Planning Unit, Putrajaya; the Department of Orang Asli Development, Kuala Lumpur; Prof. Dr. Kamila Ghazali, University of Malaya; and Prof. Dr. Rogayah A. Razak, Universiti Kebangsaan Malaysia. We are grateful to Geoffrey Benjamin, Marianne Gullberg, Nicole Kruspe, Asifa Majid, Ewelina Wnuk, and an anonymous reviewer for their insightful comments on earlier drafts of this article, and to Jens Larsson for preparing the images. Joanne Yager's research was supported by the Joint Faculties of Humanities and Theology at Lund University, and her field-based data collection was carried out with support from the Lars Hierta Memorial Foundation, The Birgit Rausing Language Programme, Fil. Dr. Uno Otterstedt Foundation, and The Knut and Alice Wallenberg Foundation. Niclas Burenhult's research was supported by a Volkswagen Foundation DOBES grant and a Bank of Sweden Tercentenary Foundation Jubilee Initiative Grant NHS14-1665:1.

Abbreviations: 1/2/3 = 1st/2nd/3rd person; BECK = beckoning particle; CAUS = causative; CLF = classifier; COLL = collective plural; CURSE = cursing word; DEI = deictic; DET = determiner; DISTR = distributive; DU = dual; EMPH = emphatic particle; EXCL = exclusive; CONTR = contrast-marking preposition; GOAL = goal-marking preposition; IMP = imperative; INCL = inclusive; INSTR = instrument-marking; INV = invitational particle; IPFV = imperfective; IRR = irrealis; ITER = iterative; LOC = location-marking preposition; N = noun; NEG = negator; NMZ = nominalizer; NOM MOD = nominal modifier; OBJ = object-marking preposition; PL = plural; PRO = pronominal; PROG = progressive; PROP = property; PSTDIST = distant past; PSTPROX = recent past; Q = question proclitic; QNT = quantifier; REL = relative; RT = relational tense; SG = singular; SOURCE = source-marking preposition; SUB = subject-marking preposition; UNIT = unitization.

References

Anderson, Gregory D. S. & Ganesh Murmu. 2010. Preliminary notes on Koro: A 'hidden' language of Arunachal Pradesh. *Indian Linguistics* 71. 1–32.

Babaev, Kirill. 2010. *Zialo: The newly-discovered Mande language of Guinea*. München: Lincom Europa.

Benjamin, Geoffrey. 1976. Austroasiatic subgroupings and prehistory in the Malay Peninsula. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds.), *Austroasiatic studies, Part I*, 37–128. Honolulu: University of Hawai'i Press.

Benjamin, Geoffrey. 1985. In the long term: Three themes in Malayan cultural ecology. In Karl L. Hutterer, A. Terry Rambo & George Lovelace (eds.), *Cultural values and human ecology in Southeast Asia*, 219–278. Ann Arbor, MI: Center for South and Southeast Asian Studies.

Benjamin, Geoffrey. 1987. Ethnohistorical perspectives on Kelantan's prehistory. In Nik Hassan Shuhaimi bin Nik Abdul Raman (ed.), *Kelantan zaman awal: Kajian arkeologi dan sejarah di Malaysia*, 108–153. Kota Bharu: Perpaduan Muzium Negeri Kelantan.

Benjamin, Geoffrey. 2001. Orang Asli languages: From heritage to death? In Razha Rashid & Wazir Jahan Karim (eds.), *Minority cultures of Peninsular Malaysia: Survivals of indigenous heritage*, 101–122. Penang: Academy of Social Sciences.

Benjamin, Geoffrey. 2012. The Aslian languages of Malaysia and Thailand: An assessment. *Language Documentation and Description* 11. 136–230.

Benjamin, Geoffrey. 2016. A new outline of Temiar grammar, Part 1. Manuscript. https://www.researchgate.net/publication/291246946_A_new_outline_of_Temiar_grammar_Part_1

Bishop, Nancy M. & Mary M. Peterson. 1993. *Maniq language survey report*. Bangkok: Thammasat University.

Bishop, Nancy M. & Mary M. Peterson. 2003. Northern Aslian language survey: Trang, Satul and Phatthalung provinces, Thailand. Unpublished report. Bangkok: Thammasat University.

Burenhult, Niclas. 2003. Attention, accessibility and the addressee: The case of the Jahai demonstrative *ton*. *Pragmatics* 13. 363–379.

Burenhult, Niclas. 2005. *A grammar of Jahai* (Pacific Linguistics 566). Canberra: Australian National University.

Burenhult, Niclas. 2006. Body part terms in Jahai. *Language Sciences* 28. 162–180.

Burenhult, Niclas. 2008. Spatial coordinate systems in demonstrative meaning. *Linguistic Typology* 12. 142–199.

Burenhult, Niclas. 2011. The coding of reciprocal events in Jahai. In Nick Evans, Alice Gaby, Stephen C. Levinson & Asifa Majid (eds.), *Reciprocals and semantic typology*, 163–176. Amsterdam: Benjamins.

Burenhult, Niclas. In press. The Jahai multi-term demonstrative system – what's spatial about it? In Stephen C. Levinson, Sarah Cutfield, Michael Dunn, Nick Enfield, Sérgio Meira & David Wilkins (eds.), *Demonstratives in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Burenhult, Niclas, Nicole Kruspe & Michael Dunn. 2011. Language history and culture groups among Austroasiatic-speaking foragers of the Malay Peninsula. In Nick Enfield (ed.), *Dynamics of human diversity: The case of Mainland Southeast Asia* (Pacific Linguistics 627), 257–275. Canberra: Australian National University.

Burenhult, Niclas & Claudia Wegener. 2009. Preliminary notes on the phonology, orthography and vocabulary of Semnam (Austroasiatic, Malay Peninsula). *Journal of the Southeast Asian Linguistics Society* 1. 283–312.

Dahl, Östen. 2008. Älvdalska – eget språk eller värsting bland dialekter? *Språktidningen* December 2008. http://spraktidningen.se/artiklar/2008/11/alvdalska-eget-sprak-eller-varsting-bland-dialekter

Diffloth, Gérard. 1975. Les langues mon-khmer de Malaisie, classification historique et innovations. *Asie du sud-est et monde insulinde* 6(4). 1–19.

Diffloth, Gérard. 1976. Jah Hut, an Austroasiatic language of Malaysia. In Nguyễn Đăng Liêm (ed.), *South-East Asian linguistic studies, Vol. 2* (Pacific Linguistics 42), 73–112. Canberra: Australian National University.

Dunn, Michael, Niclas Burenhult, Nicole Kruspe, Sylvia Tufvesson & Neele Becker. 2011. Aslian linguistic prehistory: A case study in computational phylogenetics. *Diachronica* 28. 291–323.

Dunn, Michael, Nicole Kruspe & Niclas Burenhult. 2013. Time and place in the prehistory of the Aslian languages. *Human Biology* 85. 383–399.

Endicott, Kirk M. 1997. Batek history, inter-ethnic relations and subgroup dynamics. In Robert L. Winzeler (ed.), *Indigenous peoples and the state: Politics, land, and ethnicity in the Malayan Peninsula and Borneo*, 30–50. New Haven, CT: Yale Southeast Asia Studies.

Evans, Nick, Stephen C. Levinson, Nick Enfield, Alice Gaby & Asifa Majid. 2004. Reciprocal constructions and situation type. In Asifa Majid (ed.), *Field manual*, Vol. 9, 25–30. Nijmegen: Max Planck Institute for Psycholinguistics.

Fleming, Luke. 2014. Australian exceptionalism in the typology of affinal avoidance registers. *Anthropological Linguistics* 56. 115–158.

Gomes, Alberto G. 2007. *Modernity and Malaysia: Settling the Menraq forest nomads*. London: Routledge.

Hajek, John. 2003. Evidence of tonal activity in Aslian. *Mon-Khmer Studies* 33. 59–66.

Jenny, Mathias & Paul Sidwell (eds.). 2015. *The handbook of Austroasiatic languages*. 2 vols. Leiden: Brill.

Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61. 95–120.

Kruspe, Nicole. 2004. *A grammar of Semelai*. Cambridge: Cambridge University Press.

Kruspe, Nicole. 2010. *A dictionary of Mah Meri, as spoken at Bukit Bangkong* (Oceanic Linguistics Special Publication 36). Honolulu: University of Hawai'i Press.

Kruspe, Nicole, Niclas Burenhult & Ewelina Wnuk. 2015. Northern Aslian. In Jenny & Sidwell (eds.) 2015, Vol. 1, 419–474.

Kruspe, Nicole & Niclas Burenhult. Submitted. Pronouns in affinal avoidance registers: Evidence from the Aslian languages (Austroasiatic, Malay Peninsula). Manuscript.

Matisoff, James A. 2003. Aslian: Mon-Khmer of the Malay Peninsula. *Mon-Khmer Studies* 33. 1–58.

Mintz, Malcolm W. 1994. *A student's grammar of Malay and Indonesian*. Singapore: EPB.

Phaiboon Duangchand. 2006. The Northern Aslian languages of southern Thailand. *Mon-Khmer Studies* 36. 207–224.

Phillips, Timothy C. 2012. *Proto-Aslian: Towards an understanding of its historical linguistic systems, principles and processes*. Bangi, Malaysia: Universiti Kebangsaan Malaysia doctoral dissertation.

Rentse, Anker. 1937. Panganerne: Malayas jungledværge. *Geografisk Tidsskrift* 40. 110–136.

Schebesta, Paul. 1928. *Among the forest dwarfs of Malaya*. London: Hutchinson. Reprinted, Oxford: Oxford University Press, 1973.

Sidwell, Paul & Felix Rau. 2015. Austroasiatic comparative-historical reconstruction: An overview. In Jenny & Sidwell (eds.) 2015, Vol. 1, 221–363.

Wnuk, Ewelina. 2016. *Semantic specificity of perception verbs in Maniq*. Nijmegen: Max Planck Institute for Psycholinguistics doctoral dissertation.

Wnuk, Ewelina & Niclas Burenhult. 2014. Contact and isolation in hunter-gatherer language dynamics: Evidence from Maniq phonology (Aslian, Malay Peninsula). *Studies in Language* 38. 956–981.

Yager, Joanne. 2013. Northern Aslian prehistory: Tracing the effects of contact. Lund: Lunds Universitet MA thesis.

Study II

Study II: Outcomes of small-scale egalitarian contact in closely-related languages

Evidence from Southeast Asian foragers

Abstract

Aims and objectives: The study explores the linguistic outcomes of small-scale, egalitarian language contact in closely-related language varieties. Through comparison of language production data from individual speakers in phylogenetic networks, the study investigates language contact outcomes in the Northern Aslian (Austroasiatic) language varieties Jedek and Jahai.

Design/methodology: Lexical language production data were elicited from a number of Jedek and Jahai speakers. The distances between individual speakers are visualized in phylogenetic networks, revealing patterns of contact effects without the need for a priori definition of languages as abstract and discrete entities.

Data and analysis: Data came from a total of 71 adult native speakers, in four domains: basic vocabulary, topological relations, caused motion events, and reciprocal events. Unrooted phenogram networks for each dataset were constructed using the NeighborNet algorithm.

Findings/conclusions: The networks reveal a large degree of lexical convergence in the language production of cohabitant Jedek and Jahai speakers – the majority of the networks show no differentiation between speakers of the two varieties. The patterns point to convergence of ways of speaking within geographically-defined communities and divergence of ways of speaking between communities.

Originality: Situations of small-scale egalitarian contact between closely-related language varieties such as that examined in the study represent a type of contact scenario that is severely underrepresented in the literature. In addition, the study presents a novel approach to investigating the linguistic outcomes of language contact.

Significance: The study illustrates a way in which the challenges of studying the linguistic outcomes of language contact in small-scale and closely-related languages can be overcome, by adopting an approach that allows for the assessment of contact effects without operating with pre-defined groups.

Keywords: egalitarian language contact, small-scale multilingualism, lexical convergence, topological relations, caused motion events, reciprocal events, phenogram networks, NeighborNets, foragers, Jedek, Jahai

Introduction

Situations of egalitarian language contact between small-scale, genetically-related languages are common worldwide, yet are poorly represented in the literature (Epps, Huehnergard & Pat-El, 2013; Lüpke, 2016; Kalyan & François, 2018). While shared genealogical origin or similarity of language varieties is said to facilitate the spread of linguistic features over language boundaries (Weinreich, 1953), small-scale language communities are argued to have a tendency to enhance divergence of varieties over time (Evans, 2010). Situations of contact between small-scale, closely-related language varieties thus harbor important insights into the dynamics of language change, but this type of setting also represents a number of methodological challenges for the researcher. For example, small-scale language varieties may pose a challenge for the definition of language boundaries, and shared genealogical origin of language varieties leads to challenges in the identification of contact features. In response to these challenges, the current study presents a novel approach to exploring the outcomes of language contact, through comparison of language production data from individual speakers in phylogenetic networks. The study investigates the linguistic consequences of language contact in a small-scale egalitarian contact setting among speakers of the closely-related Northern Aslian (Austroasiatic) language varieties Jedek and Jahai in northern Peninsular Malaysia. Jedek and Jahai are spoken by small bands of Semang foragers who have long been engaged in complex webs of contact with speakers of other Aslian languages (Benjamin, 1985). This setting thus represents an ideal opportunity for investigation of egalitarian language contact between small-scale, closely-related language varieties.

Background

A common type of language contact situation worldwide is that in which languages originating from a common ancestor have remained in contact, or come into renewed contact after a period of separation. In addition, contact between small-scale language communities with egalitarian (as defined by Haudricourt, 1961: 9; François, 2011, 2012) prestige relations is common. Contemporary situations of small-scale language contact are also the closest contemporary analogues available to the type of situations in which language contact took place for the bulk of human history (Evans, 2010: 10). Indeed, Evans (2018) proposes that human language itself may have evolved through a process of small-scale egalitarian language contact in exogamous hunter-gatherers. Perhaps surprisingly, given their importance for understanding processes of language contact and change, situations of small-scale egalitarian language contact between related languages are profoundly underrepresented in the literature. Much of the research has focused on contact between large-scale languages, and where studies involve small-scale languages, they have most often focused on the effects of larger-scale (often colonial) languages on small-scale (often indigenous) languages (Vaughan & Singer, 2018). This means that we know relatively little about the dynamics and outcomes of language contact between small-scale languages. It is important that theories of language contact are informed by research from a range of contact settings, not least considering that language contact outcomes are greatly affected by the social features of the setting in which the contact occurs (Muysken, 2013; Thomason & Kaufman, 1988). For example, while unequal prestige relations between languages tend to lead to asymmetric contact outcomes, egalitarian language contact is thought to lead to more symmetric influence (Muysken, 2013). The task of describing small-scale egalitarian contact settings is also somewhat urgent, given that many such settings are today endangered (Trudgill, 2011: 185; Lüpke, 2016).

Small-scale communities speaking closely-related language varieties have properties that have been argued in the literature to result in both homogenizing and differentiating processes. On the one hand, since closely-related language varieties more recently diverged from a common ancestor they tend to share more structural features, often argued to facilitate contact-induced spread of linguistic material over language boundaries (e.g. Weinreich, 1953). Since shared features also facilitate communication between speakers, this may also lead to increased contact between groups, in turn leading to increased contact effects (Thomason & Kaufman, 1988). On the other hand, a tendency towards divergence has been suggested for small-scale language communities, due to a bias toward "constructive fostering of variegation" (Evans, 2010: 14). In this way, situations of small-scale contact between closely-related languages represent the meeting of two supposedly opposing forces of language change. It is not yet well understood the extent to which the size of language communities and the

genetic proximity or similarity of their ways of speaking may affect the outcomes of language contact (Epps et al., 2013).

The value and challenge of studying small-scale language contact in related languages

Situations of small-scale egalitarian language contact between related language varieties present us with opportunities for gaining important insight into processes of language change. François' (2011) discussion of the contact between the Oceanic languages of the northern Vanuatu linkage illustrates the value of studies of this kind. The 17 languages spoken across the Torres and Banks islands are closely related and share a high degree of structural similarity, but their lexica have diverged to the point that they are no longer mutually intelligible. François describes the simultaneous processes of divergence and convergence between the languages, and shows how these seemingly opposing forces are driven by a single process, that is, the diffusion of linguistic features. François argues that since innovations affecting word forms tend to diffuse among smaller circles of speakers, usually coinciding with a village or group of villages, over time this leads to lexical divergence between the ways of speaking of spatially-anchored communities. At the same time, homogeneity of structural features across the speech communities is a result of widespread multilingualism and continued contact of speakers across language boundaries. There has been little study of the causes and processes of language change in small-scale languages, and thus our understanding of these kinds of dynamics is not well developed (Evans, 2010: 14).

To date, research involving small-scale egalitarian language contact between related language varieties has presented researchers with a number of methodological challenges. First, since related language varieties share many features as a result of their shared genealogy, it can be difficult to identify which features are shared as a result of genealogy and which are shared as a result of contact (Epps et al., 2013; Bowern, 2013). One approach to identifying contact features in small-scale communities is through comparison with other groups of speakers that are not in contact (see Meakins & Pensalfini, forthcoming for an example). But this approach may not be feasible where the language varieties in contact are related, or where there are no non-contact groups available for comparison (Bowern, 2013: 425). Second, the researcher may encounter challenges in connection with the definition of language boundaries in small-scale language communities. Many small-scale language varieties are non-literate and lack political institutions or media, the existence of which tends to promote standardization of language varieties (Lüpke, 2010). Thus the researcher who wishes to study language contact in this kind of context is often faced with the task of first defining the boundaries between the languages in contact, a task which requires in-depth knowledge of the contact situation, and of speakers and their ways of speaking. The task of defining

language boundaries is made more difficult where language varieties have relatively recently diverged from a common ancestor (see e.g. François, Franjieh, Lacrampe & Schnell, 2015 for examples). In addition, language contact itself can further frustrate efforts to define entities for comparison. Many small-scale language varieties are relatively newly described by linguists, or lack proper description altogether, thus in many cases there is simply not enough information available to define language boundaries. Indeed, since the concept of a language boundary is a gradient one (Kalyan & François, 2018), the validity of such an endeavor is uncertain.

A novel approach

Despite the widely-acknowledged idea that it is not languages that are in contact but speakers, and thus that the transfer of linguistic material is the result of speaker behavior (Milroy, 1992: 199), very little research has explored how the behavior of speakers translates into contact-induced language change (Bowern, 2013). Language contact studies tend to conceptualize contact-induced change as change in linguistic 'systems', abstracting away from the role of individual speakers as the agents of language contact (Lüpke, 2016). Research in the multilingualism field (see e.g. Cenoz, 2012 for an overview) tends to place greater emphasis on the role of the individual speaker, but there is a lack of exchange of findings between the two fields (see e.g. Muysken, 2013 for discussion of this). Similarly, research into the dynamics of language change tends to focus either on factors affecting the spread of language-internal change among individual speakers (sociolinguistics), or on connections between languages conceived of as distinct entities (language contact). Investigation of language contact between closely-related language varieties "suggests a logical bridge between the two approaches" (Epps et al., 2013: 211), illustrating the ways in which these two phenomena are connected.

The current study adopts an approach in which individual speech production data are compared in phylogenetic networks, allowing us to observe patterns of language-internal and -external change in the context of the same study. It is commonly assumed in language contact research that we must define the features of languages, conceived of as discrete, abstract codes, before we may begin to investigate the effects of contact between them (e.g. Weinreich, 1953). However as we saw above, in some contexts this may be problematic for a number of reasons. The use of phylogenetic networks to directly compare language production data from individual speakers allows for investigation of contact effects without a priori definition of language boundaries. The study uses unrooted phenogram networks, made using the NeighborNet algorithm (Bryant & Moulton, 2004), a tool that is often used in historical linguistic studies comparing languages or dialects (see e.g. Holden & Gray, 2006). Since the NeighborNet algorithm is capable of reflecting conflicting or ambiguous signals in the

data, it is ideal for representing complex patterns of shared features and variation among speakers within and between varieties.

The contact setting

In order to begin to address the issues outlined above, we must begin to examine the linguistic outcomes of language contact in contact situations involving small-scale and closely-related language varieties. The current study does this through investigation of contact outcomes in the village of Rual in northern Peninsular Malaysia. Rual is home to six bands of exogamous foragers belonging to the cluster of peninsular forager populations known ethnographically as the Semang. Semang language communities are characterized by high levels of idiolectal variation and multilingualism, and complex webs of contact between closely-related language varieties (Benjamin, 1985). Linguistic variation in the Semang context is said to be "as much idiolectal as dialectal" (Benjamin, 2009: 20; see also Benjamin, 1976), and Endicott (1997) reports that the speech varieties of Semang groups are in a constant state of change, rapidly absorbing new words from one another and from Malay, the majority language of the region. Similarly, Semang speech communities are in constant flux, as bands split and reform with different membership, and as individuals move between communities for marriage. Rual is part of a resettlement site established in the 1970s, and is home to four Jedek-speaking bands and one Jahai-speaking band who settled there from the surrounding region of the Pergau river (see maps in Figure 1). A second Jahai-speaking band from the area surrounding the resettlement site formed the hamlet of Manok approximately 2 km upstream of Rual, and a further Jahai-speaking band moved to Rual from the neighboring state of Perak in the 1980s (Gomes, 2007; Yager & Burenhult, 2017).

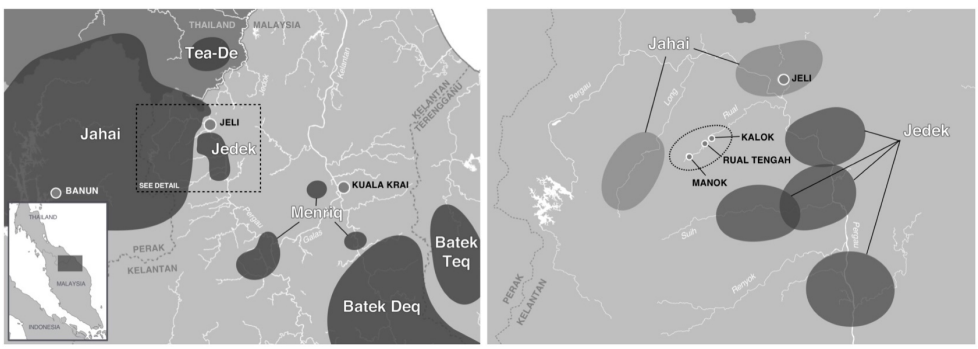


Figure 1. Maps adapted from Yager & Burenhult (2017) showing (left) the approximate historical distribution of Jedek- and Jahai-speaking groups and speakers of the surrounding Northern Aslian varieties, and (right) the approximate distribution of the Jedek- and Jahai-speaking bands of Rual prior to resettlement (adapted from Gomes, 2007: 77) as well as the Rual resettlement site with its three hamlets – Rual Tengah and Kalok (referred to collectively in the current study as Rual) and Manok, located approx. 2 km upstream of Rual.

Jedek and Jahai speakers engage in band exogamy (out-marriage) and have long been in contact with speakers of other Aslian languages (see Benjamin, 1985, 2001). Jedek and Jahai are also closely related – they belong to the same subbranch of Northern Aslian (referred to as the Menraq-Batek branch in Dunn, Burenhult, Kruspe, Tufvesson & Becker, 2011, encompassing Batek, Menriq, Jedek and Jahai). While Jedek and Jahai are not considered mutually intelligible by speakers, they share many typological and lexical features due in part to their shared genetic origin and in part to long-term traditions of contact among their speakers. The speaker populations of both Jedek and Jahai are small in size: Jedek has just over 100 adult speakers and is only spoken at Rual, while Jahai has roughly 1000 speakers, most of whom reside in other geographical areas. Both Jedek and Jahai are non-literate, and both are only recently described by linguists (Burenhult, 2005; Yager & Burenhult, 2017). The typological and lexical similarity of Jedek and Jahai coupled with the intensity of their long-term and ongoing contact means that the task of defining the boundaries between the varieties is a challenging one, exacerbated further by the relatively limited amount of linguistic information available. The study makes comparison of lexical production data from Jedek and Jahai speakers at Rual as well as from Jahai speakers in two Jahai-majority communities: Banun, located approximately 60 km West of Rual, and Manok, located approximately 2 km upstream of Rual (see the maps in Figure 1).

The domains studied

The current study uses lexical data. Existing description of the grammatical and lexical features of Jedek and Jahai (Burenhult, 2005; Yager & Burenhult, 2017) suggests that the varieties share many structural features, but are more divergent in their lexicon (see Dunn et al., 2011 for phylogenetic comparison of Aslian basic vocabulary data, including Jahai and Jedek – Jedek is there referred to as 'Menriq Rual'). The lexicon is generally thought to be the part of language that is the most susceptible to influence through contact (e.g. Muysken, 1984), and is thought to be the most rapidly changing part of language. For these reasons lexical data are likely to be appropriate for investigation of contact effects at the micro-level within and between varieties and speech communities, and are thus used in the current study. While many studies investigating lexical contact effects use basic vocabulary data, data from additional domains are also used in the current study in order to allow for comparison of contact effects between domains. The lexical data used in the study come from four domains: basic vocabulary, topological relations descriptions, caused motion descriptions, and reciprocal event descriptions. The basic vocabulary data were collected through questionnaire-style Swadesh list elicitation, while the three additional datasets were elicited through description of pictures or film clips in director-matcher tasks. This approach was used in order to provide more natural language production data than is typically elicited through questionnaire-style elicitation, while still allowing for a high

degree of comparability between the responses of individuals. The two types of data provide an interesting point of comparison for the purposes of the study since it is thought that more natural or spontaneous speech tends to show signs of contact-induced change at earlier stages than more careful speech (Milroy & Milroy, 1985: 340).

Method

Participants

Language production data from a total of 71 adult native speakers were included in the study, including 41 Jedek speakers, 17 Rual Jahai speakers, 8 Banun Jahai speakers, and 5 Manok Jahai speakers. The datasets include basic vocabulary data from 12 participants (8 Jedek, 2 Rual Jahai, 2 Banun Jahai speakers), topological relations descriptions from 45 participants (23 Jedek, 16 Rual Jahai, 3 Banun Jahai, 3 Manok Jahai speakers), caused motion event descriptions from 60 participants (35 Jedek, 16 Rual Jahai, 5 Banun Jahai, 4 Manok Jahai speakers), and reciprocal event descriptions from 57 participants (32 Jedek, 16 Rual Jahai, 5 Banun Jahai, 4 Manok Jahai speakers). Table 1 presents the participants. Speakers are treated as residents of Rual, Manok or Banun if they have resided at that location for the majority of their adult life.

Table 1. The participants

	Jedek	Rual Jahai	Banun Jahai	Manok Jahai	Total
Basic vocabulary	n=8	n=2	n=2		n=12
Topological relations	n=23	n=16	n=3	n=3	n=45
Caused motion events	n=35	n=16	n=5	n=4	n=60
Reciprocal events	n=32	n=16	n=5	n=4	n=57
All datasets	n=41	n=17	n=8	n=5	n=71

Materials and tasks

Data were collected using 4 elicitation tasks. The basic vocabulary data were collected using a 237-item wordlist based on Swadesh's (1952) 200-word list adapted for use in Aslian contexts. The topological relations data were elicited using the Topological Relations Picture Series task (TRPS; Bowerman & Pederson, 1992), the caused motion data with the PUT project task (Bowerman, Gullberg, Majid & Narasimhan, 2004), and the reciprocal event data with the Reciprocal Constructions and Situation Type task (Evans, Levinson, Enfield, Gaby & Majid, 2004). These three tasks have been used extensively in semantic typological comparison of a large number of languages (see e.g. Kopecka & Narasimhan, 2012; Evans, Gaby, Levinson & Majid, 2011), and are

thus well-established elicitation tools. While the tasks are designed for semantic analysis, they also yield lexical corpora that may be analyzed in a similar way to basic vocabulary data. But since the tasks use picture- and video-based elicitation, they allow us to avoid potential problems associated with questionnaire-style Swadesh list elicitation.

The TRPS consists of a series of 71 line drawings depicting a range of spatial relationships between physical objects, designed to elicit descriptions of topological relations (as expressed in prepositions and relational nouns in Jedek and Jahai, or prepositions in English, e.g. 'The cup is on the table'). The PUT project task consists of 63 short video clips depicting scenes in which a human agent places or removes objects in relation to other objects in space, designed to elicit descriptions of caused motion events (as expressed in verbs in Jedek and Jahai, and in English as in 'She puts a cup on a table'). The Reciprocal Constructions and Situation Type task is a series of 64 short video clips depicting a range of reciprocal and non-reciprocal events, designed to probe the semantics of the constructions used to describe reciprocal events (e.g. as expressed in verbs marked with the distributive morpheme in Jedek and Jahai, and in constructions such as 'each other' in English, as in 'They are talking to each other'). The three tasks are similar in length and structure, yet probe different semantic domains and parts of speech. The TRPS focuses on prepositions and other words used to express static spatial relations, while the PUT project task focuses on verbs used to express the caused movement of objects, and the Reciprocal Constructions and Situation Type task focuses on constructions used to express reciprocal and non-reciprocal events. See Yager & Gullberg (published online 2019) for a description of Jedek and Jahai topological relation markers and their semantics, Yager (submitted) for a description of Jedek and Jahai placement (caused motion) verbs and their semantics, Burenhult (2012) for a description of Jahai placement and removal (caused motion) verbs and their semantics, and Burenhult (2011) for a description of Jahai reciprocal constructions and their semantics.

Procedure

Following common linguistic fieldwork practice, the basic vocabulary data were elicited through a combination of translation from Malay, oral description in Jedek or Jahai, and pointing (in the case of e.g. words for body parts). List items were arranged according to semantic field rather than alphabetically (see Dixon, 2009: 299 for discussion of problems related to alphabetically-ordered wordlist elicitation). The topological relations, caused motion and reciprocal event data were elicited through a director-matcher task in which two Jedek or Jahai participants interacted with one another. This was done to create an environment in which participants could engage in dialogue with a communicative goal, with the aim of eliciting more natural speech.

The 'director' viewed each picture or video clip on a laptop screen and described it to the 'matcher', who located the scene described on the pages of a folder containing the scenes of the task. Each page of the folder included five to six target scenes and six fillers. Each participant completed each task only once, in their identity variety. Several of the participants participated in more than one of the tasks. Jedek speakers performed the task with a Jedek matcher and Jahai speakers performed the task with a Jahai matcher, in an attempt to control for the potential influence of the language background of matchers in participants' responses. The Jedek, Rual Jahai, Manok Jahai and parts of the Banun Jahai data were collected at Rual between 2014 and 2017, and the remainder of the Banun Jahai data were collected at Banun by Jahai researcher Niclas Burenhult in 2004 and 2016.

Data treatment and analyses

The full responses of each participant were transcribed. The topological relation markers (in the case of the topological relations data) and verbs (in the case of the caused motion and reciprocal event data) used by participants in their descriptions were entered into a TAB-separated CSV file and coded using a python script. Only the first response of each participant to each target stimulus was used in the analyses. Responses to 12 of the scenes of the TRPS task were removed due to a high number of missing responses. Transcriptions conform to the spelling used in Yager's (unpublished) Jedek lexicon and Burenhult's (unpublished) Jahai lexicon[1]; forms not found in the Jedek and Jahai lexica were transcribed phonemically. The datasets and the python script used to code the data are found in Appendices A–E. The Jedek and Jahai lexica are available at http://hdl.handle.net/10050/00-0000-0000-0004-012C-2@view and http://hdl.handle.net/10050/00-0000-0000-0003-FAEF-E@view.

The script generated sequences of character data in which lexical items exactly identical in orthographic form received the same character, and items whose orthographical form differed received different characters. Since the study aims to capture micro-level variation in the responses of participants, any difference at the level of phonemes led to coding as non-identical. Thus for example Jedek-Jahai cognate candidates such as *hãc* and *hɛ̃c* 'to spill' were coded as non-identical, as were responses that differed in the form of derivations such as *kmʔəm* 'to hug' and *kimʔəm* 'to hug (one another/distributed in space)'. The coding was thus based on lexical sets, rather than cognate sets as is usual in phylogenetic comparison of lexical data at the level of languages. Measures of phonological similarity such as Levenshtein distance were not used since the cognacy of the forms cannot be assumed (see Dunn, 2014 for discussion of this). In addition,

[1] The one exception to this was the transcription of the Jahai caused motion verb *ʔaŋkit* 'to take' which conformed to the spelling of the Jedek orthographic form of the verb (*ʔaŋket*) since the two forms are often indistinct in the recordings.

phonological similarity measures were not considered appropriate for the purposes of the study's focus on micro-variation between individual speakers; for example, this kind of measure would tend to underrepresent distances between cognate Jedek and Jahai forms while overrepresenting distances between synonyms that differ more in form.

The sequences of coded responses for each dataset were used to construct unrooted phenogram networks using the NeighborNet algorithm (Bryant & Moulton, 2004) implemented within the SplitsTree4 package (Huson & Bryant, 2006). Table 2 contains an example of lexical sets taken from the basic vocabulary data. The cells of the table contain the responses given by participants (columns) to the Swadesh meaning prompts (rows).

Table 2. An excerpt from the basic vocabulary data

	Jedek5	Jedek32	JRual6	JRual13	JBanun2	JBanun6
foot	*can*	*can*	*can*	*can*	*can*	*can*
belly	*cɨŋ*	*cɨŋ*	*cɨŋ*	*cɨŋ*	*ʔɛc*	*ʔɛc*
egg	*makɔʔ*	*makɔʔ*	*makɔʔ*	*makɔʔ*	*ktɨt*	*ktɨt*
eye	*mɛt*	*mɛt*	*mɨt*	*mɛt*	*mɨt*	*mɨt*
snake	*jkɔp*	*jkɔp*	*jkɔp*	*tajuʔ*	*tajuʔ*	*tajuʔ*
I	*ʔiŋ*	*ʔiŋ*	*yɛʔ*	*yɛʔ*	*yɛʔ*	*yɛʔ*
you (sing.)	*bɛʔ*	*bɛʔ*	*mɔh*	*mɔh*	*mɔh*	*mɔh*
mouth	*hãŋ*	*hãŋ*	*hɛ̃ŋ*	*hɛ̃ŋ*	*hɛ̃ŋ*	*hɛ̃ŋ*
younger.sibling	*bɛ*	*bɛ*	*bɛr*	*bɛr*	*bɛr*	*bɛr*

Results

The distances between individual speakers based on their responses to the tasks are presented in NeighborNet networks in Figures 2–5. The color coding reflects the language identity and place of residence of speakers: Jedek speakers are shown in green, Rual Jahai speakers are shown in red, Banun Jahai speakers are shown in yellow, and Manok Jahai speakers are shown in purple. In the networks, line length corresponds to the distance between the nodes connected by a line, thus the sum of the lengths of the edges connecting two speakers represents the distance between those speakers based on their responses. Reticulated patterns (box-like shapes in the networks) represent ambiguous splits.

The basic vocabulary data

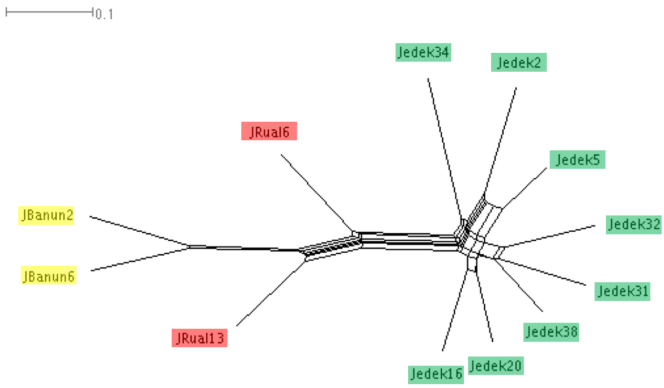


Figure 2. NeighborNet visualization of the basic vocabulary data

The network in Figure 2 shows clear separation of the Jedek speakers, the Rual Jahai speakers, and the Banun Jahai speakers, with the Rual Jahai speakers positioned roughly halfway along a series of splits between the Jedek and Banun Jahai speakers. Apart from these primary splits, some secondary splits can be seen in the network. First, three clusters of individuals within the Jedek group are separated by more reticulated splits; these clusters roughly correspond to the age of speakers. One cluster contains speakers aged around 20 years (Jedek2 and Jedek5), another contains speakers aged around 30 years (Jedek16 and Jedek20), and a third cluster contains speakers aged 40 years and over (Jedek38, Jedek31 and Jedek32). Second, a relatively long edge separates the two Rual Jahai speakers, with one of the speakers positioned at a shorter distance from the Banun Jahai speakers, and the other at a shorter distance from the Jedek speakers. This positioning has interesting parallels in the language histories of the two speakers. While the speaker that is placed closer to the Jedek speakers in the network belongs to one of the Jahai bands originating from the Rual area, the speaker who is placed closer to the Banun Jahai speakers in the network belongs to the band of Jahai speakers who moved to Rual from the Banun area in the 1980s.

The topological relations data

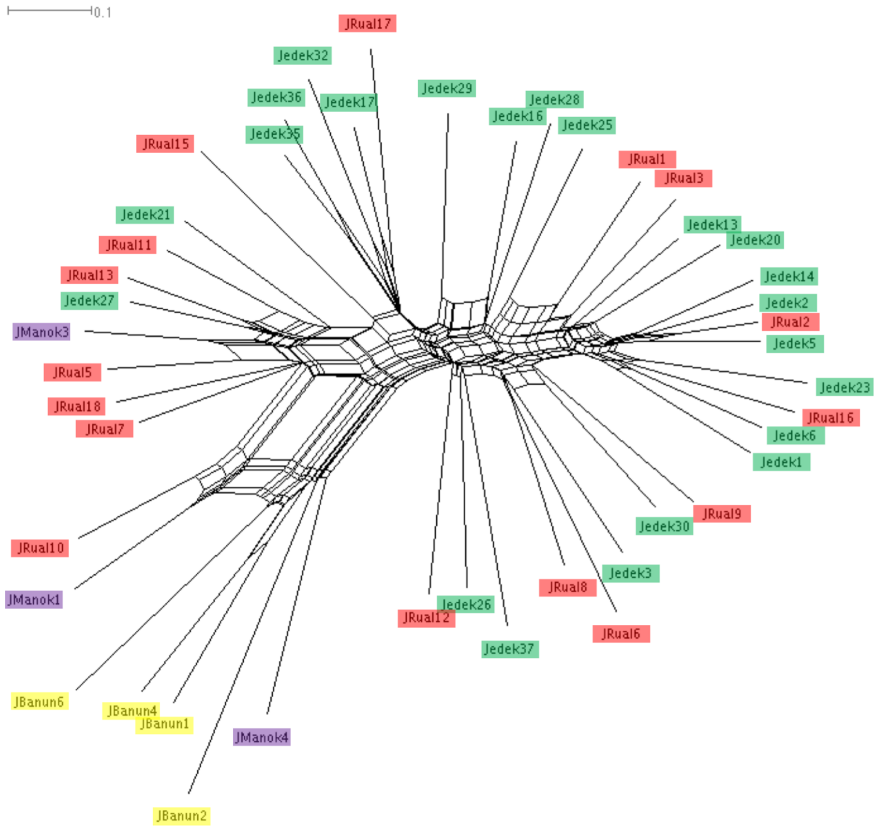


Figure 3. NeighborNet visualization of the topological relations data

In the network in Figure 3, a major split separates the Banun Jahai speakers together with two of the three Manok Jahai speakers and one Rual Jahai speaker from the remaining participants. No major splits separate the Jedek speakers from the Rual Jahai speakers. Rather, both Jedek- and Jahai-speaking Rual residents are spread out across the remainder of the network. Overall, the network in Figure 3 contains many more reticulations than the network in Figure 2, indicating that the topological relations data contain a greater amount of conflicting signal than the basic vocabulary data. Some of the conflicting splits with short branch lengths seen throughout the network appear to be potentially meaningful, while others do not. The majority of the youngest speakers of the Rual sample (both Jedek and Jahai speakers, aged 15-25 years) are found toward the righthand end of the series of splits that make up the network, that is, at the greatest distance from the Banun and Manok Jahai speakers (positioned at the lefthand end of

the network). Age is however not a reliable predictor of individuals' positioning in the network – Rual residents are not arranged according to ascending age from right to left. Some splits appear to reflect kinship relations between individuals, while in other cases, large distances separate individuals connected by kinship. For example, a father–daughter pair (JBanun4 and JBanun1) form a cluster within the Banun–Manok Jahai group. A pair of Jedek-speaking sisters (Jedek35 and Jedek36) also form a cluster in the network, but three of their siblings who are also included in the sample do not cluster with them. A pair of sisters originating from Manok (JManok3 and JRual5) also form a cluster, despite one of them being a long-term resident of Rual (this is however not systematic across the datasets – see Figure 4 below). Other sets of siblings in the sample do not cluster together in the network (e.g., JRual12 and JRual13; JRual11, JRual8 and JRual18; JRual6 and JRual2).

The caused motion data

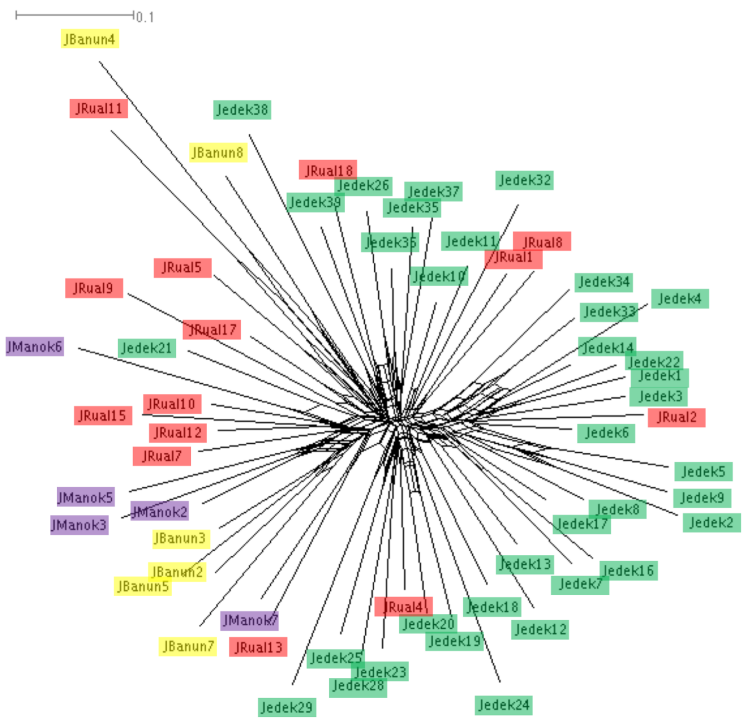


Figure 4. NeighborNet visualization of the caused motion data

In the network in Figure 4, splits separate four of the Banun Jahai speakers and three of the Manok Jahai speakers from the remainder of the participants. Banun Jahai and

Manok Jahai speakers form separate clusters within this Banun–Manok group. The two remaining Banun Jahai speakers (JBanun4, JBanun8) form a conflicting cluster together with one of the Rual Jahai speakers, and the remaining two Manok Jahai speakers (JManok6, JManok7) are placed in smaller clusters with Jedek and Rual Jahai speakers. Many of the Rual Jahai speakers are found in the same region of the network as the Manok Jahai and Banun Jahai speakers, but there are no clear splits separating this area of the network from the remainder of the Rual participants. Overall, the network contains a large amount of conflict, the edges joining individuals are short in length and tend to radiate out from the center of the network.

As was the case in Figure 3 above, the youngest speakers of the dataset tend to show up in the regions of the network that are at the greatest distance from the Banun and Manok Jahai speakers. But as in Figure 3, age is not a systematic predictor of speakers' placement in the network. Clusters representing kinship relations are not evident in this network. For example, the pair of sisters with origins in Manok who formed a cluster in Figure 3 (JManok3, JRual5) do not cluster together in Figure 4.

The reciprocal event data

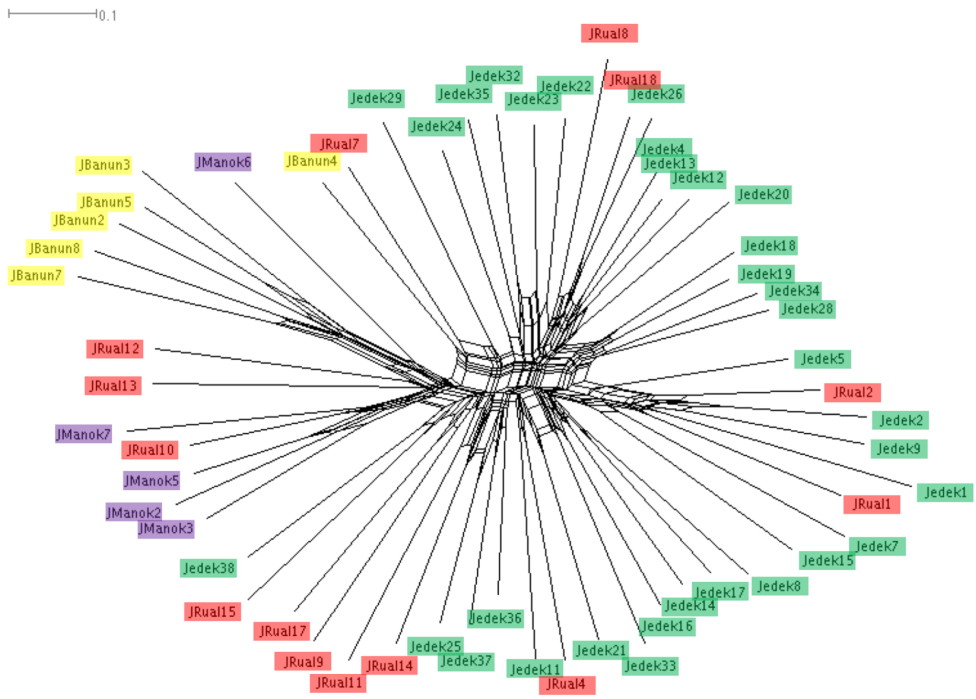


Figure 5. NeighborNet visualization of the reciprocal event data

The Banun Jahai speakers are clearly separated from the remainder of the speakers of the sample in Figure 5. A clear split also forms a cluster containing four of the five Manok Jahai speakers together with one Rual Jahai speaker. A conflicting split of relatively short length suggests clustering of the Banun and Manok Jahai speakers along with four of the Rual Jahai speakers, while the remaining Rual Jahai speakers are interspersed with the Jedek speakers in the remainder of the network. As was the case in the networks in Figures 3 and 4, in Figure 5 the youngest speakers of the sample are positioned at the point in the network most distant from the Banun and Manok Jahai speakers. Interestingly, the two Manok Jahai speakers who show some degree of separation from the Manok Jahai group in Figure 5, JManok6 and JManok7, are also positioned outside the Manok Jahai cluster in the network in Figure 4 above. The Banun Jahai speaker JBanun4 is placed outside the Banun Jahai split in Figure 5, reflecting similarity of this participant's responses to those of the Rual residents. This may reflect the fact that this particular speaker has spent significant amounts of time in both Rual and Banun, however, this is not reflected systematically across the datasets – this same speaker is placed firmly within the Banun Jahai cluster in Figure 3 above, while he forms a smaller cluster with one other Banun Jahai speaker and one Rual Jahai speaker in Figure 4 above. Splits reflecting kinship or other background parameters of participants are not evident in Figure 5.

Summary of the networks

While the basic vocabulary network shows separation of the Jedek and Rual Jahai speakers, in the topological relations, caused motion and reciprocal event networks, Jedek and Jahai speakers at Rual do not form two groups with distinct ways of speaking. Meanwhile, the networks show clear separation of the Banun Jahai and Manok Jahai speakers from Jedek and Jahai speakers at Rual. The three datasets containing Manok Jahai data reveal different patterns of separation of these speakers: while the majority of the Manok Jahai speakers cluster very clearly with the Banun Jahai speakers in the topological relations network, the Banun and Manok Jahai speakers form separate clusters connected by splits with much shorter branch lengths in the caused motion and reciprocal event networks. In all four networks, the greatest distances are seen between the youngest participants of the Rual sample (both Jedek- and Jahai-speaking individuals) and the Manok and Banun Jahai speakers.

Discussion

The study investigates small-scale, egalitarian contact between the closely-related language varieties Jedek and Jahai in northern Peninsular Malaysia, and presents a novel approach to investigating the linguistic outcomes of language contact. Unrooted phenogram networks were constructed using lexical language production data elicited from a number of Jedek and Jahai speakers in the village of Rual and Jahai speakers in two other geographical locations. The patterns seen in the networks suggest that the contact between Jedek and Jahai speakers at Rual has resulted in a large degree of lexical convergence in their ways of speaking. Indeed patterns were seen in the networks that are consistent with the contact histories of the speaker groups in the three locations. Banun Jahai speakers are geographically well-separated from Rual, and have historically had less contact with Jedek speakers. This is reflected in their separation from the Rual participants in the networks. Meanwhile, Manok Jahai speakers have long lived in geographical proximity to the Jedek- and Jahai-speaking bands of Rual, and this is reflected in slightly less clear-cut separation between the ways of speaking of Manok and Rual residents in the networks. Finally, the contact between Jedek and Jahai speakers living at Rual is close and intense, and the effects of this were seen in the non-separation of these speakers in the networks.

The four networks reflected differences between the four datasets. In particular, the basic vocabulary network differed in several ways from the networks based on the topological relations, caused motion event, and reciprocal event datasets. First, this network contained fewer conflicting splits than the remaining three networks, reflecting more straightforward grouping of individuals. Further, in contrast to the remaining three networks, Jedek and Rual Jahai speakers were separated by a clear split in the basic vocabulary network. How can this difference be explained? One potential explanation is that since the Swadesh list is designed to elicit the most stable parts of the vocabulary, the items in this dataset should be less susceptible to contact influence than those of the other datasets. According to this explanation, the basic vocabulary data capture the parts of Jedek and Rual Jahai speakers' ways of speaking that have not yet been affected by the intense contact between them. On the other hand, due to the nature of Swadesh list elicitation in which lexemes are elicited in isolation through translation from a contact language, what we see in this network is likely to in part be a reflection of speakers' metalinguistic knowledge of the forms associated with 'Jedek' or 'Jahai' rather than necessarily reflecting the forms that these speakers would use in natural speech. In contrast, the remaining three datasets were designed to elicit more natural speech and thus are less likely to be affected by this issue. Thus, this difference between the networks might be interpreted as representing differential contact effects in different parts of the lexicon, or serve as a reminder that basic vocabulary data elicited in this way may not necessarily be representative of actual language use.

The networks also reflected the profiles of individual speakers to some extent. In all four networks, the greatest distances were seen between the youngest speakers of the Rual sample and the Jahai speakers in other geographical areas. This pattern may reflect a change in progress at Rual, since generational differences tend to reflect change in languages over time. Thus, at the same time as the ways of speaking of Rual residents are converging with one another, these ways of speaking appear to be diverging from those of Jahai speakers in other areas. Interestingly, the pattern is reminiscent of François' (2011) discussion of the processes affecting language change in the closely-related Oceanic languages of the Torres and Banks islands in Vanuatu. François describes how the spread of innovations within speech communities (generally coinciding with a single village or groups of villages in this area) leads to lexical convergence within communities and lexical divergence between communities. Similarly, although Rual is made up of individuals who identify with two different language varieties, the speech community at Rual encompasses speakers of both varieties – thus ways of speaking within Rual converge across varieties, while diverging from ways of speaking in other locations. This aspect of the results of the current study illustrates that examining language contact outcomes at the level of the individual speaker may help us to capture patterns of variation that inform us about how contact-induced change spreads through a community of speakers.

Other aspects of the backgrounds of individual speakers, such as kinship relations and residence patterns, also appeared to be reflected in the networks, although not in a systematic way. Further exploration of the ways in which these kinds of metadata variables may be reflected in the positioning of individuals in networks created using this type of approach is likely to reveal interesting insights into the factors affecting micro-variation and language change in small-scale communities. In general, the networks show a great deal of variation between individual speakers, as shown by the length of the individual branches separating individual speakers from the group as a whole. Whether this level of micro-variation between speakers is unusually high (as claimed by Benjamin, 2009, 1976 for Northern Aslian speech communities) remains an empirical question, since we lack points of comparison from other contexts using similar methods (although see Slaska, 2005, who found high levels of between-speaker variation in Swadesh list responses in French and Polish speakers). Indeed, since quantitative variationist sociolinguistic research in settings like Rual is lacking (Stanford & Preston, 2009), more thorough description of the ways in which individual ways of speaking differ at Rual, as well as comparison with other contexts, would be highly valuable.

In recent years researchers have begun calling for research into societal and psycholinguistic aspects of multilingualism in small-scale contexts (Gullberg, 2012; Lüpke, 2016). The patterns seen in the current study raise a number of important questions for researchers interested in studying multilingualism in contexts like Rual.

For example, how should multilingualism be defined in a context in which language contact has resulted in convergence to the point that named ways of speaking are to a large extent undifferentiated – at what point does it become inappropriate to study multilingualism in the context of such convergence? And where speakers self-identify as multilingual in these named varieties, what do they mean by this? If we wish to study multilingualism in this kind of contact setting, it is important to consider how ideologies around language identity, language boundaries and language naming relate to actual language production practices, as well as the ways in which the ideologies of speakers in small-scale communities may diverge from those of the researcher. It may not be appropriate in this type of setting to take for granted the discreteness of language varieties and divide speakers into groups based on language identity without reflection over what these identities may represent in practice. In order to successfully study the dynamics of multilingualism in contact settings like Rual, it is imperative that we work to find solutions to issues of this kind.

Conclusion

The current study demonstrates a novel approach to investigating the outcomes of language contact through comparison of lexical language production data from individual speakers in phylogenetic networks, and investigates the outcomes of language contact in speakers of the closely-related Northern Aslian (Austroasiatic) language varieties Jedek and Jahai. We saw that long-term, intense contact between Jedek and Jahai speakers has resulted in a large degree of lexical convergence in their ways of speaking, suggesting that language contact in this kind of setting can have a profound impact on the lexicon of language varieties. Phylogenetic networks comparing lexical production data from a number of individual speakers revealed patterns of convergence within, and divergence between, geographically-bound communities, in line with the contact histories of the groups. The study demonstrates how a methodological approach in which language contact is conceptualized as a process driven by individual speakers, using phylogenetic networks for comparison of individual language production data, allows us to gain insight into the linguistic consequences of language contact involving closely-related, non-standardized, newly-documented language varieties in small-scale, egalitarian contact settings. Such an approach allows us to explore processes of language change both within and across language boundaries. The contact outcomes seen in the ways of speaking of Jedek and Jahai speakers in the current study raise a number of questions with implications for a number of linguistic fields, including studies involving bilingualism and multilingualism, language ideologies, language documentation and description, and quantitative variationist sociolinguistics. Situations of small-scale egalitarian contact between closely-related language varieties such as the one explored in the current study

are greatly underrepresented in the literature, and represent an important point of comparison if we are interested in letting our theories be informed by studies from a broader range of contact settings. It is thus crucial that we work to develop appropriate methodologies for research in this kind of contact setting.

References

Benjamin, G. (1976). Austroasiatic subgroupings and prehistory in the Malay Peninsula. In P. N. Jenner, L. C. Thompson & S. Starosta (eds.), *Austroasiatic studies, Part I,* 37–128. Honolulu: The University of Hawai'i Press.

Benjamin, G. (1985). In the long term: Three themes in Malayan cultural ecology. In K. L. Hutterer, A. T. Rambo & G. Lovelace (eds.), *Cultural values and human ecology in Southeast Asia,* 219–278. Ann Arbor, MI: Center for South and Southeast Asian Studies.

Benjamin, G. (2001). Orang Asli languages: From heritage to death? In R. Rashid & W. J. Karim (eds.), *Minority cultures of Peninsular Malaysia: Survivals of indigenous heritage,* 101–122. Penang: Academy of Social Sciences.

Benjamin, G. (2009). The current situation of the Aslian languages. (Paper presented at Symposium on Language and Diversity, Nanyang Technological University, Singapore, March 2009).

Bowerman, M., Gullberg, M., Majid, A., & Narasimhan, B. (2004). Put project: The crosslinguistic encoding of placement events. In A. Majid (ed.), *Field Manual, Vol. 9,* 10–24. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.492916.

Bowerman, M., & Pederson, E. (1992). Topological relations picture series. In S. C. Levinson (ed.), *Space stimuli kit 1.2: November 1992,* 51. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.883589.

Bowern, C. (2013). Relatedness as a factor in language contact. *Journal of Language Contact* 6(2). 411–432.

Bryant, D., & Moulton, V. (2004). Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21(2). 255–265.

Burenhult, N. (2005). *A grammar of Jahai* (Pacific Linguistics 566). Canberra: Australian National University.

Burenhult, N. (2011). The coding of reciprocal events in Jahai. In N. Evans, A. Gaby, S. C. Levinson & A. Majid (eds.), *Reciprocals and semantic typology,* 163–176. Amsterdam: Benjamins.

Burenhult, N. (2012). The linguistic encoding of placement and removal events in Jahai. In A. Kopecka & B. Narasimhan (eds.), *Events of putting and taking: A crosslinguistic perspective,* 21–36. Amsterdam: John Benjamins Publishing.

Cenoz, J. (2012). Multilingualism. In C. Chapelle (ed.) *The Encyclopedia of Applied Linguistics.* John Wiley & Sons.

Dixon, R. M. W. (2009). *Basic linguistic theory, Volume 1: Methodology.* Oxford University Press, Oxford.

Dunn, M. (2014). Language phylogenies. In C. Bowern & B. Evans (eds.), *The Routledge handbook of historical linguistics,* 190–211. Routledge.

Dunn, M., Burenhult, N., Kruspe, N., Tufvesson, S., & Becker, N. (2011). Aslian linguistic prehistory: A case study in computational phylogenetics. *Diachronica* 28. 291–323.

Endicott, K. (1997). Batek history, inter-ethnic relations and subgroup dynamics. In R. Winzeler (ed.), *Indigenous peoples and the state: Politics, land, and ethnicity in the Malayan Peninsula and Borneo,* 30–50. New Haven: Yale Southeast Asia Studies Monograph, no. 46.

Epps, P., Huehnergard, J., & Pat-El, N. (2013). Introduction: Contact among genetically related languages. *Journal of Language Contact* 6(2). 209–219.

Evans, N. (2010). *Dying words: Endangered languages and what they have to tell us.* Chichester: Wiley-Blackwell.

Evans, N. (2018). Did language evolve in multilingual settings? *Biology & Philosophy* 32(6). 905–933.

Evans, N., Gaby, A., Levinson, S.C., & Majid, A. (eds.). (2011). *Reciprocals and semantic typology.* Amsterdam: Benjamins.

Evans, N., Levinson, S. C., Enfield, N. J., Gaby, A., & Majid, A. (2004). Reciprocal constructions and situation type. In A. Majid (ed.), *Field Manual Volume 9,* 25–30. Nijmegen: Max Planck Institute for Psycholinguistics. doi:10.17617/2.506955.

François, A. (2011). Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1. 175–246.

François, A. (2012). The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language* 214. 85–110.

François, A., Franjieh, M., Lacrampe, S., & Schnell, S. (2015). The exceptional linguistic density of Vanuatu. In A. François, M. Franjieh, S. Lacrampe, & S. Schnell (eds.), *The languages of Vanuatu: Unity and diversity,* 1–22. Canberra: Australian National University.

Gomes, A. G. (2007). *Modernity and Malaysia: Settling the Menraq forest nomads.* London: Routledge.

Gullberg, M. (2012). Bilingual multimodality in language documentation data. In F. Seifart, G. Haig, N. P. Himmelmann, D. Jung, A. Margetts & P. Trilsbeek (eds.), *Language Documentation & Conservation 6* (Special issue *"Potentials of Language Documentation: Methods, Analyses, Utilization"*), 54-63. Honolulu: University of Hawai'i Press.

Haudricourt, A-G. (1961). Richesse en phonèmes et richesse en locuteurs. *L'Homme* 1(1). 5–10.

Holden, C. J., & Gray, R. D. (2006). Rapid radiation, borrowing and dialect continua in the Bantu languages. In P. Forster & C. Renfrew (eds.), *Phylogenetic methods and the prehistory of languages,* 19–31. Cambridge, UK: McDonald Institute for Archaeological Research.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2). 254–267.

Kalyan, S., & François, A. (2018). Freeing the comparative method from the tree model: A framework for Historical Glottometry. In R. Kikusawa & L. A. Reid (eds.), *Let's talk about trees: Genetic relationships of languages and their phylogenetic representation* (Senri Ethnological studies 98), 59–89. Osaka: National Museum of Ethnology

Kopecka, A., & Narasimhan, B. (eds.) (2012). *Events of putting and taking: A crosslinguistic perspective.* Amsterdam: John Benjamins Publishing.

Lüpke, F. (2010). Language and identity in flux: In search of Baïnounk. *Journal of Language Contact* THEMA 3. 155–174.

Lüpke, F. (2016). Uncovering small-scale multilingualism. *Critical Multilingualism Studies* 4.2. 35–74.

Meakins, F., & Pensalfini, R. (forthcoming). Holding the mirror up to converted languages: Two grammars, one lexicon.

Milroy, J. (1992). Social network and prestige arguments in sociolinguistics. In K. Bolton & H. Kwok (eds.), *Sociolinguistics today: International perspectives*, 146–162. London: Routledge.

Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2). 339–384.

Muysken, P. (1984). Linguistic dimensions of language contact: the state of the art in interlinguistics. *Revue québécoise de linguistique* 14(1). 49–76.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16. 709–730.

Slaska, N. (2005). Lexicostatistics away from the armchair: Handling people, props and problems. *Transactions of the Philological Society* 103(2). 221–242.

Stanford, J. N., & Preston, D. R. (2009). The lure of a distant horizon: Variation in indigenous minority languages. In J. N. Stanford & D. R. Preston (eds.), *Variation in Indigenous Minority Languages,* 1–20. John Benjamins.

Swadesh, M. (1952). Lexicostatistic dating of prehistoric ethnic contacts. *Proceedings of the American Philosophical Society* 96. 452–63.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics.* Berkeley & Los Angeles, CA: University of California Press.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity.* Oxford: Oxford University Press.

Vaughan, J., & Singer, R. (2018). Indigenous multilingualisms past and present. *Language and Communication* 62. 83–90.

Weinreich, U. (1953). *Languages in contact: Findings and problems.* Walter de Gruyter.

Yager, J. (submitted). Highly similar semantic systems become more similar in egalitarian bilingual foragers.

Yager, J., & Burenhult, N. (2017). Jedek: A newly discovered Aslian variety of Malaysia. *Linguistic Typology* 21(3). 493–545.

Yager, J., & Gullberg, M. Asymmetric semantic interaction in Jedek-Jahai bilinguals: Spatial language in a small-scale, non-standardized, egalitarian, long-term multilingual setting in Malaysia. *International Journal of Bilingualism.* doi:10.1177/1367006918814378. Published online by SAGE journals, May 6, 2019.

Appendix A: Basic vocabulary data

	Jedek2	Jedek34	Jedek5	Jedek31	Jedek32	Jedek38
hand	cyas	cas	cyas	cyas	cyas	cyas
arm	cyas	bliŋ	cyas	bliŋ	bliŋ	bliŋ
upper arm	?	?	?	klapəh	klapəh	klapəh
claw/nail	klkɔʔ	klkɔʔ cas	klkɔʔ cas	klkɔʔ cas	klkɔʔ cas	klkɔʔ cas
foot	can	can	can	can	can	can
(lower) leg	can	?	can	kteŋ	kteŋ	kteŋ
thigh	?	bliʔ	can	bliʔ	bliʔ	bliʔ
knee	kaltoŋ	kaltoŋ	kaltoŋ	kaltoŋ	kaltoŋ	kaltoŋ
head	kuy	kuy	kuy	kuy	kuy	kuy
ear	ʔnteŋ	ʔnteŋ	ʔnteŋ	ʔnteŋ	ʔnteŋ	ʔnteŋ
eye	mɛt	mɛt	mɛt	mɛt	mɛt	mɛt
nose	mɔh	mɔh	mɔh	mɔh	mɔh	mɔh
mouth	hãɲ	hãɲ	hãɲ	hãɲ	tnit	hãɲ
tongue	lntek	lntek	lntek	lntek	lntek	lntek
tooth (front)	hãɲ	hãɲ	hãɲ	hãɲ	hãɲ	hãɲ
hair	sɔk kuy	sɔk kuy	sɔk kuy	sɔk kuy	sɔk kuy	sɔk kuy
skin (of person)	ktiʔ	ktiʔ	ktiʔ	ktiʔ	ktiʔ	ktiʔ
neck	ʔŋit	ʔŋit	ʔŋit	ʔŋit	ʔŋit	ʔŋit
back of neck	?	?	?	tŋkɔk	tŋkɔk	tŋkɔk
breast	klaŋis	dkduk	ʔãm	ʔãm	cnʔah	ʔãm
back (of person)	krəʔ	krəʔ	krəʔ	krəʔ	krəʔ	krɔʔ
belly	ciŋ	ciŋ	ciŋ	ciŋ	ciŋ	ciŋ
bone	jʔeŋ	jʔeŋ	jʔeŋ	jʔeŋ	jʔeŋ	jʔeŋ
heart	klaŋes	klaŋes	klaŋes	klaŋes	klaŋes	klaŋes
liver	klaŋes	klaŋes	klaŋes	ros	klaŋes	klaŋes
guts	ʔɛcwɛ̃c	ʔɛcwɛ̃c	ʔɛc ciŋ	ʔɛc ciŋ	ktiʔ ciŋ	ʔɛc ciŋ
blood	ɲap	ɲap	ɲap	bhĩm	bhĩm	bhĩm
I	ʔiɲ	ʔiɲ	ʔiɲ	ʔiɲ	ʔiɲ	ʔiɲ
you (PL)	bɛʔ mnraʔ	bɛʔ	smpay	tigaʔ oraŋ	smpay	smpay
thou/you	bɛʔ	bɛʔ	bɛʔ	bɛʔ	bɛʔ	bɛʔ
he	ʔoʔ	ʔoʔ	ʔoʔ	ʔoʔ	ʔoʔ	ʔoʔ
we (incl.)	hiʔ	?	hiʔ	ʔapay	hiʔ	hiʔ
we (excl.)	hiʔ	ʔapay	ʔapay	ʔapay	ʔapay	ʔapay
they	ʔoʔ mnraʔ	gin	gin	gin	gin	gin
person	mnraʔ	?	mnraʔ	mnraʔ	mnraʔ	mnraʔ
child	wɔŋ kɛn	kɛn	wɔŋ kɛn	kɛn	kɛn	wɔŋ kɛn
woman	baboʔ	baboʔ	baboʔ	baboʔ	baboʔ	baboʔ
man (male)	tmkal	tmkal	tmkal	tmkal	tmkal	tmkal
mother	mɛ̃ʔ	naʔ	naʔ	naʔ	naʔ	naʔ
father	ʔɛy	ʔɛy	ʔɛy	ʔɛy	ʔɛy	ʔɛy
sibling, younger	bɛ	bɛ	bɛ	bɛ	bɛ	bɛ
wife	kneh	biʔ	kneh	kneh	kneh	kneh
husband	ksiy	ksiy	ksiy	ksiy	ksiy	ksiy
grandmother	yaʔ	naʔ	yaʔ	yaʔ	yaʔ	yaʔ
grandfather	taʔ	taʔ	taʔ	dɔn	dɔn	taʔ
great-grandparent	moyaŋ	moyaŋ	moyaŋ	moyaŋ	yik	dɔn
younger uncle	?	ʔɛy sudaraʔ	?	bɛh	bɛh	bɛh
older uncle	kiɲ	ʔɛy sudaraʔ	kiɲ	?	?	kiɲ
younger aunt	mɔʔ	naʔ sudaraʔ	?	mɔʔ	mɔʔ	mɔʔ
older aunt	?	naʔ sudaraʔ	toy	?	?	toy

siblings	ʔadiʔ bradiʔ	ʔadiʔ bradiʔ	smibɛ	bɛ	mirbɛr	ʔadiʔ bradiʔ
elder brother	toʔ	?	toʔ	toʔ	toʔ	toʔ
elder sister	toʔ	?	toʔ	toʔ	toʔ	toʔ
younger cousin	spupuʔ	spupuʔ	spupuʔ	bɛ sudaraʔ	bɛ sudaraʔ	sbradiʔ
older cousin	spupuʔ	spupuʔ	spupuʔ	bɛ sudaraʔ	bɛ sudaraʔ	sbradiʔ
niece/nephew	?	wɔŋ sudaraʔ	wɔŋ sudaraʔ	kmən	wɔŋ sdaraʔ	kmən
child	wɔŋ	wɔŋ	wɔŋ	wɔŋ	wɔŋ	wɔŋ
grandchild	cucu	kaɲcɔʔ	kaɲcɔʔ	kaɲcɔʔ	kaɲcɔʔ	kaɲcɔʔ
great-grandchild	cicit	?	cicit	cicit	cicit	dɔn
spouse's parent	knʔac	?	knʔac	knʔac	knʔac	knʔac
spouse's sibling, sibling's spouse	lamɛy	lamɛy	lamɛy	lamɛy	lamɛy	lamɛy
child's spouse's parents	besan	bisɛn	besan	bisɛn	bisɛn	bisɛn
wife's sister's husband/husband's brother's wife	?	?	?	maduʔ	biras	biras
child's spouse	mnsaw	wɔŋ mnantu	mnsaw	mnsaw	mnsaw	mnsaw
animal	binataŋ	?	binataŋ	mnataŋ	mnataŋ	mnataŋ
bird	kawaw	kawaw	kawaw	kawaw	kawaw	kawaw
dog	ʔasuʔ	ʔasuʔ	ʔasuʔ	ʔasuʔ	ʔasuʔ	ʔasuʔ
fish	ʔikan	ʔikan	ʔikan	ʔikan	ʔikan	ʔikan
snake	jkɔp	jkɔp	jkɔp	jkɔp	jkɔp	jkɔp
worm	caciŋ	caciŋ	caciŋ	caciŋ	caciŋ	caciŋ
louse	tomaʔ	tomaʔ	tomaʔ	tomaʔ	tomaʔ	tomaʔ
feather	bulu	sɔk	sɔk kawaw	sɔk	sɔk	sɔk
tail	hatɛ̃ʔ	?	hatɛ̃ʔ	hatɛ̃ʔ	hatɛ̃ʔ	hatɛ̃ʔ
wing	sayap	sayap	sayap	sayap	sayap	kɲɛŋ
egg	makɔʔ	makɔʔ	makɔʔ	makɔʔ	makɔʔ	makɔʔ
meat	dagiŋ	sɛc	sɛc	sɛc	sɛc	sɛc
fat (substance)	lʔəs	lʔəs	lmaʔ	toc	toc	toc
rice	nasiʔ	nasiʔ	nasiʔ	nasiʔ	nasiʔ	nasíʔ
salt	garam	gaʔəm	gaʔəm	garəm	garəm	garəm
knife	wɛ̃s	wɛ̃s	wɛ̃s	wɛ̃s	wɛ̃s	wɛ̃s
blowpipe	blaw	blaw	blaw	blaw	blaw	blaw
quiver	baniʔ	snlɔc dɔk	baniʔ	baniʔ	baniʔ	baniʔ
spear	?	rabuk	mataʔ	mataʔ	mataʔ	mataʔ
rope	talɛŋ	talɛŋ	talɛŋ	talɛŋ	ʔawɛy	talɛŋ
road	ha	?	ha	ha	ha	ha
woods	hɨp	hɨp	hɨp	hɨp	hɨp	hɨp
river	tɔm	tɔm	tɔm	tɔm	tɔm	tɔm
lake	tasiʔ	tasiʔ	tasíʔ	tasiʔ	tasiʔ	tasiʔ
mountain	gunuŋ	cbaʔ	gunuŋ	cbaʔ	cbaʔ	cbaʔ
tree	kayuʔ	kayuʔ	kayuʔ	phon	phon	phon kayuʔ
bark (of tree)	kulit kayuʔ	ktiʔ	ktiʔ kayuʔ	ktiʔ kayuʔ	ktiʔ kayuʔ	ktiʔ
flower	buɲəh	buɲaʔ	buɲaʔ	buɲaʔ	buɲaʔ	buɲaʔ
leaf	haliʔ	haliʔ	haliʔ	haliʔ	haliʔ	haliʔ
root	jʔes	ʔawɛy	jʔes	jʔes	ʔawɛy	ʔawɛy
fruit	kbɨʔ	kbɨʔ	kbɨʔ	kbɨʔ	kbɨʔ	kbɨʔ
seed	biji	wəh	mɛt	mɛt	mɛt	mɛt
rice (plant)	padiʔ	padiʔ	padiʔ	padiʔ	padiʔ	padiʔ
bamboo	buloʔ	buloʔ	buloʔ	buloʔ	buloʔ	buloʔ
stick (of wood)	kayuʔ	kayuʔ	tuŋkat	tnuŋkin	tuŋkat	?
stone	batuʔ	batuʔ	batuʔ	batuʔ	batuʔ	batuʔ
grass	rompot	rompot	rompot	rompot	rompot	rompot
earth (soil)	teʔ	teʔ	teʔ	teʔ	teʔ	teʔ

water	tɔm	tɔm	tɔm	tɔm	tɔm	tɔm
fire	ʔɔs	ʔɔs	ʔɔs	ʔɔs	ʔɔs	ʔɔs
smoke	ʔasap	ʔasɛp	ʔasɛp	ʔasɛp	ʔasɛp	ʔasɛp
ashes	ʔabuʔ	pip	ʔabuʔ	ʔabuʔ	pip	ʔabuʔ
dust	dbuʔ	?	ʔabuʔ	dbuʔ	haraŋ	?
wind	ʔaŋin	ʔaŋin	hlhul	ʔaŋin	ʔaŋin	ʔaŋin
cloud	kabut	kabut	kabut	kabut	kabut	ʔawãn
fog	kabut	?	kabut	?	?	kabut
sky	ktɔʔ	mɛt ktɔʔ	laŋit	mɛt ktɔʔ	laŋit	laŋit
sun	matəhari	mɛt ktɔʔ	mɛt ktɔʔ	mɛt ktɔʔ	mɛt ktɔʔ	mɛt ktɔʔ
moon	bulan	bulan	bulan	bulan	bulan	bulan
star	bintaŋ	bintaŋ	bintaŋ	bintaŋ	bintaŋ	mintãŋ
day (opp. of night)	ktɔʔ	?	ktɔʔ	hariʔ	hariʔ	hariʔ
night	gyə̃m	gyə̃m	gyə̃m	gyə̃m	gyə̃m	gyə̃m
year	tahun	tahun	tahun	tahun	tahun	tahun
sand	pasi	pasi	pasi	pasi	pasi	?
sea (ocean)	lawot	lawot	lawot	lawot	lawot	?
name	namɔ	knmɔh	knmɔh	knmɔh	knmɔh	knmɔh
good	bayik	?	ʔabən	ʔabən	ʔabən	ʔabən
bad (unsuitable)	blap ʔm=ʔabən	tidaʔ bayiʔ	blap ʔm=ʔabən	braʔ ʔm=ʔabən	blap ʔm=ʔabən	jbec
big	bəw	?	bəw	bəw	bəw	bəw
small	ʔhəy	knɔm	ʔhəy	wɔŋ	ʔhəy	ʔhəy
warm (weather)	bkit	bkit	bkit	bkit	bkit	bkit
cold (weather)	sjuʔ	blhɛ̃t	blhɛ̃t	blhɛ̃t	blhɛ̃t	sjuʔ
sharp	tajam	makan	tajam	makan	tajam	makan
dull (blunt)	blap ʔm=makan	?	blə̃n	blə̃n	blə̃n	blə̃n
near	pndɛh	pndəh	pndɛh	pndəh	pdəh	pdəh
far	mɲjiʔ	mɲjiʔ	mɲjiʔ	mɲjiʔ	mɲjiʔ	mɲjiʔ
heavy	hɲjut	brat	hɲjut	hɲjut	hɲjut	hɲjut
light (weight)	hatɔ̃p	riŋan	hampɔ̃l	hampɔ̃l	hampɔ̃l	hatɔ̃p
wide	lbar	lbar	mntec	bəw	tbɛl	?
narrow	suntoʔ	smpit	suntoʔ	kpic	nipis	?
long	mntec	btec	mntec	btec	btec	?
short	cnhɛ̃t	cnhãt	cnhɛ̃t	cnhə̃t	cnhə̃t	?
new	baruʔ	baruʔ	baruʔ	baruʔ	baruʔ	?
old	kbit	?	kbit	kbit	bakɛs	bidan
wet	pcãʔ	?	pcãʔ	pcãʔ	pcãʔ	pcãʔ
dry	kriŋ	kriŋ	kriŋ	kriŋ	kriŋ	bʔɔ̃ŋ
thick	tbɛl	?	tbɛl	tbɛl	tbɛl	tbɛl
thin	nipis	nipis	nipis	nipis	nipis	?
full	kɔm	?	sbem	kɔm	sbem	?
straight	lurus	?	btul	btul	btul	?
smooth	blac	?	blac	blac	blac	blac
dirty	jʔɔm	?	jbec	jʔɔm	jʔɔm	jʔɔm
rotten (of wood)	ʔnsɔʔ	ʔnsɔʔ	rput	?	hancɔ	?
right (correct)	btul	btul	btul	btul	btul	?
green	hijaw	hijaw	hijaw	blʔãɲ	hijaw	?
yellow	kuniŋ	kuniŋ	kuniŋ	kuniŋ	kuniŋ	?
black	hitam	gwɛ̃c	gwɛ̃c	gwɛ̃c	bltiŋ	?
white	puteh	puteh	puteh	bkɔŋ	puteh	?
red	merah	merah	merah	btɛh	btɛh	?
to walk	cip can	cip can	cip can	cip can	cip can	cip can
to fly	kapiy	kapiy	kapiy	kapiy	kapiy	kapiy

to swim	kicwac	brnaŋ	kicwac	kickac	brnaŋ	kayoh
to float	lampoŋ	ʔ	timbul	timbul	timbul	ʔ
to stand	hɲjaɲ	ʔ	hɲjaɲ	hɲjaɲ	hɲjaɲ	hɲjaɲ
to sit	ŋɔk	ʔ	ŋɔk	ŋɔk	ŋɔk	ŋɔk
to lie (on side)	pʔjiʔ	ʔ	pʔjiʔ	pʔjiʔ	tek	pʔjiʔ
to come	lɔ̃c	ʔ	lɔ̃c	lɛw	lɛw	ʔ
to turn (veer)	brilɛʔ	ʔŋat	pusiŋ	pusiŋ	pusiŋ	ʔ
to eat	ciʔ	ciʔ	ciʔ	ciʔ	ciʔ	ciʔ
to drink	ʔām	ʔām	ʔām	ʔām	ʔām	ʔām
to bite	kap	ʔ	kap	kap	tgiɲ	kap
to cut	kɛc	get	kɛc	kɛc	get	kɛc
to dig	bay	ʔawɛy	bay	bay	bay	bay
to hunt	buru	rabuk	buru	sam	sam	haluh
to kill	bunuh	bunuh	bunuh	cek	bunuh	ʔ
to shoot (blowpipe)	bdɛl	bdɛl	bdɛl	bdɛl	bdɛl	ʔ
to die	kmɔc	sarɔʔ	kmɔc	kmɔc	kmɔc	kmɔc
to live	gɔs	gɔs	gɔs	gɔs	gɔs	gɔs
to sleep	pʔjiʔ	pʔjiʔ	pʔjiʔ	pʔjiʔ	pʔjiʔ	pʔjiʔ
to wash (bathe)	ʔnay	ʔnay	ʔnay	ʔnay	ʔnay	ʔnay
to breathe	ʔ	nkjɨk	jkjɨk	jkjɨk	nkjɨk	jkjɨk
to vomit	kiʔ	ʔ	kiʔ	kiʔ	kiʔ	kiʔ
to suck	ʔisap	ʔisap	ʔisap	ʔisap	ʔisap	ʔisap
to hit	pukul	cɔ̃ŋ	pal	pal	pal	rpɛc
to hold	pgɛŋ	pgɛŋ	pgɛŋ	pgɛŋ	pgɛŋ	ʔ
to give	ʔɛk	bagiʔ	ʔɛk	ʔɛk	bagiʔ	ʔ
to split	blah	ʔ	blah	pɛk	blah	ʔ
to pull	tarik	tarik	tarik	tarik	tarik	ʔ
to push	toklɛk	ʔ	tolɔʔ	toklɛk	tolɛk	ʔ
to squeeze	hɔ̃t	ʔ	cpɛ̃t	cpɛ̃t	cpɛ̃t	ʔ
to stab	tikam	tikam	cek	cek	cek	ʔ
to spit	kbɛc	maɲ	kbɛc	kbɛc	kbɛc	ʔ
to throw	paŋkaʔ	baliŋ	paŋkaʔ	paŋkaʔ	puŋal	ʔ
to tie	ʔikɛt	ʔikɛt	ʔikɛt	ʔikɛt	ʔikɛt	ʔ
to rub	gosɔ̃ʔ	gosɔ̃ʔ	gosɔ̃ʔ	gosɔ̃ʔ	gosɔ̃ʔ	ʔ
to wipe	sapuh	sapuh	sapuh	sapuh	sapuh	ʔ
to scratch	kac	kac	kac	kac	kac	ʔ
to sew	jahit	jahit	jahit	jahit	jahit	ʔ
to laugh	lkluk	lkluk	lkluk	lkluk	lkluk	ʔ
to sing	pɲlɔɲ	pɲlɔɲ	pɲlɔɲ	pɲlɔɲ	pɲlɔɲ	ʔ
to dance	jogɛʔ	ʔ	kŋsiŋ	kŋsiŋ	kŋsiŋ	ʔ
to play	mayin	mayin	pnguh	mayin	mayin	ʔ
to fight	lawən	lawən	lawən	lawən	lawən	ʔ
to fear	ʔntiŋ	ʔntiŋ	ʔntiŋ	ʔntiŋ	ʔntiŋ	ʔ
to say	cɔl	cɔl	cɔl	cɔl	ckwik	ʔ
to hear	kjeŋ	ʔnteŋ	kjeŋ	kjeŋ	kjeŋ	ʔ
to smell	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔ
to know	ltʔet	ʔtʔet	ltʔet	ltʔet	ltʔet	ʔ
to count	bilaŋ	bilaŋ	bilaŋ	bilaŋ	bilaŋ	ʔ
to think	piki	biyət	piki	ʔiŋɛt	piki	ʔ
to fall (drop)	dahɛs	dahɛs	dahɛs	dahɛs	dahɛs	ʔ
to flow	ʔayũt	ʔ	ʔali	ʔayũt	ʔayũt	ʔ
to rain	hujan	hujan	hujan	hujan	hujan	ʔ
to burn	baka	tʔɔp	baka	cɔm	baka	ʔ
to blow (of wind)	bos	ʔ	hilhul	ʔaŋin bəw	hilhul	ʔ
to swell	ʔɔpɔ̃h	kmbaŋ	kmbaŋ	plkuc	kmbaŋ	ʔ

what?	cʔay	?	cʔay	cabap	caybap	?
when?	?	kaduy	bilaʔ	bilaʔ	bilaʔ	?
where?	ʔnah pān	ʔnah pān	ʔnah pān	ʔnah gɛl	ʔnah pān	?
who?	kɛn	?	makɛn	cʔay bap	makɛn	?
how?	maʔancin	?	maʔancin	cʔay bap	maʔancin	?
left (side)	kiri	wɛ̃ʔ	wɛ̃ʔ	wɛ̃ʔ	wɛ̃ʔ	?
right (side)	kanan	tem	tem	tem	tem	?
this	nuy	ʔə̃h	ʔũh	tudeh	ʔũh	?
that	tũn	tũn	tũn	taniʔ	tũn	?
here	ʔũh	ʔə̃h	ʔũh	ʔũh	ʔũh	?
there	tũn	?	taníʔ	taniʔ	tũn	?
many	kɔm	kɔm	kɔm	kɔm	kɔm	?
few, some	ʔhəy	ʔhəy	ʔhəy	ʔajɔʔ	ʔhəy	?
all	kɔm	kɔm	kɔm	kɔm	smwaʔ	?
and	dan	?	lɔʔ	dan	dan	?
because	sbap	sbap	sbap	cʔay bap	sbap	?
if	kaluʔ	kalaw	kaluʔ	kaluʔ	kaluʔ	?
not	blap	blap	blap	blap	blap	?
other	pãw	layin	pãw	layin	da=ʔasiŋ	?
at	?	?	lɛŋ	da=	da=	?
in	daləm	daləm	daləm	daləm	daləm	?
with	lɔʔ	?	lɔʔ	lɔʔ	lɔʔ	?
one	satuʔ	satuʔ	nay	nay	nay	?
two	dwaʔ	dwaʔ	dwaʔ	dwaʔ	dwaʔ	?
three	tigaʔ	tigaʔ	tigaʔ	tigaʔ	tigaʔ	?
four	ʔmpət	ʔmpət	ʔmpət	ʔmpət	ʔmpət	?
five	limaʔ	limaʔ	limaʔ	limaʔ	limaʔ	?

	Jedek20	**Jedek16**	**JRual13**	**JRual6**	**JBanun2**	**JBanun6**
hand	cyas	cyas	cyas	cyas	cyas	cyas
arm	bliŋ	bliŋ	bliŋ	bliŋ	bliŋ	cnreŋ
upper arm	bliŋ	klapəh	bliŋ	klapəh	?	?
claw/nail	klkɔʔ cyas	klkɔʔ cyas	klkɔʔ cyas	klkɔʔ cyas	cnrɔs	cnrɔs
foot	can	can	can	can	can	can
(lower) leg	kteŋ	kteŋ	kteŋ	bliʔ	?	?
thigh	bliʔ	bliʔ	bliʔ	bliʔ	bliʔ	bliʔ
knee	kaltoŋ	kaltoŋ	kaltoŋ	kaltoŋ	kaltoŋ	kaltoŋ
head	kuy	kuy	kuy	kuy	kuy	kuy
ear	ʔnteŋ	ʔnteŋ	ʔnteŋ	ʔnteŋ	ʔnteŋ	ʔnteŋ
eye	mɛt	mɛt	mɛt	mit	mit	mit
nose	məh	məh	məh	məh	məh	məh
mouth	hãŋ	hãŋ	hɛ̃ŋ	hɛ̃ŋ	hɛ̃ŋ	hɛ̃ŋ
tongue	lntek	lntek	lntek	lntek	lntek	lntek
tooth (front)	hãŋ	hãŋ	hɛ̃ŋ	hɛ̃ŋ	hɛ̃ŋ	hɛ̃ŋ
hair	sɔk kuy	sɔk kuy	sɔk kuy	sɔk kuy	sɔk	sɔk
skin (of person)	ktiʔ	ktiʔ	ktiʔ	ktiʔ	ktiʔ	ktiʔ
neck	ʔŋit	ʔŋit	ʔŋit	ʔŋit	ʔŋut	ʔŋit
back of neck	tŋkɔk	tŋkɔk	?	?	?	?
breast	ʔãm	ʔãm	ʔɛ̃m	ʔɛ̃m	dkduk	ʔɛ̃m
back (of person)	krəʔ	krəʔ	krəʔ	krəʔ	krɔʔ	krɔʔ
belly	ciŋ	ciŋ	ciŋ	ciŋ	ʔɛc	ʔɛc
bone	jʔeŋ	jʔeŋ	jʔeŋ	jʔeŋ	jʔeŋ	jʔeŋ
heart	klaŋes	klaŋes	klaŋis	klaŋis	klaŋis	klaŋis
liver	?	klaŋes	ros	klaŋis	ros	ros

guts	ʔɛc ciŋ	ʔɛcwɛ̃c	ʔɛc ciŋ	ʔɛc	ʔɛcwĩc	ʔɛcwĩc
blood	bhĩm	bhĩm	bhĩm	darah	darah	darah
I	ʔiŋ	ʔiŋ	yɛʔ	yɛʔ	yɛʔ	yɛʔ
you (PL)	smpay	smpay	gin	gin	gin	gin
thou/you	bɛʔ	bɛʔ	mɔh	mɔh	mɔh	mɔh
he	ʔoʔ	ʔoʔ	ʔoʔ	ʔoʔ	ʔoʔ	ʔoʔ
we (incl.)	hiʔ	hiʔ	heʔ	heʔ	heʔ	heʔ
we (excl.)	ʔapay	?	yapɛ̃h	?	yapɛ̃h	yapɛ̃h
they	gin	?	gin	?	gin	gin
person	mnraʔ	mnraʔ	?	mnraʔ	mnraʔ	mnraʔ
child	kɛn	kɛn	wɔŋ kɛn	wɔŋ kɛn	kɛn	kɛn
woman	baboʔ	baboʔ	baboʔ	baboʔ	baboʔ	baboʔ
man (male)	tmkal	tmkal	tmkal	tmkal	tmkal	tmkal
mother	naʔ	naʔ	biʔ	biʔ	biʔ	biʔ
father	ʔɛy	ʔɛy	ʔɛy	ʔɛy	ʔɛy	ʔɛy
sibling, younger	bɛ	bɛ	bɛr	bɛr	bɛr	bɛr
wife	kneh	kneh	kneh	kneh	kneh	kneh
husband	ksiy	ksiy	ksiy	ksiy	ksiy	ksiy
grandmother	yaʔ	yaʔ	yaʔ	yaʔ	yaʔ	yaʔ
grandfather	taʔ	taʔ	taʔ	taʔ	taʔ	toy
great-grandparent	dɔn	dɔn	dɔn	dɔn	dɔn	dɔn
younger uncle	bɛh	bɛh	bɛh	bɛh	bɛh	bɛh
older uncle	kiŋ	kiŋ	kiŋ	kiŋ	?	?
younger aunt	mɔʔ	mɔʔ	mɔʔ	mɔʔ	biʔ mɔʔ	mɔʔ
older aunt	toy	toy	toy	toy	?	?
siblings	smibɛ	?	smɛrbɛr	bɛr	smrbɛr	bɛr samɛʔ
elder brother	toʔ	toʔ	pɛ̃ʔ	pɛ̃ʔ	pɛ̃ʔ	pɛ̃ʔ
elder sister	toʔ	toʔ	pɛ̃ʔ	pɛ̃ʔ	pɛ̃ʔ	pɛ̃ʔ
younger cousin	bɛ	?	bɛr	bɛr	smrwɔŋ	bɛr
older cousin	toʔ	?	?	pɛ̃ʔ	smrwɔŋ	pɛ̃ʔ
niece/nephew	kmən	wɔŋ sudaraʔ	wɔŋ	kmən	kaɲcɔʔ	wɔŋ
child	wɔŋ	wɔŋ	wɔŋ	wɔŋ	wɔŋ	wɔŋ
grandchild	kaɲcɔʔ	kaɲcɔʔ	kaɲcɔʔ	kaɲcɔʔ	kaɲcɔʔ	kaɲcɔʔ
great-grandchild	kaɲcɔʔ	kaɲcɔʔ	cucit	kaɲcɔʔ	?	kaɲcɔʔ
spouse's parent	knʔac	knʔac	knʔac	knʔac	knʔac	knʔac
spouse's sibling, sibling's spouse	lamɛy	lamɛy	lamɛy	lamiy	lamiy	lamiy
child's spouse's parents	bisɛn	bisɛn	?	?	bisɛn	bisɛn
wife's sister's husband/husband's brother's wife	lamɛy	lamɛy	?	?	pɛ̃ʔ	?
child's spouse	mnsaw	mnsaw	mnsaw	mnsaw	mnsaw	mnsaw
animal	binataŋ	binataŋ	binataŋ	hɛywan hip	ʔay	ʔay
bird	kawaw	kawaw	kawɔ̃t	kawaw	kawɔ̃t	kawɔ̃t
dog	ʔasuʔ	ʔasuʔ	ʔɔ̃t	ʔasuʔ	ʔɔ̃t	ʔɔ̃t
fish	ʔikan	ʔikan	ʔikəʔ	ʔikan	ʔikəʔ	ʔikəʔ
snake	jkɔp	jkɔp	tajuʔ	jkɔp	tajuʔ	tajuʔ
worm	taciŋ	caciŋ	caciŋ	caciŋ	caciŋ	caciŋ
louse	tomaʔ	tomaʔ	ciʔ	ciʔ	ciʔ	ciʔ
feather	sɔk kawaw	sɔk kawaw	sɔk kawɔ̃t	sayap	sɔk kawɔ̃t	sayap
tail	hatɛ̃ʔ	hatɛ̃ʔ	hatĩʔ	hatĩʔ	hatĩʔ	hatĩʔ
wing	sayap	sayap	sayap	sayap	sayap	sayap
egg	makɔʔ	makɔʔ	makɔʔ	makɔʔ	ktit	ktit
meat	sɛc	dagiŋ	sɛc	sɛc	sɛc	sɛc
fat (substance)	lʔəs	lʔəs	lʔəs	lʔəs	lʔəs	lʔəs

rice	nasiʔ	nasiʔ	nasiʔ	nasiʔ	nasiʔ	nasiʔ
salt	gaʔəm	gaʔəm	gaʔəm	garəm	gaʔəm	garəm
knife	wɛ̄s	wɛ̄s	tajiʔ	wɛ̄s	tajiʔ	tajiʔ
blowpipe	blaw	blaw	blaw	blaw	blaw	blaw
quiver	baniʔ	baniʔ	baniʔ	baniʔ	baniʔ	baniʔ
spear	mataʔ	mataʔ	mataʔ	mataʔ	mataʔ	mataʔ
rope	talɛŋ	talɛŋ	ʔawɛy	talɛŋ	ʔawɛy	ʔawɛy
road	ha	ha	jalən	jalən	jalən	glɔŋ
woods	?	hip	hip	hip	hip	hip
river	tɔm	tɔm	tɔm	tɔm	tɔm	tɔm
lake	tasiʔ	tasiʔ	tasiʔ	tasiʔ	paŋ	paŋ
mountain	cbaʔ	gunuŋ	cbaʔ	gunuŋ	pɔʔ	tul jlmɔl
tree	phon kayuʔ	kayuʔ	tom kayuʔ	phon kayuʔ	jhūʔ	jhūʔ
bark (of tree)	ktiʔ kayuʔ	ktiʔ kayuʔ	ktiʔ	ktiʔ	ktiʔ	ktiʔ jhūʔ
flower	buŋaʔ	buŋaʔ	buŋɛʔ	buŋaʔ	bkaw	buŋɛʔ
leaf	haliʔ	haliʔ	haliʔ	haliʔ	haliʔ	haliʔ
root	jʔes	jʔes	ʔawɛy	ʔawɛy	jʔes	jʔes
fruit	kbiʔ	kbiʔ	kmɔʔ	kbiʔ	bɔh	kbiʔ
seed	mɛt	biji	mit	mit	mit	mit
rice (plant)	padiʔ	padiʔ	padɛy	padiʔ	padɛy	padɛy
bamboo	buloʔ	buloʔ	?	buloʔ	buloʔ	buloʔ
stick (of wood)	tnuŋkat	tuŋkat	tnuŋkin	tnuŋkat	ʔat	ʔat
stone	batuʔ	batuʔ	batuʔ	batuʔ	batuʔ	batuʔ
grass	rompot	rompot	rompot	rompot	rompot	rompot
earth (soil)	teʔ	teʔ	teʔ	teʔ	teʔ	teʔ
water	tɔm	tɔm	tɔm	tɔm	tɔm	tɔm
fire	ʔɔs	ʔɔs	ʔɔs	ʔɔs	ʔɔs	ʔɔs
smoke	ʔasɛp	ʔasɛp	ʔyʔey	ʔasɛp	ʔyʔey	ʔyʔey
ashes	ʔabuʔ	ʔabuʔ	ʔabuʔ	ʔabuʔ	ʔabuʔ	ʔabuʔ
dust	?	ʔabuʔ	ʔabuʔ	ʔabuʔ	habuk	ʔabuʔ
wind	ʔaŋin	ʔaŋin	ʔaŋin	ʔaŋin	bgiw	bgiw
cloud	kabut	kabut	kabut	kabut	sagup	sagup
fog	?	?	?	?	sagup	sagup
sky	laŋit	mɛt ktɔʔ	mit ktɔʔ	ktɔʔ	ktɔʔ	mit ktɔʔ
sun	mɛt ktɔʔ	mɛt ktɔʔ	mit ktɔʔ	matəhari	mit ktɔʔ	kit ktɔʔ
moon	bulan	bulan	bulan	bulan	bulan	bulan
star	bintaŋ	bintaŋ	bintaŋ	bintaŋ	bintɛŋ	bintɛŋ
day (opp. of night)	hariʔ	hariʔ	hariʔ	ktɔʔ	ktɔʔ	peŋ
night	gyə̄m	gyə̄m	hrkit	gyə̄m	hrkit	hrkit
year	tahun	tahun	cnrɔy	tahun	tawon	tawon
sand	pasi	pasi	pasir	pasi	pantɛy	pasir
sea (ocean)	lawot	lawot	lawot	lawot	lawot	lawot
name	klmɔh	namɔ	knmɔh	knmɔh	knmɔh	knmɔh
good	ʔabən	ʔabən	btʔɛt	ʔabən	btʔɛt	btʔɛt
bad (unsuitable)	jbec	jbec	jbec	jbec	braʔ wa=btʔɛt	braʔ wa=btʔɛt
big	bəw	bəw	bəw	bəw	bəw	bəw
small	ʔhəy	ʔajɔʔ	ʔhəy	ʔhəy	ʔajɔʔ	ʔajɔʔ
warm (weather)	bkit	bkit	bkit	bkit	bkit	bkit
cold (weather)	sjuʔ	sjuʔ	sjuʔ	blhɛ̄t	sjuʔ	sjuʔ
sharp	makan	tajam	cmɛʔ	makan	cmɛʔ	cmɛʔ
dull (blunt)	blān	blān	blān	tumpul	blān	tlkūh
near	pndəh	pndəh	pndəh	pndɛh	pdəh	pndəh
far	mɲjiʔ	mɲjiʔ	mɲjiʔ	mɲjiʔ	mɲjiʔ	mɲjiʔ
heavy	hɲjut	hɲjut	hɲjut	hɲjut	hɲjut	hɲjut

light (weight)	hampõl	hampõl	hampõl	hampõl	haɲɔʔ	hatõp
wide	lwəs	bəw	tbɛl	tbɛl	bəw	bəw
narrow	suntoʔ	nipis	hrtẽl	nipis	lipis	simpet
long	mntec	btec	mntec	mntec	btec	btec
short	cnhẽt	cnhẽt	cnhə̃t	cnhẽt	cnhə̃t	cnhə̃t
new	baruʔ	baruʔ	baruʔ	baruʔ	baruʔ	baruʔ
old	manah	manah	manɛh	manɛh	manɛh	manɛh
wet	pcãʔ	pcãʔ	pcɛ̃ʔ	pcɛ̃ʔ	pcɛ̃ʔ	pcɛ̃ʔ
dry	kriŋ	kriŋ	kriŋ	kriŋ	bʔõŋ	kriŋ
thick	lwəs	tbɛl	?	?	tbɛl	tbɛl
thin	nipis	?	?	?	nipis	phɛr
full	sbem	kɔm	sbem	kɔm	sbem	sbem
straight	btul	btec	btul	lurus	btul	btul
smooth	blac	blac	bclac	krcə̃p	bclac	bclac
dirty	jʔɔm	jbec	jʔɔm	jʔɔm	jʔɔm	brkyɛk
rotten (of wood)	sɔʔ	rapuh	ʔnsɔʔ	ʔnsɔʔ	ʔnsɔʔ	ʔnsɔʔ
right (correct)	btul	btul	btul	?	btul	btul
green	hijaw	hijaw	hijaw	hija	hijɔw	hijɔw
yellow	kuniŋ	kuniŋ	kuniŋ	kuniŋ	kuniŋ	kuniŋ
black	gwɛ̃c	gwɛ̃c	gwɛ̃c	gwɛ̃c	gcĩh	gcĩh
white	?	bkɔɲ	puteh	puteh	puteh	puteh
red	brtɛh	brtɛh	merah	merah	merah	rhik
to walk	cip can	cip can	cip can	cip can	cip can	cip can
to fly	kapiy	kapiy	kapiy	kapiy	kapiy	kapiy
to swim	kicwac	kicwac	kicwac	brnaŋ	kayoh	syər
to float	ʔayũt	timbul	lampoŋ	timbul	timəl	bliwĩs
to stand	hɲjaɲ	hɲjaɲ	hɲjaɲ	hɲjaɲ	hɲjaɲ	hɲjaɲ
to sit	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk
to lie (on side)	tek	pʔjiʔ	pʔjiʔ	pʔjiʔ	clŋhĩŋ	tek
to come	lɛw	cip	cip	cip	cip	dɛs
to turn (veer)	pusiŋ	pusiŋ	pusiŋ	pusiŋ	brilɛʔ	pusiŋ
to eat	ciʔ	ciʔ	gey	gey	gey	gey
to drink	ʔãm	ʔãm	ʔɛ̃m	ʔɛ̃m	ʔɛ̃m	ʔɛ̃m
to bite	kap	kap	kap	kap	kap	kap
to cut	katəm	katəm	katəm	katəm	get	katəm
to dig	bay	bay	bay	bay	bay	bay
to hunt	sam	hamɛt	buru	bdɛl	sam	sam
to kill	bunuh	bunuh	rmpət	bdɛl	tbɔh	cek
to shoot (blowpipe)	bdɛl	bdɛl	bdɛl	bdɛl	bdɛl	bdɛl
to die	kmɔc	kmɔc	ʔnsɔʔ	ʔnsɔʔ	kbis	hapaʔ
to live	gɔs	gɔs	gɔs	gɔs	gɔs	gɔs
to sleep	pʔjiʔ	pʔjiʔ	pʔjiʔ	pʔjiʔ	tek	tek
to wash (bathe)	ʔnay	ʔnay	ʔnay	ʔnay	ʔnay	ʔnay
to breathe	nkjik	tarik klaŋes	jkjik	jkjik	jkjik	jkjik
to vomit	kiʔ	kiʔ	kiʔ	kiʔ	kiʔ	kiʔ
to suck	ʔisap	jɔt	sruc	sruc	rhə̃t	krhak
to hit	pal	pal	pal	pal	tbɔh	tbɔh
to hold	pgɛŋ	pgɛŋ	pgɛŋ	pgɛŋ	pgɛŋ	pgɛŋ
to give	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
to split	blah	pɛk	pɛk	pɛk	pɛk	pɛk
to pull	keŋ	tarik	tarik	tarik	keŋ	keŋ
to push	tolɔʔ	toklɛk	tolɔʔ	tolɔʔ	tolɛk	tolɛk
to squeeze	cpẽt	cpẽt	cpĩt	pulas	cpĩt	cpĩt
to stab	cek	cek	cek	cek	cek	cek
to spit	?	kbɛc	kbɛc	khɛʔ	kbɛc	kbɛc

to throw	paŋkaʔ	paŋkaʔ	paŋkaʔ	paŋkaʔ	bdal	hakɔk
to tie	ʔikɛt	ʔikɛt	ʔikɛt	ʔikɛt	rbɛt	ʔikɛt
to rub	gosɔ̃ʔ	gosɔ̃ʔ	gosɔ̃ʔ	gosɔ̃ʔ	jit	gosɔ̃ʔ
to wipe	sapuh	hilap	sapuh	sapuh	sapuh	sapuh
to scratch	kac	kac	kac	kac	kac	kac
to sew	jahit	jahit	jahit	jahit	jayit	jajit
to laugh	lkluk	lkluk	sitkət	lkluk	lkluk	lkluk
to sing	pɲlɔɲ	pɲlɔɲ	pɲlɔɲ	pɲlɔɲ	pɲlɔɲ	pɲlɔɲ
to dance	jogɛʔ	jogɛʔ	kŋsiŋ	jogɛʔ	sisɛʔ	sisɛʔ
to play	mayin	pnguh	mayin	mnmin	mnmin	mnmin
to fight	balɔh	balɔh	guhcɔh	guhcɔh	tumoʔ	cipcɛp
to fear	ʔntiŋ	ʔntiŋ	ʔntiŋ	ʔntiŋ	hgik	hgik
to say	cɔl	cɔl	cɔl	cɔl	cɔl	pŋsɛŋ
to hear	kjeŋ	kjeŋ	kjeŋ	kjeŋ	kjeŋ	kjeŋ
to smell	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔɔ̃ɲ	ʔɔ̃ɲ
to know	ltʔet	ltʔet	ltʔet	ltʔet	ʔtʔet	ʔtʔet
to count	bilaŋ	bilaŋ	bilaŋ	bilɛŋ	bilɛŋ	bilɛŋ
to think	piki	piki	piki	pikir	nim	pikir
to fall (drop)	dahɛs	dahɛs	rɛs	tbəl	tbəl	rɛs
to flow	ʔayũt	ʔayũt	wet	mŋalir	wet	wet
to rain	hujan	hujan	hĩc	hĩc	hĩc	hĩc
to burn	baka	cos	tʔɔp	syiŋ	syiŋ	syiŋ
to blow (of wind)	hilhul	hilhul	pit	thup	kapiy	pit
to swell	kmbaŋ	kmbaŋ	kmbaŋ	kmbaŋ	kmbaŋ	kmbaŋ
what?	cabap	cabap	mɛy ʔayoʔ	mamɛy	lbah	mamɛy
when?	bilaʔ	bilaʔ	mapuʔ bilaʔ	bilaʔ	mapuʔ	mapuʔ
where?	ʔnah pãn	ʔnah pãn	ma=teʔ ba=ʔɛ̃n	teʔ ba=ʔɛ̃n	ma=teʔ	ma=teʔ
who?	makɛn	makɛn	makɛn	kɛn	makɛn	makɛn
how?	maʔancin	paʔancin	maʔacin	maʔancin	maʔacin	maʔacin
left (side)	wɛ̃ʔ	ʔndaŋ	wɛ̃ʔ	wĩʔ	wĩʔ	wĩʔ
right (side)	tem	ʔadeh	tem	tem	tem	tem
this	tudeh	tudeh	tãh	tãh	tãh	ʔãh
that	tũn	ton	taniʔ	tũn	taniʔ	ton
here	ʔũh	ʔũh	pdəh	ʔãh	ʔãh	ʔãh
there	taniʔ	ʔaniʔ	taniʔ	ʔaniʔ	ʔaniʔ	ʔaniʔ
many	kɔm	kɔm	kɔm	kɔm	kɔm	kɔm
few, some	ʔhəy	ʔajɔʔ	ʔajɔʔ	ʔhəy	ʔajɔʔ	ʔhəy
all	kɔm	kɔm	kɔm	kɔm	kɔm	smwɛʔ
and	ʔalɔʔ	lɔʔ	ʔalɔʔ	lɔʔ	?	?
because	sbap	sbap	sbap	sbap	sbap	sbap
if	kaluʔ	kaluʔ	kaluʔ	kaluʔ	kalɔw	kalɔw
not	blap	blap	braʔ	blap	braʔ	braʔ
other	pãw	layin	layin	layin	pɛ̃w	pɛ̃w
at	da=	da=	k=la=	la=	k=	k=la=
in	?	daləm	daləm	?	klɛŋ	klɛŋ
with	lɔʔ	lɔʔ	ʔalɔʔ	lɔʔ	samɛʔ	dŋɛn
one	nay	nɛy	nɛy	nɛy	nɛy	nɛy
two	dwaʔ	dwaʔ	dwaʔ	dwaʔ	dwaʔ	dwaʔ
three	tigaʔ	tigaʔ	tigaʔ	tigaʔ	tigaʔ	tigaʔ
four	ʔmpət	ʔmpət	ʔmpət	ʔmpət	ʔmpat	ʔmpat
five	limaʔ	limaʔ	limaʔ	limaʔ	limɛʔ	limɛʔ

Appendix B: Topological relations data

	Jedek 1	Jedek3	Jedek2	Jedek13	Jedek6	Jedek5	Jedek37	Jedek27
1	lɛŋ	lɛŋ	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	k=ʔatɛs	klɛŋ
2	daləm	lɛŋ	lɛŋ	daləm	daləm	daləm	daləm	klɛŋ
3	lɛŋ	klɛŋ	ʔatɛs	lɛŋ	k=ʔatɛs	lɛŋ	?	klɛŋ
4	lɛŋ	klɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ tŋah	klɛŋ
5	lɛŋ	klɛŋ	ʔatɛs	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
6	lɛŋ	pndəh	?	clah	clah	lɔʔ	klɛŋ	klɛŋ
7	lɛŋ	k=ʔatɛs	ʔatɛs	lɛŋ	ʔatɛs	ʔatɛs	klɛŋ	klɛŋ
8	lɛŋ	k=ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	klɛŋ	klɛŋ
9	lɛŋ	ʔatɛs	lɛŋ	lɛŋ	lɛŋ	ʔatɛs	klɛŋ	klɛŋ
10	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɔʔ	lɛŋ	klɛŋ
11	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
12	lɛŋ	klɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
13	lɛŋ	lɛŋ	?	ʔatɛs	?	ʔatɛs	?	ʔatɛs
14	lɛŋ	lɛŋ	daləm	daləm	lɛŋ	lɛŋ daləm	?	daləm
16	lɛŋ kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	klɛŋ
17	tŋah	lɛŋ	ʔatɛs	lɛŋ	lɛŋ kyəm	ʔatɛs	klɛŋ	klɛŋ
19	lɛŋ	klɛŋ	lɛŋ	lɛŋ	ʔatɛs	ʔatɛs	klɛŋ	klɛŋ
20	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
21	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
23	lɛŋ	klɛŋ	lɛŋ	ʔatɛs	ʔatɛs	ʔatɛs	kyəm	ʔatɛs
24	kyəm	?	daləm	kyəm	kyəm	kyəm	?	kyəm
25	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	ʔatɛs
27	lɛŋ	k=ʔatɛs	ʔatɛs	klɛŋ	lɛŋ	ʔatɛs	kyəm	kyəm
29	lɛŋ	lɛŋ	ʔatɛs	klɛŋ	ʔatɛs	ʔatɛs	klɛŋ	da=
30	lɛŋ	?	lɛŋ	?	lɛŋ	klɛŋ	?	klɛŋ
31	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm
32	lɛŋ	klɛŋ	daləm	daləm	lɛŋ	daləm	daləm	klɛŋ
34	ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs
35	lɛŋ	samaʔ pdeʔ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
36	ʔatɛs	klɛŋ	lɛŋ ʔatɛs	ʔatɛs	pndəh	lɛŋ	kyəm	klɛŋ
37	lɛŋ	lɛŋ	lɛŋ	da=	lɛŋ	lɛŋ	?	klɛŋ
38	lɛŋ	clah	clah	clah	lɛŋ	?	clah	klɛŋ
39	lɛŋ	klɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	da=	klɛŋ
40	lɛŋ	klɛŋ	ʔatɛs	ʔatɛs	lɛŋ	lɛŋ	daləm	klɛŋ
41	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	?	klɛŋ
42	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
43	lɛŋ	ʔatɛs	ʔatɛs	lɛŋ	lɛŋ	ʔatɛs	kyəm	klɛŋ
44	lɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
45	daləm	ʔatɛs	lɛŋ	lɛŋ	k=ʔatɛs	ʔatɛs	k=ʔatɛs	klɛŋ
46	lɛŋ	lɛŋ	ʔatɛs	lɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
47	daləm	klɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
48	lɛŋ	?	lɛŋ	lɛŋ	lɛŋ	lɛŋ	k=la=	klɛŋ
49	lɛŋ	pndəh	ʔatɛs	clah	pndəh	clah	lɛŋ	clah
50	daləm	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
51	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
52	lɛŋ	ʔatɛs	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
53	kyəm	kyəm	kyəm	kyəm	lɛŋ	lɛŋ kyəm	k=ʔatɛs	kyəm
54	daləm	lɛŋ	daləm	daləm	lɛŋ	daləm	daləm	klɛŋ
55	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	kyəm	klɛŋ
58	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	?	?

59	lɛŋ ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ	klɛŋ
60	daləm	?	daləm	daləm	daləm	daləm	klɛŋ	lɛŋ
63	lɛŋ	ʔatɛs	ʔatɛs	?	?	ʔatɛs	?	klɛŋ
64	ʔndaŋ	ʔndaŋ	lɛŋ	lɛŋ	lɛŋ ʔndaŋ	ʔndaŋ	da=	klɛŋ
65	ʔatɛs	lɛŋ	lɛŋ	ʔatɛs	ʔatɛs	ʔatɛs	klɛŋ	klɛŋ
67	daləm	klɛŋ	lɛŋ	daləm	lɛŋ	daləm	klɛŋ	klɛŋ
68	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
69	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
71	daləm	daləm	daləm	daləm	lɛŋ	lɛŋ	lɛŋ	daləm

	Jedek29	**Jedek17**	**Jedek35**	**Jedek14**	**Jedek30**	**Jedek21**	**Jedek32**	**Jedek25**
1	ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ	lɛŋ	klɛŋ ʔatɛs	ʔatɛs	ʔatɛs
2	daləm	klɛŋ	daləm	lɛŋ	lɛŋ	klɛŋ	daləm	lɛŋ
3	klɛŋ	klɛŋ	daləm	?	?	lɛŋ	ʔatɛs	?
4	klɛŋ samaʔ tŋah	klɛŋ	?	lɛŋ	?	klɛŋ	k=tŋah	klɛŋ tŋah tŋah
5	ʔatɛs	klɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	?
6	clah	klɛŋ	?	ʔndaŋ	lɛŋ	klɛŋ	clah	da=
7	lɛŋ	ʔatɛs	?	ʔatɛs	lɛŋ	klɛŋ	clah	ʔatɛs
8	ʔatɛs	ʔatɛs	k=ʔatɛs	klɛŋ	lɛŋ	klɛŋ	ʔatɛs	ʔatɛs
9	lɛŋ	klɛŋ	?	lɛŋ	?	klɛŋ	klɛŋ	klɛŋ
10	lɛŋ	klɛŋ	?	lɛŋ	lɛŋ	lɛŋ	klɛŋ	?
11	klɛŋ	klɛŋ	?	lɛŋ	lɛŋ	klɛŋ	ʔatɛs	tŋah tŋah
12	klɛŋ	klɛŋ	klɛŋ	lɛŋ	?	klɛŋ	lɛŋ	klɛŋ
13	ʔatɛs	?	k=ʔatɛs	?	ʔatɛs	klɛŋ	ʔatɛs	?
14	daləm	daləm	daləm	daləm	daləm	daləm	daləm	klɛŋ
16	?	kyəm	kyəm	?	kyəm	kyəm	kyəm	kyəm
17	kyəm tŋah tŋah	klɛŋ	tŋah	lɛŋ	klɛŋ	klɛŋ	clah	tpiʔ
19	klɛŋ	ʔatɛs	klɛŋ	ʔatɛs	klɛŋ	lɛŋ	ʔatɛs	klɛŋ
20	klɛŋ	klɛŋ	?	lɛŋ	lɛŋ	lɛŋ	lɛŋ	da=sɛŋ
21	lɛŋ	klɛŋ	?	lɛŋ	lɛŋ	lɛŋ	daləm	lɛŋ
23	klɛŋ	ʔatɛs	ʔatɛs	lɛŋ	lɛŋ	klɛŋ	lɛŋ	ʔatɛs
24	kyəm	kyəm	?	daləm	klɛŋ	?	clah	kyəm
25	klɛŋ	klɛŋ	?	lɛŋ	lɛŋ	lɛŋ	ʔatɛs	klɛŋ
27	klɛŋ kyəm	klɛŋ	k=tŋah	ʔatɛs	lɛŋ	lɛŋ	kyəm	kyəm
29	k=ʔatɛs	ʔatɛs	?	lɛŋ	klɛŋ	lɛŋ	clah	ʔatɛs
30	tŋah	?	?	?	lɛŋ	?	?	tŋah
31	kyəm	kyəm	kyəm	kyəm	kyəm	klɛŋ kyəm	kyəm	kyəm
32	daləm	daləm	klɛŋ	daləm	klɛŋ	klɛŋ	daləm	lɛŋ
34	lɛŋ	klɛŋ	ʔatɛs	lɛŋ	lɛŋ	ʔatɛs	ʔatɛs	ʔatɛs
35	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
36	ʔatɛs	ʔatɛs	?	ʔatɛs	?	?	ʔatɛs	?
37	lɛŋ	klɛŋ	?	lɛŋ	lɛŋ	klɛŋ	kyəm	lɛŋ
38	clah	can=ʔndaŋ	?	clah	lɛŋ	?	clah	?
39	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ
40	klɛŋ	ʔatɛs	klɛŋ	ʔatɛs	klɛŋ	klɛŋ	ʔatɛs	klɛŋ
41	klɛŋ	klɛŋ	?	lɛŋ	lɛŋ	klɛŋ	ʔatɛs	klɛŋ
42	k=	klɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ
43	klɛŋ	ʔatɛs	klɛŋ	ʔatɛs	klɛŋ	klɛŋ	ʔatɛs	ʔatɛs
44	klɛŋ	clah	klɛŋ	lɛŋ	lɛŋ	klɛŋ	clah	lɛŋ
45	klɛŋ	ʔatɛs	?	ʔatɛs	lɛŋ	klɛŋ	ʔatɛs	klɛŋ
46	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	lɛŋ
47	klɛŋ	klɛŋ	klɛŋ	ʔatɛs	klɛŋ	klɛŋ	klɛŋ	klɛŋ
48	klɛŋ	?	?	lɛŋ	lɛŋ	klɛŋ	clah	klɛŋ

49	?	clah	clah	clah	klɛŋ	klɛŋ	clah	clah
50	klɛŋ	klɛŋ	?	lɛŋ	lɛŋ	?	kyəm	klɛŋ
51	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	lɛŋ
52	klɛŋ	klɛŋ	?	lɛŋ	klɛŋ	klɛŋ	klɛŋ	lɛŋ
53	kyəm	kyəm	kyəm	kyəm	lɛŋ	lɛŋ	kyəm	kyəm
54	daləm	daləm	daləm	daləm	lɛŋ	klɛŋ	daləm	daləm
55	lɛŋ	klɛŋ	klɛŋ kyəm	lɛŋ	klɛŋ	klɛŋ	klɛŋ	tŋah tŋah
58	lɛŋ	clah	?	lɛŋ	lɛŋ	lɛŋ	clah	ʔndaŋ
59	lɛŋ ʔatɛs	ʔatɛs	k=ʔatɛs	lɛŋ	ʔatɛs	klɛŋ	ʔatɛs	ʔatɛs
60	tŋah kliliŋ	daləm	?	daləm	daləm	k=daləm	daləm	daləm
63	?	klɛŋ ʔatɛs	?	?	?	klɛŋ	kyəm	?
64	klɛŋ krəʔ	ʔndaŋ	?	ʔndaŋ	kyɔm	?	krəʔ	ʔndaŋ
65	klɛŋ	ʔatɛs	k=ʔatɛs	ʔatɛs	ʔatɛs	?	ʔatɛs	ʔatɛs
67	daləm	daləm	k=daləm	daləm	klɛŋ	klɛŋ	lɛŋ	k=daləm
68	klɛŋ	klɛŋ	?	?	?	lɛŋ	klɛŋ	da=
69	lɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ
71	daləm	daləm	daləm	daləm	daləm	klɛŋ	daləm	daləm

	Jedek20	**Jedek16**	**Jedek26**	**Jedek28**	**Jedek36**	**Jedek23**	**JRual18**	**JRual7**
1	ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ	klɛŋ	lɛŋ	klɛŋ	krpiŋ
2	daləm	daləm	klɛŋ	lɛŋ	daləm	daləm	klɛŋ	klɛŋ
3	?	ʔatɛs	?	ʔatɛs	?	da=	lɛŋ	klɛŋ
4	?	klɛŋ	klɛŋ	lɛŋ	?	lɛŋ	klɛŋ	klɛŋ
5	ʔatɛs	lɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
6	clah	lɛŋ	klɛŋ	clah	da=ʔndaŋ	lɛŋ	daləm	clah
7	lɛŋ	?	ʔatɛs	lɛŋ	ʔatɛs	ʔatɛs	lɛŋ	ʔatɛs
8	ʔatɛs	ʔatɛs	ʔatɛs	?	?	lɛŋ	klɛŋ	krpiŋ
9	lɛŋ	ʔatɛs	k=ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ	klɛŋ	lɛŋ ʔatɛs
10	?	lɛŋ	lɛŋ	klɛŋ	samaʔ pdeʔ	lɛŋ	klɛŋ	klɛŋ
11	lɛŋ	lɛŋ	lɛŋ	lɛŋ	?	lɛŋ	klɛŋ	klɛŋ
12	lɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	?	klɛŋ
13	?	?	?	ʔatɛs	?	lɛŋ	ʔatɛs	?
14	daləm	daləm	daləm	daləm	daləm	lɛŋ	klɛŋ	daləm
16	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm
17	lɛŋ	klɛŋ	ʔatɛs	klɛŋ ʔatɛs	?	lɛŋ	klɛŋ	lɛŋ
19	ʔatɛs	lɛŋ	lɛŋ	lɛŋ	clah	lɛŋ	klɛŋ	krpiŋ
20	lɛŋ	klɛŋ	lɔʔ	klɛŋ	?	lɛŋ	klɛŋ	klɛŋ
21	?	lɛŋ	lɛŋ	lɛŋ	?	lɛŋ	lɛŋ	klɛŋ
23	ʔatɛs	klɛŋ	lɛŋ	klɛŋ	k=ʔatɛs	lɛŋ	klɛŋ	krpiŋ
24	kyəm	kyəm	kyəm	kyəm	?	lɛŋ kyəm	kyəm	kyəm
25	lɛŋ	lɛŋ	k=ʔatɛs	lɛŋ	?	lɛŋ	klɛŋ	klɛŋ
27	klɛŋ	lɛŋ	klɛŋ	lɛŋ	?	lɔʔ	klɛŋ	kyəm
29	lɛŋ	ʔatɛs	lɛŋ	da=	k=ʔatɛs	ʔatɛs	klɛŋ	klɛŋ
30	?	?	?	daləm	?	?	klɛŋ	klɛŋ
31	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm bawɔh	kyəm
32	daləm	lɛŋ	lɛŋ	lɛŋ	daləm	daləm	klɛŋ	klɛŋ
34	klɛŋ	ʔatɛs	klɛŋ	klɛŋ ʔatɛs	ʔatɛs	ʔatɛs	klɛŋ ʔatɛs	krpiŋ
35	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
36	ʔatɛs	?	ʔatɛs	lɛŋ	ʔatɛs	k=ʔatɛs	klɛŋ	krpiŋ
37	lɛŋ	ʔatɛs	lɛŋ	klɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
38	can=ser	lɛŋ	can=ʔndaŋ	mŋjiʔ can=	?	clah	ser	clah
39	lɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ

40	ʔatɛs	klɛŋ	klɛŋ	klɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
41	lɛŋ	klɛŋ	ʔndaŋ ʔndaŋ	klɛŋ	?	lɛŋ	klɛŋ	klɛŋ
42	lɛŋ	klɛŋ	klɛŋ	klɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
43	ʔatɛs	klɛŋ	ʔatɛs	lɛŋ	klɛŋ	ʔatɛs	klɛŋ	klɛŋ
44	lɛŋ	ʔatɛs	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ
45	klɛŋ	lɛŋ	?	can=kyəm	k=ʔatɛs	lɛŋ	klɛŋ	klɛŋ
46	lɛŋ	klɛŋ	klɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ
47	lɛŋ	ʔndaŋ	klɛŋ	klɛŋ	klɛŋ	lɛŋ	daləm	klɛŋ
48	da=	lɛŋ	lɛŋ	lɛŋ	klɛŋ	da=	lɛŋ	ʔatɛs klɛŋ
49	clah	krəʔ	can=ʔndaŋ	clah	?	clah	ser	clah
50	klɛŋ	klɛŋ	lɛŋ	?	sblah	lɛŋ	klɛŋ	klɛŋ
51	lɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
52	lɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ
53	kyəm	kyəm	kyəm	kyəm	kyəm	lɛŋ kyəm	klɛŋ	kyəm
54	daləm	daləm	klɛŋ daləm	daləm	daləm	lɛŋ daləm	klɛŋ	klɛŋ
55	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
58	da	lɛŋ	k=ʔatɛs	lɛŋ	ʔatɛs	lɛŋ	k=ser	klɛŋ
59	ʔatɛs	ʔatɛs	ʔatɛs	k=ʔatɛs	k=ʔatɛs	lɛŋ	klɛŋ ʔatɛs	klɛŋ
60	?	?	?	klɛŋ	daləm	lɛŋ	daləm	k=daləm
63	ʔatɛs	ʔatɛs	k=ʔatɛs	?	?	lɛŋ	klɛŋ	ʔatɛs
64	ʔndaŋ	kyəm	ʔndaŋ	klɛŋ	kyəm	lɛŋ krəʔ	ʔndaŋ	pndəh
65	ʔatɛs	ʔatɛs	k=ʔatɛs	klɛŋ	k=ʔatɛs	lɛŋ	klɛŋ	ʔatɛs
67	lɛŋ	daləm	lɛŋ	klɛŋ	klɛŋ	lɛŋ	klɛŋ	daləm
68	?	klɛŋ	?	lɛŋ	klɛŋ	lɛŋ	lɛŋ	ba=
69	lɛŋ	lɛŋ	klɛŋ	lɛŋ	?	lɛŋ	klɛŋ	lɛŋ
71	lɛŋ	daləm	daləm	daləm	daləm	lɛŋ daləm	?	daləm

	JRual8	**JRual5**	**JRual10**	**JRual2**	**JRual1**	**JRual15**	**JRual12**	**JRual13**
1	klɛŋ	krpiŋ	klɛŋ	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs
2	lɛŋ	klɛŋ	klɛŋ	daləm	lɛŋ	k=daləm	daləm	klɛŋ
3	lɛŋ	klɛŋ	klɛŋ	ʔatɛs	lɛŋ	k=	?	ʔatɛs
4	klɛŋ	?	klɛŋ	lɛŋ	lɛŋ	k=	samaʔ pdeʔ	klɛŋ
5	lɛŋ	klɛŋ	k=	ʔatɛs	lɛŋ	klɛŋ	daləm	klɛŋ
6	clah	klɛŋ	klɛŋ	lɛŋ clah	lɛŋ	clah	luwa	klɛŋ
7	lɛŋ	lɛŋ	krpiŋ	ʔatɛs	lɛŋ	clah	ʔatɛs	klɛŋ
8	lɛŋ	klɛŋ	krpiŋ	ʔatɛs	ʔatɛs	clah	ʔatɛs	klɛŋ
9	lɛŋ	?	klɛŋ	lɛŋ	klɛŋ	klɛŋ clah	klɛŋ	klɛŋ
10	lɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
11	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ
12	lɛŋ	klɛŋ	klɛŋ	lɛŋ	ʔatɛs	klɛŋ	ba=	?
13	klɛŋ ʔatɛs	klɛŋ	klɛŋ	?	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs
14	lɛŋ	klɛŋ	klɛŋ	daləm	daləm	daləm	daləm	ʔatɛs
16	kyəm	?	kyəm	k=kyəm	kyəm	kyəm	kyəm	klɛŋ kyəm
17	lɛŋ khkɛh	klɛŋ	krpiŋ	lɛŋ	lɛŋ	ʔatɛs	ʔatɛs	klɛŋ
19	klɛŋ	klɛŋ	klɛŋ	ʔatɛs	lɛŋ	klɛŋ	ʔatɛs	klɛŋ
20	lɛŋ	?	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	klɛŋ
21	lɛŋ	?	?	lɛŋ	lɛŋ	k=	klɛŋ	klɛŋ
23	klɛŋ	klɛŋ	k=	lɛŋ	ʔatɛs	klɛŋ	lɛŋ	klɛŋ
24	kyəm	kyəm	kyɔm	kyəm	kyəm	kyəm	kyəm	kyəm
25	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	clah	daləm	klɛŋ
27	klɛŋ	klɛŋ	k=	ʔatɛs	lɛŋ	kyəm	ba=	klɛŋ
29	lɛŋ	klɛŋ	klɛŋ	ʔatɛs	lɛŋ ʔatɛs	krpiŋ	ʔatɛs	klɛŋ

30	lɛŋ daləm	ʔ	ʔ	ʔ	lɛŋ	daləm	daləm	daləm
31	kyəm	kyɔm	kyɔm	kyəm	kyəm	kyəm	kyəm	kyəm
32	klɛŋ	klɛŋ	klɛŋ	daləm	daləm	k=daləm	klɛŋ	klɛŋ
34	k=	klɛŋ	krpiŋ	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs
35	lɛŋ	klɛŋ	ʔ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	klɛŋ
36	k=ʔatɛs	lɛŋ	klɛŋ	lɛŋ	lɛŋ ʔatɛs	ba=	ʔatɛs	klɛŋ
37	klɛŋ	ʔ	k=	lɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ
38	ʔ	ʔ	ser	klɛŋ clah	lɛŋ	klɛŋ clah	pndəh	tpiʔ
39	lɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ
40	klɛŋ	krpiŋ	krpiŋ	k=ʔatɛs	klɛŋ	klɛŋ	klɛŋ	ʔatɛs
41	lɛŋ	ʔ	krpiŋ	lɛŋ	klɛŋ	krəʔ	ʔatɛs	klɛŋ
42	lɛŋ	lɛŋ	k=	lɛŋ	lɛŋ	klɛŋ	klɛŋ	klɛŋ
43	klɛŋ	klɛŋ	krpiŋ	ʔatɛs	k=ʔatɛs	ʔatɛs	lɛŋ	ʔatɛs
44	lɛŋ	lɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
45	lɛŋ	ʔ	klɛŋ	ʔatɛs	klɛŋ	k=kyəm	kyəm	ʔatɛs
46	lɛŋ	klɛŋ	klɛŋ	ʔatɛs	lɛŋ	lɛŋ	klɛŋ	klɛŋ
47	klɛŋ	klɛŋ	krpiŋ	klɛŋ	klɛŋ	klɛŋ	klɛŋ	klɛŋ
48	ba=	ʔ	klɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
49	ʔndaŋ	ʔ	ʔ	k=ʔatɛs	lɛŋ	clah	pndəh	klɛŋ
50	lɛŋ	ʔ	klɛŋ	lɛŋ	klɛŋ	clah	klɛŋ	klɛŋ
51	lɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	klɛŋ
52	lɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	pndəh	lɛŋ	klɛŋ
53	kyəm	kyəm	kyɔm	kyəm	kyəm	kyəm	kyəm	kyəm
54	daləm	lɛŋ	klɛŋ	daləm	lɛŋ	klɛŋ	daləm	klɛŋ
55	lɛŋ	lɛŋ	ʔ	lɛŋ	klɛŋ kyəm	klɛŋ	lɛŋ	klɛŋ
58	lɛŋ	ʔ	ser	lɛŋ	lɛŋ	clah	ʔ	klɛŋ
59	lɛŋ	krpiŋ	krpiŋ	ʔatɛs	ʔatɛs	k=ʔatɛs krpiŋ	ʔatɛs	klɛŋ
60	daləm	ʔ	klɛŋ	ʔ	daləm	tŋah	daləm	klɛŋ
63	klɛŋ	klɛŋ	krpiŋ	ʔ	lɛŋ ʔatɛs	ʔ	kyəm	ʔ
64	ʔndaŋ	ba=krəʔ	ʔ	ʔndaŋ	ʔndaŋ	ʔndaŋ	kyəm	kyəm
65	klɛŋ ʔatɛs	klɛŋ	k=	ʔatɛs	lɛŋ	ʔatɛs	lɛŋ	klɛŋ
67	daləm	klɛŋ	klɛŋ	daləm	lɛŋ	lɛŋ	daləm	daləm
68	lɛŋ	ʔ	k=	lɛŋ	klɛŋ	k=	lɛŋ	klɛŋ
69	lɛŋ	klɛŋ	k=	lɛŋ	lɛŋ	klɛŋ	daləm	klɛŋ
71	daləm	klɛŋ	klɛŋ	ʔ	daləm	klɛŋ	daləm	klɛŋ

	JRual11	**JRual9**	**JRual17**	**JRual6**	**JRual3**	**JRual16**	**JManok3**	**JManok1**
1	ʔatɛs	ʔatɛs	klɛŋ	k=ʔatɛs	lɛŋ	ʔatɛs	klɛŋ	krpiŋ
2	daləm	klɛŋ	klɛŋ	lɛŋ	lɛŋ	daləm	klɛŋ	daləm
3	klɛŋ	ʔ	daləm	k=ʔatɛs	k=ʔatɛs	daləm	klɛŋ	klɛŋ
4	klɛŋ	lɛŋ	ʔ	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ
5	klɛŋ	lɛŋ	klɛŋ	k=ʔatɛs	lɛŋ	lɛŋ	k=	k=
6	klɛŋ	klɛŋ	lɔʔ	lɛŋ	lɛŋ	lɛŋ	k=	klɛŋ
7	klɛŋ	ʔatɛs	klɛŋ	k=ʔatɛs	ʔ	lɛŋ	klɛŋ	krpiŋ
8	clah	lɛŋ	ʔatɛs	k=ʔatɛs	ʔatɛs	lɛŋ ʔatɛs	klɛŋ	klɛŋ
9	klɛŋ	klɛŋ	da=	ʔatɛs	lɛŋ	lɛŋ	klɛŋ	klɛŋ
10	ʔ	lɛŋ	k=	lɛŋ	lɛŋ	lɛŋ	k=	klɛŋ
11	daləm	daləm	daləm	ʔatɛs	lɛŋ	daləm	klɛŋ	klɛŋ
12	klɛŋ	lɛŋ	daləm	ba=	lɛŋ	lɛŋ	klɛŋ	ʔ
13	ba=	ʔatɛs	ʔ	k=ʔatɛs	ʔatɛs	ʔ	ba=	krpiŋ
14	daləm	daləm	daləm	lɛŋ	daləm	daləm	klɛŋ	klɛŋ
16	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyəm	kyɔm
17	klɛŋ	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	ʔndaŋ	k=

19	klɛŋ	daləm	klɛŋ	k=ʔatɛs	lɛŋ	lɛŋ	klɛŋ	klɛŋ
20	klɛŋ	lɛŋ	klɛŋ	lɛŋ	?	lɛŋ	klɛŋ	klɛŋ
21	klɛŋ	lɛŋ	da=	lɛŋ kyəm	lɛŋ	lɛŋ	k=	klɛŋ
23	klɛŋ	ʔatɛs	klɛŋ	k=ʔatɛs	ʔatɛs	lɛŋ	klɛŋ	krpiŋ
24	kyəm	kyəm	?	kyəm	kyəm	kyɔm	kyəm	kyɔm
25	clah	lɛŋ	klɛŋ	klɛŋ	ʔatɛs	lɛŋ	klɛŋ	klɛŋ
27	klɛŋ	ʔatɛs	k=ʔatɛs	lɛŋ	lɛŋ	kyɔm	klɛŋ	k=
29	da=	ʔatɛs	ʔatɛs	k=ʔatɛs	ʔatɛs	ʔatɛs	klɛŋ	klɛŋ
30	daləm	?	daləm	lɛŋ	daləm	?	klɛŋ	?
31	kyəm	kyəm	kyəm	kyəm	kyəm	kyɔm	klɛŋ kyəm	kyɔm
32	klɛŋ	lɛŋ	daləm	lɛŋ	daləm	daləm	klɛŋ	klɛŋ
34	klɛŋ	ʔatɛs	k=ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ ʔatɛs	klɛŋ	krpiŋ
35	?	?	klɛŋ	lɛŋ	klɛŋ	lɛŋ	klɛŋ	k=
36	?	ʔatɛs	k=ʔatɛs	k=ʔatɛs	?	lɔʔ	ba=	krpiŋ
37	klɛŋ	lɛŋ	k=ʔatɛs	k=ʔatɛs	lɛŋ	lɛŋ	klɛŋ	klɛŋ
38	tpiʔ	ba=kahkɛh	klɛŋ	lɛŋ clah	ser	?	ba=kahkɛh	?
39	lɛŋ	ba=sɛŋ	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	k=
40	klɛŋ	klɛŋ	ʔatɛs	klɛŋ	ʔatɛs	lɛŋ	krpiŋ	krpiŋ
41	klɛŋ	lɛŋ	?	ʔatɛs	klɛŋ	lɛŋ	kyəm	klɛŋ
42	klɛŋ	ba=	klɛŋ	lɛŋ	lɛŋ	klɛŋ	klɛŋ	k=
43	ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ	ʔatɛs	lɛŋ	klɛŋ	?
44	klɛŋ	lɛŋ	?	lɛŋ	klɛŋ	lɛŋ	klɛŋ	klɛŋ
45	klɛŋ	klɛŋ	ʔatɛs	lɛŋ	lɛŋ	lɛŋ	klɛŋ	k=
46	klɛŋ	ʔatɛs	klɛŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	k=
47	klɛŋ	lɛŋ	daləm	klɛŋ	daləm	lɛŋ	klɛŋ	klɛŋ
48	klɛŋ	ba=	klɛŋ	ba=	?	lɛŋ	ba=	?
49	tpiʔ	klɛŋ	klɛŋ	klɛŋ clah	tpiʔ	?	clah	ser
50	klɛŋ	?	klɛŋ	lɛŋ	klɛŋ	?	klɛŋ	k=
51	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	ʔatɛs	klɛŋ	k=
52	klɛŋ	lɛŋ	da=	lɛŋ	lɛŋ	lɛŋ	klɛŋ	k=
53	kyəm	kyəm	daləm	kyəm	kyəm	kyəm	kyəm	kyɔm
54	klɛŋ	daləm	daləm	lɛŋ	daləm	lɛŋ	klɛŋ	klɛŋ
55	klɛŋ	daləm	klɛŋ	lɛŋ	kyəm	?	klɛŋ	k=
58	klɛŋ	klɛŋ	klɛŋ	lɛŋ	lɛŋ	?	klɛŋ	ser
59	ʔatɛs	ʔatɛs	ʔatɛs	ʔatɛs	lɛŋ	ʔatɛs	krpiŋ	krpiŋ
60	?	daləm	daləm	daləm	daləm	?	klɛŋ	?
63	ʔatɛs	ʔatɛs	?	k=ʔatɛs	ʔatɛs	?	klɛŋ	klɛŋ
64	ʔndaŋ	krɔʔ	ʔndaŋ	lɛŋ	lɛŋ	lɛŋ	klɛŋ	krpiŋ
65	klɛŋ	lɛŋ	k=ʔatɛs	lɛŋ	lɛŋ	ʔatɛs	krpiŋ	k=
67	klɛŋ	daləm	daləm	daləm	daləm	lɛŋ	klɛŋ	klɛŋ
68	ba=	ba=	klɛŋ	ba=	?	?	klɛŋ	k=
69	klɛŋ	?	klɛŋ	daləm	lɛŋ	?	klɛŋ	k=
71	klɛŋ	daləm	daləm	daləm	daləm	lɛŋ	klɛŋ	klɛŋ

	JManok4	JBanun6	JBanun2	JBanun1	JBanun4
1	krpiŋ	krpiŋ	krpiŋ	k=	ʔatɛs
2	lɛŋ	klɛŋ	ba=klɛŋ	klɛŋ	klɛŋ
3	?	?	?	k=	?
4	?	k=	klɛŋ	k=	k=
5	k=	k=	ba=	k=	k=
6	kyəm	can=hip	ba=	?	pndəh
7	krpiŋ	kyɔm	ba=	krpiŋ	krpiŋ
8	klɛŋ krpiŋ	krpiŋ	ba=	krpiŋ	krpiŋ
9	?	krpiŋ	ba=sir	krpiŋ	k=

10	k=	k=	ba=	k=	k=
11	klɛŋ	k=	ba=krpiŋ	klɛŋ	k=
12	k=	k=la=	ba=krpiŋ	k=	k=
13	klɛŋ krpiŋ	krpiŋ	krpiŋ	?	krpiŋ
14	lɛŋ	klɛŋ	ba=klɛŋ	klɛŋ	klɛŋ
16	klɛŋ kyəm	kyɔm	kyɔm	kyəm	kyəm
17	k=	k=	ba=	k=krpiŋ	k=
19	lɛŋ	krpiŋ	ba=klɛŋ	k=	krpiŋ
20	klɛŋ	k=	k=	k=	ba=
21	lɛŋ	k=la=	?	?	k=
23	k=	k=	ba=	krpiŋ	k=
24	?	kyɔm	ba=kyɔm	kyəm	kyəm
25	lɛŋ	klɛŋ	ba=	k=	k=
27	lɛŋ	kyɔm	ba=kyɔm	krpiŋ	k=
29	krpiŋ	krpiŋ	ba=	krpiŋ	krpiŋ
30	lɛŋ	klɛŋ	klɛŋ	ba=	klɛŋ
31	kyəm	kyɔm	kyɔm	kyəm	kyəm
32	klɛŋ	k=klɛŋ	klɛŋ	klɛŋ	k=
34	piŋ	krpiŋ	krpiŋ	krpiŋ	krpiŋ
35	?	k=	k=	k=	k=
36	ba=krpiŋ	krpiŋ	ba=krpiŋ	ba=	k=
37	lɛŋ	?	klɛŋ	k=	k=
38	?	ʔndaŋ	ba=sir	?	k=la=ser
39	ba=	ka=la=	ba=	k=	k=la=
40	lɛŋ	krpiŋ	ba=krpiŋ	?	krpiŋ
41	?	k=	ba=ʔndaŋ	k=	k=la=
42	k=	ba=	k=	k=	k=
43	klɛŋ	krpiŋ	ba=krpiŋ	krpiŋ	k=
44	krpiŋ	klɛŋ	ba=	k=	k=
45	klɛŋ	klɛŋ	ba=klɛŋ	k=	ka=
46	k=	ba=	k=	k=	k=
47	lɛŋ	klɛŋ	k=klɛŋ	lɛŋ	k=
48	?	k=	ba=klɛŋ	ba=	k=
49	klɛŋ ser	k=	ʔndaŋ	lɔʔ	pndəh
50	?	kyɔm	ba=	krpiŋ	k=
51	k=	k=	k=	k=	k=
52	klɛŋ	krɔʔ	ka krpiŋ	krpiŋ	k=
53	kyɔm	kyɔm	ba=kyɔm	kyəm	kyəm
54	klɛŋ	klɛŋ	ba=klɛŋ	klɛŋ	klɛŋ
55	k=	kyɔm	k=	k=	k=
58	?	ba=	ba=	ba=	k=
59	krpiŋ	?	krpiŋ	k=	krpiŋ
60	klɛŋ	klɛŋ	klɛŋ	klɛŋ	klɛŋ
63	krpiŋ kyɔm	?	k=	lɛŋ	kyəm
64	klɛŋ krɔʔ	kyɔm	ba=krɔʔ	ʔndaŋ	ʔndaŋ
65	lɛŋ	k=	k=	krpiŋ	krpiŋ
67	klɛŋ	k=klɛŋ	klɛŋ	klɛŋ	klɛŋ
68	?	k=	k=	ba=	k=
69	k=	kyɔm	k=	k=	k=
71	klɛŋ	ba=klɛŋ	k=	klɛŋ	klɛŋ

Appendix C: Caused motion event data

	Jedek1	Jedek3	Jedek4	Jedek2	Jedek18	Jedek13	Jedek39	Jedek38	Jedek7
1	boh	boh	boh	boh	boh	boh	boh	boh	boh
2	boh	?	?	boh	boh	boh	boh	boh	boh
3	boh	boh	boh	boh	boh	boh	?	boh	boh
4	boh	boh	boh	boh	boh	boh	boh	boh	boh
5	boh	boh	boh	boh	boh	boh	boh	boh	boh
6	boh	boh	simpən	boh	boh	boh	boh	susun	boh
7	boh	boh	boh	boh	boh	boh	boh	boh	boh
8	hmpɛs	hmpɛs	hmpɛs	boh	hmpɛs	hmpɛs	yɔh	hwit	hmpɛs
9	dahɛs	hltuh	hltuh	hltuh	hltuh	hltuh	hltuh	hltuh	hltuh
10	hɔk	hwit	paŋkaʔ	hɔk	paŋkaʔ	hmpɛs	yɔh	hwit	paŋkaʔ
11	boh	boh	boh	boh	boh	boh	boh	boh	boh
12	ʔisiʔ	p-hltuh	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
13	boh	tuh	boh	boh	boh	boh	boh	tuh	boh
14	boh	lon	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh
15	boh	lon	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh	boh
16	ʔisiʔ	ʔisiʔ	saŋgruk	boh	ʔisiʔ	ʔisiʔ	saŋgruk	ʔisiʔ	boh
17	dlit	lon	ʔisiʔ	dlit	ktel	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
18	caduk	caduk	caduk	caduk	boh	caduk	caduk	caduk	boh
19	boh	boh	boh	boh	boh	boh	boh	boh	boh
20	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	tuh	cʔiʔ
21	hāc	hāc	hāc	hāc	hāc	?	?	?	hāc
22	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
23	lkaʔ	lkaʔ	lon	lon	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ
24	jrloʔ	lon	jklɔk	jruk	jlɔ̄k	jlɔ̄k	saŋgruk	jruk	boh
25	pakɛy	pakɛy	pakɛy	pakɛy	bləh	pakɛy	boh	boh	pakɛy
26	pakɛy	pakɛy	pakɛy	pakɛy	bləh	bləh	saŋgruk	lɔp	bləh
27	boh	boh	boh	boh	gantoŋ	boh	gantoŋ	saŋkut	boh
28	tipet	timpɛʔ	timpɛʔ	tipet	timpɛʔ	tipet	tibɛŋ	tipet	boh
31a	?	?	?	?	?	ʔaŋket	ʔaŋket	?	?
31b	boh	boh	boh	boh	boh	boh	boh	boh	boh
33	pakɛy	pakɛy	pakɛy	pakɛy	bləh	bləh	bləh	bləh	bləh
35	boh	boh	boh	boh	boh	boh	ʔisiʔ	ʔisiʔ	boh
50a	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket	?	ʔaŋket	?	ʔaŋket
50b	boh	boh	boh	boh	boh	boh	boh	boh	boh
51a	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket
51b	boh	boh	boh	boh	boh	boh	boh	boh	boh
52	ŋuʔ	ŋuʔ	ɲuh	tolɔʔ	boh	ɲuh	tkuŋ	boh	ɲuh
101	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
102	ʔaŋket	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ʔaŋket
103	kacep	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp
104	ʔaŋket	ʔaŋket	ʔaŋket	haŋkat	kmʔəm	ʔaŋket	kilik	kompɔʔ	ʔaŋket
105	ʔaŋket	ʔaŋket	ʔaŋket	ktɛp	ckam	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
106	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
107	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	yəw	ʔaŋket
111	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
112	tuh	tuh	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	tuh	tuh	cʔiʔ
113	snitkut	sipat	tnisdes	tnisdes	sipat	tnisdes	hāc	hāc	?
114	sintak	sintak	sintak	ʔoy	ʔoy	sintak	sintak	bwaŋ	tilɛh
115	ʔaŋket	ʔaŋket	ʔaŋket	ʔoy	sintak	sintak	sintak	sintak	sintak
116	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	saŋgruk	ʔaŋket	lkaʔ

117	sintak	sintak	tarik	ʔoy	ʔoy	sumat	ʔaŋket	ʔaŋket	ʔaŋket
118	ʔaŋket	sintak	ʔaŋket	ʔoy	ʔaŋket	ʔoy	sɲɔh	sintak	ʔaŋket
119	ʔaŋket	ʔaŋket	ʔaŋket	tjik	ʔaŋket	ʔaŋket	ʔaŋket	jruk	ʔaŋket
120	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	tuh	tuh
122	?	ʔaŋket	?	?	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket
123	kluwə	?	?	ʔoy	?	tarik	?	?	tarik
124	buŋkɛ	haŋkat	buŋkɛ	buŋkɛ	buŋkɛ	kibuk	?	?	sintak
125	ʔoy	ʔoy	bukaʔ	ʔoy	ʔoy	ʔoy	ʔoy	ʔaŋket	ʔoy
126	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	pcət	ʔoy
127	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
128	kuyak	ʔoy	ʔaŋket	ʔoy	ʔaŋket	ʔoy	ʔaŋket	bukaʔ	kuyak
129a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	tijak	ʔaŋket	ʔaŋket
129b	boh	boh	?	boh	?	boh	boh	boh	boh
130	ʔaŋket	ʔaŋket	ʔaŋket	haŋkat	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
131a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
131b	boh	boh	?	boh	?	?	?	boh	?
133	ʔoy	ʔoy	bukaʔ	ʔoy	ʔoy	ʔoy	ʔoy	bwaŋ	ʔoy
135	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	lkaʔ	ʔaŋket	ʔaŋket	ʔaŋket

	Jedek8	**Jedek12**	**Jedek11**	**Jedek6**	**Jedek5**	**Jedek37**	**Jedek20**	**Jedek25**	**Jedek26**
1	boh	boh	boh	boh	boh	boh	boh	boh	boh
2	boh	boh	boh	boh	boh	boh	boh	boh	boh
3	boh	boh	boh	boh	boh	boh	boh	boh	boh
4	boh	boh	boh	boh	boh	boh	boh	boh	boh
5	boh	boh	boh	boh	boh	boh	boh	boh	boh
6	boh	boh	ʔisiʔ	boh	boh	boh	boh	simpən	boh
7	boh	boh	boh	boh	boh	boh	boh	boh	boh
8	hmpɛs	hmpɛs	hmpɛs	hmpɛs	boh	boh	hmpɛs	hmpɛs	hmpɛs
9	hltuh	hltuh	hltuh	hltuh	hltuh	hltuh	tbəl	hltuh	tbəl
10	paŋkaʔ	paŋkaʔ	hɔk	hɔk	paŋkaʔ	paŋkaʔ	paŋkaʔ	paŋkaʔ	limpah
11	boh	boh	boh	boh	boh	boh	boh	boh	boh
12	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ
13	p-guliŋ	tuh	ʔisiʔ	hɔk	boh	cʔiʔ	boh	masuk	boh
14	boh	ʔisiʔ	ʔisiʔ	boh	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ
15	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
16	boh	boh	ʔisiʔ	ʔisiʔ	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ
17	sumat	ʔisiʔ	sumat	lon	ktel	ʔisiʔ	ʔisiʔ	masuk	ʔisiʔ
18	caduk	caduk	caduk	caduk	caduk	caduk	caduk	caduk	ʔisiʔ
19	boh	boh	boh	boh	boh	jruk	ʔisiʔ	boh	boh
20	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	tuh
21	hãc	hãc	hãc	hãc	hãc	hãc	hãc	plpɛ̃l	hãc
22	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
23	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ
24	rndəm	smlɛm	ʔisiʔ	boh	boh	jruk	cilɔ̃m	jlɔp	saŋgruk
25	pakɛy	pakɛy	pakɛy	pakɛy	pakɛy	boh	boh	boh	pakɛy
26	bləh	bləh	pakɛy	pakɛy	pakɛy	saŋgruk	bləh	bləh	saŋgruk
27	boh	boh	hipɛt	boh	boh	saŋkut	saŋkut	boh	hipɛt
28	tipet	timpɛʔ	tipet	tipet	tampãl	tipet	timpɛʔ	tipet	pet
31a	ʔaŋket	?	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
31b	boh	tuduŋ	boh	boh	boh	boh	boh	boh	boh
33	bləh	bləh	bləh	pakɛy	pakɛy	bləh	bləh	bləh	bləh
35	boh	ʔisiʔ	ʔisiʔ	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
50a	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
50b	boh	boh	boh	boh	boh	boh	boh	boh	boh

51a	ʔaŋket	?	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	?	ʔaŋket
51b	boh	boh	boh	boh	boh	boh	boh	boh	boh
52	tolɔʔ	tkuŋ	tkuŋ	tolɔʔ	tolɔʔ	ɲuh	tolɔʔ	tolɔʔ	ɲuh
101	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
102	ktɛp	ktɛp	ktɛp	ktɛp	ʔaŋket	ktɛp	ktɛp	ktɛp	jaŋgaŋ
103	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	tjik	ktɛp	ktɛp	ktɛp
104	ʔaŋket	ʔaŋket	sʔun	ʔaŋket	ʔaŋket	sʔun	haŋkat	ʔaŋket	haŋkat
105	rahop	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	rahop	ckam	rahop	ʔaŋket
106	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
107	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
111	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
112	cʔiʔ	tuh	cʔiʔ	cʔiʔ	cʔiʔ	tuh	sɔh	cʔiʔ	tuh
113	tnisdes	sipat	sipat	sipat	snitkut	tbalik	sipat	sipat	sipat
114	sintak	ʔoy	sintak	sintak	cabut	sintak	sintak	sintak	sintak
115	sintak	ʔoy	ʔaŋket	sintak	ʔaŋket	sintak	sintak	sintak	ʔaŋket
116	ʔaŋket	lkaʔ	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	lkaʔ	pridɛŋ	stlot
117	ʔoy	ʔoy	ʔaŋket	sintak	ʔoy	?	sintak	ʔaŋket	ʔaŋket
118	sintak	ʔoy	sintak	sintak	ʔoy	ʔaŋket	sintak	sintak	sintak
119	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket
120	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
122	ʔaŋket	ʔaŋket	?	ʔɛk	ʔaŋket	ʔaŋket	?	?	?
123	sintak	?	?	sintak	kluwə	?	tarik	sintak	sintak
124	bliwīs	buŋkɛ	haŋkat	buŋkɛ	kluwə	?	buŋkɛ	buŋkɛ	buŋkɛ
125	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	sintak	ʔoy	ʔoy	ʔoy
126	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
127	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔoy
128	ʔoy	kuyak	ʔoy	kuyak	kuyak	sintak	kuyak	suyak	ʔoy
129a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	haŋkat	ʔaŋket	ʔaŋket	tjik	ʔaŋket
129b	boh	boh	boh	boh	boh	boh	boh	?	boh
130	ʔaŋket	yiy	ʔaŋket	ʔaŋket	haŋkat	ʔaŋket	jīŋ	ʔaŋket	ʔaŋket
131a	ʔaŋket	bukaʔ	haŋkat	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
131b	?	?	?	boh	boh	?	boh	boh	boh
133	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	bukaʔ	bukaʔ	ʔoy
135	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket

	Jedek33	**Jedek32**	**Jedek16**	**Jedek22**	**Jedek14**	**Jedek23**	**Jedek34**	**Jedek28**	**Jedek21**
1	boh	boh	boh	boh	boh	boh	boh	boh	boh
2	?	boh	boh	boh	boh	boh	?	boh	boh
3	boh	boh	boh	boh	boh	boh	?	boh	?
4	boh	boh	boh	boh	boh	?	boh	boh	boh
5	boh	boh	boh	boh	boh	boh	boh	boh	boh
6	boh	boh	boh	boh	boh	susun	boh	simpən	ʔisiʔ
7	boh	boh	boh	boh	boh	boh	boh	boh	boh
8	hmpɛs	hmpɛs	hmpɛs	hmpɛs	hmpɛs	boh	hmpɛs	boh	hmpɛs
9	tbəl	tbəl	dahɛs	dahɛs	hltuh	hltuh	tbəl	hltuh	hltuh
10	paŋkaʔ	hwit	hɔk	paŋkaʔ	paŋkaʔ	paŋkaʔ	hɔk	boh	hɔk
11	boh	boh	boh	boh	boh	ʔisiʔ	boh	ʔisiʔ	boh
12	boh	yɔh	boh	boh	boh	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ
13	cʔiʔ	hwit	boh	boh	boh	boh	boh	tuh	cʔiʔ
14	kintup	ʔisiʔ	boh	boh	ʔisiʔ	pel	ʔisiʔ	lon	pasaŋ
15	ʔisiʔ	ʔisiʔ	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
16	boh	boh	boh	boh	boh	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ
17	ʔisiʔ	sumat	boh	sumat	ʔisiʔ	pel	ʔisiʔ	ktel	ʔisiʔ

18	caduk	cɔk	boh	boh	caduk	caduk	caduk	caduk	caduk
19	boh	boh	boh	boh	boh	ʔisiʔ	boh	boh	yɔh
20	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
21	?	?	?	hãc	hãc	hãc	hãc	hãc	hãc
22	?	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
23	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ
24	ʔisiʔ	jlɔp	jruk	jruk	ʔisiʔ	ʔɛk	ʔisiʔ	rndəm	kibuk
25	boh	pakɛy	pakɛy	saŋgruk	pakɛy	pakɛy	boh	bləh	boh
26	pakɛy	pakɛy	bləh	pakɛy	pakɛy	boh	pakɛy	bləh	bləh
27	boh	boh	boh	boh	boh	boh	hipɛt	boh	sipɛt
28	pet	pet	boh	tipet	timpɛʔ	plet	timpɛʔ	pet	timpɛʔ
31a	ʔaŋket	ʔaŋket	?	?	ʔaŋket	?	?	ʔaŋket	ʔaŋket
31b	boh	boh	boh	boh	?	boh	boh	boh	?
33	pakɛy	bləh	bləh	pakɛy	pakɛy	bləh	pakɛy	bləh	bləh
35	boh	ʔisiʔ	boh	boh	boh	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ
50a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
50b	boh	boh	boh	boh	boh	boh	boh	boh	boh
51a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	?	?
51b	boh	?	boh	boh	boh	boh	boh	boh	boh
52	ɲuh	ɲuh	ɲuh	ɲuh	boh	ɲuh	tolɔʔ	simpən	ɲuh
101	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
102	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	jaŋgaŋ	ktɛp
103	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp
104	ʔaŋket	sʔun	ʔaŋket	ʔaŋket	sʔun	ʔaŋket	kmʔəm	ʔaŋket	kilik
105	ʔaŋket	rahop	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ckam	ʔaŋket
106	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
107	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
111	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
112	cʔiʔ	cʔiʔ	tuh	tuh	tuh	tuh	cʔiʔ	tuh	cʔiʔ
113	sipat	sipat	tŋkək	sipat	snitkut	sipat	tiŋkuh	sipat	sipat
114	sintak	sintak	sintak	sintak	sintak	sintak	sintak	sintak	sintak
115	ʔaŋket	sintak	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	sintak	sintak	sɲɔh
116	ʔaŋket	ʔaŋket	lkaʔ	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
117	sintak	sintak	ʔaŋket	sintak	sintak	ʔaŋket	sintak	sintak	sɲɔh
118	ʔaŋket	tarik	ʔoy	bwaŋ	ʔaŋket	ʔoy	sintak	sintak	sɲɔh
119	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
120	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
122	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	?
123	haŋkat	tarik	sintak	kluwə	?	sintak	?	?	?
124	buŋkɛ	rndəm	buŋkɛ	kluwə	buŋkɛ	buŋkɛ	tijak	buŋkɛ	buŋkɛ
125	ʔoy	bukaʔ	ʔoy	ʔoy	ʔoy	ʔoy	tijak	sintak	ʔoy
126	ʔoy	bwaŋ	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
127	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	tanjĩŋ
128	ʔaŋket	bukaʔ	ʔoy	kuyak	kuyak	ʔaŋket	ʔaŋket	ʔoy	ʔaŋket
129a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
129b	boh	?	boh	boh	boh	?	boh	simpən	boh
130	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
131a	ʔaŋket	bukaʔ	ʔaŋket	?	ʔaŋket	?	tijak	ʔaŋket	ʔaŋket
131b	boh	?	?	boh	boh	boh	boh	boh	boh
133	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	bukaʔ	ʔoy	ʔoy	ʔoy
135	pgɛŋ	ʔaŋket	lkaʔ	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket

	Jedek9	Jedek10	Jedek35	Jedek17	Jedek36	Jedek19	Jedek24	Jedek29	JRual2
1	boh	boh	boh	boh	boh	boh	boh	boh	boh
2	boh	boh	boh	boh	boh	boh	boh	boh	boh
3	boh	boh	boh	boh	boh	boh	boh	boh	boh
4	boh	boh	boh	boh	boh	boh	boh	?	boh
5	boh	boh	boh	boh	boh	boh	boh	boh	boh
6	boh	boh	boh	boh	boh	boh	ʔisiʔ	simpən	boh
7	boh	boh	boh	boh	boh	boh	boh	boh	boh
8	yoh	hmpɛs	boh	boh	hmpɛs	hmpɛs	hɔk	hmpɛs	hmpɛs
9	hltuh	?	tbəl	hltuh	hltuh	hltuh	hltuh	hltuh	hltuh
10	hɔk	limpah	paŋkaʔ	limpah	hmpɛs	paŋkaʔ	paŋkaʔ	limpah	paŋkaʔ
11	boh	boh	boh	boh	boh	boh	boh	boh	boh
12	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	boh	ʔisiʔ	masuk	boh	boh
13	boh	ʔisiʔ	cʔiʔ	boh	cʔiʔ	boh	boh	boh	boh
14	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	pet	boh	ʔisiʔ
15	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	masuk	boh	masuk
16	boh	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
17	ktel	sumat	ktel	ʔisiʔ	ʔisiʔ	sumat	masuk	ʔisiʔ	ʔisiʔ
18	caduk	caduk	caduk	boh	ʔisiʔ	caduk	cɔk	caduk	boh
19	boh	boh	jruk	boh	boh	boh	masuk	boh	boh
20	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
21	hãc	?	hãc	hãc	hãc	hãc	hãc	?	hãc
22	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
23	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	saŋgruk	lkaʔ	masuk
24	jlɔ̃k	jruk	jruk	jruk	cilɔ̃m	jlɔp	jlɔ̃k	jruk	ʔisiʔ
25	pakɛy	pakɛy	boh	pakɛy	boh	boh	pakɛy	saŋgruk	pakɛy
26	pakɛy	bləh	bləh	pakɛy	saŋgruk	pakɛy	bləh	bləh	pakɛy
27	boh	gantoŋ	hipɛt	boh	hipɛt	saŋkut	saŋkut	saŋkut	hipɛt
28	pet	pet	tibɛŋ	timpɛʔ	tipet	timpɛʔ	timpɛʔ	boh	tipet
31a	?	ʔaŋket	ʔaŋket	?	ʔaŋket	?	ʔaŋket	?	?
31b	boh	boh	boh	boh	boh	boh	boh	boh	boh
33	pakɛy	bləh	bləh	bləh	bləh	bləh	bləh	bləh	pakɛy
35	boh	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
50a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	haŋkat	ʔaŋket	?	ʔaŋket
50b	boh	boh	boh	boh	boh	boh	?	boh	boh
51a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	?	?	ʔaŋket
51b	boh	boh	boh	boh	boh	boh	boh	boh	boh
52	tolɔʔ	tolɔʔ	ɲuh	ɲuh	ɲuh	tolɔʔ	dlũʔ	surut	tolɔʔ
101	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	tjik
102	ʔaŋket	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ʔaŋket	ktɛp	ktɛp
103	ʔaŋket	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp
104	haŋkat	ʔaŋket	ʔaŋket	sʔun	sʔun	haŋkat	kmʔəm	ʔaŋket	ʔaŋket
105	rahop	ʔaŋket	rahop	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
106	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
107	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
111	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
112	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	tuh
113	tiyə	sipat	sipat	sipat	sipat	sipat	sipat	sipat	sipat
114	ʔoy	sintak	sintak	sintak	sintak	sintak	sintak	sintak	sintak
115	ʔoy	sintak	sintak	ʔaŋket	sintak	sintak	sintak	sintak	ʔaŋket
116	lkaʔ	ʔaŋket	sintak	ʔaŋket	ʔaŋket	ʔaŋket	pridɛŋ	ʔaŋket	ʔaŋket
117	ʔoy	ʔaŋket	sintak	ʔaŋket	sintak	sintak	sintak	ktel	sintak
118	ʔaŋket	ʔoy	sintak	caduk	sintak	sintak	sintak	tarik	sintak
119	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	tjik	ʔaŋket
120	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ

122	ʔaŋket	ʔaŋket	?	?	?	ʔaŋket	?	?	ʔaŋket
123	?	sintak	?	?	?	?	?	tarik	tarik
124	kluwə	?	?	?	buŋkɛ	?	?	biŋgaŋ	sintak
125	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
126	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
127	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	tarik	ʔaŋket	saŋkut	ʔaŋket
128	ʔaŋket	kuyak	ʔoy	ʔoy	ʔoy	kuyak	ʔoy	ʔaŋket	kuyak
129a	?	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	haŋkat	ʔaŋket	tjik	ʔaŋket
129b	boh	boh	boh	boh	boh	boh	boh	boh	boh
130	haŋkat	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	haŋkat	ʔaŋket	tjik	ʔaŋket
131a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	tjik	tijak	ʔaŋket	ʔaŋket
131b	boh	?	boh	boh	boh	boh	boh	?	?
133	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
135	ʔaŋket	ʔaŋket	lkaʔ	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	sintak	sintak

	JRual4	**JRual1**	**JRual15**	**JRual12**	**JRual13**	**JRual11**	**JRual9**	**JRual5**	**JRual10**
1	boh	boh	boh	boh	boh	boh	boh	boh	boh
2	boh	?	?	boh	boh	boh	?	boh	boh
3	boh	boh	boh	boh	boh	boh	?	boh	boh
4	boh	boh	boh	boh	boh	boh	ʔisiʔ	boh	boh
5	boh	boh	boh	boh	boh	boh	ckam	boh	boh
6	boh	boh	ʔisiʔ	boh	boh	boh	boh	boh	boh
7	boh	boh	boh	boh	boh	boh	boh	boh	boh
8	paŋkaʔ	hɔk	yɔh	yɔh	yɔh	hntɛk	yɔh	plihtuh	yɔh
9	hltuh	hltuh	tbəl	hltuh	hltuh	tbəl	hltuh	tbəl	tbəl
10	paŋkaʔ	hɔk	hɔk	hɔk	hɔk	paŋkaʔ	hɔk	paŋkaʔ	hɔk
11	boh	boh	boh	boh	boh	boh	boh	boh	boh
12	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ
13	boh	yɔh	boh	boh	tuh	toklɛk	sɔh	tuh	boh
14	ʔisiʔ	ʔisiʔ	masuk	ʔisiʔ	ʔisiʔ	sumɛt	ʔisiʔ	sumɛt	boh
15	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	pel	ʔisiʔ
16	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	boh	ʔisiʔ
17	sumɛt	ktel	ktel	sumɛt	ʔisiʔ	ʔisiʔ	ʔisiʔ	ktel	ʔisiʔ
18	lon	cɔk	caduk	caduk	caduk	lon	saŋgruk	caduk	caduk
19	ʔisiʔ	boh	boh	boh	boh	boh	boh	boh	boh
20	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
21	hɛ̃c	hāc	?	?	hɛ̃c	?	hɛ̃c	hɛ̃c	hɛ̃c
22	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
23	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ
24	masuk	jruk	ypyɔp	cilɔ̃m	cilɔ̃m	toŋgɛŋ	jruk	jruk	jruk
25	ʔisiʔ	boh	pakɛy	pakɛy	saŋgruk	pakɛy	bləh	pakɛy	bləh
26	bləh	pakɛy	pakɛy	bləh	bləh	pakɛy	bləh	bləh	bləh
27	boh	boh	saŋkut	saŋkut	gantoŋ	hipɛt	saŋkut	saŋkut	hipɛt
28	tipet	tipet	boh	pridɛŋ	timpɛʔ	timpɛʔ	timpɛʔ	tipet	tibɛŋ
31a	?	?	?	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket
31b	boh	boh	tuduŋ	boh	liŋkuk	tutup	boh	boh	boh
33	bləh	bləh	bləh	bləh	bləh	bləh	pakɛy	bləh	bləh
35	ʔisiʔ	ʔisiʔ	masuk	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	boh	ʔisiʔ
50a	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
50b	boh	boh	boh	boh	?	boh	boh	boh	boh
51a	ʔaŋket	?	ʔaŋket	ʔaŋket	?	jĩŋ	?	ʔaŋket	ʔaŋket
51b	boh	boh	boh	boh	boh	boh	boh	boh	boh
52	tolɔʔ	hidit	boh	ɲuh	surut	kisɔ̃t	tolɔʔ	ɲuh	tkuŋ
101	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket

102	ktɛp	ktɛp	ktɛp	jaŋgaŋ	ktɛp	ktɛp	ktɛp	jaŋgaŋ	ktɛp
103	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp
104	ʔaŋket	ʔŋsuŋ	sʔun	ʔaŋket	cdum	cdum	ʔaŋket	sʔun	cduŋ
105	ʔaŋket	ktɛp	ʔaŋket	ʔaŋket	rahop	gŋgam	ckam	ckam	ckam
106	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	sʔun	jĩŋ	ʔaŋket	jĩŋ	ʔaŋket
107	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket
111	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
112	tuh	tuh	cʔiʔ	tuh	cʔiʔ	cʔiʔ	sɔh	cʔiʔ	tuh
113	sipat	sipat	sipat	sipat	tnisdes	sipat	sipat	sipat	sipat
114	sintak	sintak	sɲɔh	sɲɔh	sintak	sintak	ʔoy	sintak	sɲɔh
115	sintak	ʔaŋket	sɲɔh	sɲɔh	ʔaŋket	sintak	sɲɔh	sintak	sɲɔh
116	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	lkaʔ	sɲɔh	ʔaŋket	ʔaŋket
117	sintak	ʔaŋket	sɲɔh	sɲɔh	ʔaŋket	jĩŋ	sɲɔh	?	ʔaŋket
118	ʔaŋket	ʔoy	ʔaŋket	ʔoy	ʔaŋket	sintak	sintak	ʔoy	ʔaŋket
119	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	jĩŋ	ʔaŋket
120	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
122	?	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	ʔaŋket	?
123	?	?	sɲɔh	sɲɔh	?	?	?	?	?
124	bliwĩs	?	jŋɔʔ	?	?	?	?	?	?
125	ʔoy	ʔoy	ʔoy	ʔoy	ʔaŋket	bukaʔ	ʔoy	ʔoy	?
126	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
127	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket
128	kuyak	kuyak	sɲɔh	ʔaŋket	ʔoy	kuyak	kuyak	kuyak	ʔaŋket
129a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	jĩŋ	ʔaŋket
129b	boh	boh	boh	?	boh	boh	boh	boh	boh
130	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket
131a	?	ʔaŋket	?	ʔaŋket	bukaʔ	ʔaŋket	ʔoy	ʔaŋket	ʔaŋket
131b	boh	boh	boh	boh	?	boh	?	boh	boh
133	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	bukaʔ	ʔoy	ʔoy	ʔoy
135	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	lkaʔ	ʔaŋket	ʔaŋket

	JRual18	**JRual17**	**JRual7**	**JRual8**	**JManok6**	**JManok5**	**JManok2**	**JManok3**	**JManok7**
1	boh	boh	boh	boh	boh	boh	boh	boh	boh
2	boh	boh	boh	boh	boh	?	boh	boh	boh
3	?	boh	boh	boh	boh	boh	boh	boh	?
4	boh	boh	boh	boh	boh	boh	boh	boh	boh
5	boh	boh	boh	boh	boh	boh	boh	boh	boh
6	boh	boh	bah	boh	susun	susun	boh	boh	boh
7	boh	boh	boh	boh	boh	boh	boh	boh	boh
8	yɔh	yɔh	boh	boh	yɔh	hɔk	praʔ	boh	yɔh
9	tbəl	tbəl	tbəl	hltuh	tbəl	tbəl	tbəl	tbəl	tbəl
10	hwit	paŋkaʔ	hɔk	hɔk	hɔk	hɔk	hɔk	hɔk	hɔk
11	boh	boh	boh	boh	boh	boh	boh	boh	boh
12	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	yɔh
13	boh	cʔiʔ	boh	boh	tuh	sɔh	sɔh	sɔh	boh
14	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	kitit	ʔisiʔ	ʔisiʔ	boh	boh
15	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh
16	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	?	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh
17	ʔisiʔ	ʔisiʔ	dlit	sumat	ʔisiʔ	klit	lon	klit	ʔisiʔ
18	caduk	caduk	saŋgruk	caduk	caduk	caduk	caduk	lon	caduk
19	boh	boh	boh	boh	yɔh	yɔh	boh	yɔh	yɔh
20	tuh	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
21	hãc	hẽc	hẽc	hãc	hẽc	hẽc	hẽc	hẽc	hẽc
22	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk

23	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ
24	jklɔk	ʔnay	cilɔ̃m	blәh	jiŋәʔ	jlɔp	jlɔp	ʔnɔm	cilɔ̃m
25	boh	boh	pakɛy	boh	boh	lon	pakɛy	boh	boh
26	blәh	blәh	pakɛy	saŋgruk	blәh	pakɛy	blәh	blәh	blәh
27	boh	srpɛk	saŋkut	boh	saŋkut	hipɛt	hipɛt	hipɛt	gantoŋ
28	pet	tibɛŋ	tipet	timpɛʔ	tibɛŋ	tipet	timpɛʔ	pet	boh
31a	ʔaŋket	?	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	?	?
31b	boh	boh	boh	boh	boh	?	boh	boh	boh
33	blәh	blәh	blәh	blәh	blәh	blәh	blәh	blәh	blәh
35	ʔisiʔ	ʔisiʔ	ʔisiʔ	boh	boh	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
50a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	?	ʔaŋket	ʔaŋket	?
50b	boh	boh	boh	boh	boh	boh	boh	boh	boh
51a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket	?
51b	boh	boh	boh	boh	boh	boh	boh	boh	boh
52	surut	ɲuh	surut	tkuŋ	ɲuh	ʔaŋket	surut	surut	surut
101	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
102	ktɛp	ktɛp	ktɛp	jaŋgaŋ	ktɛp	ktɛp	ktɛp	ktɛp	ktɛp
103	ktɛp	ktɛp	ktɛp	ʔaŋket	ktɛp	ktɛp	ktɛp	ʔaŋket	ktɛp
104	sʔun	sʔun	ʔaŋket	ʔŋsuŋ	jĩŋ	kmʔәm	ʔaŋket	ʔaŋket	ʔaŋket
105	rahop	ckam	gŋgam	rahop	jĩŋ	rahop	ʔaŋket	ʔaŋket	rahop
106	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
107	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
111	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
112	tuh	sɔh	cʔiʔ	cʔiʔ	tuh	sɔh	tuh	sɔh	sɔh
113	sipat	sipat	sipat	tnisdes	snitkut	tiŋkuh	tŋkәk	sipat	sipat
114	sintak	sintak	sintak	sintak	sɲɔh	sɲɔh	sɲɔh	sɲɔh	sintak
115	sintak	sintak	tarik	sɲɔh	tarik	ʔaŋket	sɲɔh	sɲɔh	sintak
116	saŋgruk	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	lkaʔ	ʔaŋket
117	sintak	sintak	sintak	sɲɔh	sɲɔh	ʔoy	sintak	tarik	ʔaŋket
118	sintak	sintak	ʔaŋket	ʔoy	sɲɔh	sɲɔh	sɲɔh	sɲɔh	ʔaŋket
119	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	tjik	ʔaŋket
120	tuh	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
122	?	?	?	ʔaŋket	jĩŋ	ʔaŋket	?	?	ʔaŋket
123	?	sintak	tarik	?	?	?	?	?	?
124	?	tjik	buŋkɛr	bliwĩs	?	buŋkɛr	?	tjik	?
125	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	sɲɔh	ʔoy	ʔaŋket
126	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
127	saŋkut	ʔaŋket	ʔaŋket	ʔaŋket	slәpɛʔ	ʔaŋket	ʔaŋket	ʔoy	ʔaŋket
128	ʔoy	kuyak	ʔaŋket	suyak	ʔaŋket	kuyak	suyak	ʔoy	ʔoy
129a	haŋkat	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
129b	?	boh	boh	boh	boh	boh	boh	boh	boh
130	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
131a	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket
131b	?	boh	?	boh	boh	?	?	boh	boh
133	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy
135	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket

	JBanun2	**JBanun5**	**JBanun3**	**JBanun4**	**JBanun8**	**JBanun7**
1	boh	boh	boh	boh	boh	boh
2	boh	boh	boh	boh	boh	boh
3	boh	boh	boh	boh	boh	boh
4	boh	boh	boh	boh	boh	boh
5	boh	boh	boh	boh	boh	boh
6	boh	boh	boh	simpɛn	boh	boh

7	boh	boh	boh	boh	boh	boh
8	praʔ	yɔh	praʔ	boh	boh	boh
9	tbəl	tbəl	tbəl	tbəl	tbəl	tbəl
10	hɔk	hakɔk	hakɔk	hɔk	hɔk	hakɔk
11	boh	boh	boh	boh	boh	boh
12	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
13	ʔisiʔ	tuh	sɔh	sɔh	tuh	boh
14	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	tanɛm	lon
15	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ	ʔisiʔ
16	ʔisiʔ	ʔisiʔ	boh	boh	ʔisiʔ	ʔisiʔ
17	klit	klit	klit	sumɛt	ʔisiʔ	sel
18	caduk	?	caduk	boh	caduk	caduk
19	ʔisiʔ	boh	boh	boh	boh	?
20	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
21	hɑ̃c	cohɛʔ	hɑ̃c	hɑ̃c	?	cohɛʔ
22	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
23	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ	lkaʔ
24	ʔisiʔ	ʔisiʔ	cikũm	rnəm	?	bləh
25	bləh	tol	topih	bləh	tanɛm	boh
26	bləh	bləh	bləh	pakɛy	bləh	bləh
27	boh	hipɛt	boh	boh	boh	boh
28	pet	pet	tipɛt	tipɛt	tipet	pet
31a	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	?	?
31b	boh	tutop	boh	boh	boh	boh
33	bləh	bləh	bləh	pakɛy	bləh	bləh
35	boh	ʔisiʔ	ʔisiʔ	boh	ʔisiʔ	boh
50a	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	?	?
50b	boh	boh	boh	boh	boh	boh
51a	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	?	ʔaŋket
51b	boh	boh	boh	boh	boh	boh
52	jlũʔ	surut	surut	tkuŋ	boh	boh
101	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket
102	ktɛp	ʔaŋket	ʔaŋket	ktɛp	ʔaŋket	ʔaŋket
103	ktɛp	ktɛp	ktɛp	ktɛp	ʔaŋket	ʔaŋket
104	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	cduŋ	ʔaŋket
105	ʔaŋket	kpic	ʔaŋket	jĩŋ	jɛk	ʔaŋket
106	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket
107	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	jĩŋ	ʔaŋket
111	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	jĩŋ	ʔaŋket
112	tuh	tuh	sɔh	sɔh	tuh	tuh
113	tisdes	laŋah	tides	tnsdes	srkon	sipat
114	yɔk	rəc	sintak	sintak	sintak	bukaʔ
115	yɔk	sintak	sintak	sintak	sintak	ʔaŋket
116	ʔaŋket	lkaʔ	ʔaŋket	jĩŋ	ʔaŋket	ʔaŋket
117	yɔk	sintak	sintak	rəc	sintak	?
118	ʔaŋket	ʔaŋket	ʔaŋket	sintak	sintak	ʔaŋket
119	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	?	ʔaŋket
120	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ	cʔiʔ
122	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	?
123	kjũt	sintak	sintak	sintak	?	?
124	snrɛk	buŋkɛr	buŋkɛr	rnəm	?	buŋkɛr
125	ʔaŋket	ʔaŋket	ʔoy	sɔh	sintak	bukaʔ
126	ʔoy	ʔoy	ʔoy	sɔh	ʔoy	bukaʔ
127	ʔaŋket	ʔaŋket	ʔaŋket	jĩŋ	jĩŋ	ʔaŋket
128	ʔaŋket	lok	ʔaŋket	ʔaŋket	ʔaŋket	ʔaŋket

129a	ʔaŋket	ʔaŋket	ʔaŋket	jĭŋ	?	ʔaŋket
129b	boh	boh	boh	boh	boh	boh
130	ʔaŋket	ʔaŋket	ʔaŋket	jĭŋ	jĭŋ	yiy
131a	ʔaŋket	ʔaŋket	ʔaŋket	jĭŋ	?	?
131b	boh	boh	boh	boh	boh	boh
133	ʔoy	ʔoy	ʔoy	ʔoy	ʔoy	bukaʔ
135	ʔaŋket	ʔaŋket	ʔaŋket	jĭŋ	ʔaŋket	ʔaŋket

Appendix D: Reciprocal event data

	Jedek1	Jedek4	Jedek2	Jedek18	Jedek13	Jedek38	Jedek7	Jedek8	Jedek12
1	ʔ	cilcɔl	ʔ	cilcɔl	cilcɔl	kicɛʔ	cilcɔl	cilcɔl	cilcɔl
2	kmʔəm	kimʔəm	knʔəm	kimʔəm	kimʔəm	ʔ	kimʔəm	kmʔəm	kimʔəm
3	bihlɔh	pilpal	pal	pilpal	pilpal	guhcɔh	pilpal	pilpal	pilpal
4	ʔ	ʔikʔɛk	bagiʔ	bagiʔ	ʔikʔɛk	ʔikʔɛk	bagiʔ	ʔ	biʔgiʔ
5	pal	pilpal	pilpal	pilpal	pilpal	prikluk	lmpah	lihpah	lmpah
6	diʔ	miymay	may	may	may	may	may	miymay	miymay
7	knʔəm	kimʔəm	knʔəm	kimʔəm	kimʔəm	kmʔəm	knʔəm	kimʔəm	kimʔəm
8	ŋɔk	ŋikŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk	ʔ	ŋɔk
9	cɔl	taɲaʔ	cilcɔl	tiʔɲaʔ	taɲaʔ	ʔ	ʔ	miʔmɛʔ	cilcɔl
10	brilɛʔ	may	miymay	may	miymay	ʔɛl	may	may	miymay
11	cɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cɔl	cilcɔl	cilcɔl	cilcɔl
12	ʔɛk	ʔindɨn	ʔindɨn	ciktuk	ʔindɨn	ʔikʔɛk	cuntuk	ʔikʔɛk	ʔikʔɛk
13	salam	mɲamut	salam	salam	mɲamut	mɲamut	salam	salam	salam
14	lɔy	hitmɛt	pmpem	pmpem	hitmɛt	pmpem	pmpem	hitmɛt	hitmɛt
15	ŋɔk	ŋikŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk
16	ʔ	kimʔəm	ʔ	kimʔəm	ʔ	ʔ	ʔ	ʔ	kimʔəm
17	guhcɔh	gucɔh	gucɔh	gucɔh	gucɔh	gucɔh	ktɨm	tumboʔ	gucɔh
18	ŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk	ŋikŋɔk	ŋikŋɔk
19	bagiʔ	bagiʔ	ʔikʔɛk	biʔgiʔ	biʔgiʔ	ʔikʔɛk	bagiʔ	toka	biʔgiʔ
20	knʔəm	kmʔəm	knʔəm	kimʔəm	kmʔəm	kmʔəm	knʔəm	kmʔəm	kmʔəm
21	bagiʔ	ʔikʔɛk	bagiʔ	biʔgiʔ	biʔgiʔ	ʔ	bagiʔ	toka	biʔgiʔ
22	ʔ	gilsil	tiŋkoh	gilsil	gilsil	ʔ	tmpuh	ʔ	ʔ
23	knʔəm	kmʔəm	knʔəm	kmʔəm	kmʔəm	kmʔəm	knʔəm	kmʔəm	kmʔəm
24	pjaʔ	pjaʔ	ʔ	pjaʔ	pjaʔ	pjaʔ	ʔ	pjaʔ	ʔ
25	ʔɛl	ʔɛl	ʔ	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl
26	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
27	ʔ	ʔ	cŋcoŋ	bitkut	coŋ	ʔ	cŋcoŋ	hay	coŋ
28	bagiʔ	ʔikʔɛk	bagiʔ	biʔgiʔ	biʔgiʔ	ʔ	biʔgiʔ	toka	ʔikʔɛk
29	knʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kmʔəm	knʔəm	kmʔəm	kimʔəm
30	bagiʔ	ʔikʔɛk	baɲci	biʔgiʔ	ʔikʔɛk	ʔikʔɛk	bagiʔ	ʔikʔɛk	ʔikʔɛk
31	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔilʔɛl	ʔɛl	ʔ
32	ʔɛk	ʔ	ʔikʔɛk	ʔ	ʔ	ʔ	ʔ	ʔikʔɛk	ʔikʔɛk
33	ʔ	ʔ	ʔɛl	ʔ	ʔilʔɛl	ʔilʔɛl	ʔ	ʔilʔɛl	ʔilʔɛl
34	bagiʔ	ʔikʔɛk	baɲci	bagiʔ	biʔgiʔ	ʔɛk	bagiʔ	pisləs	bagiʔ
35	hatut	susun	picin	susun	boh	ʔ	ʔ	toŋkɛt	boh
36	hamɛt	hitmɛt	ʔ	ʔ	hitmɛt	ʔ	ʔ	hitmɛt	hitmɛt
37	ʔɛk	ʔikʔɛk	baɲci	biʔgiʔ	biʔgiʔ	ʔɛk	bagiʔ	b-gilih	biʔgiʔ
38	salam	mɲamut	salam	salam	cipcɛp	mɲamut	salam	salam	ʔ
39	ʔ	hitmɛt	hitmɛt	ʔ	hitmɛt	ʔ	ʔ	hitmɛt	hitmɛt
40	cɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cɔl	cilcɔl	cilcɔl	cilcɔl
41	toka	ʔikʔɛk	toka	toka	toka	toka	ʔikʔɛk	toka	ʔɛk
42	guhcɔh	pilpal	ʔ	guhcɔh	guhcɔh	guhcɔh	ktɨm	gucɔh	guhcɔh
43	ʔ	ʔ	hitmɛt	ʔ	hitmɛt	hitmat	hitmat	hamɛt	hitmɛt
44	pilpal	pilpal	pilpal	pal	pilpal	guhcɔh	pilpal	ʔ	ʔ
45	pimpõm	miymay	may	miymay	miymay	may	miymay	tikrik	miymay
46	ʔ	ʔilʔɛl	ʔ	ʔ	ʔilʔɛl	ʔilʔɛl	ʔilʔɛl	ʔ	ʔilʔɛl
47	ʔ	kimʔəm	knʔəm	kimʔəm	kmʔəm	ʔ	ʔ	kmʔəm	ʔ
48	pal	pilpal	piŋkuh	tiŋkoh	guhcɔh	gucɔh	kimtɨm	gucɔh	guhcɔh
49	salam	mɲamut	salam	salam	cipcɛp	mɲamut	cipcɛp	salam	salam
50	pal	pilpal	pilpal	pilpal	pilpal	lmpah	tampah	lmpah	pilpal

51	diʔ	may	may	may	may	may	may	may	may
52	knʔəm	kimʔəm	knʔəm	kimʔəm	kimʔəm	kimʔəm	knʔəm	kmʔəm	kimʔəm
53a	paŋkah	kalaŋ	paŋkah	palaŋ	kalaŋ	jriwāŋ	paŋkah	kalaŋ	kalaŋ
53b	diʔ	kalaŋ	pincin	boh	boh	b-cuntuk	cuntuk	toŋket	kalaŋ
54	pilpal	pilpal	pilpal	pilpal	lmpah	lihpah	plpal	lmpah	tpoʔ
55	?	tmpuh	tiŋkoh	tihpuh	gilsil	tŋkik	tmpuh	tmpuh	tlisih
56	may	miymay	may	may	miymay	miymay	miymay	miymay	miymay
57	plpal	pilpal	pilpal	pilpal	lmpah	gucɔh	pal	pilpal	pilpal
58	?	kimʔəm	knʔəm	kimʔəm	kimʔəm	kmʔəm	kimʔəm	kmʔəm	kimʔəm
59	tnisdes	tmpuh	tmpuh	tmpuh	tnisdes	gilsil	silgil	snitkut	tlisih
60	pryə	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	pryə	pryə
61	knʔəm	kimʔəm	knʔəm	kimʔəm	kimʔəm	?	kimʔəm	kmʔəm	kimʔəm
62	b-salam	mŋamut	salam	salam	cipcɛp	mŋamut	salam	miʔmɛʔ	kumpul
63	salam	mŋamut	salam	salam	mŋamut	mŋamut	cipcɛp	salam	salam
64	pmpem	?	pmpem	pmpem	hitmɛt	bruŋal	hitmat	hitmɛt	?

	Jedek11	**Jedek5**	**Jedek37**	**Jedek15**	**Jedek29**	**Jedek32**	**Jedek25**	**Jedek26**	**Jedek33**
1	?	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	?	cɔl
2	kmʔəm	knʔəm	kmʔəm	kinʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kmʔəm
3	limpah	pilpal	lihpah	pilpal	b-baloh	lmpah	lihpah	piŋgɛŋ	pilpal
4	baɲci	ʔikʔɛk	ʔɛk	ʔikʔɛk	ʔikʔɛk	bagiʔ	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk
5	lmpah	bihlɔh	guhcɔh	?	b-baloh	lihpah	lihpah	lihpah	pilpal
6	miymay	may	may	may	miymay	may	may	miymay	may
7	kimʔəm	knʔəm	kmʔəm	kinʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm
8	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	ŋɔk
9	?	miʔmɛʔ	daɲin	jumpaʔ	cilcɔl	miʔmɛʔ	miʔmɛʔ	?	cɔl
10	miymay	miymay	miymay	miymay	kihwɔh	miymay	miymay	miymay	miymay
11	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cɔl
12	tipet	kirəʔ	ʔikʔɛk	cuntuk	ŋikŋɔk	kiʔrəʔ	kiʔrəʔ	ŋɔk	ktpit
13	mŋamut	salam	piŋgɛŋ	salam	mŋamut	salam	piŋgɛŋ	salam	cipcɛp
14	hitmat	hitmat	hitmɛt	hitmɛt	pmpem	lɔy	hitmat	hitmɛt	pmpem
15	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk
16	?	kimʔəm	?	knʔəm	kimʔəm	?	kimʔəm	kimʔəm	?
17	gucɔh	guhcɔh	gucɔh	p-disdəs	guhcɔh	?	gucɔh	tiʔmɛʔ	ktim
18	ŋikŋɔk	ŋɔk	ŋɔk	?	ŋɔk	ŋɔk	ŋɔk	?	ŋikŋɔk
19	baɲci	biʔgiʔ	ʔikʔɛk	bagiʔ	bagiʔ	biʔgiʔ	biʔgiʔ	ʔikʔɛk	bagiʔ
20	kmʔəm	knʔəm	kimʔəm	knʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm
21	baɲci	ʔikʔɛk	ʔikʔɛk	bagiʔ	biʔgiʔ	biʔgiʔ	ʔikʔɛk	ʔikʔɛk	biʔgiʔ
22	gilsil	tnihpuh	tihpuh	?	gilsil	gilsil	kiɲĩŋ	gilsil	gilsil
23	kmʔəm	knʔəm	kmʔəm	knʔəm	kmʔəm	kmʔəm	kmʔəm	kmʔəm	kmʔəm
24	pjaʔ	pjaʔ	pjaʔ	?	pjaʔ	pjaʔ	pjaʔ	?	?
25	ʔɛl	ʔɛl	ʔɛl	ʔɛl	?	cilkil	ʔɛl	?	ʔɛl
26	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
27	?	cŋcoŋ	?	?	hamat	coŋ	coŋ	cŋcoŋ	hay
28	baɲci	ʔikʔɛk	ʔɛk	ʔɛk	biʔgiʔ	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk
29	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm
30	baɲci	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔɛk
31	ʔɛl	ʔɛl	?	?	ʔɛl	ʔɛl	ʔɛl	ʔilʔɛl	?
32	?	ʔikʔɛk	ʔikʔɛk	?	biʔgiʔ	ʔikʔɛk	ʔikʔɛk	bagiʔ	ʔikʔɛk
33	ʔilʔɛl	ʔilʔɛl	ʔilʔɛl	ticmic	ʔilʔɛl	ʔilʔɛl	ʔilʔɛl	?	?
34	baɲci	ʔikʔɛk	ʔikʔɛk	bagiʔ	bagiʔ	biʔgiʔ	ʔikʔɛk	biʔgiʔ	?
35	himpon	susun	boh	boh	picin	cuntuk	kacep	susun	kapit
36	hitmat	?	htmɛt	?	?	hitmɛt	hitmat	hitmɛt	?
37	baɲci	biʔgiʔ	ʔikʔɛk	bagiʔ	bagiʔ	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	biʔgiʔ

38	salam	salam	ʔaŋket	pŋgɛŋ	mɲamut	pimpɔ̃m	piŋgɛŋ	pimpɔ̃m	cipcɛp
39	hitmat	lɔy	?	?	hamɛt	hamɛt	hitmat	hitmɛt	hamɛt
40	cilcɔl	cilcɔl	cilcɔl	cilcɔl	kicɛʔ	likluk	cilcɔl	cilcɔl	cilcɔl
41	toka	ʔikʔɛk	?	ʔikʔɛk	biʔgiʔ	toka	biʔgiʔ	biʔgiʔ	ʔikʔɛk
42	guhcɔh	?	guhcɔh	guhcɔh	guhcɔh	guhcɔh	guhcɔh	guhcɔh	tŋik
43	hitmat	hitmat	?	?	piŋraŋ	lɔy	hitmat	hitmɛt	hitmɛt
44	pilpal	pal	?	lihpah	prikluk	lihpah	lihpah	lihpah	tŋkik
45	miymay	miymay	miymay	miymay	miymay	miymay	miymay	miymay	miymay
46	?	ʔilʔɛl	?	?	?	ʔilʔɛl	ʔɛl	ʔilʔɛl	ʔilʔɛl
47	?	knʔəm	?	knʔəm	?	kimʔəm	kmʔəm	?	?
48	guhcɔh	guhcɔh	pilpal	guhcɔh	tiɲuŋ	tmpis	?	gihcɔh	kimtim
49	mɲamut	salam	piŋgɛŋ	pgɛŋ	mɲamut	mɲamut	piŋgɛŋ	pimpɔ̃m	?
50	pilpal	pilpal	pilpal	plpal	tihpɛh	lihpah	lihpah	pilpal	pilpal
51	may	may	may	may	may	may	may	may	may
52	kimʔəm	knʔəm	kimʔəm	knʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	?
53a	himpon	kalaŋ	gantoŋ	kalaŋ	kalaŋ	jrmtim	kalaŋ	kalaŋ	kalaŋ
53b	blato	kalaŋ	kalaŋ	boh	susun	cuntuk	hɲjaŋ	cuntuk	p-hajil
54	lmpah	pilpal	pilpal	lmpah	lmpah	lihpah	pikpak	pilpal	?
55	tnihpuh	tnihpuh	?	tmpuh	b-tŋka	slihsih	kiŋlĩŋ	gilsil	?
56	may	miymay	may	may	may	miymay	may	miymay	miymay
57	lmpah	pilpal	lihpah	pilpal	lihpah	lihpah	lihpah	pilpal	pilpal
58	kimʔəm	kimʔəm	kimʔəm	kinʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kmʔəm
59	tmpuh	tmpuh	tmpuh	tiŋkoh	tiŋkoh	tnisdes	kiŋlĩŋ	gilsil	tnisdes
60	pryə	ʔɛl	ʔɛl	pryə	puyə	ʔɛl	ʔɛl	puyə	ʔɛl
61	kimʔəm	kimʔəm	kmʔəm	kinʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	?
62	mɲamut	salam	sitmut	pŋgɛŋ	mɲamut	salam	piŋgɛŋ	pimpɔ̃m	cipcɛp
63	mɲamut	salam	mɲitmut	pgɛŋ	mɲamut	salam	piŋgɛŋ	pimpɔ̃m	cipcɛp
64	?	?	?	pmpem	pmpem	?	hitmat	hitmɛt	?

	Jedek16	**Jedek22**	**Jedek23**	**Jedek34**	**Jedek28**	**Jedek21**	**Jedek9**	**Jedek36**	**Jedek20**
1	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl
2	kmʔəm	kimʔəm	kmʔəm	?	kmʔəm	kmʔəm	knʔəm	kimʔəm	kmʔəm
3	pilpal	pilpal	lihpah	pilpal	pilpal	pilpal	plpal	lihpah	b-baloh
4	?	ʔikʔɛk	ʔikʔɛk	?	?	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	biʔgiʔ
5	pilpal	pilpal	lihpah	pilpal	baloh	pilpal	lmpah	lmpah	lmpah
6	may	miymay	may	may	may	miymay	diʔ	may	miymay
7	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm	knʔəm	kimʔəm	kimʔəm
8	ŋikŋɔk	ŋikŋɔk	?	ŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋɔk
9	jumpaʔ	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	?	cilcɔl
10	miymay	miymay	miymay	miymay	may	miymay	may	miymay	miymay
11	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cɔl	cilcɔl	cilcɔl	cilcɔl
12	ʔindin	siŋboŋ	kiʔrəʔ	cuktuk	ʔikʔɛk	ʔikʔɛk	ʔindin	ʔikʔɛk	cuntuk
13	piŋgɛŋ	salam	pimpɔ̃m	salam	salam	mɲamut	salam	sambut	piŋgɛŋ
14	hitmɛt	hitmɛt	hitmɛt	hamɛt	hamɛt	hitmɛt	pmpem	pmpem	pmpem
15	ŋikŋɔk	ŋɔk	?	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk
16	?	kimʔəm	kmʔəm	?	kimʔəm	kmʔəm	knʔəm	?	?
17	gucɔh	?	gucɔh	gucɔh	?	gucɔh	?	gucɔh	gucɔh
18	ŋikŋɔk	ŋikŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋikŋɔk
19	ʔikʔɛk	biʔgiʔ	biʔgiʔ	bagiʔ	bagiʔ	ʔɛk	ʔikʔɛk	ʔikʔɛk	biʔgiʔ
20	kimʔəm	kimʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	knʔəm	kimʔəm	kmʔəm
21	?	biʔgiʔ	biʔgiʔ	bagiʔ	ʔikʔɛk	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	biʔgiʔ
22	?	timpuh	kilsil	kilkil	?	krilsil	timpuh	gilsil	gilsil
23	kmʔəm	kmʔəm	kmʔəm	kmʔəm	kmʔəm	kmʔəm	knʔəm	kmʔəm	kmʔəm
24	?	pjaʔ	pjaʔ	pjaʔ	pjaʔ	pjaʔ	?	pjaʔ	pjaʔ

25	ʔɛl	?	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl
26	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
27	?	coŋ	cŋcoŋ	?	coŋ	hamɛt	cŋcoŋ	?	hiyhay
28	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	biʔgiʔ	bagiʔ	ʔikʔɛk	ʔikʔɛk	ʔɛk	biʔgiʔ
29	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm	kʔəm	kimʔəm	kimʔəm
30	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	?	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk
31	ʔɛl	ʔɛl	ʔɛl	?	?	?	ʔɛl	ʔɛl	ʔɛl
32	?	?	biʔgiʔ	ʔɛk	?	ʔikʔɛk	toka	ʔɛk	bagiʔ
33	?	ʔɛl	ʔilʔɛl	ʔilʔɛl	?	ʔilʔɛl	ʔɛl	ʔɛl	ʔilʔɛl
34	ʔikʔɛk	biʔgiʔ	biʔgiʔ	bagiʔ	baɲci	ʔɛk	bagiʔ	ʔikʔɛk	bagiʔ
35	kalaŋ	hatow	cuktuk	diʔ	diʔ	hatow	hatow	boh	susun
36	?	mayin	hitmɛt	hitmɛt	hitmɛt	hamɛt	prikluk	?	hitmɛt
37	ʔɛk	ʔikʔɛk	ʔɛk	ʔɛk	baɲci	ʔikʔɛk	bagiʔ	ʔɛk	biʔgiʔ
38	cipcɛp	salam	salam	salam	?	ʔaŋket	salam	mɲitmut	cipcɛp
39	hitmɛt	hitmɛt	?	?	hitmɛt	hamɛt	hitmɛt	?	hamɛt
40	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	kiʔcɛʔ	cilcɔl	cilcɔl	cilcɔl
41	ʔɛk	toka	toka	ʔikʔɛk	ʔɛk	ʔikʔɛk	toka	ʔikʔɛk	biʔgiʔ
42	guhcɔh	gustih	guhcɔh	guhcɔh	pilpal	guhcɔh	gucɔh	guhcɔh	guhcɔh
43	?	hitmat	hitmɛt	hitmɛt	?	hamɛt	hamɛt	?	hitmɛt
44	pilpal	lihpah	?	pilpal	pilpal	pilpal	pilpal	lihpah	pilpal
45	miymay	miymay	miymay	miymay	miymay	may	bayɛʔ	may	tikrik
46	?	ʔilʔɛl	ʔilʔɛl	ʔilʔɛl	?	?	ʔilʔɛl	?	?
47	?	?	kmʔəm	?	?	kmʔəm	knʔəm	?	kmʔəm
48	guhcɔh	guhcɔh	guhcɔh	guhcɔh	tiʔmboʔ	guhcɔh	guhcɔh	guhcɔh	guhcɔh
49	cipcɛp	?	salam	salam	salam	mɲamut	salam	mɲamut	sitmut
50	pilpal	lihpah	tiʔpoʔ	pilpal	pilpal	pilpal	pilpal	pilpal	lihpah
51	may	may	may	may	may	may	may	may	may
52	?	kimʔəm	kimʔəm	kmʔəm	kimʔəm	kmʔəm	knʔəm	kmʔəm	kimʔəm
53a	kalaŋ	jrawãŋ	cpek	diʔ	kalaŋ	kalaŋ	paŋkah	kalaŋ	b-kalaŋ
53b	tŋkat	cpek	tuŋkuʔ	diʔ	diʔ	saŋkut	ŋɔk	boh	tŋkat
54	pilpal	lihpah	lmpah	pilpal	lmpah	pilpal	pal	lmpah	pal
55	tmpuh	tihpuh	thpuh	silgil	gilsil	tlisih	tumpuh	tmpuh	kiŋlĩŋ
56	miymay	miymay	may	miymay	miymay	miymay	may	may	may
57	pilpal	lihpah	lihpah	pilpal	pilpal	pilpal	pal	pilpal	pilpal
58	kimʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kmʔəm	knʔəm	kimʔəm	?
59	tnisdes	tmpuh	klisil	silgil	tiŋkoh	krilsil	?	gilsil	gilsil
60	ʔɛl	puyə	puyə	puyə	ʔɛl	ʔɛl	ʔɛl	ʔɛl	pryə
61	kmʔəm	kimʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	knʔəm	kmʔəm	kimʔəm
62	cipcɛp	salam	salam	salam	salam	mɲamut	salam	mɲamut	cipcɛp
63	cipcɛp	salam	salam	salam	?	pimpõm	salam	?	sambut
64	pmpem	hitmat	hitmɛt	hitmɛt	hitmɛt	hamɛt	pimpem	pmpem	?

	Jedek14	**Jedek35**	**Jedek17**	**Jedek19**	**Jedek24**	**JRual11**	**JRual9**	**JRual13**	**JRual7**
1	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl
2	knʔəm	kmʔəm	kmʔəm	kimʔəm	kmʔəm	kimʔəm	kmʔəm	ʔiɲʔõɲ	?
3	pilpal	pilpal	pilpal	pilpal	bihlɔh	?	pikpak	lihpah	?
4	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔɛk	ʔikʔɛk	?
5	pilpal	lmpah	lmpah	pilpal	lihpah	lihpah	pilpal	prikluk	guhcɔh
6	lawat	miymay	may	may	may	miymɛy	mɛy	mɛy	may
7	kmʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm	kimʔəm
8	ŋɔk	ŋikŋɔk	ŋikŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	?
9	cilcɔl	?	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	?
10	may	miymay	may	may	miymay	miymɛy	mɛy	mɛy	mɛy
11	cilcɔl	cilcɔl	bihlɔh	cilcɔl	cilcɔl	cilcɔl	cɔl	cilcɔl	cilcɔl

12	ʔndɨn	ciktuk	cuntuk	cuntuk	ʔandɨn	ʔikʔɛk	cuntuk	pet	ŋɔk
13	pgɛŋ	sitmut	salam	salam	cipcɛp	ʔitkit	mɲamut	sambut	piŋgɛŋ
14	hitmɛt	pmpem	liylɔy	hitmɛt	pmpem	?	hitmɛt	hamat	pmpem
15	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk
16	kmʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kimʔəm	ʔiɲʔɔ̃ɲ	?	kimʔəm
17	tumboʔ	tumboʔ	guhcɔh	gucɔh	?	tumboʔ	?	?	tumboʔ
18	ŋikŋɔk	ŋɔk	?	ŋɔk	ŋɔk	?	ŋɔk	ŋɔk	ŋɔk
19	ʔikʔɛk	biʔgiʔ	bagiʔ	bagiʔ	ʔikʔɛk	biʔgiʔ	biʔgiʔ	ʔɛk	biʔgiʔ
20	kmʔəm	kimʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm
21	bagiʔ	?	bagiʔ	bagiʔ	ʔikʔɛk	biʔgiʔ	biʔgiʔ	ʔikʔɛk	biʔgiʔ
22	tmpuh	?	tmpuh	gilsil	tŋkik	tihkoh	gilsil	?	?
23	knʔəm	kmʔəm	kmʔəm	kmʔəm	kmʔəm	kmʔəm	kmʔəm	ʔmʔɔm	kmʔəm
24	pjaʔ	?	pjaʔ	pjaʔ	pjaʔ	pjaʔ	?	pjaʔ	pjaʔ
25	ʔɛl	?	ʔɛl	ʔɛl	?	ʔɛl	?	?	ʔɛl
26	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	?	ʔɛk	ʔɛk
27	cŋcoŋ	cŋcoŋ	pdɛp	hamɛt	?	hamat	hay	hamɛt	?
28	ʔikʔɛk	?	toka	bagiʔ	ʔikʔɛk	biʔgiʔ	bagiʔ	ʔɛk	biʔgiʔ
29	kmʔəm	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm
30	ʔikʔɛk	ʔikʔɛk	biʔgiʔ	biʔgiʔ	ʔɛk	biʔgiʔ	bagiʔ	ʔikʔɛk	biʔgiʔ
31	ʔɛl	?	ʔɛl	ʔɛl	ʔɛl	?	?	?	ʔɛl
32	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔɛk	ʔikʔɛk	ʔikʔɛk	ʔɛk	ʔɛk	?
33	?	ʔɛl	ʔilʔɛl	ʔilʔɛl	?	titlət	?	?	ʔilʔɛl
34	ʔikʔɛk	ʔikʔɛk	biʔgiʔ	biʔgiʔ	biʔgiʔ	ʔikʔɛk	bagiʔ	ʔɛk	biʔgiʔ
35	cuntuk	picin	cuntuk	boh	diʔ	plagaʔ	diʔ	boh	boh
36	hitmɛt	?	hitmɛt	hitmɛt	?	hitmat	hitmɛt	hamɛt	hamɛt
37	ʔikʔɛk	ʔikʔɛk	bagiʔ	biʔgiʔ	bagiʔ	biʔgiʔ	biʔgiʔ	ʔɛk	biʔgiʔ
38	piŋgɛŋ	sitmut	salam	salam	?	sitmut	salam	sambut	piŋgɛŋ
39	?	?	hitmɛt	hitmɛt	hitmɛt	hitmat	hitmɛt	hitmat	hitmɛt
40	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cɔl
41	ʔikʔɛk	bagiʔ	toka	toka	ʔikʔɛk	ʔikʔɛk	toka	torkar	biʔgiʔ
42	guhcɔh	guhcɔh	guhcɔh	guhcɔh	guhcɔh	tiʔmboʔ	gihcɔh	guhcɔh	guhcɔh
43	hitmɛt	hitmɛt	hitmɛt	hitmɛt	?	hitmat	hamɛt	hamɛt	hamɛt
44	pilpal	lmpah	lmpah	tampəh	tiɲuɲ	lihpah	kimtim	lihpah	?
45	miymay	miymay	miymay	may	miymay	miymɛy	mɛy	mɛy	mɛy
46	ʔilʔɛl	?	?	?	?	?	?	?	?
47	?	?	kmʔəm	kimʔəm	?	?	?	?	?
48	tiʔmboʔ	?	guhcɔh	guraw	tiɲuɲ	tiʔmboʔ	gihcɔh	pihkoh	guhcɔh
49	piŋgɛŋ	sitmut	sambut	salam	sitmut	simlam	mɲitmut	sambut	piŋgɛŋ
50	pilpal	lihpah	pilpal	pilpal	tampəh	lmpah	lihpah	pikyok	tiʔpoʔ
51	may	may	may	may	may	mɛy	?	mɛy	mɛy
52	?	kimʔəm	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	?
53a	kalaŋ	kalaŋ	kalaŋ	kalaŋ	diʔ	palaŋ	diʔ	boh	kalaŋ
53b	picin	picin	cuntuk	p-hajɨl	diʔ	diʔ	diʔ	cuntuk	boh
54	lmpah	lahɔm	lmpah	pilpal	lihpah	lihpah	lihpah	lmpah	lmpah
55	tmpuh	tmpuh	tmpuh	gilsil	?	gilsil	tihpuh	tiŋndaŋ	guhcɔh
56	may	miymay	miymay	may	miymay	miymɛy	mɛy	mɛy	mɛy
57	pilpal	pilpal	pilpal	pilpal	pikyok	lmpah	pikpak	lihpah	lihpah
58	kmʔəm	kmʔəm	kmʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kimʔəm	kmʔəm
59	tmpuh	tmpuh	gilsil	sipat	tnisdes	gilsil	glsil	gilsil	gilsil
60	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	pryər	ʔɛl	ʔɛl	ʔɛl
61	kmʔəm	?	kmʔəm	kimʔəm	kimʔəm	?	kimʔəm	ʔiɲʔɔ̃ɲ	kimʔəm
62	piŋgɛŋ	diʔ	salam	salam	pimpɔ̃m	sitmut	mɲamut	sambut	piŋgɛŋ
63	piŋgɛŋ	sitmut	sambut	salam	sitmut	pimpɔ̃m	mɲamut	sambut	piŋgɛŋ
64	pmpem	?	hitmɛt	hitmɛt	pmpem	hitmat	hitmɛt	hamɛt	hamat

	JRual14	JRual4	JRual15	JRual2	JRual12	JRual1	JRual8	JRual10	JRual18
1	cilcɔl	ʔ	kicɛʔ	cilcɔl	cilcɔl	ʔ	cilcɔl	cilcɔl	cilcɔl
2	kimʔəm	kmʔəm	kimʔəm	knʔəm	kmʔəm	knʔəm	kimʔəm	kmʔəm	kimʔəm
3	lihpah	bihlɔh	lmpah	pilpal	tampəh	pilpal	limpah	lihpah	pilpal
4	ʔikʔɛk	ʔɛk	ʔ	bagiʔ	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk
5	lihpah	lmpah	guhcɔh	pilpal	lihpah	pilpal	pilpal	pak	pilpal
6	miymay	lawat	mɛy	may	mɛy	may	miymay	mɛy	miymay
7	kimʔəm	kimʔəm	kimʔəm	kmʔəm	ʔmʔəm	knʔəm	kmʔəm	kmʔəm	kimʔəm
8	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk
9	ʔilʔɛl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	?	cilcɔl	cilcɔl
10	mɛy	lawat	may	may	?	miymay	?	mɛy	miymɛy
11	cilcɔl	misnɛs	kicɛʔ	cilcɔl	cilcɔl	bihlɔh	cilcɔl	cilcɔl	cilcɔl
12	cuktuk	cuntuk	ʔɛk	ʔindin	ʔindin	cuntuk	ciktuk	siŋboŋ	cuntuk
13	sambut	salam	sambut	salam	sambut	salam	pimpõm	mŋamut	pimpõm
14	hitmat	hitmat	pmpem	hitmɛt	pmpem	hamɛt	hitmɛt	hitmat	hitmɛt
15	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋɔk
16	?	kmʔəm	kimʔəm	?	?	knʔəm	kimʔəm	?	kimʔəm
17	gucɔh	katuŋ	?	guhcɔh	gucɔh	?	gucɔh	gucɔh	?
18	ŋɔk	ŋikŋɔk	ŋɔk	ŋikŋɔk	ŋikŋɔk	ŋikŋɔk	ŋikŋɔk	ŋɔk	?
19	ʔikʔɛk	biʔgiʔ	ʔikʔɛk	bagiʔ	ʔikʔɛk	biʔgiʔ	biʔgiʔ	ʔikʔɛk	biʔgiʔ
20	kimʔəm	kimʔəm	kmʔəm	knʔəm	kmʔəm	knʔəm	kimʔəm	kmʔəm	kimʔəm
21	biʔgiʔ	bagiʔ	ʔɛk	bagiʔ	?	ʔikʔɛk	biʔgiʔ	ʔɛk	biʔgiʔ
22	?	gilsil	?	?	tnihpuh	tmpuh	kilhil	gilsil	gilsil
23	kmʔəm	kmʔəm	?	?	?	knʔəm	kmʔəm	kmʔəm	kmʔəm
24	pjaʔ	pjaʔ	pjaʔ	pjaʔ	?	pjaʔ	pjaʔ	?	?
25	ʔilʔɛl	?	ʔɛl	?	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔilʔɛl
26	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
27	htmat	hay	b-ʔikut	dɔp	hay	hamɛt	hamɛt	hay	?
28	ʔikʔɛk	bagiʔ	bagiʔ	bagiʔ	biʔgiʔ	biʔgiʔ	ʔikʔɛk	ʔɛk	bagiʔ
29	kimʔəm	?	kmʔəm	kimʔəm	kmʔəm	knʔəm	kimʔəm	kmʔəm	kimʔəm
30	ʔikʔɛk	ʔikʔɛk	bagiʔ	bagiʔ	ʔikʔɛk	bagiʔ	ʔikʔɛk	ʔikʔɛk	biʔgiʔ
31	ʔɛl	ʔɛl	ʔɛl	?	?	ʔɛl	ʔɛl	?	ʔɛl
32	biʔgiʔ	ʔikʔɛk	bagiʔ	?	ʔikʔɛk	biʔgiʔ	bagiʔ	ʔɛk	biʔgiʔ
33	ʔilʔɛl	ʔilʔɛl	?	ʔɛl	ʔilʔɛl	ʔilʔɛl	?	ʔilʔɛl	?
34	bagiʔ	ʔikʔɛk	bagiʔ	bagiʔ	ʔɛk	biʔgiʔ	biʔgiʔ	ʔɛk	biʔgiʔ
35	cuntuk	boh	taŋkok	boh	boh	tihpɛh	boh	giʔ	cuntuk
36	hitmat	?	hoŋal	prikluk	?	hitmɛt	?	hamat	hitmɛt
37	bagiʔ	bagiʔ	baŋci	bagiʔ	ʔikʔɛk	bagiʔ	biʔgiʔ	ʔikʔɛk	biʔgiʔ
38	sitmut	salam	mŋamut	salam	sambut	salam	piŋgɛŋ	mŋamut	pimpõm
39	?	?	?	?	?	hitmɛt	?	?	hitmɛt
40	cilcɔl	kiʔcɛʔ	kicɛʔ	cilcɔl	cɔl	cilcɔl	clcɔl	cilcɔl	cilcɔl
41	b-gilir	toka	ʔikʔɛk	toka	tokar	toka	biʔgiʔ	ʔikʔɛk	ʔikʔɛk
42	guhcɔh	guhcɔh	guhcɔh	guhcɔh	guhcɔh	guhcɔh	gihcɔh	lihpah	guhcɔh
43	hitmat	hitmat	?	?	hitmat	?	hitmɛt	?	hitmɛt
44	lihpah	pilpal	guhcɔh	pal	kimtim	lmpah	gihcɔh	?	pilpal
45	mɛy	miymɛy	mɛy	may	drəc	miymay	ricrẽc	mɛy	miymay
46	ʔilʔɛl	?	ʔilʔɛl	?	ʔilʔɛl	ʔilʔɛl	?	?	?
47	kimʔəm	?	?	?	kmʔəm	?	?	kmʔəm	?
48	kiʔtoʔ	guhcɔh	guhcɔh	tiŋkok	tictɔc	gilsil	plpal	gucɔh	pilpal
49	sitmut	salam	mŋamut	salam	sambut	salam	pimpõm	mŋamut	pimpõm
50	lihpah	pilpal	tihpuh	pilpal	lmpah	pilpal	pilpal	pah	pilpal
51	may	mɛy	?	may	mɛy	may	miymay	mɛy	may
52	kimʔəm	kmʔəm	kimʔəm	knʔəm	kmʔəm	knʔəm	kimʔəm	kmʔəm	kimʔəm

53a	boh	kalaŋ	kalaŋ	kalaŋ	sandɛŋ	hatuh	kalaŋ	kalaŋ	pihkəh
53b	toŋkat	toŋkat	picin	diʔ	boh	diʔ	boh	diʔ	piwek
54	lihpah	pilpal	tihpuh	pilpal	tampəh	pal	?	lmpah	pilpal
55	tlisih	kilsil	tlihsih	?	?	pal	khil	gilsil	?
56	may	mɛy	may	may	mɛy	may	miymay	mɛy	miymay
57	lihpah	pilpal	?	pilpal	tampəh	pilpal	plpal	lihpah	pilpal
58	kimʔəm	kimʔəm	kimʔəm	?	ʔmʔəm	kinʔəm	kmʔəm	kmʔəm	kimʔəm
59	gilsil	tiŋkoh	tlihsih	tmpuh	tmpuh	gilsil	gisɛʔ	gilsil	glsil
60	pryər	ʔɛl	ʔɛl	ʔɛl	?	pryə	ʔilʔɛl	ʔɛl	ʔilʔɛl
61	kimʔəm	kimʔəm	kimʔəm	knʔəm	knʔəm	knʔəm	kmʔəm	kmʔəm	kimʔəm
62	sitmut	mŋamut	krumun	salam	sambut	salam	pimpɔ̃m	mŋamut	pimpɔ̃m
63	sitmut	piŋgɛŋ	jabat	salam	?	salam	cipcip	mŋamut	pimpɔ̃m
64	hitmat	hitmat	hoŋal	hitmɛt	hitmat	prikluk	?	hamat	hitmɛt

	JRual17	**JManok2**	**JManok3**	**JManok6**	**JManok5**	**JManok7**	**JBanun8**	**JBanun7**	**JBanun5**
1	cɔl	cilcɔl	?	cilcɔl	cilcɔl	bhariʔ	cilcɔl	cilcɔl	cilcɔl
2	ʔiŋʔɔ̃ŋ	kmʔəm	ʔimʔəm	?	kmʔəm	kmʔəm	?	?	ʔimʔəm
3	guhcɔh	tihbɔh	tihbɔh	lmpah	tihbɔh	lihpah	tpoʔ	sihlah	tihbɔh
4	ʔɛk	biʔgiʔ	prihleh	ʔɛk	bagiʔ	ʔikʔɛk	?	?	ʔikʔɛk
5	pilpal	lmpah	tihbɔh	lmpah	tihbɔh	lihpah	sayaŋ	tihbɔh	tihbɔh
6	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mamɛy
7	kimʔəm	kmʔəm	kmʔəm	ʔimʔəm	kmʔəm	kmʔəm	ʔimʔəm	ʔimʔəm	ʔimʔəm
8	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	ŋɔk	ŋɔk	ŋɔk
9	?	lihmah	cilcɔl	?	cilcɔl	taɲɛʔ	sihsəh	taɲɛʔ	cilcɔl
10	mɛy	mɛy	mɛy	miymɛy	mɛy	mɛy	?	mɛy	mamɛy
11	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	simbaŋ	simbaŋ	cilcɔl
12	ŋɔk	cntuk	ŋikŋɔk	kitcit	ʔɛk	cuntuk	ŋɔk	ŋɔk	cundoʔ
13	ʔaŋket	mŋamut	mŋamut	piŋgɛŋ	mŋamut	mŋamut	slamat	slamat	slamat
14	hamɛt	hitmat	hitmɛt	tampɛŋ	hamat	hamat	hitmɛt	hitmat	tampɛŋ
15	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk
16	kmʔəm	kmʔəm	?	?	kimʔəm	?	ʔimʔəm	ʔmʔəm	ʔmʔəm
17	gucɔh	?	?	?	?	?	tumboʔ	?	tumboʔ
18	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋɔk	ŋikŋɔk	?	ŋɔk	ŋɔk
19	ʔikʔɛk	bagiʔ	biʔgiʔ	ʔikʔɛk	bagiʔ	?	ʔikʔɛk	ʔɛk	ʔikʔɛk
20	kimʔəm	kmʔəm	kmʔəm	kimʔəm	kmʔəm	kmʔəm	ʔimʔəm	ʔmʔəm	ʔmʔəm
21	biʔgiʔ	bagiʔ	bagiʔ	?	bagiʔ	ʔɛk	?	?	ʔikʔɛk
22	?	tŋkik	tnikik	tnikik	kiswẽs	dɛs	?	b-gadoh	laŋah
23	ʔɔ̃ŋ	kmʔəm	kmʔəm	ʔtʔɔ̃t	kmʔəm	kmʔəm	ʔmʔəm	ʔmʔəm	ʔmʔəm
24	pjaʔ	?	pjaʔ	?	pjaʔ	pjaʔ	?	pjaʔ	pjaʔ
25	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	?	ʔɛl	ʔɛl
26	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk	ʔɛk
27	htmɛt	hay	hay	?	hay	?	hamat	hamat	hay
28	biʔgiʔ	bagiʔ	bagiʔ	ʔikʔɛk	bagiʔ	ʔɛk	?	?	ʔikʔɛk
29	kimʔəm	kmʔəm	kimʔəm	ʔtʔɔ̃t	kmʔəm	kmʔəm	ʔimʔəm	ʔmʔəm	ʔimʔəm
30	biʔgiʔ	bagiʔ	biʔgiʔ	?	bagiʔ	ʔɛk	ʔikʔɛk	ʔɛk	ʔikʔɛk
31	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	?	?	ʔɛl
32	?	ʔikʔɛk	bagiʔ	ʔikʔɛk	?	ʔɛk	ʔikʔɛk	?	ʔikʔɛk
33	ʔɛl	ʔɛl	?	ʔilʔɛl	?	?	ʔilʔɛl	ʔɛl	ʔilʔɛl
34	bagiʔ	bagiʔ	biʔgiʔ	ʔɛk	bagiʔ	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk	ʔɛk
35	boh	blatur	boh	picin	boh	diʔ	cundoʔ	boh	tiltul
36	hitmɛt	hamat	hamat	hitmat	hitmat	hamat	hitmat	hitmat	tampɛŋ
37	bagiʔ	bagiʔ	biʔgiʔ	bagiʔ	bagiʔ	ʔikʔɛk	ʔikʔɛk	baɲci	ʔikʔɛk
38	ʔɛk	mŋamut	?	piŋgɛŋ	mŋamut	mŋamut	slamat	slamat	slamat
39	?	hitmat	hitmat	?	hamat	hamat	?	hitmat	tampɛŋ

40	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	cilcɔl	?	simbaŋ	cilcɔl
41	ʔikʔɛk	tokar	prileh	biʔgiʔ	bagiʔ	ʔɛk	ʔikʔɛk	ʔɛk	ʔikʔɛk
42	guhcɔh	guhcɔh	gihcɔh	?	guhcɔh	plagɔh	tumboʔ	?	tiʔmboʔ
43	?	plagɔh	hitmat	hitmat	hamat	hamat	hamat	?	tampɛŋ
44	kusɛʔ	lmpah	guhcɔh	lihpah	lihpah	tihbɔh	?	tbɔh	tihbɔh
45	mɛy	miymɛy	miymɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mamɛy
46	ʔɛl	ʔilʔɛl	?	ʔilʔɛl	?	?	?	?	ʔɛl
47	?	?	?	?	?	?	?	?	ʔmʔəm
48	pilpal	tiɲuŋ	guhcɔh	crigɔʔ	tihbɔh	guhcɔh	kimkəm	?	tihbɔh
49	ʔaŋket	mŋamut	mŋamut	piŋgɛŋ	mŋamut	mŋamut	slamat	slamat	slamat
50	lihpah	tihbɔh	tihbɔh	bihlɔh	tihbɔh	gitnit	wiʔliʔlɛʔ	tihbɔh	tihbɔh
51	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy
52	kimʔəm	kmʔəm	kmʔəm	kimʔəm	kmʔəm	kmʔəm	ʔimʔəm	?	ʔimʔəm
53a	kalaŋ	boh	kalaŋ	sripɛŋ	kalaŋ	cpɛ̄k	diʔ	ŋɔk	tuŋkar
53b	piwek	toŋkal	boh	piwek	diʔ	toŋkat	boh	ŋɔk	?
54	lihpah	tpoʔ	tbɔh	lmpah	tihbɔh	tihbɔh	tpoŋ	tihbɔh	tihbɔh
55	klĩŋ	tŋkik	?	slitkut	sikəh	dʔɛ̄ʔ	?	sihlah	laŋah
56	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	mɛy	miymɛy
57	bihlɔh	tihbɔh	tihbɔh	bihlɔh	tihbɔh	plagɔh	?	tihbɔh	tihbɔh
58	kimʔəm	kmʔəm	kmʔəm	kimʔəm	kmʔəm	kmʔəm	ʔimʔəm	ʔimʔəm	ʔmʔəm
59	tnisdes	tŋkik	tŋkik	tŋkik	kwɛ̄s	səh	puk	gutsĭt	laŋah
60	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	ʔɛl	pryər
61	kimʔəm	kmʔəm	?	ʔtʔɔ̄t	kmʔəm	kmʔəm	ʔmʔəm	ʔimʔəm	ʔimʔəm
62	ʔaŋket	mŋamut	mŋamut	mŋamut	mŋamut	mŋamut	slamat	slamat	slamat
63	ʔaŋket	mŋamut	sambut	cipcɛp	mŋamut	cɛp	piŋgɛŋ	slamat	slamat
64	lɔy	pmpem	hitmat	tampɛŋ	?	hamat	hitmat	hitmat	tampɛŋ

	JBanun3	**JBanun2**	**JBanun4**
1	simbaŋ	cilcɔl	cilcɔl
2	ʔmʔəm	ʔimʔəm	kimʔəm
3	tampəh	lmpah	lmpah
4	?	ʔaʔɛk	ʔikʔɛk
5	tampəh	lihpah	lmpah
6	miymɛy	mɛy	miymay
7	ʔimʔəm	ʔmʔəm	kimʔəm
8	ŋɔk	ŋɔk	ŋɔk
9	b-cilcɔl	cilcɔl	?
10	mɛy	mamɛy	mɛy
11	cilcɔl	cilcɔl	cilcɔl
12	ʔɛk	kicit	snaman
13	slamat	slamat	ngamat
14	tampɛŋ	tampɛŋ	hitmɛt
15	ŋɔk	ŋɔk	ŋɔk
16	ʔmʔəm	ʔmʔəm	kimʔəm
17	tumboʔ	tumboʔ	tumboʔ
18	ŋɔk	ŋɔk	ŋɔk
19	ʔikʔɛk	ʔaʔɛk	bagiʔ
20	ʔmʔəm	ʔmʔəm	kimʔəm
21	ʔikʔɛk	ʔaʔɛk	ʔikʔɛk
22	laŋah	kiŋniŋ	gilsil
23	ʔmʔəm	ʔimʔəm	kmʔəm
24	hay	pjaʔ	pjaʔ
25	ʔɛl	ʔɛl	ʔɛl
26	ʔɛk	ʔɛk	ʔɛk

27	hay	?	hay
28	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk
29	ʔimʔəm	ʔimʔəm	kimʔəm
30	ʔikʔɛk	ʔikʔɛk	biʔgiʔ
31	ʔɛl	ʔɛl	ʔɛl
32	ʔɛk	ʔikʔɛk	bagiʔ
33	ʔɛl	ttlət	?
34	ʔikʔɛk	ʔɛk	?
35	picin	boh	cundoʔ
36	tampɛŋ	hitmat	tampɛŋ
37	ʔikʔɛk	ʔikʔɛk	ʔikʔɛk
38	slamat	slamat	?
39	tampɛŋ	tampɛŋ	tampɛŋ
40	cilcɔl	sitkət	cilcɔl
41	tokar	tokar	trkar
42	tuʔmoʔ	tumboʔ	guhcɔh
43	hamat	tampɛŋ	?
44	tampəh	lmpah	tuʔmoʔ
45	ʔmʔəm	tək	miymay
46	ʔɛl	ʔɛl	?
47	ʔmʔəm	ʔmʔəm	kmʔəm
48	ktim	tuʔmoʔ	?
49	cɛp	slamat	piŋgɛŋ
50	tampəh	lihpah	lihpah
51	mɛy	mɛy	?
52	ʔimʔəm	ʔmʔəm	kmʔəm
53a	pi-wek	cundoʔ	kalaŋ
53b	?	?	p-hɲjaɲ
54	lmpah	lihpah	teʔpoʔ
55	laŋah	kiŋniŋ	taŋkis
56	mɛy	mamɛy	mɛy
57	lihpah	lmpah	lihpah
58	ʔimʔəm	ʔmʔəm	kmʔəm
59	laŋah	duk	tlihsih
60	ʔɛl	ʔɛl	ʔɛl
61	ʔimʔəm	ʔmʔəm	kmʔəm
62	cɛp	slamat	b-jabat
63	cɛp	slamat	b-jabat
64	tampɛŋ	hitmat	pmpem

Appendix E: Python script used for coding the data

```
#!/usr/bin/env python2
# -*- coding: utf-8 -*-
"""
Created on Mon Sep 23 09:28:24 2019
@author: Joanne Yager
"""
import pandas as pd
import numpy as np

### Define nexus() function
def nexus(raw_data,my_alphabet,filename_out):
    """
    nexus() automatically codes a dataset, column by column, replacing each
    word in a column with a unique code (a-z) for that word.
    It outputs a nexus file (.txt) containing the coded data along with other
    information needed in order to be read in SplitsTree to conduct NeighborNet
    analysis of the distances between the speakers of the sample.
    It takes three elements as input: raw_data is a dataframe containing the
    responses of each speaker to each item, my_alphabet is an alphabet where the
    symbols to be used for coding are specified, and a filename_out is the filename
    with which to name the nexus file produced by the function.

    The nexus file contains three sections: the taxlabels section that contains
    the names of the speakers, the symbols section that contains the symbols
    used in the coding, and the characters section that contains the coded data.
    """
```

```
def replace_all(text,dictionary):
    """
    replace_all() takes 2 arguments: an original text and a dictionary
    that tells the function the code to replace the words in the
    original text with.
    """
    for i,j in dictionary.iteritems():
        text = text.replace(i,j)
    return text

# coding loop for multiple sequence alignment:
raw_data_t = raw_data.T
data = []
for col in raw_data_t:
    unique1 = list(set(raw_data_t[col]))
    unique2 = [s.strip('?') for s in unique1]
    bowped_dict = dict(zip(unique2, my_alphabet))
    bowped_dict['?'] = '?'
    columns = replace_all(raw_data_t[col], bowped_dict)
    data.append(columns)
df = pd.DataFrame(data).T
headers = list(df.index)

# begin nexus file

f_out = open(filename_out,'w')

# taxlabels section
```

```
f_out.write('#NEXUS'+'\n'+
        'BEGIN TAXA;'+'\n'+
        'DIMENSIONS ntax='+str(len(headers))+';'+'\n'+
        'TAXLABELS'+'\n')
for i in range(len(headers)):
   labels = ''.join("'"+headers[i]+"'"+'\n')
   f_out.write(labels)

# symbols section
f_out.write(';'+'\n'+
        'END;'+'\n'+
        'BEGIN Unaligned;'+'\n'+
        'DIMENSIONS ntax='+str(len(headers))+';'+'\n'+
        'FORMAT'+'\n'+
        'datatype=STANDARD'+'\n'+
        'MISSING=?'+'\n')

symbols = list(np.unique(df.values.ravel()))
symbols2 = symbols.remove('?')

f_out.write('symbols="'+(' '.join(symbols))+'"'+'\n'+'labels=left'+'\n'+
        ';'+'\n'+'MATRIX'+'\n')

# characters section
taxa = [''.join("'"+headers[i]+"'") for i in range(len(headers))]
alignment = [''.join(df.iloc[r])+','+'\n' for r in range(len(headers)-1)]

for i in range(len(headers)-1):
```

```
        a_b = ''.join(taxa[i]+alignment[i])
        f_out.write(a_b)

    f_out.write(''.join("'"+headers[-1]+"'")+''.join(df.iloc[-1])+';'+'\n')

    f_out.write('END;')
    f_out.close()

### Define my_alphabet
my_alphabet = ['a','b','c','d','e','f','g','h','i','j','k','l','m','n','o','p','q','r','s','t','u','v','w','x','y',
'z','1','2','3','4']

### Read in the datafiles
basic_vocabulary_data = pd.read_csv('basic_vocabulary_data.txt',sep="\t",usecols=lambda x: 'Unnamed' not in x)
topological_relations_data = pd.read_csv('topological_relations_data.txt',sep="\t",usecols=lambda x: 'Unnamed' not in x)
caused_motion_data = pd.read_csv('caused_motion_data.txt',sep="\t",usecols=lambda x: 'Unnamed' not in x)
reciprocal_data = pd.read_csv('reciprocal_data.txt',sep="\t",usecols=lambda x: 'Unnamed' not in x)

### Output NEXUS files
filename_out_nex_basic_vocabulary = 'basic_vocabulary_nexus.txt'
nexus(basic_vocabulary_data,my_alphabet,filename_out_nex_basic_vocabulary)
filename_out_nex_topological_relations = 'topological_relations_nexus.txt'
nexus(topological_relations_data,my_alphabet,filename_out_nex_topological_relations)
```

```
filename_out_nex_caused_motion = 'caused_motion_nexus.txt'
nexus(caused_motion_data,my_alphabet,filename_out_nex_caused_motion)
filename_out_nex_reciprocal = 'reciprocal_nexus.txt'
nexus(reciprocal_data,my_alphabet,filename_out_nex_reciprocal)
```

Study III

Article

The International Journal of Bilingualism

International Journal of Bilingualism
1–16

DOI: 10.1177/1367006918814378
journals.sagepub.com/home/ijb
SAGE

Asymmetric semantic interaction in Jedek-Jahai bilinguals: Spatial language in a small-scale, non-standardized, egalitarian, long-term multilingual setting in Malaysia

Joanne Yager and Marianne Gullberg
Lund University, Sweden

Abstract

Aims and objectives/purpose/research questions: We investigate semantic interaction in bilinguals' topological relations descriptions in a small-scale, non-standardized, egalitarian, long-term multilingual setting in Malaysia.

Design/methodology/approach: Two groups of bilingual speakers of Jedek and Jahai (8 Jedek-identifying, 6 Jahai-identifying bilinguals) and two groups of monolingual Jedek and Jahai speakers (15 Jedek, 3 Jahai speakers) described the Topological Relations Picture Series in a director-matcher task, the bilinguals completing the task in both Jedek and Jahai.

Data and analysis: We compare the semantic boundaries of Jedek and Jahai topological relation markers (TRMs) as used by the monolingual and bilingual groups in extension maps and congruence analyses. The analyses focus on the TRM *klɛŋ*, which is identical in form but semantically different in the two varieties.

Findings/conclusions: We find evidence for asymmetric interaction in the expression of topological relations in Jedek and Jahai, with bidirectional influences in the Jahai-identifying bilinguals and a unidirectional influence of Jedek on Jahai in the Jedek-identifying bilinguals. This is commensurate with predictions based on Muysken's framework of bilingual optimization strategies.

Originality: The analyses shed new light on the nature of semantic interaction in bilingual systems by providing evidence from hitherto understudied bilingual language production in small-scale, non-standardized, egalitarian settings.

Significance/implications: The results suggest that Muysken's model is useful for understanding different bilingual outcomes, and highlight the complexity and connectedness of bilingual semantic systems. They also stress the need for more work in a variety of bilingual settings if we are to more fully understand the nature of bilingual systems.

Corresponding author:
Joanne Yager, Centre for Languages and Literature, Lund University, Sweden.
Email: joanne.yager@ling.lu.se

Keywords
Egalitarian bilingualism, topological relations, Jahai, Jedek, semantic interaction, crosslinguistic influence

Introduction

There is a great deal of variation in how monolingual speakers of different languages divide semantic space in different domains such as color, taste, odor, objects, spatial relations, kinship, and actions (see Malt & Majid, 2013 for an overview). For example, while English distinguishes between containment ('in') and support from below ('on'), Korean distinguishes between tight versus loose fit, cross-cutting the two English categories (Choi, 2006). Individuals who speak more than one language must therefore deal with differences in form-meaning associations in their different languages. Many studies have shown that bilinguals' languages are subject to interaction, rather than being separate, at all levels of language (see Pavlenko, 2011 for an overview). The current study specifically addresses the effects of contact between different semantic systems in the minds of bilinguals. A range of studies have already shown that interaction between bilinguals' semantic systems is often reflected in increased similarity between the languages, and in differences between the semantic systems of bilinguals and monolinguals. However, most research has studied bilingualism involving large-scale, largely monolingual speech communities with standardized languages. It therefore remains unclear how semantic interaction is affected when speech communities are small, mostly bilingual, with non-standardized languages in long-term contact. Therefore, this study investigates bilingual semantic interaction in spoken production in a small-scale bilingual speech community in northern Peninsular Malaysia.

Background

A substantial body of research shows that bilinguals' semantic systems interact in both comprehension and production. Such interaction or crosslinguistic influence is discussed in terms of the direction of the influence, the processes underlying it, and its outcomes (Jarvis & Pavlenko, 2008 for an overview). Several studies have investigated the phenomenon of convergence, generally defined as an increase in resemblance between bilinguals' languages in some respect, as compared to monolingual speakers. Some researchers consider convergence as an 'in-between' pattern between two monolingual patterns (Pavlenko, 2011; Treffers-Daller & Tidball, 2015). Others view it as unique patterns different both from the source and the target languages (Basseti & Cook, 2011). Recent work discussing the direction and effects of convergence has shown that convergence may affect each language differently in different linguistic domains within the same language pair (Alferink, 2015). In the current study we use the relatively neutral term 'interaction' to refer to the influence of two or more languages on each other regardless of direction or pattern of outcome (cf. Gathercole & Moawad, 2010).

Evidence of semantic interaction in language production is attested in various domains. For example, studies of object naming in Belgian Dutch (Flemish)-French early simultaneous bilinguals show evidence of partially merged systems, diverging from those of monolingual Dutch and French speakers (e.g. Ameel, Malt, Storms, & Van Assche, 2009; Ameel, Storms, Malt, & Sloman, 2005; White, Malt, & Storms, 2017). For instance, the influence of Dutch on French was seen in bilinguals' use of the French term *bouteille* 'bottle,' which became more similar to the use of Dutch *fles* 'bottle.' Conversely, the influence of French on Dutch was revealed in bilinguals' use of Dutch *bus* 'can,' which became more similar to the use of French *spray* 'spray bottle.' Evidence for shifting of semantic boundaries has also been seen in late consecutive bilinguals; for example, in

naming of household objects among Chinese speakers of second language (L2) English (e.g. Malt & Lebkuecher, 2017; Malt, Li, Pavlenko, Zhu, & Ameel, 2015; Zinszer, Malt, Ameel, & Li, 2014).

Similar results are found in studies of motion with evidence of interaction in the encoding of manner and path seen in shifts of frequency of use of certain verbs, leading to increased similarity between bilinguals' languages (Alferink, 2015). Interaction effects are present even at low levels of proficiency. Brown and Gullberg (2008) found that Japanese speakers of L2 English encoded manner of motion in speech more frequently in Japanese than monolingual Japanese speakers, and less frequently in English than monolingual English speakers, resulting in an in-between bilingual pattern. Even semantic categories that involve a semi-obligatory binary choice may show evidence of semantic interaction. Alferink and Gullberg (2014) found that Belgian Dutch-French bilinguals consistently dropped one of the two Dutch placement verbs (*zetten*/*leggen* 'set/lay'), making Dutch more similar to French with its one main placement verb (*mettre* 'put'). Similar results are found in early simultaneous bilingual speakers of Swiss German and Romansch (Berthele, 2015).

The effects of proficiency, language dominance, and age of acquisition (early or late) on semantic interaction are often discussed. For example, Athanasopoulos, Danjanovic, Burnand, and Bylund (2015) found that L1 speakers of English (an aspect language) learning L2 German (a non-aspect language) were more prone to base similarity judgments on endpoint saliency than English monolinguals as a function of L2 proficiency. In a classic study of color naming among Navajo-English bilinguals, Ervin (1961) showed that where the languages differed in color category boundaries, speakers' dominant language determined the category boundary for both languages. Gathercole and Moawad (2010) found evidence of interaction between the semantic systems of both early and late English-Arabic bilinguals' L1 (Arabic) and L2 (English). Crucially, the outcome of the interaction was affected by age of acquisition: in early bilinguals, the semantic structure of both languages was affected, while in late bilinguals only an effect of the L1 on the L2 was seen.

Although the literature thus provides much evidence for semantic interaction in bilinguals, the data (and therefore the theorizing) draws predominantly on settings where bilingualism involves standardized languages, literate, instructed speakers, and where societal and individual bilingualism is the result of instruction, migration, language shift or diglossia. We know much less about bilingual semantic interaction in non-Western settings where speech communities are small, mostly bilingual, and egalitarian (i.e. where neither language enjoys higher status, see Evans, 2017; François, 2012), where the languages spoken are non-standardized and in contact long term, and where speakers are non-literate and no formal acquisition is involved. This kind of bilingual setting is common in the world, yet underrepresented in research (see François, 2012; Gullberg, 2012). This is unfortunate since the specific characteristics of the settings in which bilingualism occurs may influence the nature of the outcomes of bilingualism. Muysken's (2013) framework of *bilingual optimization strategies* attempts to account for the differing documented outcomes of bilingualism, predicting that the outcome of language contact is influenced by social factors, perceived language distance, and processing constraints. Muysken proposes four 'outcome types': (1) high L1 prestige, low L2 proficiency, and limited access to L2 speakers are predicted to favor an 'L1-type outcome' ("maximize structural coherence of the first language"); (2) high L2 prestige, high L2 proficiency, and large numbers of L2 speakers are instead predicted to favor an 'L2-type outcome' ("maximize structural coherence of the second language"); (3) typological and lexical similarity of the languages and low normativity are predicted to favor an 'L1/L2-type outcome' ("match between L1 and L2 patterns where possible"); and finally, (4) political, typological, and lexical distance and a short contact period are predicted to favor a 'Universal Principle-type' outcome ("rely on universal principles of language processing"). The current study probes the framework's predictions for semantic interaction in the domain of topological relations in a non-Western setting, namely the village of Rual in northern Peninsular Malaysia.

The current study

We examine speakers of two Northern Aslian (Austroasiatic) varieties, Jedek and Jahai at Rual, near Jeli, Kelantan, Malaysia. Rual was established in the 1970s as a resettlement site for four Jedek-speaking and two Jahai-speaking mobile bands of hunter-gatherers then living along the mid-section of the Pergau valley (Gomes, 2007). Of the 169 adult residents of Rual today, 106 identify as primarily Jedek speakers, 53 as primarily Jahai speakers, and 10 as speakers of other Aslian varieties. Individuals who grew up in Rual tend to identify as primarily Jedek-speaking if they have at least one Jedek-speaking parent, while individuals with two Jahai parents or one Jahai parent and one parent who speaks another Aslian variety tend to identify as Jahai-speaking. Whereas it is still unclear how language identity interacts with actual language practices, Jedek appears to be the more widely spoken variety at Rual.

In accordance with the tradition of band exogamy, there has been a high degree of intermarriage among the Jedek- and Jahai-speaking bands at Rual. Around 50% of Rual residents are of mixed Jedek-Jahai parentage, and around half of Rual couples are mixed Jedek-Jahai. In such cases, both parties tend to retain their language identity as a speaker of either Jedek or Jahai, sometimes claiming proficiency in the other language, sometimes not. In some cases Rual residents have knowledge of both varieties from childhood, and in others acquisition occurs in adulthood. Some adult Jahai speakers moving into Rual claim to have gained proficiency in Jedek while others do not. The Rual speech community is small, as are the numbers of speakers of the two varieties. Jedek is spoken by around 280 speakers (including children), and is spoken only at Rual; Jahai has roughly 1000 speakers, most of whom live in the neighboring state of Perak. Jedek and Jahai are not standardized or written varieties, and no formal instruction exists in either variety.

The semantic test domain in the current study is that of topological relations, a sub-domain of spatial language in which something (a 'Figure') is located in relation to something else (a 'Ground'), typically expressed by prepositions in English as in 'The cup is *on* the table.' It has been argued that spatial cognition plays a central role in human thinking and reasoning and that spatial language is a basic part of human language (Gentner & Bowerman, 2009; Levinson, 2003). Importantly, there is substantial crosslinguistic variation in spatial language and specifically in the expression of topological relations (Levinson & Meira, 2003; papers in Levinson & Wilkins, 2006). This is therefore a rich domain in which to explore interaction in bilinguals. Only one study has investigated bilingual semantic interaction in this domain. Indefrey, Şahin, and Gullberg (2017) examined Turkish-Dutch bilinguals' topological relations descriptions and found evidence of semantic interaction between Turkish and Dutch in the descriptions of Dutch-dominant early simultaneous bilinguals, but not in the descriptions of Turkish-dominant later bilinguals who learned Dutch as adolescents or young adults. The semantic boundaries of translation-equivalent Turkish and Dutch topological relation markers (TRMs) were more congruent for the Dutch-dominant bilinguals than for Turkish and Dutch monolinguals (with unidirectional influence of Dutch on Turkish). Meanwhile, the Turkish-dominant bilinguals increased their use of a topologically neutral locative marker. The results were in line with Muysken's (2013) predictions: the Dutch-dominant early simultaneous bilinguals used an 'L2-type' strategy, predicted by the dominance of Dutch for this group, whereas the Turkish-dominant later bilinguals used a 'Universal Principle-type' strategy, predicted by their relatively shorter contact period.

Jedek and Jahai topological relation markers

Jedek and Jahai have a number of TRMs for describing spatial relations. Table 1 presents 14 Jedek TRMs (nine relational nouns, five prepositions; Yager & Burenhult, 2017), and 10 Jahai TRMs (seven relational nouns, three prepositions; Burenhult, 2005).

Table 1. Jedek and Jahai TRMs.

Gloss	Jedek	Jahai	Part of speech
'upper side'	*ʔatɛs*	*krɸiŋ*	relational noun
'underside'	*kyəm*	*kyɔm*	relational noun
'inside'	*daləm*	*klɛŋ*	relational noun
'outside'	*hɨɸ*	*ʔnaŋ*	relational noun
'front'	*sɛɲ*	*sɛɲ*	relational noun
'back'	*krəʔ*	*tkɨh*	relational noun
'side'	*ʔndaŋ, ʔirah*	*sir*	relational noun
'other side'	*ditɛɸ*		relational noun
location (LOC)	*lɛŋ, klɛŋ, la=*	*k=*	preposition
goal (GOAL)	*da=*	*ba=*	preposition
source (SOURCE)	*can=*	*can=*	preposition

While prepositions are used only in combination with noun phrases to form prepositional phrases, relational nouns are nominal forms that refer to spatial zones. They may be used alone or in combination with prepositions, and may also combine with other nominal elements in compound-like constructions to refer to a spatial zone in relation to another nominal referent, thus taking on prepositional properties. While Jedek and Jahai relational nouns refer to specific spatial zones and thus encode specific information about the relations between Figure and Ground, the prepositions are only general markers of the location, goal, or source of a referent.

The primary focus in the present study is on the use of the TRM *klɛŋ*, which has quite different functions in the two varieties. The Jahai relational noun *klɛŋ* 'inside' denotes containment (1), whereas the Jedek locative preposition *klɛŋ* (LOC) encodes only general information about the location of a Figure in the vicinity of a Ground (2).

Jahai

(1) *ʔarnap* *klɛŋ* *ragaʔ*
rabbit inside cage
'a rabbit in a cage'

Jedek

(2) *talipon* *klɛŋ* *dindiŋ*
telephone LOC wall
'a telephone on a wall'

Research question and predictions

The current study investigates semantic interaction in the Jedek and Jahai topological relations descriptions of Jedek-Jahai bilinguals. The key issue is whether there is evidence of semantic bilingual interaction effects, as found in other settings, in this untutored, non-Western, non-standardized, non-literate, highly bilingual setting; and whether effects manifest themselves in the same ways. The Rual speech community is characterized by low normativity and long-term contact, and Jedek and Jahai are lexically and typologically similar varieties, conditions predicted to lead to Muysken's 'L1/L2-type' strategy ("match between L1 and L2 patterns where possible"). We therefore expect bidirectional influence or increased congruence in the semantic boundaries of the Jedek and Jahai

Table 2. The participants.

	Jedek	Jahai
Monolinguals	n=15	n=3
Bilinguals	n=8	n=6

TRMs in the bilinguals as compared to monolinguals. Importantly, this influence is predicted to occur "where possible", that is, structural constraints are predicted to also play an important role in the outcomes. We predict that the TRM *klɛŋ* is particularly sensitive to bilingual semantic interaction since it is identical in form but not in semantic content in the two varieties.

Method

Participants

Two groups of functional bilingual Jedek and Jahai speakers participated in the study: eight bilinguals who identify as primarily Jedek-speaking, and six bilinguals identifying as primarily Jahai-speaking (labeled id-Jedek and id-Jahai respectively). Eighteen monolingual speakers of Jedek and Jahai also participated: 15 self-identifying monolingual Jedek speakers from Rual, and three monolingual Jahai speakers from Banun, a Jahai-majority area 60 km west of Rual (data by Burenhult in 2016). This group of Jahai monolinguals was chosen since it is impossible to find Jahai speakers at Rual who do not have knowledge of Jedek. Since Jedek is only spoken at Rual, the monolingual Jedek speakers are necessarily Rual residents. Table 2 presents the groups.

The labels monolingual and bilingual need some clarification. No Rual resident is a functional monolingual, since both Jedek and Jahai monolinguals are intensively exposed to the other variety. Moreover, Jedek and Jahai speakers are in contact with speakers of other Aslian varieties, and all participants also speak the majority language Malay. The label 'Jedek/Jahai monolingual' therefore does not refer to functional monolinguals but rather to individuals who self-identify as speakers of only one of the varieties. Similarly, the label 'bilingual' does not reflect proficiency, language dominance, or usage, but rather self-identification as a speaker of both Jedek and Jahai.

Materials and task

Data were collected using the Topological Relations Picture Series (Bowerman & Pederson, 1992), consisting of 71 line drawings depicting various topological relations between Figures and Grounds. Picture 1 of the series (Figure 1) can be described in English as 'the cup [Figure] is on the table [Ground].' We focus on the element expressing the spatial relation between Figure and Ground (the preposition *on* in the example), referred to as a topological relation marker (TRM; cf. Indefrey et al., 2017).

Procedure

The task was administered as a director-matcher task where both the director and the matcher were local participants. The director described each picture to the matcher by answering the question "Where is the (Figure)?" The matcher then selected the picture described among a set of twelve. The task was administered as a director-matcher task for three reasons. First, speakers generally speak more in a dialogic situation where they have a communicative reason for precision. Second,

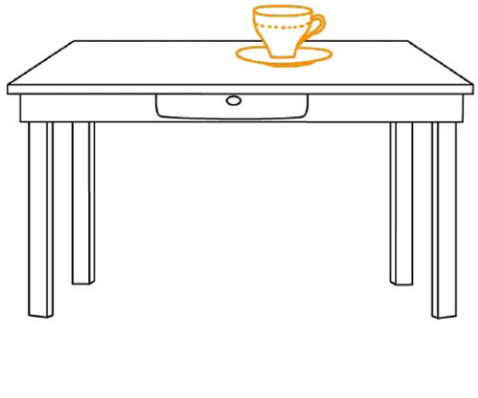


Figure 1. Picture 1 of the topological relations picture series (Bowerman & Pederson, 1992).

if their interlocutor is from their own speech community rather than an outsider with limited proficiency in the varieties, this minimizes the risk of so-called Foreigner Talk (e.g. Ferguson, 1975), thus contributing to data validity. Finally, this design allowed us to control participants' so-called language mode (Grosjean, 1998). To stack the odds against our hypothesis, we aimed to keep participants in 'monolingual mode' in either Jedek or Jahai. Participants therefore did the task in Jedek with a Jedek matcher (a Rual resident), and in Jahai with a Jahai matcher (a Jahai speaker from outside the Rual community).

The director sat in front of a laptop displaying each picture of the task one by one, and the matcher was given a folder with 12 pages, opened to the first page. The 71 pictures of the set were spread out over the 12 pages; each page contained five to six relevant items and six fillers. The bilinguals performed the task once in Jedek and once in Jahai, while the monolinguals performed the task only once, in their identity variety. The bilinguals' two sessions were conducted at least 24 hours apart, and variety order was balanced across participants as far as possible. The pictures were presented in three different orders assigned randomly. Where a participant performed the task more than once, the order was kept constant between their two sessions to minimize within-speaker order effects.

Data treatment and analyses

The full responses of each session were transcribed – the Rual data by the first author and the Banun data by Jahai researcher Niclas Burenhult. Invalid responses (i.e. those not containing the target Figure, Ground, and a TRM) were discarded, and if more than one response was given, the first valid response was selected for analysis. Responses to 12 pictures (15, 18, 22, 26, 28, 33, 56, 57, 61, 62, 66, 70) were removed from analysis due to a low number of valid responses (at least 50% of the participants gave a non-valid response). These were pictures of situations typically not expressed using TRM constructions in Jedek and Jahai, or of objects not recognizable to participants for culture-specific reasons.

The analyses focus on the TRM *klɛŋ* and its related TRMs across the varieties. We first created extension maps of the scenes labeled with the TRMs by more than 33% of speakers (following Bowerman, 1996; Levinson & Meira, 2003; cf. Indefrey et al., 2017). In the maps the scenes are

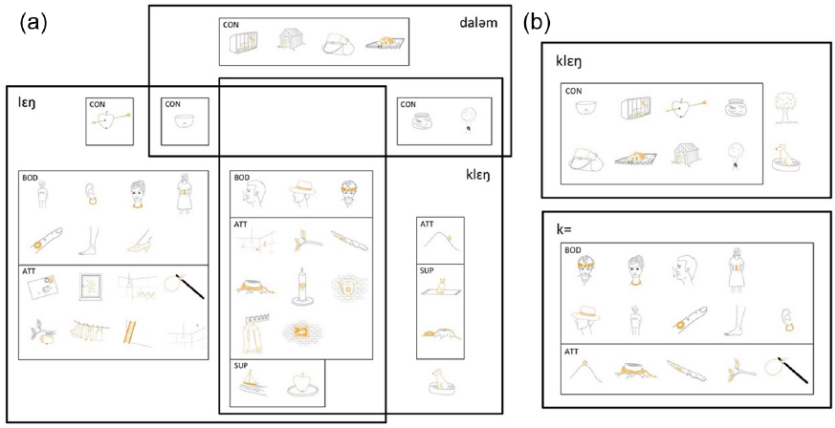


Figure 2. (a) Extension map of the use of *klɛŋ, lɛŋ*, and *daləm* by the Jedek monolinguals. (b) Extension map of the use of *klɛŋ* and *k=* by the Jahai monolinguals. CON=containment scenes; BOD=body-part-as-Ground scenes; ATT=attachment scenes; SUP=superjacent scenes. The bold lines indicate scenes labeled with a TRM by more than 33% of speakers.

coded for five semantic subdomains or situation types relevant in Jedek and Jahai. In *superjacency* scenes the Figure is positioned above the Ground, with or without contact. In *containment* scenes the Figure is fully or partially contained by the Ground, including full/partial containment, containment by encirclement (without contact), or by piercing. In *adjacent* scenes the Figure is beside the Ground, with or without contact. In *body-part-as-Ground* scenes the Ground is a part of the body. Finally, *attachment* scenes include relations of 'clingy attachment,' 'point-to-point attachment,' 'encirclement with contact,' and 'hanging against' situations (following Gentner & Bowerman, 2009).

Second, we quantified the degree of extension overlap in the monolingual and bilingual groups, respectively. Following Indefrey et al., 2017, for all pairs of Jedek-Jahai TRMs we calculated the proportion of congruent pictures, where congruence was defined as the number of pictures for which a pair of Jedek-Jahai TRMs were most frequently used in Jedek *and* Jahai divided by the number of pictures for which at least one of the two TRMs was most frequently used in Jedek *or* Jahai.

We first present the extension maps and congruence analysis for the monolingual data. Next, we present extension maps and congruence analyses for the two bilingual groups, and then compare the congruence in the two bilingual groups with that of the monolinguals.

Results

The monolinguals

Figure 2 shows extension maps of the use of *klɛŋ* in monolingual Jedek and Jahai. The Jedek map also includes the TRMs *lɛŋ* (LOC), with which Jedek *klɛŋ* overlaps semantically, and *daləm* 'inside,' the Jedek semantic equivalent of Jahai *klɛŋ*. The Jahai map includes the TRM *k=* (LOC), that is, the Jahai semantic equivalent of Jedek *lɛŋ*.

The maps reveal that Jedek *klɛŋ* is semantically very general: its extension includes nine attachment scenes, four superjacency scenes, three body-part-as-Ground scenes, and two containment

Table 3. Matrix of congruence values of the Jedek and Jahai monolingual TRMs. C=congruence value (0–1).

		Jahai									
		krpiŋ		***klɛŋ***		***k=***		*kyɔm*		other	all
		N	C	N	C	N	C	N	C	N	N
Jedek	*ʔatɛs*	7	0.44	1	0.05	2	0.07	0	0	0	13
	daləm	0	0	7	0.70	0	0	0	0	0	7
	lɛŋ	2	0.06	1	0.03	15	0.52	1	0.03	0	26
	klɛŋ	1	0.07	1	0.07	1	0.04	0	0	0	6
	kyəm	0	0	0	0	0	0	4	0.80	0	4
	other	0		0		0		0		0	3
	all	10		10		18		5		0	

The relevant TRMs are shown in bold face.

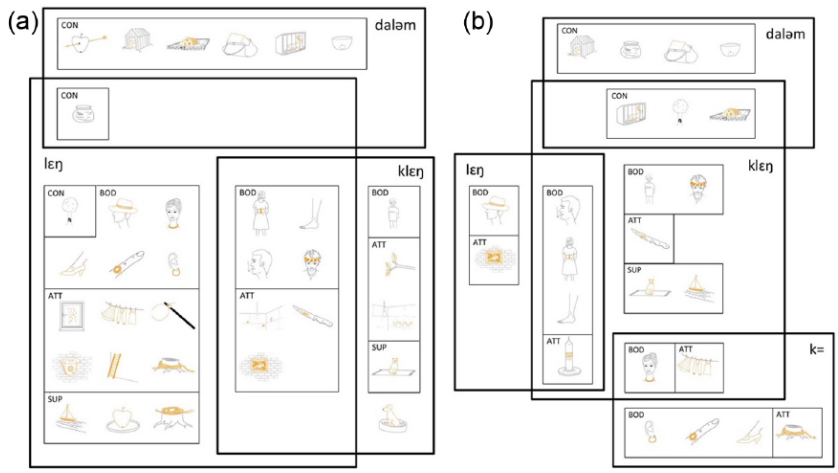


Figure 3. (a) Id-Jedek bilinguals' use of *klɛŋ, lɛŋ*, and *daləm* in Jedek.
(b) Id-Jedek bilinguals' use of *klɛŋ, lɛŋ, daləm*, and *k=* in Jahai.
CON=containment scenes; BOD=body-part-as-Ground scenes; ATT=attachment scenes; SUP=superjacent scenes. The bold lines indicate scenes labeled with a TRM by more than 33% of speakers.

scenes. The extension of Jedek *klɛŋ* also overlaps with that of *lɛŋ*, and that of *daləm*, which includes only containment scenes. Meanwhile, Jahai *klɛŋ* includes in its extension all eight containment scenes and two uncategorized scenes that may be construed as containment.

Table 3 presents congruence values for the TRMs that were most frequently used in descriptions of at least three scenes in monolingual Jedek and Jahai. The congruence values in Table 3 reflect the patterns of the extension maps. The extension of *klɛŋ* is very different in the two varieties, with a very low congruence value (*klɛŋ – klɛŋ*, 0.07). Jahai *klɛŋ* is most congruent with Jedek *daləm* (0.70), while Jedek *klɛŋ* has low levels of congruence (range 0.04–0.07) with Jahai TRMs *krpiŋ, klɛŋ* and *k=*.

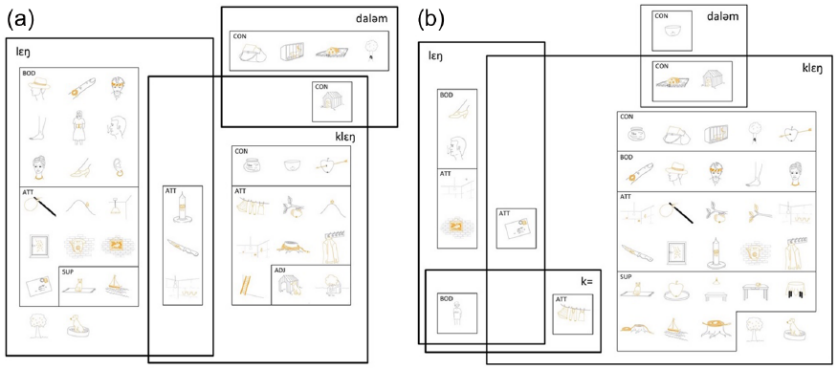


Figure 4. (a) Id-Jahai bilinguals' use of *klɛŋ, lɛŋ*, and *daləm* in Jedek.
(b) Id-Jahai bilinguals' use of *klɛŋ, lɛŋ, daləm*, and *k=* in Jahai.
CON=containment scenes; BOD=body-part-as-Ground scenes; ATT=attachment scenes; SUP=superjacent scenes; ADJ=adjacent scenes. The bold lines indicate scenes labeled with a TRM by more than 33% of speakers.

In summary, monolingual Jahai *klɛŋ* is used only for containment, whereas monolingual Jedek *klɛŋ* is primarily used for other scene types, expressing location.

The bilinguals

The id-Jedek bilinguals in Jedek and Jahai. Figure 3 presents extension maps of the id-Jedek bilinguals' use of *klɛŋ* and related TRMs in Jedek and Jahai. The maps show that in both Jedek and Jahai this group uses the Jedek TRM *daləm* for containment, and *klɛŋ/lɛŋ* for a variety of other scenes. Thus, in both varieties, the id-Jedek bilinguals' use of *klɛŋ* resembles monolingual Jedek in that it includes a variety of scene types. Importantly, however, their extension of *klɛŋ* in Jedek contains no containment scenes, while their extension of *klɛŋ* in Jahai does, in resemblance to the monolingual Jahai pattern. Overall, we see a unidirectional influence of Jedek on the extension of Jahai *klɛŋ* in this group, in its broader, more general extension as compared to the extension of Jahai monolingual *klɛŋ*.

The id-Jahai bilinguals in Jedek and Jahai. Figure 4 presents extension maps of the id-Jahai bilinguals' use of *klɛŋ* and related TRMs in Jedek and Jahai. The maps show that in Jedek, this group uses *daləm* for containment, and *klɛŋ/lɛŋ* for a variety of other scenes, including containment. In Jahai, they use *daləm* for a few containment scenes, but *klɛŋ* for a broad set of scenes including containment. Thus, in both varieties, the id-Jahai bilinguals' use of *klɛŋ* resembles monolingual Jedek in that it includes a variety of scene types. But their broad extension of *klɛŋ* in Jahai is noteworthy in that it appears to be used as a kind of cover-all term. Overall, we see a bidirectional influence of the varieties, in the broadening of the extension of Jahai *klɛŋ*, and in the expanded use of Jedek *klɛŋ* for containment.

Congruence across id-Jedek and id-Jahai bilingual usage. Table 4 presents congruence values for the TRMs most frequently used to describe at least three scenes by the id-Jedek bilinguals in Jedek and Jahai. The analysis shows that the extensions of *klɛŋ* in the two varieties are more congruent for the id-Jedek bilinguals (0.30) than for the monolingual groups (0.07, cf. Table 3). Jahai *klɛŋ* also has some congruence with Jedek *lɛŋ* (0.10) for this group.

Table 4. Matrix of congruence values of the Jedek and Jahai TRMs as used by the id-Jedek bilinguals. C=congruence value (0–1).

		Jahai													
		ʔatɛs		*krɒiŋ*		***dalǝm***		***klɛŋ***		***k=***		*kyǝm*		other	all
		N	C	N	C	N	C	N	C	N	C	N	C	N	N
Jedek	*ʔatɛs*	3	0.20	5	0.36	0	0	0	0	0	0	0	0	0	13
	dalǝm	0	0	0	0	3	0.50	0	0	0	0	0	0	0	5
	lɛŋ	0	0	1	0.05	0	0	2	0.10	4	0.20	0	0	0	7
	klɛŋ	0	0	0	0	0	0	3	0.30	0	0	0	0	0	17
	kyǝm	0	0	0	0	0	0	0	0	0	0	4	0.80	0	4
	other	0		0		0		0		0		0		2	1
	all	5		6		4		6		5		5			

The TRMs in focus are in bold face.

Table 5. Matrix of congruence values of the Jedek and Jahai TRMs as used by the id-Jahai bilinguals. C=congruence value (0–1).

		Jahai									
		krɒiŋ		***klɛŋ***		***lɛŋ***		*kyǝm*		other	all
		N	C	N	C	N	C	N	C	N	N
Jedek	*ʔatɛs*	3	0.33	5	0.13	0	0	0	0	0	9
	dalǝm	0	0	3	0.08	0	0	0	0	1	4
	lɛŋ	0	0	13	0.32	3	0.15	0	0	0	19
	klɛŋ	0	0	6	0.15	0	0	0	0	1	11
	kyǝm	0	0	0	0	0	0	4	1	0	4
	other	0		0		0		0		1	2
	all	3		35		4		4			

The TRMs in focus are in bold face.

Table 5 presents congruence values for the TRMs most frequently used to describe at least three scenes by the id-Jahai bilinguals in Jedek and Jahai. The analysis shows that the extensions of *klɛŋ* in the two varieties are more congruent for the id-Jahai bilinguals (0.15) than for the two monolingual groups (0.07, cf. Table 3). Jahai *klɛŋ* also has some congruence with Jedek *lɛŋ* (0.32) for this group.

To illustrate the comparison of the monolingual and bilingual analyses (Tables 3–5), Figure 5 presents the congruence values for *klɛŋ* and related TRMs for the two bilingual groups and for the monolinguals of each variety. Pairs with relatively high congruence values for the monolingual groups, *dalǝm – klɛŋ* and *lɛŋ – k=*, were much less congruent for the bilinguals (*dalǝm – klɛŋ* 0 for the id-Jedek bilinguals, 0.08 for the id-Jahai bilinguals; *lɛŋ – k=* 0.22 for the id-Jedek bilinguals, 0 for the id-Jahai bilinguals). Thus, in bilingual usage there was little overlap in the extensions of these TRMs.

Conversely, there was an increase in congruence of form-based TRM pairs: the Jedek-Jahai TRM pairs *dalǝm – dalǝm, klɛŋ – klɛŋ, lɛŋ – lɛŋ* and *lɛŋ – klɛŋ* were more congruent for the bilingual groups than for monolinguals.

Rather than an overall increase in the congruence of translation-equivalent TRMs for the bilinguals, we find increased form-based congruence. Similar forms are used for relatively congruent extensions in both varieties, suggesting semantic interaction in the bilinguals.

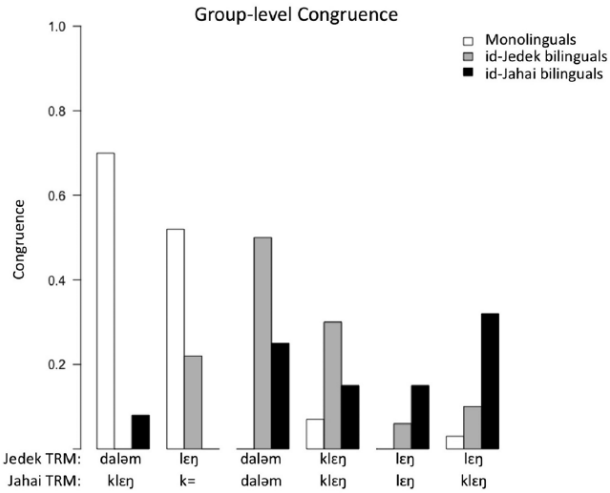


Figure 5. Group-level congruence of TRM pairs for the three groups.

To explore whether the group-level analyses corresponded to individual bilingual speakers' use of the TRMs, we calculated the congruence of each TRM pair in Jedek and Jahai for each bilingual speaker. Individual congruence for a TRM pair was the number of scenes for which the speaker used both TRMs of the pair divided by the number of scenes for which the speaker used one of the TRMs of the pair. Figure 6 shows the mean individual congruence values for the TRM pairs in Jedek and Jahai for the two bilingual groups.

The same four Jedek-Jahai TRM pairs were relatively congruent (all > 0.1) for both bilingual groups in the individual analysis as in the group-level analysis: *daləm – daləm, klɛŋ – klɛŋ, lɛŋ – lɛŋ*, and *lɛŋ – klɛŋ*; the pairs *daləm – klɛŋ* and *lɛŋ – k=* had low congruence values also in the individual analysis. An independent samples t-test revealed no difference between the bilingual groups in mean individual congruence ($t(150) = 0.843$, $p = .401$).

Discussion

This study investigated semantic interaction in the topological relation markers (TRMs) of two groups of Jedek-Jahai bilinguals in the small-scale, non-standardized, egalitarian, long-term multilingual setting of Rual in northern Peninsular Malaysia. Based on predictions from Muysken (2013), the typological and lexical similarity of Jedek and Jahai and the low normativity of the Rual speech community were predicted to lead to bidirectional influence in bilinguals ("match between L1 and L2 patterns where possible"). We expected increased congruence of the semantic boundaries of Jedek and Jahai TRMs in bilinguals compared to those of monolinguals. We focused on the use of *klɛŋ*, a TRM with identical form but different semantics in monolingual Jedek and Jahai. The Jahai relational noun *klɛŋ* denotes containment, whereas the Jedek preposition *klɛŋ* is a general locative marker indicating only non-specific information about the relation between Figure and Ground. Contrary to predictions, there was no general increase in congruence of Jedek and Jahai bilingual extensions. Instead, there was an increase in congruence only where there was also

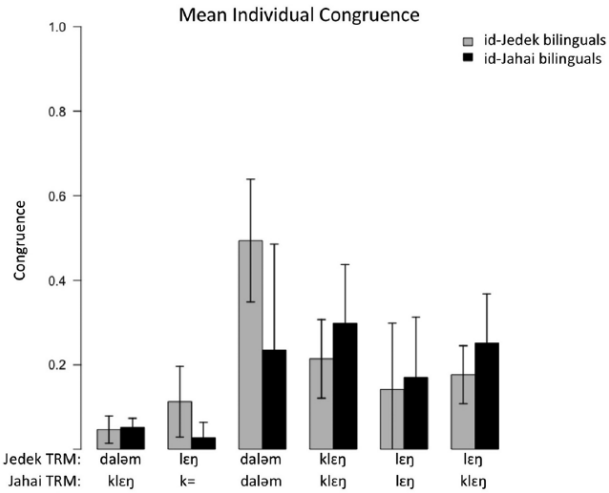


Figure 6. Mean individual congruence of TRM pairs for the two bilingual groups.

form overlap in the two varieties. While the semantics of *klɛŋ* differed greatly for the monolinguals, the bilinguals used the TRM more similarly in the two varieties.

Following Muysken (2013), we predicted a bidirectional influence of Jedek and Jahai in the bilinguals' use of *klɛŋ*. But this prediction was not borne out for both bilingual groups. Both groups used *klɛŋ* in a way congruent with the Jedek pattern in both Jedek and Jahai descriptions, using it as a general location marker, primarily for non-containment scenes. But the influence of Jahai, wherein the bilinguals expanded their *klɛŋ* extension to include more containment scenes, was seen in both varieties only for the id-Jahai bilinguals, while for the id-Jedek bilinguals this pattern was only seen in Jahai. In sum, the results reveal a bidirectional influence in the id-Jahai bilinguals' use of *klɛŋ*, while in the id-Jedek bilinguals, we see only a unidirectional influence of Jedek on Jahai. Our prediction is thus only partially supported, since we find an asymmetric semantic interaction in the two bilingual groups. This result is similar to the findings of Indefrey et al. (2017), who also found asymmetric effects across Turkish-Dutch bilinguals.

How can we account for this asymmetry? In Muysken's (2013) framework, differential access to speakers of the languages of bilinguals is predicted to affect bilingual outcomes. While the id-Jahai bilinguals have extensive access to Jedek speakers, id-Jedek bilinguals' exposure to Jahai speakers is more limited. This would predict Muysken's 'L1-type outcome,' or "maximize structural coherence of the first language." This is indeed what the data for the id-Jedek bilinguals show: they use a bilingual strategy in which both varieties are structured around Jedek. Conversely, the id-Jahai bilinguals, who have ready access to Jedek speakers, use a bilingual strategy wherein Jedek and Jahai bidirectionally affect each other, leading to the predicted 'L1/L2-type outcome,' or "match between L1 and L2 patterns where possible" (Muysken, 2013).

It is important to consider what "where possible" means in practice, that is, the role of structural aspects of the languages involved. The identical form of *klɛŋ* in Jedek and Jahai and its semantically general nature in Jedek means that it can be used by bilinguals to encompass the extensions of both monolingual varieties. Its use as a general spatial marker in Jedek may simply be expanded

to incorporate more (containment) scenes, thus increasing its congruence with Jahai *klɛŋ*. It has been proposed that the effects of interaction in bilinguals may be greatest where there are options in the languages that allow one language to easily move closer to the other (Alferink, 2015). Other findings suggest that bilinguals tend to adopt the use of broader semantic categories over more specific ones in lexical areas where their languages differ in semantic granularity, at least for some semantic categories (cf. Gathercole, Stadthagen-González, Pérez-Tattam, & Yavas, 2017). *Klɛŋ* is a fine example of both these tendencies. These findings point both to the effects of the nature of semantic systems in contact and to the effects of the features of the setting in which contact occurs, suggesting a complexity and connectedness in bilingual systems that must be reflected in our theoretical models (cf. Alferink, 2015).

Conclusion

Semantic interaction in bilingual systems was found in a non-Western setting in Malaysia, in a small egalitarian bilingual speech community speaking non-standardized varieties and characterized by long-term contact. The results from this study provide evidence for asymmetric interaction in the expression of topological relations in Jedek and Jahai, with bidirectional influences evident in one case, and a unidirectional influence found in the other. The findings support Muysken's (2013) framework for bilingual optimization strategies, proposing effects of variations in social factors, perceived language distance, and processing constraints. However, they also indicate that outcomes may not always cluster as in Muysken's framework. Even in settings with low normativity and high typological and lexical similarity between varieties (predicted to lead to bidirectional influence), limited access to 'L2 speakers' can result in an 'L1-type' outcome. This finding suggests that we need to pay even closer attention to the details of bilingual conditions, if we are to understand how languages interact in bilinguals. We must broaden the empirical basis to include settings that have not yet been explored in the literature, such as small and functionally bilingual speech communities, where the language varieties spoken are non-standardized and non-literate. We must also begin to acknowledge multilingualism beyond the use of two languages, given that multilingualism may be more widespread than bilingualism. And finally, we must broaden analyses to examine various parts of the language system if we are to truly understand the complex ways in which languages interact in bilinguals.

Acknowledgements

We wish to extend thanks to the community at Sungai Rual without whom the research would not have been possible. We also wish to acknowledge the Economic Planning Unit, Putrajaya; the Department of Orang Asli Development, Kuala Lumpur; and Prof. Dr. Kamila Ghazali, University of Malaya. We are deeply grateful to Virginia Gathercole and to anonymous reviewers for their insightful comments on earlier drafts of this paper.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Joint Faculties of Humanities and Theology at Lund University, The Lars Hierta Memorial Foundation, The Birgit Rausing Language Programme, Fil. Dr. Uno Otterstedt Foundation, and The Knut and Alice Wallenberg Foundation.

Supplemental material

Supplemental material for this article is available online.

References

Alferink, I. (2015). *Dimensions of Convergence in Bilingual Speech and Gesture* (Unpublished PhD dissertation). Radboud Universiteit Nijmegen, Nijmegen. Utrecht: LOT (nr 391).

Alferink, I., & Gullberg, M. (2014). French–Dutch bilinguals do not maintain obligatory semantic distinctions: Evidence from placement verbs. *Bilingualism: Language and Cognition*, *17*(1), 22–37.

Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, *60*(2), 270–290.

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, *53*(1), 60–80.

Athanasopoulos, P., Damjanovic, L., Burnand, J., & Bylund, E. (2015). Learning to think in a second Language: Effects of proficiency and length of exposure in English learners of German [Supplemental material]. *Modern Language Journal*, *99*, 138–153.

Bassetti, B., & Cook, V. (2011). Relating language and cognition: The second language user. In V. Cook & B. Bassetti (Eds.), *Language and bilingual cognition* (pp. 143–190). Hove, UK: Psychology Press.

Berthele, R. (2015). Convergence in the domains of static spatial relations and events of putting and taking. Evidence from bilingual speakers of Romansh and German. *International Journal of Bilingual Education and Bilingualism*, *18*(5), 624–642.

Bowerman, M. (1996). Learning how to structure space for language: A cross-linguistic perspective. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.

Bowerman, M., & Pederson, E. (1992). Topological relations picture series. In S. C. Levinson (Ed.), *Space stimuli kit 1.2: November 1992* (p. 51). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from http://fieldmanuals.mpi.nl/volumes/1992/bowped/.

Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, *30*(2), 225–251.

Burenhult, N. (2005). *A grammar of Jahai* (Pacific Linguistics 566). Canberra: Australian National University.

Choi, S. (2006). Influence of language-specific input on spatial cognition: Categories of containment. *First Language*, *26*(2), 207–232.

Ervin, S. M. (1961). Semantic shift in bilingualism. *The American Journal of Psychology*, *74*(2), 233–241.

Evans, N. (2017). Did language evolve in multilingual settings? *Biology and Philosophy*, *32*, 905–933.

Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics*, *17*(1), 1–14.

François, A. (2012). The dynamics of linguistic diversity: Egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language*, *214*, 85–110.

Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late bilinguals: All words are not created equally. *Bilingualism: Language and Cognition*, *13*(4), 385–408.

Gathercole, V. C. M., Stadthagen-González, H., Pérez-Tattam, R., & Yavaş, F. (2017). Semantic and conceptual factors in Spanish-English bilinguals' processing of lexical categories in their two languages. *Second Language Research*, *32*(4), 537–562.

Gentner, D., & Bowerman, M. (2009). Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, & S. Ozcaliskan (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 465–480). Psychology Press.

Gomes, A. G. (2007). *Modernity and Malaysia: Settling the Menraq forest nomads*. London: Routledge.

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition*, *1*(3), 175–176.

Gullberg, M. (2012). Bilingual multimodality in language documentation data. *Language Documentation & Conservation, Special Publication*, (3), 46–53.

Indefrey, P., Şahin, H., & Gullberg, M. (2017). The expression of spatial relationships in Turkish/Dutch bilinguals. *Bilingualism: Language and Cognition*, *20*(3), 473–493. doi:10.1017/S1366728915000875

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity* (Vol. 5). Cambridge: Cambridge University Press.

Levinson, S. C., & Meira, S. (2003). 'Natural concepts' in the spatial topological domain—Adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language*, *79*(3), 485–516.

Levinson, S. C., & Wilkins, D. P. (Eds.). (2006). *Grammars of space: Explorations in cognitive diversity* (Vol. 6). Cambridge: Cambridge University Press.

Malt, B. C., & Lebkuecher, A. L. (2017). Representation and process in bilingual lexical interaction. *Bilingualism: Language and Cognition*, *20*(5), 867–885. doi:10.1017/S1366728916000584.

Malt, B. C., Li, P., Pavlenko, A., Zhu, H., & Ameel, E. (2015). Bidirectional lexical interaction in late immersed Mandarin-English bilinguals. *Journal of Memory and Language*, *82*, 86–104.

Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(6), 583–597.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition*, *16*, 709–730.

Pavlenko, A. (2011). Thinking and speaking in two languages: Overview of the field. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 237–257). Bristol: Multilingual Matters.

Treffers-Daller, J., & Tidball, F. (2015). Can L2 learners learn new ways to conceptualise events? Evidence from motion event construal among English-speaking learners of French. In P. Guijarro-Fuentes, K. Schmitz, & N. Müller (Eds.) *The acquisition of French in multilingual contexts* (pp. 145–184). Bristol: Multilingual Matters.

White, A., Malt, B. C., & Storms, G. (2017). Convergence in the bilingual lexicon: A pre-registered replication of previous studies. *Frontiers in Psychology*, *7*, 2081.

Yager, J., & Burenhult, N. (2017). Jedek: A newly discovered Aslian variety of Malaysia. *Linguistic Typology 21*(3), 493–545.

Zinszer, B. D., Malt, B., Ameel, E., & Li, P. (2014). Nativelikeness in second language lexical categorization reflects individual language history and linguistic community norms. *Frontiers in Psychology: Language Sciences*, *5*, 1–16.

Author biographies

Joanne Yager is a PhD Candidate in linguistics at Lund University. Her research project focuses on egalitarian small-scale bilingualism among Jedek and Jahai speakers in Rual in northern Peninsular Malaysia.

Marianne Gullberg is full Professor of Linguistics at Lund University, and Chair/Director of Lund University Humanities Lab, an independent dept. for research technology and method. Her research focuses on acquisition and real time language use in adult second language users or bi-/multilingual speakers with a particular focus on semantics and discourse. She also studies the bimodal nature of language, targeting gesture production and comprehension in acquisition and multilingualism.

Study IV

Study IV: Highly similar semantic systems become more similar in egalitarian bilingual foragers

Abstract

A number of studies show that bilinguals do not perform like two monolinguals in one body. Rather, the languages of bilinguals tend to interact with one another and this interaction leads to properties in bilingual systems that are not seen in monolinguals. One particularly interesting line of research explores the interaction between the semantic systems of bilinguals. Findings of bilingual semantic interaction come almost exclusively from Western settings, and involve large-scale languages with key differences in semantic granularity. The current study asks how bilingual semantic interaction may play out in small-scale egalitarian multilingual communities and in highly similar semantic systems. The study investigates bilingual semantic interaction in the domain of placement events in Jedek-Jahai (Aslian, Austroasiatic) bilinguals in a small-scale multilingual setting in northern Peninsular Malaysia. Extension maps and congruence analyses compare the semantic extensions of Jedek and Jahai verbs in bilinguals to those of Jedek and Jahai monolinguals. The study finds evidence of symmetric semantic interaction in Jedek-Jahai bilinguals, suggesting that even highly similar semantic systems may become more similar through contact. The results point to a complexity in the interplay of factors influencing bilingual outcomes and highlight the need for research in small-scale egalitarian multilingual settings in order to further our understanding of the dynamics of bilingualism in a wider range of bilingual contexts.

Keywords: semantic interaction, bilingual optimization strategies, egalitarian bilingualism, small-scale multilingualism, crosslinguistic influence, placement events

Introduction

A large number of studies have shown that the languages of bilinguals are subject to interaction with one another. In the domain of semantics, interaction has been seen for example in shifts in the boundaries of semantic categories in one or both of the languages of bilinguals (e.g., Alferink & Gullberg, 2014; Ameel, Storms, Malt & Sloman, 2005; Ervin, 1961; Gathercole & Moawad, 2010). Bilingual semantic interaction studies have dealt almost exclusively with Western settings and with speakers of large-scale, standardized languages, and so we know little about what semantic interaction looks like in non-standardized language varieties and in small-scale, egalitarian multilingual speech communities. Similarly, little is known about bilingual semantic interaction in typologically and lexically very similar language varieties, and most studies focus on language pairs with key differences in semantic granularity. The current study asks how these variables might impact bilingual outcomes. The study investigates bilingual semantic interaction in the domain of placement events, among bilinguals of the typologically and lexically similar Northern Aslian (Austroasiatic) language varieties Jedek and Jahai in a small-scale egalitarian multilingual setting in northern Peninsular Malaysia.

Background

Given the great amount of crosslinguistic variation that has been found in how speakers of different languages map words onto referents or events (see Malt & Majid, 2013 for an overview), how do bilinguals and second language learners navigate these differences in form-meaning mappings when using more than one language? There is a great deal of evidence for semantic interaction in bilinguals, that is, that the languages of bilinguals are more similar as compared to the languages when used by monolinguals. The interaction can be symmetric or asymmetric, that is, it may involve changes in one or both of the languages of bilinguals (see e.g. Jarvis & Pavlenko, 2008 for an overview), and different linguistic domains may reveal differing directionality of influence within the same language pair (Alferink, 2015). The effects of bilingualism on the languages of bilinguals are usually discussed using the terms convergence, crosslinguistic influence, transfer or interaction. Following Gathercole and Moawad (2010), the term 'interaction' is used in the current study as a neutral term referring to the influence of the languages of bilinguals on one another, irrespective of the directionality of influence or specific nature of interaction patterns.

The current study investigates bilingual semantic interaction in the domain of caused motion, specifically in placement events. This type of events involve basic human sensory-motor patterns and are a universal part of the human experience (Levinson,

2012), yet a great deal of diversity has been found in the ways in which placement events are expressed in different languages (Narasimhan, Kopecka, Bowerman, Gullberg & Majid, 2012). Placement verb inventories vary in size and in terms of the semantic specificity of verbs. For example, some languages have semantically general placement verbs that cover a range of different kinds of placement events, such as the English placement verb *put*, while others categorize placement events according to more fine-grained distinctions, such as the caused posture verbs of Polish (Kopecka, 2012), Yélî Dnye (Levinson & Brown, 2012) and the Germanic languages (with the exception of English; see Alferink & Gullberg, 2014). Such crosslinguistic differences in the form-meaning mappings of placement verbs make the domain a fruitful one both in the context of semantic typology and for studies of semantic interaction in bilinguals.

A number of studies have found evidence for bilingual semantic interaction in the placement domain. The findings of these studies have shown that where the semantic systems of the languages of bilinguals differ in semantic granularity, bilinguals tend to simplify the more fine-grained system, affording increased overlap with the semantic boundaries of the less fine-grained system. For example, functional bilinguals of Dutch (with two caused posture verbs, *zetten/leggen*, 'set/lay') and French (with the more general placement verb *mettre*, 'put') were found to drop one of the semi-obligatory categories in Dutch, extending the use of *leggen* 'lay' to include scenes which were described by Dutch monolinguals with *zetten* 'set' (Alferink & Gullberg, 2014). Similarly, Romansch-dominant functional bilinguals of Swiss German (with three caused posture verbs, *stellen/legen/setzen*, 'stand/lay/set') and Romansch (with the semantically general placement verb *metter*, 'put') were found to overextend the use of one of the German caused posture verbs, *legen* 'lay', and used a less-common, semantically general German verb, *tun* 'do' (Berthele, 2015). In both of these studies, bilinguals used a simplified version of the more complex system, allowing for increased similarity with the other, less complex system.

Similar interaction effects have been seen in the second language (L2) of second language learners. Where the placement verbs of the L2 are more semantically fine-grained than those of the first language (L1), learners tend to use a simplified version of the L2 system, increasing its similarity to the less fine-grained L1 system (Gullberg, 2009a; Viberg, 1998). Meanwhile, learners whose L1 is more semantically fine-grained than their L2 appear to have no trouble using the simpler L2 system in language production (although they may use the more fine-grained distinctions of the L1 in their speech-associated gestures; Gullberg, 2009b, 2011). An example of L2 learning in both directions comes from a study involving Spanish learners of L2 Danish and Danish learners of L2 Spanish (Cadierno, Ibarretxe-Antuñano & Hijazo-Gascón, 2016). The placement verbs used by Danish native speakers mark orientation of the object placed, with three caused posture verbs, *sætte/stille/lægge*, 'set/stand/lay'. Meanwhile, the verbs used by Spanish native speakers do not mark orientation but instead mark a

containment/support distinction, with the verbs *dejar* 'leave [in a place]' and *meter* 'put in'. The Spanish learners of L2 Danish were found to use a simplified system with regard to the orientation of objects placed, overgeneralizing the Danish caused posture verb *lægge*, 'lay', while the Danish learners of L2 Spanish did not make a distinction between support and containment. These learners tended to use the general Spanish verb *poner*, 'put' (a verb not used by Spanish native speakers in their descriptions of these scenes) for both types of scene. Similar results were found for Spanish and Danish L2 learners by Ibarretxe-Antuñano, Cadierno and Hijazo-Gascón (2016). A recent study suggests that effects of differences in semantic granularity can not only be seen in language production but also in memory effects (Koster & Cadierno, 2019). Native speakers of German, a caused posture verb language, had more accurate recognition memory of object orientation than speakers of Spanish, a non-posture verb language. But this effect of fine-grained semantic distinctions can be acquired through focused instruction: Spanish learners of L2 German who received instruction on German caused posture verbs had more accurate recognition memory for object orientation than German native speakers.

The findings discussed above come almost exclusively from Western, tutored settings involving relatively large-scale languages. In contrast, the current study examines bilingual semantic interaction in a small-scale egalitarian multilingual setting, in two non-standardized, non-literate, typologically and lexically similar language varieties. This type of bilingual scenario represents a large gap in our knowledge about bilingual semantic interaction. In addition, while research to date has focused on semantic interaction in language pairs between which there are key differences in semantic granularity, the current study deals with languages with highly similar placement verb inventories. Burenhult's (2012) description of the Jahai placement verb lexicon suggests that Jahai has a mix of semantically very specific verbs (such as *lka?* 'to insert one's hand' and *caduk* 'to skewer an oblong object in one's hair') and relatively general verbs (such as *boh* 'put'). The data collected in connection with the current study suggest that the Jedek placement verb inventory is similar in verb forms and semantics. Thus, the current study represents an opportunity to study bilingual semantic interaction in very similar semantic systems.

We have seen that the kind of small-scale, egalitarian setting investigated in the current study represents a substantial gap in our understanding of bilingual semantic interaction. But to what extent should we expect to see different bilingual outcomes in different bilingual settings? It has been suggested that the outcomes of language contact and bilingualism may vary greatly depending on the features of the social setting in which the interaction occurs (e.g. Muysken, 2013; Thomason & Kaufman, 1988). If this is the case, it is not enough to consider the features of language systems and of bilingual speakers – we must also take into account the social features of bilingual settings when making predictions about bilingual outcomes. Muysken's framework of

bilingual optimization strategies proposes that a combination of social factors, processing constraints and perceived language distance leads to four types of bilingual strategies. (1) High L1 prestige, low L2 proficiency and limited access to L2 speakers are predicted to favor an 'L1-type outcome' ("maximize structural coherence of the first language"); (2) high L2 prestige, high L2 proficiency and large numbers of L2 speakers are predicted to favor an 'L2-type outcome' ("maximize structural coherence of the second language"); (3) typological and lexical similarity of the languages as well as low normativity are predicted to favor an 'L1/L2-type outcome' ("match between L1 and L2 patterns where possible"); and finally, (4) political, typological and lexical distance as well as a short contact period are predicted to favor a 'Universal Principle-type' outcome ("rely on universal principles of language processing").

The predictions of Muysken's model have been tested in two studies investigating bilingual semantic interaction (Indefrey, Şahin & Gullberg, 2017; Yager & Gullberg, published online 2019). Both studies focus on lexical semantics in the domain of topological relations descriptions, that is, static spatial relationships often expressed in adpositions. Indefrey et al. (2017) found evidence in support of the predictions of the model in the topological relations descriptions of two groups of Turkish-Dutch bilinguals (Dutch-dominant and Turkish-dominant bilinguals) in the Netherlands. The nature of the outcomes of semantic interaction was seen to differ between the two groups of bilinguals. The Dutch-dominant group showed a unidirectional influence of Dutch on the semantic boundaries of translation-equivalent Turkish topological relation markers (maximizing the structural coherence of Dutch, Muysken's 'L2-type' outcome). Meanwhile, the Turkish-dominant bilinguals did not show evidence of widespread shifts in semantic boundaries, instead increasing their use of a topologically neutral Turkish locative marker (akin to Muysken's 'Universal Principle-type' outcome). In line with predictions, the higher prestige of Dutch in the Netherlands, combined with high proficiency in Dutch led to a unidirectional influence of Dutch on Turkish in the Dutch-dominant bilinguals, while the political, typological and lexical distance between Turkish and Dutch and a shorter contact period led to a more neutral outcome in the Turkish-dominant bilinguals.

Yager and Gullberg (published online 2019) also tested the predictions of the model for the topological relations domain, for two groups of Jedek-Jahai bilinguals (Jedek-identifying and Jahai-identifying bilinguals) in a small-scale egalitarian setting in northern Peninsular Malaysia. While Jahai-identifying bilinguals showed evidence of symmetric semantic interaction (Muysken's 'L1/L2-type outcome'), Jedek-identifying bilinguals showed evidence of asymmetric interaction, with a unidirectional influence of Jedek on Jahai (Muysken's 'L1-type' outcome, in which both varieties are patterned around one of the languages). Thus, the typological and lexical similarity of Jedek and Jahai and the low normativity present in the community favored symmetric interaction in Jahai-identifying bilinguals, while the relatively low level of access to Jahai speakers

in the speech community favored asymmetric interaction in the direction of Jedek in Jedek-identifying bilinguals.

These two studies suggest that Muysken's framework is useful in making predictions about the effects of the social features of bilingual settings on bilingual outcomes, for bilingual semantic interaction in the topological domain. The current study extends the investigation to the semantic domain of placement events. There is evidence to suggest that motion verbs (such as those used in placement event descriptions) and more static language (such as that used in topological relations descriptions) may be processed in fundamentally different ways (Wallentin, Ellegaard, Østergaard, Østergaard & Roepstorff, 2005; Wallentin et al., 2011). Thus motion verbs represent an important point of comparison with the findings from the topological domain. Placement events involve the kinds of spatial distinctions relevant in the topological domain and thus build on this earlier work, yet incorporate the element of motion, taking the investigation into the verb lexicon. The placement event domain has been found in previous studies to be a fruitful domain for studies of semantic interaction in bilinguals. For these reasons, the placement event domain was chosen as the test domain in the current study.

The Rual setting

The setting of the current study is the small-scale, egalitarian multilingual setting of Rual in northern Peninsular Malaysia. Rual is a resettlement site established in the 1970s, made up of six bands of mobile hunter-gatherers: four Jedek-speaking bands and one Jahai-speaking band who were living along the mid-section of the Pergau valley at that time, and one Jahai-speaking band who relocated to Rual from the neighboring state of Perak in the 1980s (Gomes, 2007). Of roughly 170 current adult Rual residents, around 105 identify primarily as Jedek speakers, 55 primarily as Jahai speakers, and 10 as speakers of other Aslian varieties. A high rate of intermarriage among the Jedek and Jahai bands at Rual (in accordance with the tradition of band exogamy practiced by the groups) means that roughly half of the Rual population is of mixed Jedek-Jahai parentage, and around half of Rual couples are mixed Jedek-Jahai couples. In mixed marriages, in some cases one or both parties claim to have proficiency in their spouse's variety, while in other cases they do not. Where a Jahai speaker from outside Rual comes to reside in Rual through marriage, in some cases these Jahai speakers claim to have gained proficiency in Jedek (in at least one case to the point of claiming to have forgotten Jahai) while in other cases, they do not claim proficiency in Jedek.

Jedek and Jahai are typologically and lexically similar, and both are small in terms of the number of their speakers. Jedek is spoken only at Rual, by around 280 speakers in total (including children), while Jahai has roughly 1,000 speakers, the greater part of

whom live in the neighboring state of Perak. Neither of the varieties are standardized or written, and no formal instruction exists in either. The contact between Jedek and Jahai at Rual is long-term and pervasive, and normativity levels in the community are low.

Predictions

The prediction in the current study is for semantic interaction in Jedek-Jahai bilinguals in the placement domain. On the basis of previous findings for bilingual semantic interaction in the placement domain it is predicted that evidence of interaction will be seen in the form of simplification in parts of the domain in which one of the varieties makes more fine-grained distinctions. However, since the Jedek and Jahai placement verb inventories do not differ in semantic granularity, it is unclear what the prediction should be for semantic interaction in Jedek-Jahai bilinguals. We might predict that areas of minor difference between Jedek and Jahai will reveal greater similarity in bilinguals' use of Jedek and Jahai verbs as compared to monolinguals.

Further predictions about the pattern of interaction expected for Jedek-Jahai bilinguals can be made on the basis of Muysken's (2013) framework of bilingual optimization strategies. Two predictions are possible for the Rual setting. The typological and lexical similarity of Jedek and Jahai, and the low levels of normativity at Rual are predicted to favor symmetric semantic interaction in bilinguals. At the same time, the more limited access to Jahai speakers at Rual is predicted to favor asymmetric interaction with a bias toward Jedek. The findings of a previous study of Jedek-Jahai bilinguals at Rual (Yager & Gullberg, published online 2019) suggest that evidence of both patterns may be seen in the results, and that interaction patterns may differ in Jedek- and Jahai-identifying bilinguals.

Method

Participants

The study included two groups of functional bilingual Jedek and Jahai speakers: 20 bilinguals who identify as primarily Jedek-speaking (labeled id-Jedek) and 4 bilinguals who identify as primarily Jahai-speaking (labeled id-Jahai). In addition, 20 monolingual Jedek and Jahai speakers participated: 15 monolingual Jedek-speaking Rual residents, and 5 monolingual Jahai speakers residing at Banun, a Jahai-majority area approximately 60 km west of Rual. The Jahai monolingual sample includes data

from three of the speakers reported in Burenhult (2012)[1], and from two additional Banun residents collected at Rual in 2014 in connection with the current study. The data from Jahai monolinguals from Banun were used since it is impossible to find Jahai-speaking Rual residents who do not have some (at least passive) knowledge of Jedek. Since Jedek is spoken only at Rual, the Jedek monolinguals are necessarily Rual residents. Note that the small size of the id-Jahai bilingual sample is a consequence of the low number of Jahai-identifying Jedek-Jahai bilinguals in existence. The Jahai monolingual sample was kept small so as to be proportionate to the id-Jahai bilingual sample. Participants were aged between 17 and 68 years (Mean 43 years; ages are approximate due to the imprecise nature of the birth records of many Rual and Banun residents). Table 1 presents the participant groups.

Since all Rual residents are exposed to both Jedek and Jahai, and all Jedek and Jahai speakers are in contact with speakers of other Aslian varieties as well as the majority language Malay, none of the participants of the current study are functionally monolingual. The label 'monolingual' in the current study instead refers to individuals who self-identify as speakers of either Jedek or Jahai but not both. Similarly, the label 'bilingual' reflects self-identification as a speaker of both Jedek and Jahai, and does not take into account questions of proficiency, usage or language dominance. Note that due to the pervasive and long-term nature of the contact between Jedek and Jahai speakers at Rual it is not possible to report on variables such as age of acquisition.

Table 1. The participants.

	Jedek	**Jahai**
Monolinguals	n=15	n=5
Bilinguals	n=20	n=4

Materials and task

Data were collected using the PUT task (Bowerman, Gullberg, Majid & Narasimhan, 2004), a task developed at the Max Planck Institute for Psycholinguistics to explore the semantics of placement event descriptions. The PUT task represents the current best standard for eliciting placement event descriptions and has been used in a number of studies showing that languages differ in their expression of events in the placement event domain (See e.g. papers in Kopecka & Narasimhan, 2012). The task uses a series of film clips to elicit event descriptions, and thus allows researchers to cover a range of event types while controlling for reference. The task consists of 63 short video clips each depicting a caused motion event performed by a human agent, designed to cover the types of distinctions that are relevant in the expression of the domain. The clips

[1] The data from one of the four participants involved in Burenhult's (2012) study were not used since this speaker was included in the bilingual Jahai group of the current study. This speaker currently resides at Rual and has knowledge of both Jedek and Jahai.

depict events of placement and removal, and differ in the nature of the Figure and Ground depicted, the spatial configuration between them, the Manner in which the Figure is moved, and the type and degree of control the agent has over the causation. Scene 1 of the task (as seen in Figure 1) might be described in English as "she puts a cup [Figure] on a table [Ground]".

Figure 1. Screenshot of Scene 1 of the PUT task (Bowerman et al., 2004). The individuals in the images have given their permission for the use of the images in connection with the distribution of the PUT task materials.

Procedure

The data were elicited using a director-matcher task in which two Jedek or Jahai participants interacted with one another: a 'director' who described each scene, and a 'matcher' whose task was to select a screenshot of the scene described on the basis of the director's descriptions. Director-matcher tasks produce highly naturalistic yet structured and focused language, and thus are ideal for the purposes of the current study. In addition, by using a native speaker of Jedek or Jahai as interlocutor (rather than the researcher whose proficiency in the varieties is more limited), the aim was to reduce the risk of Foreigner Talk (Ferguson, 1975), as well as regulate the language-mode of participants (Grosjean, 1998). Participants performed the task in Jedek with a Jedek matcher (a Rual resident) and in Jahai with a Jahai matcher (from outside Rual).

The director was shown each clip of the task one by one, preceded by three warm-up clips, on the screen of a laptop. The matcher was given a folder with eleven pages containing screenshots of the 63 clips of the task and the three warm-up clips, each page including six relevant items and six fillers. The director described each clip by responding to the prompt (in Jedek) *cʔay ʔoʔ diʔ* or (in Jahai) *mamɛy ʔoʔ deʔ* 'What

did/does s/he do?'[2]. The matcher then located on the pages of the folder a screenshot representing the clip described. Bilingual participants performed the task twice – once in Jedek and once in Jahai, while monolingual participants performed the task once, in their identity variety. The two sessions of bilingual participants were conducted at least 48 hours apart. The clips were presented in two different orders (versions 1 and 2), and the two versions of the task were randomly assigned to participants. The task version was kept constant between the two sessions of bilingual participants.

Data treatment and analyses

The full responses of each participant in each session were transcribed. Any response not containing a description of the target event was discarded, such as when a participant described a removal event where the target event was one of placement. Where a participant gave more than one response, the first description of the target event was selected for analysis. The analyses focus on the verbs used by participants in their descriptions of the 35 scenes of the task that involve placement events (Scenes 1–28, 31, 33, 35, 50, 51, 52 and 129).

The results are presented in extension maps (cf. Bowerman, 1996; Levinson & Meira, 2003; Indefrey et al., 2017; Yager & Gullberg, published online 2019) and in congruence analyses (cf. Indefrey et al., 2017; Yager & Gullberg, published online 2019). Due to the small sample size of the study, it was not possible to perform inferential statistics, thus the results section presents descriptive statistics only. Extension maps show the range of scenes described with each placement verb in Jedek and Jahai, giving the semantic extensions of the verbs used in the descriptions. In this way the maps give an overview of the similarities and differences between the placement verb extensions in the two varieties and between the extensions of monolinguals and bilinguals. In the extension maps, boxes denote the scenes described with a particular verb by more than 33% of the speakers of each group. The congruence analyses give a measure of the overlap between the Jedek and Jahai verb extensions of the groups. Congruence is defined as the number of scenes for which a pair of Jedek-Jahai verbs were most frequently used in Jedek *and* Jahai divided by the number of scenes for which at least one of the two verbs was most frequently used in Jedek *or* Jahai. The congruence analyses allow for comparison of the amount of overlap of Jedek and Jahai verb extensions in the bilinguals as compared to the monolinguals, providing a quantitative measure of whether the Jedek and Jahai verb extensions are more congruent in bilinguals as compared to monolinguals.

[2] Note that Jedek and Jahai third person singular pronouns do not encode gender, and Jedek and Jahai verbs are not specified for tense.

The Jedek and Jahai monolingual data are reported on first. The verbs used by Jedek and Jahai monolinguals are presented, followed by extension maps showing the use of verbs over scenes by Jedek and Jahai monolinguals, and congruence analyses quantifying the overlap of Jedek and Jahai verb extensions between the monolingual groups. Next, the bilingual data are presented, in extension maps showing the Jedek and Jahai verb extensions of the bilingual groups as compared to the verb extensions of the monolinguals, and in congruence analyses quantifying the overlap of Jedek and Jahai verb extensions in bilinguals. The congruence of the extensions of Jedek and Jahai verbs in the bilingual groups is then compared with the congruence of the extensions of the verbs as used by Jedek and Jahai monolinguals, and the congruence values of the two bilingual groups are compared. Finally, the use of less widely-used verbs (used for one scene of the task each) by the monolingual and bilingual groups is presented.

Results

The monolinguals

The Jedek and Jahai verbs

This section presents the verbs used in placement event descriptions by the monolingual Jedek and Jahai speakers. Table 2 presents the verbs used by the Jedek and Jahai monolinguals to describe two or more of the placement event scenes of the task.

Table 2. Verbs used for two or more scenes by the Jedek and Jahai monolinguals.

Jedek	Jahai	Gloss
boh	*boh*	'to put', 'to keep'
ʔisiʔ	*ʔisiʔ*	'to put inside', 'to insert' (from Malay *isi* 'to fill')
pakɛy, bləh	*bləh*	'to wear', 'to don' (Jedek *pakɛy* from Malay *pakai* 'to wear/to don')

The verbs used by the Jedek and Jahai monolinguals for groups of two or more scenes form almost identical sets. In both varieties, *boh* was used to describe events of placement, and *ʔisiʔ* for events of insertion. Jahai *bləh* and Jedek *bləh* and *pakɛy* were used for dressing events.

In addition to the verbs presented in Table 2, the monolingual Jedek and Jahai speakers used a number of verbs to describe individual scenes. Verbs that were used by the monolingual groups for only one scene of the task are presented in Table 3.

Table 3. Verbs used for one scene by the Jedek and Jahai monolinguals

Scene	Jedek	Jahai	Gloss
8	*hmpɛs*	*praʔ*	'to drop deliberately'
9	*hltuh*	*tbəl*	'to fall'
10	*paŋkaʔ*	*hakɔk*	'to throw'
13	*(boh)*	*tuh*	'to dump'
17	*(ʔisiʔ)*	*klit*	'to stuff', 'to plug'
18	*caduk*	*caduk*	'to skewer an oblong object in one's hair'
20	*cʔiʔ*	*cʔiʔ*	'to pour'
21	*hãc*	*hɛ̃c, cohɛʔ*	'to spill'
22	*ʔɛk*	*ʔɛk*	'to give'
23	*lkaʔ*	*lkaʔ*	'to insert one's hand'
24	*jruk*	*(ʔisiʔ)*	'to insert one's head'
28	*tipet*	*pet*	'to fix'
52	*ɲuh, tolɔʔ*	*surut*	'to push' (Jedek *tolɔʔ* from Malay *tolok*, 'to push')

While the more widely-used Jedek and Jahai verbs of Table 2 were almost identical in form, greater differences can be seen in the Jedek and Jahai verbs of Table 3 which were each used for one scene of the task. For three of these scenes, Jedek and Jahai monolinguals used identical forms (*caduk* for Scene 18, *ʔɛk* for Scene 22, *lkaʔ* for Scene 23), but for the majority of the scenes, they used either completely different forms (*hmpɛs/praʔ* for Scene 8, *hltuh/tbəl* for Scene 9, *paŋkaʔ/hakɔk* for Scene 10 and *ɲuh*, *tolɔʔ/surut* for Scene 52) or similar, but not identical forms (*hãc/hɛ̃c* for Scene 21, *tipet/pet* for Scene 28). For three scenes, monolinguals of one variety used a verb that was specific to that scene while monolinguals of the other variety used one of the more widely-used verbs.

The next section presents the use of the placement verbs over the scenes of the task by the Jedek and Jahai monolinguals, in extension maps. The analyses focus on the more widely-used verbs presented in Table 2, returning to the less widely-used verbs in the final section of the results.

Extension maps

We have seen that the more widely-used Jedek and Jahai placement verbs are largely identical in form. In this section we examine how these verbs were used over the placement event scenes of the task. Figures 2A and 2B present extension maps visualizing the use of the verbs presented in Table 2 over scenes by the two monolingual groups. Colored boxes represent the extension of each verb, that is, the scenes for which that verb was used by more than one third of the speakers of the group.

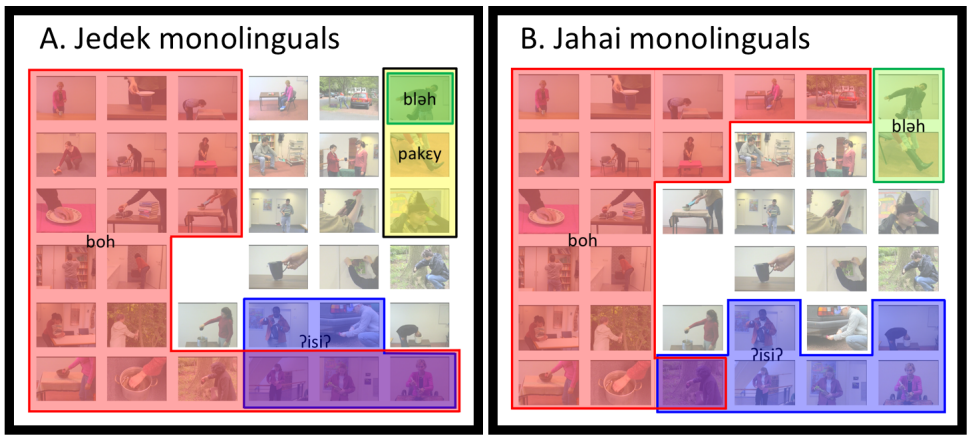


Figure 2A. Extension map showing the use of verbs over scenes by the Jedek monolinguals. **Figure 2B.** Extension map showing the use of verbs over scenes by the Jahai monolinguals.

Figures 2A and 2B show that the Jedek and Jahai monolinguals' use of placement verbs over scenes is broadly similar, but not identical. Differences between the varieties are not seen in terms of the semantic granularity of verbs, rather the differences can be seen in minor variations in the boundaries of verb extensions in the two varieties. The most striking of these differences is seen in the extension of *boh*. In addition to the extension shared by Jedek and Jahai monolinguals (scenes of controlled placement into situations of support from below, loose containment and hanging support, performed with the hands, arms, teeth or tongs), in Jedek the extension of *boh* also includes scenes of insertion (scenes of placement of a solid Figure into situations of tight containment), while in Jahai it does not. Jedek *boh* was also used for one additional scene (Scene 13 "flip block off notepad into bowl"), while Jahai *boh* was used for two additional scenes (Scene 8 "drop book deliberately on floor" and Scene 52 "push suitcase from car to tree"). The extension of *ʔisiʔ* is also similar but not identical in monolingual Jedek and Jahai. Its extension in both varieties contains the three scenes of the task involving insertion (placement of a solid Figure into situations of tight containment) as well as Scene 12 ("drop apple into bag"). The extension of Jedek *ʔisiʔ* also includes Scene 17 ("stuff rag into car exhaust pipe"), while the extension of Jahai *ʔisiʔ* includes Scenes 35 ("put pen in hole in tree trunk") and 24 ("put head into bucket").

Jedek monolinguals used one additional verb form not used by Jahai monolinguals, the dressing verb *pakɛy* (from Malay *pakai* 'to wear/to don'). The Jedek monolinguals used this verb in addition to the indigenous form *bləh* (the dressing verb also used by the Jahai monolinguals), whose extension is reduced as compared to the extension of *bləh* in Jahai.

Congruence analysis

In order to quantify the similarity of verb extensions in Jedek and Jahai monolinguals, Table 4 presents congruence values for the Jedek and Jahai verbs whose extensions were presented in Figure 2. A congruence value of 0 represents no overlap of the extensions of the Jedek and Jahai verbs and a value of 1 represents full overlap. Note that the total number of scenes for which the extension of a verb overlaps with that of other verbs in the congruence tables ('all') may be smaller than the sum of the rows or columns for that verb. This is because for some scenes more than one verb was used with equal highest frequency.

Table 4. Matrix of congruence values for the Jedek and Jahai verbs used by monolinguals. C = congruence value (0-1).

		Jahai							
		boh		*ʔisiʔ*		*bləh*		other	all
		N	C	N	C	N	C	N	N
Jedek	*boh*	14	0.7	4	0.2	0	0	1	19
	ʔisiʔ	0	0	3	0.4	0	0	1	4
	bləh	0	0	0	0	1	0.5	0	1
	pakɛy	0	0	0	0	1	0.3	1	2
	other	3		1		0			
	all	16		6		2			

Table 4 shows that each of the verbs is most congruent with the identical verb form of the other variety, but that none of the verb extensions overlap completely. One of the verbs used by the Jedek monolinguals (*boh*) and two of the verbs used by the Jahai monolinguals (*ʔisiʔ* and *bləh*) also have lower levels of congruence with one non-identical verb form in the other variety. The extension of Jedek *boh* overlaps with that of Jahai *ʔisiʔ* for four scenes (congruence value 0.2), reflecting the use of *boh* for insertion scenes in Jedek as shown in the extension map in Figure 2A. Meanwhile, the extension of Jahai *boh* does not overlap with Jedek *ʔisiʔ* – it is not used for insertion scenes, as shown in Figure 2B.

The bilinguals

Extension maps

Figure 3 presents extension maps comparing the id-Jedek and id-Jahai bilinguals' use of verbs over scenes in Jedek and Jahai to that of Jedek and Jahai monolinguals.

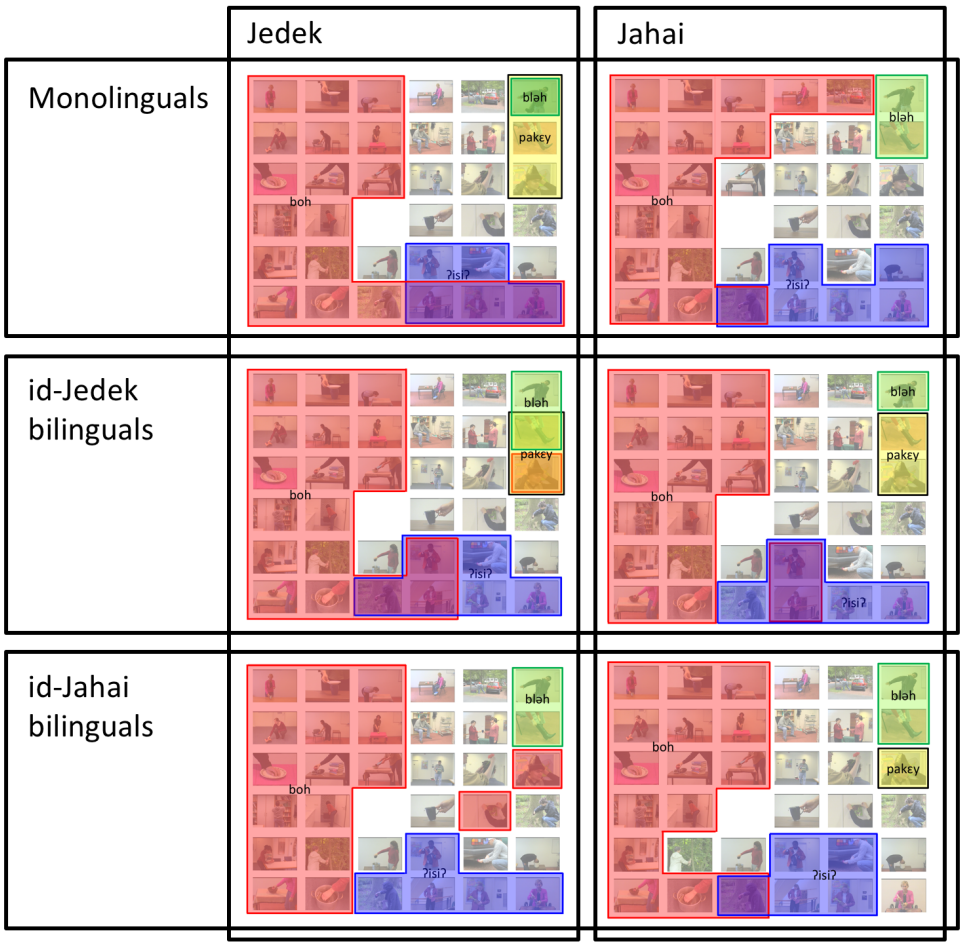


Figure 3. Extension maps showing the use of verbs over scenes by the id-Jedek and id-Jahai bilinguals in Jedek and Jahai, as compared to the Jedek and Jahai monolinguals.

The verb extensions of neither bilingual group follow the Jedek or Jahai monolingual pattern exactly, rather the Jedek and Jahai descriptions of both bilingual groups share features with both monolingual groups. The influence of Jahai can be seen in the Jedek extensions of both groups. The most obvious example of this Jahai influence is the reduction of the use of *boh* for tight containment scenes in Jedek by both bilingual groups – it is not used for any tight containment scenes by the id-Jahai bilinguals, and for only one tight containment scene by the id-Jedek bilinguals. Additional instances of Jahai influence can be seen in the expanded use of Jedek *bləh* in both bilingual groups, and the use of Jedek *ʔisiʔ* for Scene 35 ("put pen in hole in tree trunk").

Meanwhile, the influence of Jedek can also be seen in the Jahai extensions of both bilingual groups, in the use of *pakɛy* for at least one dressing scene by both groups, in

both groups' use of *boh* for Scene 13 ("flip block off notepad into bowl"), and in that neither bilingual group uses Jahai *boh* for Scene 8 ("drop book deliberately on floor") or 52 ("push suitcase from car to tree"). Finally, some features of the extensions of the bilingual groups differ from both monolingual varieties. Both bilingual groups used the placement verb *boh* for Scene 25 ("put hat on head") in Jedek, a scene for which neither monolingual group used this verb. The id-Jedek bilinguals also used *boh* for Scene 12 ("drop apple into bag") in both varieties, a scene for which it was not used by either monolingual group. And neither bilingual group used *boh* for Scene 35 ("put pen in hole in tree trunk") in their non-identity variety, a scene for which *boh* was used in both monolingual varieties. The patterns seen in the extension maps suggest symmetric semantic interaction in the descriptions of both bilingual groups.

Congruence analysis

In order to give a numeric measure of the overlap of Jedek and Jahai verb extensions in the bilinguals, Tables 5 and 6 present congruence values for the Jedek and Jahai verb extensions of the two bilingual groups. The congruence values for the id-Jedek bilinguals are presented in Table 5 and the congruence values for the id-Jahai bilinguals are presented in Table 6.

Table 5. Matrix of congruence values for the Jedek and Jahai verbs as used by the id-Jedek bilinguals. C = congruence value (0-1).

		Jahai									
		boh		*ʔisiʔ*		*bləh*		*pakɛy*		other	all
		N	C	N	C	N	C	N	C	N	N
Jedek	*boh*	15	0.9	1	0.1	0	0	0	0	0	16
	ʔisiʔ	1	0.1	5	0.8	0	0	0	0	1	6
	bləh	0	0	0	0	1	0.5	1	0.3	0	2
	pakɛy	0	0	0	0	0	0	1	0.5	0	1
	other	0		0		0		0			
	all	16		6		1		2			

For the id-Jedek bilinguals, each of the verbs is most congruent with the identical verb form in the other variety, but none of the extensions overlap completely across the varieties. Three of the verbs used by the id-Jedek bilinguals in Jedek (*boh*, *ʔisiʔ* and *bləh*) and three of the verbs used by this group in Jahai (*boh*, *ʔisiʔ* and *pakɛy*) also overlap with one non-identical verb form in the other variety, for one scene each. The overlap of the extensions of *boh* and *ʔisiʔ* is smaller for the id-Jedek bilinguals than that seen in the monolinguals – for the id-Jedek bilinguals, the extension of Jedek *boh* overlaps with Jahai *ʔisiʔ* for only one scene (compare the overlap of 4 scenes for the monolingual groups).

Table 6. Matrix of congruence values for the Jedek and Jahai verbs as used by the id-Jahai bilinguals.
C = congruence value (0-1).

		Jahai									
		boh		*ʔisiʔ*		*bləh*		*pakɛy*		other	all
		N	C	N	C	N	C	N	C	N	N
Jedek	*boh*	14	0.7	0	0	0	0	1	0.1	2	17
	ʔisiʔ	2	0.1	5	0.8	0	0	0	0	2	5
	bləh	0	0	0	0	2	1	0	0	0	2
	other	2		1		0		0			
	all	16		6		2		1			

For the id-Jahai bilinguals, as was the case for the monolinguals and for the id-Jedek bilinguals, each of the verbs is most congruent with the identical verb form in the other variety. Two of the verbs used by the id-Jahai bilinguals in Jedek (*boh* and *ʔisiʔ*) and one of the verbs they used in Jahai (*boh*) have some congruence with one non-identical verb form in the other variety. For the id-Jahai bilinguals, Jedek *boh* does not overlap with Jahai *ʔisiʔ*, while Jahai *boh* overlaps with Jedek *ʔisiʔ* for two scenes.

Congruence compared

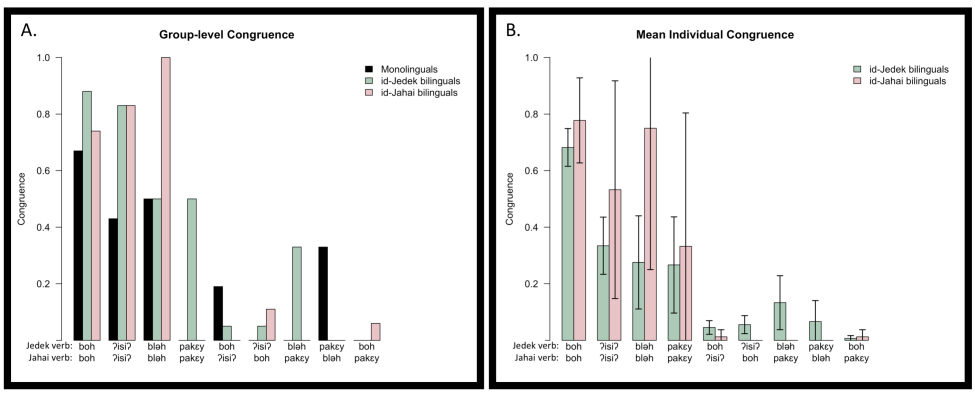


Figure 4A. Congruence of the extensions of Jedek and Jahai verbs as used by the monolingual and bilingual groups.
Figure 4B. Mean individual congruence of Jedek and Jahai verb extensions in the two groups of bilinguals.

In order to give a comparison of the monolingual and bilingual congruence analyses presented in Tables 4–6, Figure 4A presents the congruence values for the Jedek and Jahai verb extensions in the two bilingual groups and in the Jedek and Jahai monolinguals. For each identical verb pair, the congruence values of both bilingual groups are equal to or higher than the congruence values between the monolingual groups, thus revealing an overall increase in the congruence of the extensions of identical verb forms across varieties in bilinguals as compared to monolinguals.

Bilingual congruence levels are generally low for non-identical verb pairs, and reflect the use of both *bləh* and *pakɛy* for dressing scenes and of both *ʔisiʔ* and *boh* for some of the scenes involving insertion.

Since the results of the group-level analyses do not necessarily correspond to individual bilingual speakers' use of the verbs in Jedek and Jahai, the congruence of Jedek-Jahai verb extensions as used by each individual bilingual speaker was also calculated. Individual congruence was defined as the number of scenes for which a speaker used both the Jedek and Jahai verbs of a pair divided by the number of scenes for which the speaker used one of the verbs of the pair. Figure 4B shows the mean individual congruence values for the verbs in Jedek and Jahai in the two bilingual groups. As in the group congruence values, mean individual congruence values are substantially higher for identical verb forms than for non-identical verb forms. Comparing the two bilingual groups, mean individual congruence values for each identical verb pair are relatively similar for id-Jedek and id-Jahai bilinguals (but see the verb *bləh* with a congruence value of 0.75 across varieties for the id-Jahai bilinguals and 0.28 for the id-Jedek bilinguals). The id-Jahai bilinguals tend to have higher congruence values than the id-Jedek bilinguals for identical verb pairs, while the congruence of non-identical verb forms is low for both groups of bilinguals, but slightly higher for the id-Jedek bilinguals than the id-Jahai bilinguals.

The less widely-used verbs

So far, the analyses have focused on the verbs of the dataset used in descriptions of two or more scenes. A number of other verbs, used for one scene of the task each were presented in Table 3, where it was evident that there were greater differences in form between the two monolingual varieties for these verbs than for the more widely-used verbs. Table 7 presents the use of these verbs over scenes by monolinguals and bilinguals in Jedek and Jahai. The table shows the verb form used by the largest number of speakers of each group for each scene, with green indicating verb forms used by the Jedek monolinguals and red indicating verb forms used by the Jahai monolinguals. Verb forms used by both monolingual groups are coded in blue. In this way Table 7 compares the id-Jedek and id-Jahai bilinguals' use of Jedek and Jahai verb forms for these scenes.

Table 7. Use of the less widely-used verbs by monolinguals and bilinguals in Jedek and Jahai. Green=Jedek verb; red=Jahai verb; blue=Jedek/Jahai verb.

Scene	Jedek mono-linguals	id-Jedek bilinguals		id-Jahai bilinguals		Jahai mono-linguals
		Jedek	Jahai	Jedek	Jahai	
8	*hmpɛs*	*hmpɛs*	*hɔk*	*yɔh*	*yɔh*	*praʔ*
9	*hltuh*	*hltuh*	*hltuh*	*hltuh* / *tbəl*	*tbəl*	*tbəl*
10	*paŋkaʔ*	*paŋkaʔ*	*hɔk*	*hɔk*	*hɔk*	*hakɔk*
13	*(boh)*	*(boh)*	*(boh)*	*(boh)*	*(boh)*	*tuh*
17	*(ʔisiʔ)*	*(ʔisiʔ)*	*sumɛt, (ʔisiʔ)*	*sumat*	*(ʔisiʔ)*	*klit*
18	*caduk*	*caduk*	*caduk*	*caduk*	*caduk*	*caduk*
20	*cʔiʔ*	*cʔiʔ*	*cʔiʔ*	*cʔiʔ*	*cʔiʔ*	*cʔiʔ*
21	*hāc*	*hāc*	*hẽc*	*hāc*	*hẽc*	*hẽc, cohɛʔ*
22	*ʔɛk*	*ʔɛk*	*ʔɛk*	*ʔɛk*	*ʔɛk*	*ʔɛk*
23	*lkaʔ*	*lkaʔ*	*lkaʔ*	*lkaʔ*	*lkaʔ*	*lkaʔ*
24	*jruk, (boh)*	*jruk*	*jruk*	-	*jruk*	*(ʔisiʔ)*
28	*tipet*	*timpɛʔ*	*timpɛʔ*	*(boh)*	*pet*	*pet*
52	*ɲuh, tolɔʔ*	*ɲuh*	*ɲuh*	*ɲuh*	-	*surut*

The id-Jedek bilinguals used only Jedek verb forms in their Jedek descriptions of the scenes. In their Jahai descriptions, they primarily used Jedek verb forms, but used a Jahai verb form (*hẽc*) for one scene. Meanwhile, the id-Jahai bilinguals used a mix of Jedek and Jahai verb forms in their descriptions in both varieties, using more Jedek forms in their Jedek descriptions (three Jedek verb forms and one Jahai verb form) and more Jahai forms in their Jahai descriptions (three Jahai verb forms and one Jedek verb form). Although based on a very small subset of the data, the patterns seen in these less widely-used verbs suggest bidirectional influence of Jedek and Jahai on one another in the id-Jahai bilinguals and unidirectional influence of Jedek on Jahai in the id-Jedek bilinguals.

Discussion

The study investigates bilingual semantic interaction in the domain of placement events in two groups of Jedek-Jahai bilinguals in the small-scale, egalitarian multilingual setting of Rual in northern Peninsular Malaysia. The results reveal a high degree of similarity in the Jedek and Jahai placement verb lexica; Jedek and Jahai placement verbs are seen to be largely identical in form, with roughly equivalent semantic extensions. Despite this high degree of similarity in the Jedek and Jahai monolingual varieties, the bilingual data reveal semantic interaction in Jedek-Jahai bilinguals. Congruence analyses showed an increase in congruence of the extensions of identical Jedek and Jahai verb forms in bilinguals as compared to monolinguals. The prediction was for semantic interaction in bilinguals in the form of simplification where the semantic distinctions

of one of the monolingual varieties were more fine-grained. This outcome was not seen in the results, since no differences in semantic granularity were found between monolingual Jedek and Jahai. Instead, evidence of semantic interaction was seen in terms of an increase in the congruence of the extensions of form-identical Jedek and Jahai verbs in bilinguals.

While the Jedek and Jahai placement verb inventories were highly similar, identical forms were not 100% semantically equivalent across monolingual Jedek and Jahai. Parts of the domain in which there existed differences in the verb extensions of Jedek and Jahai revealed the directionality of semantic interaction. Extension maps showed evidence of symmetric semantic interaction in the placement verb extensions of both groups of bilinguals – that is, the Jedek and Jahai placement verb extensions of Jedek-identifying and Jahai-identifying bilinguals were seen to share features with both monolingual varieties. This pattern of interaction is in line with one of the predictions possible for Jedek-Jahai bilinguals at Rual on the basis of Muysken's (2013) framework of bilingual optimization strategies: that the typological and lexical similarity of Jedek and Jahai and the low level of normativity at Rual should favor symmetric semantic interaction. However, another prediction of the framework, for asymmetric semantic interaction favoring Jedek based on the relative lack of access to Jahai speakers at Rual, was not borne out in the results. The implications of the results for studies of bilingual semantic interaction and for the testing of the predictions of Muysken's framework are discussed below.

The findings of the current study have implications for studies of semantic interaction. While most studies to date focus on language pairs whose semantic systems differ in semantic granularity, the current study offers evidence of bilingual semantic interaction in the absence of such a difference. The results show that even in cases where monolingual varieties are closely related and very similar both lexically and in terms of semantics, bilinguals still manage to find ways to make the varieties yet more similar as compared to monolinguals. Further investigation of interaction phenomena in highly similar semantic systems may lead to important insights about the nature of bilingual systems. The findings also have implications for the testing of Muysken's (2013) framework of bilingual optimization strategies. We have seen that two different predictions were possible for the Rual setting on the basis of the framework. One issue in using Muysken's model to form specific predictions about bilingual outcomes is that it does not provide weights for the relative importance of the different factors predicted to play a role. For example, in a setting such as Rual in which there is low normativity in the speech community and where language varieties are typologically and lexically similar, but where there is also an asymmetry in the level of access to speakers of the varieties, how should the relative impact of these features be interpreted in order to allow us to form predictions about bilingual outcomes? As Muysken points out, a system of weighting would be necessary for the model to have enough predictive power

to be able to generate more specific predictions. The development of such a system will require rigorous testing of the model's predictions in a range of bilingual scenarios.

Comparison of the findings of the current study with those of a previous study of bilingual semantic interaction in the Rual setting offers interesting insights. While the current study finds symmetric semantic interaction in both Jedek- and Jahai-identifying bilinguals in the placement domain, in the domain of topological relations Yager and Gullberg (published online 2019) found symmetric semantic interaction in Jahai-identifying bilinguals, but asymmetric interaction in Jedek-identifying bilinguals. That is, different patterns of symmetry were found in the two studies. How can these different results be explained? One possibility is that there is something about the domains themselves that makes placement event descriptions more susceptible to symmetric interaction in bilinguals (cf. processing differences in motion verbs and more static language, see Wallentin et al., 2005; Wallentin et al., 2011). Another explanation may lie in the scope of analysis of the studies. The analyses of Yager and Gullberg (published online 2019) focus on the topological relation marker *klɛŋ*, whose form is identical but whose semantic extension differs greatly in monolingual Jedek and Jahai. Meanwhile, the current study examined the placement domain as a whole, taking into account the full set of verbs used in the placement event descriptions of the task. It is possible that broader analysis of Jedek and Jahai topological relation markers may reveal aspects of symmetry in the interaction in this domain, potentially with differing patterns of directionality in different parts of the domain.

Another factor that may play a role in the different interaction patterns seen for Jedek-Jahai bilinguals in the two studies is the degree of difference between the varieties in the domains studied. We have seen that a more symmetric outcome was found for the largely equivalent Jedek-Jahai placement verbs than for the topological relation marker *klɛŋ* whose semantics in Jedek and Jahai differed to a greater extent. Might domains in which monolingual varieties are more similar be more prone to symmetric semantic interaction? Interestingly, the less widely-used placement verbs of the current study, which revealed greater lexical differences between monolingual Jedek and Jahai, showed a similar pattern to that found for *klɛŋ* by Yager and Gullberg (published online 2019). While this result is based on a small number of verb forms used for a small number of scenes, it is striking that the pattern of asymmetry is the same across the studies where there is greater divergence between monolingual varieties. Further research should explore the effect of the degree of similarity of language varieties on the symmetry of patterns of interaction in bilinguals. In addition, since different interaction patterns may be revealed in different parts of a domain, it is important that studies of bilingual semantic interaction consider results not only from different semantic domains (cf. Alferink, 2015), but also from different parts of the domains under study.

A final aspect of the current study that warrants discussion relates to the methodological issues associated with investigating bilingual semantic interaction in speakers of non-standardized, closely-related language varieties. In the current study several issues were encountered in applying the methodologies used in Western settings with standardized, written languages, to a setting of small-scale multilingualism in closely-related language varieties. In general, the issues encountered in the current study relate to problems in the separation of language varieties. Since the language boundaries of Jedek and Jahai could not be defined in an a priori way, the current study used the self-reported language identity of speakers to define the boundaries of Jedek and Jahai. Thus individuals who self-identified as speakers of Jedek were taken to represent the language variety named 'Jedek'. But since Jedek and Jahai are non-standardized and highly similar, the question of what speakers mean when referring to Jedek and Jahai as ways of speaking is not necessarily a straightforward one. The case of Jahai is particularly perplexing, since Jahai is spoken in different geographical areas and is likely subject to areal variation, in addition to differences resulting from the influence of Jedek on the variety of Jahai spoken at Rual (see Yager, submitted). These kinds of issues represent particular complications for bilingualism researchers. For example, where language boundaries cannot be defined in a straightforward way, what do speakers mean when they claim to be bilingual? Methodological issues of this kind must be grappled with if we wish to gain an understanding of the dynamics of bilingualism in a wider range of bilingual settings. In contexts where the existence of language varieties as uniform and easily separable entities cannot be taken for granted, one potentially less problematic approach may involve analysis of bilingual outcomes at the level of individual speakers. This kind of approach should be explored in further studies dealing with this kind of bilingual settings.

Conclusion

The study provides evidence of bilingual semantic interaction in the domain of placement events in a small-scale egalitarian multilingual setting. The results suggest that even very similar semantic systems may become more similar in bilinguals. Results are in line with the predictions of Muysken's (2013) framework of bilingual optimization strategies, suggesting that the framework is useful in forming predictions about bilingual outcomes for the verbal domain of placement events. But the results also suggest a complexity in the interplay of factors influencing bilingual outcomes. Social factors such as the prestige of languages, normativity levels in the speech community and relative access to speakers may interact with features of the languages in focus such as the level of similarity of semantic systems to produce specific bilingual outcomes. The relative importance of the different factors affecting bilingual outcomes should be explored to allow for the generation of more specific predictions on the basis

of the model. Given the role of the social features of bilingual settings in affecting bilingual outcomes, it is imperative that we work to fill the large gaps that currently exist in the range of types of bilingual settings explored in the literature. The findings point to a need for more in-depth work in non-standardized, small-scale egalitarian multilingual settings, and for work that goes beyond the level of language varieties and takes into account the language production of individual speakers.

Ethics statement

This study was carried out in accordance with the recommendations of the Swedish Research Council CODEX Rules and Guidelines for Research. Subjects, who were all non-literate, gave oral informed consent according to a protocol developed in accordance with the standards of the CODEX of the Swedish Research Council and the recommendations of the Volkswagen Foundation's Documentation of Endangered Languages program.

Acknowledgements

I would foremost like to thank the community at Rual without whom this research would not have been possible. I would like to acknowledge the Economic Planning Unit, Putrajaya and the Department of Orang Asli Development, Kuala Lumpur, who granted permission for the research to be conducted, and Prof. Dr. Kamila Ghazali from the University of Malaya, who acted as my academic counterpart in Malaysia. Field-based data collection was financed with the help of travel grants from the Lars Hierta Memorial Foundation, The Birgit Rausing Language Programme, Fil. Dr. Uno Otterstedt Foundation, and The Knut and Alice Wallenberg Foundation.

References

Alferink, I. (2015). *Dimensions of convergence in bilingual speech and gesture*. Nijmegen: Radboud Universiteit Nijmegen. (Doctoral dissertation.)

Alferink, I., & Gullberg, M. (2014). French–Dutch bilinguals do not maintain obligatory semantic distinctions: Evidence from placement verbs. *Bilingualism: Language and Cognition* 17. 22–37.

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language* 53. 60–80.

Berthele, R. (2015). Convergence in the domains of static spatial relations and events of putting and taking. Evidence from bilingual speakers of Romansh and German. *International Journal of Bilingual Education and Bilingualism* 18. 624–642.

Bowerman, M. (1996). Learning how to structure space for language: A cross-linguistic perspective. In P. Bloom, M. Peterson, L. Nadel & M. Garrett (eds.), *Language and space,* 385–436. Cambridge, MA: MIT Press.

Bowerman, M., Gullberg, M., Majid, A., & Narasimhan, B. (2004). Put project: The crosslinguistic encoding of placement events. In A. Majid (ed.), *Field Manual, Vol. 9,* 10–18. Nijmegen: Max Planck Institute for Psycholinguistics.

Burenhult, N. (2012). The linguistic encoding of placement and removal events in Jahai. In A. Kopecka & B. Narasimhan (eds.), *Events of putting and taking: A crosslinguistic perspective,* 21–36. Amsterdam: John Benjamins Publishing.

Cadierno, T., Ibarretxe-Antuñano, I., & Hijazo-Gascón, A. (2016). Semantic categorization of placement verbs in L1 and L2 Danish and Spanish. *Language Learning* 66. 191–223.

Ervin, S. M. (1961). Semantic shift in bilingualism. *The American Journal of Psychology* 74. 233–241.

Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics* 17. 1–14.

Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late bilinguals: All words are not created equally. *Bilingualism: Language and Cognition* 13. 385–408.

Gomes, A. (2007). *Modernity and Malaysia: Settling the Menraq forest nomads*. London: Routledge.

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition* 1. 175–176.

Gullberg, M. (2009a). Reconstructing verb meaning in a second language: How English speakers of L2 Dutch talk and gesture about placement. *Annual Review of Cognitive Linguistics* 7. 222–245.

Gullberg, M. (2009b). Gestures and the development of semantic representations in first and second language acquisition. *AILE Lia* 1. 117–139.

Gullberg, M. (2011). Thinking, speaking and gesturing about motion in more than one language. In A. Pavlenko (ed.), *Thinking and speaking in two languages,* 143–169. Clevedon, UK: Multilingual Matters.

Ibarretxe-Antuñano, I., Cadierno, T., & Hijazo-Gascón, A. (2016). The role of force dynamics and intentionality in the reconstruction of L2 verb meanings: A Danish-Spanish bidirectional study. *Review of Cognitive Linguistics* 14. 136–160.

Indefrey, P., Şahin, H., & Gullberg, M. (2017). The expression of spatial relationships in Turkish/Dutch bilinguals. *Bilingualism: Language and Cognition* 20. 473–493.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Kopecka, A. (2012). Semantic granularity of placement and removal expressions in Polish. In A. Kopecka & B. Narasimhan (eds.), *Events of putting and taking: A crosslinguistic perspective*, 327–348. Amsterdam: John Benjamins Publishing.

Kopecka, A., & Narasimhan, B. (2012). *Events of putting and taking: A crosslinguistic perspective*. Amsterdam: John Benjamins Publishing.

Koster, D., & Cadierno, T. (2019). The effect of language on recognition memory in L1 and L2 speakers: The case of placement events. *International Journal of Bilingualism* 23. 651–669.

Levinson, S. C. (2012). Preface. In A. Kopecka & B. Narasimhan (eds.), *Events of putting and taking: A crosslinguistic perspective*, xi–xv. Amsterdam: John Benjamins Publishing.

Levinson, S. C., & Brown, P. (2012). Put and Take in Yélî Dnye, the Papuan language of Rossel Island. In A. Kopecka & B. Narasimhan (eds.), *Events of putting and taking: A crosslinguistic perspective*, 273–296. Amsterdam: John Benjamins Publishing.

Levinson, S. C., & Meira, S. (2003). 'Natural concepts' in the spatial topological domain—Adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79. 485–516.

Malt, B., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science* 4. 583–597.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16. 709–730.

Narasimhan, B., Kopecka, A., Bowerman, M., Gullberg, M., & Majid, A. (2012). Putting and taking events: A crosslinguistic perspective. In A. Kopecka & B. Narasimhan (eds.), *Events of putting and taking: A crosslinguistic perspective*, 1–20. Amsterdam: John Benjamins Publishing.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley & Los Angeles, CA: University of California Press.

Viberg, Å. (1998). Crosslinguistic perspectives on lexical acquisition: The case of language-specific semantic differentiation. In K. Haastrup & Å Viberg (eds.), *Perspectives on lexical acquisition in a second language*, 175–208. Lund, Sweden: Lund University Press.

Wallentin, M., Ellegaard, T., Østergaard, S., Østergaard, L., & Roepstorff, A. (2005). Motion verb sentences activate left posterior middle temporal cortex despite static context. *NeuroReport* 16. 649–652.

Wallentin, M., Nielsen, A. H., Vuust, P., Dohn, A., Roepstorff, A., & Ellegaard, T. (2011). BOLD response to motion verbs in left posterior middle temporal gyrus during story comprehension. *Brain and language* 119. 221–225.

Yager, J. (submitted). Outcomes of small-scale egalitarian contact in closely-related languages: Evidence from Southeast Asian foragers.

Yager, J., & Gullberg, M. Asymmetric semantic interaction in Jedek-Jahai bilinguals: Spatial language in a small-scale, non-standardized, egalitarian, long-term multilingual setting in Malaysia. *International Journal of Bilingualism*. doi:10.1177/1367006918814378. Published online by SAGE Journals, May 6, 2019.

Printed by Media-Tryck, Lund 2019 NORDIC SWAN ECOLABEL 3041 0903

Small-scale multilingualism and language contact in egalitarian foragers

Multilingualism and language contact in small-scale, egalitarian contexts are important phenomena affecting processes of language change throughout human history, yet our understanding of the outcomes of multilingualism and language contact in this kind of setting remains limited. This thesis provides insight into the linguistic consequences of interaction between closely-related, recently-described language varieties in small-scale egalitarian contexts, and works to overcome some of the methodological challenges associated with the study of language contact and multilingualism in this kind of setting. In four studies, the thesis provides the first linguistic description of the newly discovered Aslian (Austroasiatic) language variety Jedek, and investigates lexical and semantic outcomes of multilingualism and language contact in egalitarian foragers speaking the closely-related language varieties Jedek and Jahai in northern Peninsular Malaysia. The findings highlight the value of research in lesser-known linguistic settings for advancing our theories of multilingualism and language contact.

LUND UNIVERSITY

The Faculties of Humanities and Theology
Centre for Languages and Literature
ISBN 978-91-88899-73-6