Multilingual practices
Profile and significance for evaluating speakers’ plans and intentions

Brook Hefright, PhD, Michelle Morrison, PhD, Rachel Strong, Claudia M. Brugman, PhD

Table of Contents
Multilingual practice ........................................................................................................................... 3
Switching vs. mixing .......................................................................................................................... 3
Mixing, not “borrowing” ................................................................................................................... 5
Implications for data collection ....................................................................................................... 6
Implications for automatic language identification ........................................................................ 7
Excerpt 1: Inconsistent evaluation of language affiliation ........................................................... 7
Corpus development and annotation ............................................................................................... 9
Findings ........................................................................................................................................... 11
Excerpt 2: Mixing of Arabic, Somali, and English ........................................................................ 12
Excerpt 3: Confirmation .................................................................................................................... 13
Excerpt 4: Correction ....................................................................................................................... 14
Excerpt 5: Topic shift ......................................................................................................................... 15
Excerpt 6: Negotiating the floor ...................................................................................................... 16
Excerpt 7: Constructed discourse ................................................................................................... 16
Preliminary conclusions .................................................................................................................... 17
Implications for training and operations .......................................................................................... 18
References ....................................................................................................................................... 20
Appendix A. Annotation schema .................................................................................................... 22
Appendix B. Sample ELAN transcriptions ...................................................................................... 23
Appendix C. Training materials for transcription of natural conversation ....................................... 25
Language Codes .............................................................................................................................. 26
Using ELAN .................................................................................................................................... 27

During the first year of a two-year project, the University of Maryland Center for Advanced Study of Language (CASL) conducted a study of multilingual practices in a group of Somali speakers who also use Arabic, and began collecting data for a parallel study in a group of Persian speakers who
also use Arabic. This report presents interim findings in the form of a profile of multilingual practices with respect to the Somali/Arabic group, as well as an assessment of the significance of these practices for evaluating plans and intentions of multilingual speakers.

The language materials we analyze in this report exemplify how different patterns of multilingual practice in conversation can be used to infer who is in control of how the conversation proceeds, and how participants orient to the group. The key constructs for our analysis are “switching,” the use of multiple languages to convey something about the conversation other than its content, and “mixing,” the use of multiple languages with no such additional intent.

Speakers who switch demonstrate that they are attending to the differences between the languages used and the boundaries between groups. Mixing, by contrast, suggests that the borders between languages and groups are less salient to the participants. Distinguishing between switching and mixing can provide clues to group identity, an aid to discerning plans and intentions.

Our goal at the end of the study is to provide analysts with an exemplified approach that will allow them reliably to identify patterns of switching and mixing, use these patterns to characterize conversations in a target group, then use these characterizations to draw reliable inferences about group members’ relationships with each other and their orientation to the group.

In this approach, identification of the source language of individual linguistic forms can be useful, and we suggest some implications of our research for automatic language identification. The main focus, however, is on developing analysts’ ability to discern and draw inferences from patterns of linguistic practice. Since our approach is grounded in research on communicative universals, it is generalizable to any combination of languages, and in fact represents a natural extension of established research on monolingual conversation. Whatever the specific language or languages, close attention to the details of linguistic practice has the potential to yield valuable intelligence about individual speakers and the groups to which they belong.

TARGET LANGUAGES

Our research design, developed in consultation with the client, focused on the use of Arabic among speakers of languages other than Arabic. To that end, we chose to examine the practices of two groups: speakers of Somali who also use Arabic, and speakers of Persian who also use Arabic.

Somali is spoken as a first or second language by approximately 95 percent of the population of Somalia (Appleyard & Orwin, 2008), and it is the country’s official language (Federal Republic of Somalia, 2012). Arabic has been designated as a “second” national language, and many Somalis know at least some Arabic, though estimates place the number of highly proficient Arabic speakers in Somalia at less than 20 percent (Ouane, 2003).

Somali is also spoken by large diaspora populations in Djibouti, Ethiopia, and Kenya (Lewis, Simons, & Fennig, 2014) and by large populations in the Middle East, Europe, North America, India, and Australia. Banti (2013) notes that the complex relationship between the Somali and Arabic languages includes centuries of contact between their speakers on the coast of Somalia. This has resulted in significant shared vocabulary.

Persian is spoken by a total of 56 million people in Iran; major dialects spoken in Afghanistan and Tajikistan are often considered languages in their own right (Lewis, et al., 2014). The greatest Arabic language influence dates back to the seventh century CE. In this case also, the greatest influence has been in vocabulary where, by some measures, as much as 50 percent of the vocabulary of twelfth-century Persian was of Arabic origin (Encyclopedia Iranica, n.d.). This long history of contact with Arabic has also affected the morphology of modern Persian (Perry, 2005).

Persian is also spoken by large diaspora populations outside of Iran. The 2010 World Bank Migration and Remittances Factbook listed Iran among the “top ten emigration countries” of the Middle
East and North Africa. For emigrants from Iran, four of the top ten destination countries are Arabic-speaking states: Qatar, Kuwait, the United Arab Emirates, and Bahrain.

### MULTILINGUAL PRACTICE

The term “code switching” has become commonplace among linguists and laypeople alike to describe language use that combines elements of two or more languages. The Merriam-Webster dictionary defines “code switching” in this everyday sense as “The switching from the linguistic system of one language or dialect to that of another” (“Code-switching,” n.d.).

While this definition makes reference to linguists’ notion of a “linguistic system,” it does not distinguish, as many linguists do, among switching within words, between words, or at the level of the clause, the sentence, or the discourse. Auer (1999), for example, does not find switching below the level of the clause to be useful for understanding conversation or identity, while MacSwan (2005) argues that switching above the level of the clause is not switching at all. Nor does the dictionary definition address switching related to addressee or situation, which was the focus of Blom & Gumperz’s (1972) seminal work on code switching, but which Poplack (1980[1979]) backgrounds. Moreover, the dictionary definition includes switching between both languages and dialects; while Auer considers language and dialect switching within a single theoretical framework, Ramat (1995) has argued for important differences between the two.

Since our goal is to assess the significance of language use for evaluating a speaker’s plans and intentions, we follow Auer in focusing on how a speaker’s use of multiple languages can “contextualize” his or her speech.

“Contextualization” (Gumperz, 1982) refers to the way that speakers signal to hearers the most appropriate context through which to interpret their speech. For example, the sentence “It’s cold” may be a simple statement of fact, or it can be request for the addressee to close the window. When the speaker pronounces “It’s cold” with emphasis on cold and a shudder in his voice (perhaps while briskly rubbing his arms, miming shivering, or waving at the window) he may signal to the hearer that his discomfort is so great that the hearer should do something about it, and therefore signal that the sentence is a request rather than a statement.

When using multiple languages serves this contextualization function – that is, it serves as a signal to the hearer to interpret the speech in a particular way – Auer calls it “language switching”; when it does not – that is, when using multiple languages serves no function other than communicating content – Auer calls it “code mixing.” In this report we will use “multilingual practices” as a general term to cover both; since we do not want to presuppose whether the linguistic elements that we are considering constitute a “code” or a “language,” we refer to practices that serve a contextualization function simply as “switching,” and practices that do not as “mixing.”

### SWITCHING VS. MIXING

Auer’s (1988, 1998, 1999) work is in the tradition of Blom and Gumperz (1972), which examines how language users in Norway switch between standard and dialect both in response to changes in situation, or “situational” switching, as well as to create changes in situation, or “metaphorical” switching. Subsequently, Gumperz (1982, p. 131) expands on this notion of metaphorical switching as one of many kinds of “contextualization cues,” which he defines as “any feature of linguistic form that contributes to the signaling of contextual presuppositions.” In other words, contextualization cues signal how speakers intend their speech to be interpreted in context. Gumperz underscores that such cues are indeterminate: speakers may signal one interpretation, but hearers may come to another, and different hearers may come to different interpretations.
Gumperz’s social approach to multilingual practice has strong parallels in Silverstein’s semiotic approach. Silverstein builds upon Peirce’s concept of the sign, and particularly his concept of the index, a type of sign that derives its meaning from the context of its occurrence. From this, Silverstein (2003) develops the concepts of “indexical presupposition” and “indexical creativity.” Indexical presupposition (like Blom & Gumperz’s situational switching) describes how a sign may evoke previous situations that speakers and hearers believe to be typical of its use; indexical creativity (like metaphorical switching) describes how the use of a sign may signal a new situation. Like Auer, Silverstein emphasizes that this signaling is indeterminate, and frequently implicit.

Applied more concretely to multilingual practice, individual linguistic elements – whether they are sounds, words, grammatical constructions, or discourse structures – may recall the other linguistic elements with which they typically co-occur. When a speaker juxtaposes two linguistic elements in a novel way, each element may evoke its own typical context through indexical presupposition, and the juxtaposition itself may signal a new situation through indexical creativity.

For example, if a speaker of American English says, “That dress is très chic!” he or she is juxtaposing forms that most hearers recognize as occurring more typically in two separate languages, English and French. Assuming the rest of the conversation is in English, the French phrase très chic can serve not only to emphasize the utterance, but also to evoke (or indexically presuppose) values typically associated with French and France, such as quality and style. At the same time, juxtaposing that dress is with très chic may express (or indexically create) a clever, ironic tone and allow the speaker to come off as witty and cosmopolitan.

Silverstein’s approach underscores that what speakers juxtapose are not grammatical systems, but rather individual linguistic elements that index typical situations of co-occurrence, which speakers and hearers may understand to be grammatical systems. This means that a switch can consist of a single word, or even a single distinctive pronunciation. It also means that switching does not require language users to be fluent in the language into which they switch: even inconsistent, stylized uses of a stereotypical feature of a language or dialect may be enough to evoke it, along with its typical users, situations of use, and message types.

For example, some speakers of American English occasionally pronounce the name of the company Target as if it rhymed with the word chargé. As it happens, the English word target is an old loanword from French; however, most speakers are unaware of this etymology. Instead, they use the stylized French pronunciation first to evoke (or indexically presuppose) French quality and style, then to express (or indexically create) a satiric comment on the chain’s attraction to shoppers whose pretensions are out of step with their budgets.2

Auer contrasts switching with “mixing,” multilingual practice that speakers do not intend and hearers do not perceive to be novel. Previous scholars of code switching (Poplack, 1980[1979]; Myers-Scotton, 1988) have noted that, in some groups, multilingual practice becomes routinized to the point that participants no longer attend to individual juxtapositions. Instead, they come to see the overall pattern of mixing as an emblem of group identity, and frequently describe the pattern as a language or dialect in its own right. In Silverstein’s terms, they cease to be indexically creative and become indexically presupposing.

1 Peirce defines sign as “something which stands to somebody for something in some respect or capacity” (1955[1940], p. 98). Saussure’s definition of a linguistic sign as a link between a concept and a sound pattern (Saussure (1966[1916], p. 66) is more familiar to linguists. Peirce’s definition, however, is more general than Saussure’s: Saussure’s sign corresponds to Peirce’s symbol, a type of sign for which the meaning is arbitrary, or entirely conventional.

2 There are other equally plausible indexations of this practice, for example, that the store, rather than its shoppers, pretends to be something it is not. The variability of intentions and the variability of interpretations are exactly to be expected given the indeterminate relationship between individual and group intentions in language use.
For example, many English speakers in the American Southwest use the term *Spanglish* – a portmanteau of the words *Spanish* and *English* – as if it described a language in its own right. In fact, *Spanglish* describes many different practices that involve switching and mixing elements of Spanish and English. What these practices have in common, however, is that they evoke (or indexically presuppose) typical users, such as U.S. Latinos, and typical situations of use, such as transactional exchanges on the U.S.-Mexican border.

A key goal of Auer’s approach is to carefully distinguish the perspectives of conversational participants and in-group members from the perspectives of outside observers, including linguists. Previous approaches to code switching took standard languages as their reference point without considering in-group members’ own perspectives on the elements in their linguistic repertoires; Auer’s approach requires the analyst to assess language use from the language users’ perspective.

Auer does this by looking carefully at cases of multilingual practice, and in each case, evaluating whether the use of elements from different sources causes an effect on the further progression of the conversation. If a practice can be tied to a concrete effect, one can infer that language users perceived the juxtaposition as novel, and classify the practice as switching rather than mixing.

In terms of assessing the significance of multilingual practice for evaluating plans and intentions, insights from Auer’s and Silverstein’s work provide a principled way for outside observers (such as language analysts) to relate patterns of language use to the salience of group identity. Language use that exhibits an overall pattern of mixing suggests orientation toward one’s own group, or a group identity, while language use that exhibits an overall pattern of switching suggests orientation toward boundaries between groups.

Thus, being able to differentiate switching and mixing will provide different kinds of clues about the conversationalists as to whether they are members of a self-identifying group or not. The ability to discern this difference rests in understanding what ensues when one conversation participant engages in multilingual practice – and especially, but not exclusively, what other participants in the conversation do in response. In this interim report, we provide examples of the kinds of conversations that will provide clues as to the participants’ group identity.

**MIXING, NOT “BORROWING”**

The concept of mixing is useful because it suggests a way out of a difficulty many scholars (for example, Poplack, 1980[1979] and Myers-Scotton, 2003) have encountered in distinguishing switching from borrowing, or the use of loanwords.

Merriam-Webster defines “loanword” as “a word taken from another language and at least partly naturalized” (“Loanword,” n.d.). Approaches to multilingual practice that rely on the common-sense notion of borrowing, however, risk confusing language at the level of the individual user and language at the level of the group.

Since no two language users have the same vocabulary, a workable definition would need to distinguish evidence of borrowing at the individual level, such as integration into the mental lexicon, from evidence of borrowing at the group level, such as appearance in print or inclusion in dictionaries. The conclusion would be inescapable that use of a particular element might constitute borrowing for one person, but switching for another.

For example, some American English speakers know the German word *Gesundheit* (literally, ‘health’) as an alternative to *[God] bless you* to acknowledge another’s sneeze. But not all English speakers recognize the word, only some speakers use it, and no speakers use it habitually in the same contexts as English *health*. At the individual level, *Gesundheit* is integrated into the mental lexicons of some English speakers and not of others; at the group level, it is a widely recognized, but recognizably marginal, element of English vocabulary.
Therefore, from a single kind of evidence (such as one speaker’s language use, or a set of reference materials) one cannot render an accurate conclusion about the role of a pair of languages in characterizing a society generally; conversely, from a society-level characterization, we may not be able to characterize the degree of bilingual competence of a given individual (as noted above for Arabic in Somalia, for instance). This state of affairs is represented in the distinction between diglossia and bilingualism that we cite below.

Auer does not propose to demonstrate that integration has occurred; instead, he focuses on whether juxtapositions of linguistic elements from different sources appear novel to particular language users in particular conversations, based on how those juxtapositions affect the conversation itself. Whether the elements are integrated at the individual or the group level is irrelevant. The same element may be a switch or a mix to the same users at different points in the same conversation.

Returning to the example of Gesundheit above, Auer’s advice would be stop focusing on the German origins of the word (of which most speakers are only vaguely aware) or its integration at the individual and group levels (which are very difficult to demonstrate conclusively) and instead focus on what is empirically verifiable – the word’s concrete function at individual moments in conversation.

In terms of assessing the significance of multilingual practice for evaluating plans and intentions, Auer’s approach is useful because it is “bottom-up.” It constitutes a set of strategies for out-group observers to discover in-group members’ perspectives on language use and group identity, and provides an interpretive framework for observers to ground their characterizations empirically in specific examples of language use. A bottom-up approach is particularly useful when working with small, marginal groups, since top-down, country or society-level treatments of language use and group identity are not likely to pertain to them. In short, Auer’s approach provides a way for out-group observers to discover and align with the in-group perspective.

**IMPLICATIONS FOR DATA COLLECTION**

The approach to multilingual practice that we have adopted has important implications for data collection. The first is that switching, and particularly mixing, are almost always marginal practices at the national level. Since switching owes its indexical creativity to its novelty, conventionalized switching (at least in Auer’s approach) is a contradiction in terms. As for mixing, once a set of multilingual practices becomes conventionalized and mainstream enough to be used in administration and taught in schools, people usually consider it to be a language in its own right. Stable bilingualism at the national scale is surprisingly rare.

Fishman’s useful (1967, p. 30) model crosses individual multilingualism (here, simply “bilingualism”) and societal multilingualism (following Ferguson [1959], “diglossia”)\(^3\) to sketch four potential scenarios: a nation may have (1) both bilingualism and diglossia; (2) diglossia without bilingualism; (3) bilingualism without diglossia; or (4) neither diglossia nor bilingualism.

Scenario (1) describes Paraguay with respect to Spanish and Guarani speakers; (2) describes developing countries, like India, in which an elite uses a former colonial language and the broader population uses one or more indigenous languages; (3) describes industrialized countries, like the U.S., in which new immigrants are multilingual, but their other languages have no established social role; and (4),

---

\(^3\) Ferguson defines “diglossia” by observing, “[i]n many speech communities two or more varieties of the same language are used by some speakers under different conditions” (1959:232). In adopting Ferguson’s term, Fishman expands it by noting, “diglossia exists not only in multilingual societies which officially recognize several ‘languages’ but, also, in societies which are multilingual in the sense that they employ separate dialects, registers or functionally differentiated language varieties of whatever kind” (1967:30). We note that speakers of Arabic often control multiple varieties and that this can be identified as diglossia in both Ferguson’s (1959) narrow sense and Fishman’s (1967) expanded sense.
while theoretically possible, can exist only in societies small enough to maintain linguistic homogeneity and isolated enough to avoid contact with other societies. Fishman notes that scenario (1), in which social and individual bilingualism co-exist, is rare at the national level, though relatively frequent at the subnational level.

With respect to the target languages in this study, to inquire about the multilingual practices of speakers of Arabic, Persian, and Somali in general is to pose the question the wrong way. Even though many users of Somali and Persian are multilingual, they live in a number of different countries, and they speak a number of different languages. More critically for language analysts, this means that nothing on a global sociopolitical scale can be concluded about the insertion of Arabic expressions, or even stable bilingualism with Arabic, of speakers of other languages. Instead, it is necessary to examine mixing and switching within specific groups.

As we proceed with the study on Persian-Arabic multilingual practice, we will be able to compare and contrast the significance of Arabic-language use in the two communities under study. While our description of specific practices in this report may not be generalizable, our approach and methods should be generalizable not only to other groups, but also to other language combinations.

The second implication is that switching and mixing – at least as they are relevant to evaluating plans and intentions – are spontaneous practices not always subject to speakers’ conscious control or reflection. That is, speakers switch and mix for functional communicative purposes, without being consciously aware that they are mixing and switching languages and dialects from the linguist’s perspective. Speakers are communicating, or at least attempting to; linguists are analyzing structure and function. This means that we cannot directly elicit these practices without rendering the data useless for our purposes. What we can do is to work with existing groups in which we believe (based on observations or self-reporting) that switching and mixing are likely to occur, create the conditions for relatively natural conversation, and record enough speech to capture spontaneous occurrences as they occur.

IMPLICATIONS FOR AUTOMATIC LANGUAGE IDENTIFICATION

A third implication of our approach to multilingual practice is that speakers and hearers will not always agree whether a particular element “belongs” to one language or another. This is the case even in standardized languages like English and French. For example, some English speakers pronounce the word hors d’oeuvre as if it were French, while others nativize it to the English sound system. The word also has a slightly narrower range of meaning in English, where it often refers to food served outside the context of a meal, than in French, where it also refers to food served at the start of a meal, for which English speakers often use appetizer or starter. Based on these facts, speakers could make the case either that hors d’oeuvre is French word that English speakers happen to use, or that it is a French word that has been borrowed into English. This issue is even more pronounced in unstandardized languages, particularly those with a long history of contact with other languages.

Consider excerpt 1, an example from our Somali/Arabic corpus in which the speakers are talking about a child’s eating habits. (We discuss the process by which we collected, transcribed, annotated, and analyzed the data in detail below; with respect to this example, it is important to note that our native speaker consultant produced the original transcriptions and translations, which we have left unedited. In this example, the consultant identified unformatted text as Arabic and underlined text as Somali.)

Excerpt 1: Inconsistent evaluation of language affiliation

063014b, 03:58

أيوه، بس يخيلها تأكل لوحدها ساعاتً.1

© 2014 University of Maryland. All rights reserved. December 2014 7
Yeah, yeah, but I also let her eat by herself on occasions.

2. A: ﻓﻮل kaligeeda aa cuni kartaa
   She eats beans by herself.

3. B: ﺑﻌﯿﺸ؟
   With pita bread?

4. A: ﺑﯿﺾ
   Eggs-

5. A: ﺑﻲ ﺑﯿﺾ kaligeeda aa cuni kartaa.
   Yeah, she eats eggs by herself.

6. A: ﺑﺲ foolka mug soo qaadi kartoo.
   But she can’t eat with beans.

What follows is a full interlinear analysis of lines 2, 5, and 6. (Note that the second author, a Somali linguist, slightly edited our native speaker consultant’s translation to match the interlinearization more closely.⁴)

063014b, 03:58, Interlinearization

2. ﻓﻮل beans fuul kaligeeda aa cuni kartaa
   one-K-3SG.F.PSS eat-INFIN be.able-F-PRES
   ‘She can eat beans by herself.’

5. ﺑﯿﺾ eggs bayd kaligeeda aa cuni kartaa
   one-K-3SG.F.PSS eat-INFIN be.able-F-PRES
   ‘She can eat eggs by herself.’

6. ﺑﺲ foolka mug soo qaadi kartoo
   one-K-3SG.F.PSS bring-INFIN be.able-F-NEG
   ‘But she can’t eat with beans.’

In this example, the Arabic word ﻓﻮل ‘beans’ occurs in lines 2 and 6. In line 2, our consultant transcribed the word in Arabic script and labeled it as an Arabic word. In line 6, however, the word appears with the Somali definite article –ka; in this case, our consultant transcribed the word in Somali orthography and labeled it as Somali. He then made a note in the transcription: “Sometimes we tend to treat Arabic words as Somali words. For instance the word [fiuul] ﻓﻮل, which is Arabic but in this segment the speaker used it as a Somali word.” The consultant intuitively recognized mixing at the level below the

---
⁴ Glossing conventions for the interlinearized text: DEF = definite determiner, F = feminine; INFIN = infinitive; K = K grammatical gender; NEG = negative; POSS = possessive; REL = relative; SG = singular.
word, where the stem originates in a language other than the language from which the inflecting affixes originate.

This phenomenon poses a challenge for development of automatic language identification – that is, a computational solution that at one level can identify the language of a document, and at a second level also identify “other language” words or phrases within the document.

While until recently most work on automatic language identification has functioned on the text level and focused on standardized languages, more recent work has tackled identification at the word or phrase level and turned to less standardized languages that are involved in switching and mixing.

We note above that both Somali and Persian have a long history of contact with Arabic, and share many lexical items with it; therefore, speakers of these languages are particularly likely to vary in their evaluations of which items “belong” to which languages.

Automatic language identification has the potential to simplify processing by bringing the most relevant expertise and resources – such as electronic dictionaries or machine translation – to bear on individual elements of a text. It only works, however, to the extent it can model the evaluations of some user or group of users. To the extent evaluative diversity makes it necessary to develop a sampling strategy to assemble a representative sample of users, it poses an additional challenge to computational approaches.

CORPUS DEVELOPMENT AND ANNOTATION

During this period of performance, we collected the full data set foreseen by our protocol for the Somali/Arabic language pair, and a partial data set for Persian/Arabic. For Somali/Arabic this comes to six hours of recorded conversation involving three participants plus one consultant; for Persian/Arabic it comes to ninety minutes and two participants plus one consultant.

Because our approach to multilingual practice implies that data collection should occur within existing groups, for each language pair we began by recruiting a native-speaker consultant who was likely to engage in mixing and switching, and asked him to recruit further participants from among his close associates.

For Somali/Arabic, the consultant is a native Somali-speaking graduate student at a university in the Midwest who received much of his education in Egypt. Participants are family members with similar backgrounds. For Persian/Arabic, the consultant is a native Persian-speaking graduate student at a university on the East Coast who previously lived in the Middle East. Participants are friends with similar backgrounds.

The consultants scheduled and coordinated the recording sessions. In order to facilitate relaxed, naturalistic conversation, we encouraged the consultants to host sessions wherever the participants found most convenient. For Somali/Arabic, a member of the research team traveled to the consultant’s home in the Midwest; for Persian/Arabic, he traveled to the consultant’s university.

The team member began by introducing the study and securing informed consent. He then set up a hand-held digital audio recorder. Since participants had the option to stop recording at any time, he made certain the consultant and participants understood how to operate the recorder. Then, the team member started the recorder and withdrew from the room. It was important that the team member not be present, so that participants would not accommodate him by switching to English; however, he remained nearby in order to resolve any problems that might arise.

Since our protocol limited sessions to a maximum of ninety minutes, at the end of the allotted time the team member returned to the room, stopped the recorder, and checked in with the participants. At the end of each day, the team member compensated each participant $30 per hour (or pro-rated fraction) in cash.
For Somali/Arabic, after returning from the field we re-packaged the digital recordings on a DVD and returned them to the consultant by mail. We also provided the consultant with detailed instructions (appendix C) to transcribe and annotate the conversation using ELAN (https://tla.mpi.nl/tools/tla-tools/elan/), an open-source software package available from the Max Planck Institute for Psycholinguistics, Nijmegen, which allows the user to time-align multi-tiered, annotated transcription with audio (and/or video) files (viz. Lausburg & Sloetjes, 2009). The consultant provided a sample transcription of about two minutes for feedback. Once we were satisfied that the consultant understood the task, we asked him to choose three stretches of conversation, each ten minutes in length, for full transcription.

The consultant produced a time-aligned, verbatim transcription in native orthography using a simple protocol designed to capture some discourse phenomena such as pauses, hesitations, false starts, truncations, corrections, fillers, and back-channel responses. The consultant segmented the discourse into intonation units, then sub-segmented intonation units by perceived language, allowing us to see his evaluation of whether each “chunk” was in Somali, Arabic, or, as frequently turned out to be the case, English. The consultant also provided a free translation of each “chunk” into English (sample at appendix B). During transcription, the consultant identified and anonymized all personal names. After transcription was complete, we effaced the personal name with a “beep” sound over the corresponding audio segments.

Figure 1 illustrates the proportion of “chunks” the consultant assigned to each language. Somewhat surprisingly, although all of the participants identified ethnically as Somali, the consultant classified 57 percent of their discourse as Arabic, 35 percent as English, and only 8 percent as Somali.

Figure 2 breaks down these proportions down by speaker. Although the approach to multilingual practice we have outlined focuses on interactive speech in group settings, it is important to recognize individual differences between group members, as well: all group members use more Arabic than either Somali or English, but Speaker SOM03f uses very little Somali (only two utterances), while Somali makes up 13% of Speaker SOM01m’s speech.
The team coded the finished transcripts using the rubric at appendix A. Following Auer’s approach to multilingual practice, we coded for discourse phenomena that have been suggested in the literature to serve as contextualization cues in both monolingual and multilingual discourse. We then triangulated the occurrence of these discourse features with the transcriber’s coding of language affiliation and our own evaluation of language affiliation. We classified as switching those multilingual practices that appeared to serve as contextualization cues, and as mixing those that did not.

FINDINGS

In this section we present findings from our analysis of the Somali/Arabic corpus. We contrast an extended example of mixing, in which language users appear to find the juxtaposition of elements from different languages unremarkable, with a number of examples of switching, in which the structure of the conversation allows us to infer that languages users attend to the switch and use it to contextualize speech. (As noted above, we anticipate completing data collection and conducting analysis of the Persian/Arabic corpus the second year of this two-year project.)

It is important to underscore that the scope of this contextualization in most cases is limited to the context of the conversation. In keeping with Auer’s bottom-up approach, we have refrained from imposing interpretations of the social meaning of language drawn from society-level descriptions of Arabic- or Somali-speaking countries, of which these speakers are far from typical. Instead, we focus on assessing the potential of multilingual practice for evaluating plans and intentions at the individual and group scales.

Excerpt 2 illustrates an example of mixing, in which language users do not appear to attend to the juxtaposition of elements which our consultant retrospectively identified as belonging to Arabic, Somali, and English. (In this and the examples to follow, the consultant identified underlined text as Somali, bolded text as English, and unformatted text as Arabic.)
Excerpt 2: Mixing of Arabic, Somali, and English

063014b, 12:32

1. А: آه ده فيتامين
    Yes, this is vitamin C.

2. Б: ده يناع إيه؟
    What is it for?

3. А: آآ- للوش اه. افكرته بتخلطه مع-
    For the face, yes, I thought you mix it with-

4. Б: A- a- habartaa markastoo
    Your mother always buys strange stuff, I don’t know.

5. А: إيه؟
    What?

6. Б: Habartaa-
    Your mother-

7. А: بس ده- ده فعلاً كوبس. بس الأحصين اللي هو الفيتامين:
    But this- this is actually good. But what's better is the vitamin C which-

8. Б: Runta fiiri fiiri.
    Look at the truth.

9. А: مستخلص من الأكل
    Extracted from food,

10. Б: Runta fiiri
    هي كل مرة هتخلط دوه في دوه
    Look at the truth every time she mixes this with that.

11. Б: أخر مرة رحت هناك
    Last time I went there she sent me

12. Б: turmeric iyo
    عصير ليمون
    turmeric and lemon juice.

In lines 1-3, speakers A and B begin in Arabic. In line 4, speaker B moves to Somali, but returns to Arabic (between the Somali noun markastoo and the following Arabic verb phrase). In line 5, speaker A asks for clarification from speaker B in Arabic. In line 6, speaker B starts to respond in Somali, but in line 7 speaker A continues on in Arabic, briefly inserting a Somali conjunction before being cut off. In line 8 speaker B interrupts in Somali, but in line 9 speaker A continues in Arabic. In line 10 speaker B repeats the Somali interjection from line 8, but then returns to Arabic. In line 11 speaker B continues in Arabic, then uses a Somali verb phrase. Finally, in line 12 speaker B uses the English word turmeric, the Somali conjunction ‘and’ and the Arabic phrase for ‘lemon juice’.
We will not analyze the following excerpts in quite such detail. The point in doing so here is to illustrate that the speakers appear to be juxtaposing elements from different languages without particularly attending to them. Taken as a whole, it is difficult to say whether this is primarily an Arabic discourse with insertions in Somali, or a Somali discourse with insertions in Arabic. Auer (1999) points out that this lack of a clear frame or “matrix” language (see Myers-Scotton, 1993) is characteristic of mixing, rather than switching.

That said, speaker B’s repetition of the Somali expression ‘look at the truth’ in lines 8 and 10 suggests a degree of attention to the linguistic form. Speaker B seems to be mobilizing the contrast between this formulaic expression in Somali and speaker A’s Arabic statements in lines 7 and 9 in order to capture the conversational floor and press the point. This suggests switching, rather than mixing. In addition, speaker B’s use of the English word *turmeric* in line 12 appears to be a lexical insertion. While it is not possible to identify Arabic or Somali as a matrix language, it is possible to contrast the Arabic-Somali mix as a matrix with respect to the English switch, with which the speaker fills a lexical gap.

As this example demonstrates, our approach requires close attention to the function of multilingual practice primarily at the local, turn-by-turn level of discourse. It is unlikely that any multilingual discourse will consist only of switching or mixing: a general pattern of mixing may include incidents of switching, and vice-versa.

We will now look at more typical examples of switching, which we identify by focusing on discourse phenomena that act as contextualization cues in both monolingual and multilingual discourse. The conversational properties we exemplify here are: confirmation; correction; topic shift; negotiating the floor; and constructed discourse.

As we noted in our analysis of excerpt 2, repetition is one discourse phenomenon that speakers can use to direct attention to linguistic form. In excerpt 3, the speakers are discussing bank fees.

**Excerpt 3: Confirmation**

063014b, 07:34

1. B: قال الرجل there's a standard $5 charge
   The bank teller told me there's a standard $5 charge.

2. لكل مرة؟ كله مرة؟ Every time?
3. B: In addition to conversion like-
   In addition to conversion like-
4. B: Every time-
   Every time-
5. كل مرة؟ A: كل مرة؟ Every time?
6. B: مرة كل
   Every time.

Note the pattern of repetition of the phrase ‘every time.’ Speaker A first uses the phrase in Arabic in the information question in line 2. Speaker B answers with the English equivalent in line 4. But speaker
A repeats the phrase in Arabic in a confirmation question in line 5. Finally, speaker B answers the question in line 6 by repeating the phrase in Arabic.

It is worth noting that all of the speakers in this group are highly proficient in English and Arabic; it is very unlikely that speaker A simply did not understand. The fact that speaker A posed the question in line 2 in Arabic, rather than English, may not be significant – a case of mixing rather than switching. By line 5, however, speaker A seems to leverage the contrast between English and Arabic, metaphorically framing the confirmation question as a translation in order to express incredulity and outrage about the bank’s exorbitant fees.

Incidentally, in line 1 speaker B also uses English to set off the bank teller’s statement ‘there’s a standard $5 charge’ in constructed (or reported) discourse. We will consider constructed discourse in more detail in excerpt 7; at this point, we should stress that research on multilingual groups (for example, Dorian’s [1997] study among bilingual users of Scottish Gaelic and English) has not found switching in constructed discourse to reliably reproduce the original language of discourse – although it almost certainly does so in this example.

A second discourse phenomenon that can function as a contextualization cue is correction, either of oneself or others. In excerpt 4, the speakers are discussing the color of an iPod.

Excerpt 4: Correction

070114a, 02:24

1. C: أیواة silver aas ahay
   Yes, it was silver.

2. B: Silver?
   Silver?

3. A: رمادي؟ Gray?

4. A: بجد؟ Really?

5. C: أیواة silver aas ahay.
   Yes, it was silver.

6. B: آل- ال- الكبير كان بيحي <cough>
   The big one comes with- <cough>

7. C: Silver aas-
   It was silver.

8. A: It was silver back.
   It was silver back.

9. C: كان جوء case silver
   It was silver inside a case.
10. B: I think أبيض أبيض
   I think white, white.

11. B: ال- بتاعك إنت كان أسود أنا بتاعي كان أبيض:
   The- yours was black, mine was white.

12. A: خلاص من الآخر كان iPod
   Let’s just end it, it was an iPod.

Speaker C insists that the iPod was silver, repeating the phrase ‘yes, it was silver’ (mixing Arabic, Somali, and English) fully in lines 1 and 5, and partially in line 7. Speakers A and B seem to be skeptical either about the color itself, or the distribution of the colors on the case. Finally, in line 8, speaker A cuts speaker C off, switching into English to clarify ‘it was silver back.’ Speaker C takes this up to agree ‘it was silver inside a case’ (mixing English and Arabic). Speaker A’s brief switch into English serves metaphorically to redirect the discourse away from a direction that speaker A seems to find unproductive or boring, and back to the main thread of the narrative.

Switching also has the potential to serve as a contextualization cue by organizing discourse, for example by signaling a change in topic. In excerpt 5, the speakers are discussing their health and their children.

**Excerpt 5: Topic shift**

070114a, 07:37
1. A: haddii aad u noogganahay.
   If I’m very tired
2. A: Or xanuunsanahay.
   Or if I’m in pain.
   < 4 second pause >
3. A: ﻫﻮ إﻧﺖ ﺧﻮﻓﺖ اﻟﻤﻔﻌﻮﺻﺔ ﺑﺘﺎﻋﺘﻲ ﻣﻦ ﯾﻮﻣﯿﻦ؟
   Did you scare my troublemaker two days ago?
4. B: ﺍﻹﺳﺒﻮع ﺍﻟﻠﻲ ﻓﺎت
   Last week
5. A: ﺑﺣﺎجة ماشية ﻓﻲ ﺳﻔлага ﻻ ﻋنكوسة ماشية ﻓﻲ ﺳﻔлага؟
   With something on the wall or a spider on the wall?

In lines 1-2, speaker A finishes a discourse in Somali (mixed with English) describing the insomnia he or she has recently suffered. The speakers pause for four seconds. Then in line 3, speaker A initiates a new topic with a question about his or her child. The speaker leverages the contrast between Somali and Arabic to metaphorically mark the shift between topics.

A second way that switching can organize discourse is by signaling participants’ desire to take or hold the conversational floor. In excerpt 6, three speakers are speaking simultaneously, each trying to make a contribution. (In lines 2-6, we have spaced the start of each line to indicate overlapping speech.)
Excerpt 6: Negotiating the floor

063014b, 10:46

1. A: ﻫﻄﯿﻨﮭﺎ و ﻫﻲ ﻛﺪھ ﺑﺘﺼﺮخ و ﺑﺘﻘﻮل ﻛﺎن ﻓﻲ ﺣﺎﺟﺔ ﻓﻲ اﻟﺤﯿﻄﺔ
   We did but still she kept yelling and saying there's something on the wall.

2. A: ﺑﺲ
   Just-

3. B: و و دﻟﻮﻗﺘﻲ إﻣﺒﺎرح
   a- and yesterday-

4. A: fa-

5. C: It's another horror movie.
   It’s another horror movie.

6. A: no. no. no.
   No. No. No.

7. A: وﺑﻌﺪ ﻛﺪھ ﻋﺎرف ال takabaan
   And after that, you know the hanger you hang clothes on.
   ﺩھ اﻟﺒﯿﺤﻄﮫ ﻓﯿﮫ dharka la saasaarayo,

In lines 1-2, speaker A is in the midst of a discourse in Arabic about a child’s reaction to an insect on the wall (also alluded to in excerpt 5). In line 3, speaker B interrupts in Arabic. In line 4, speaker A attempts to recover the floor with a contribution in Somali, but is again interrupted, in line 5, by speaker C, who interrupts in English. Finally, in lines 6-7, speaker A recaptures the floor, rejects the interruptions with an explicit metalinguistic commentary (‘no. no. no’ in English), then resumes the discourse from line 1 in Arabic (mixed with Somali).

Throughout the negotiation, each speaker uses switching in order to metaphorically frame his or her “bid” for a new speaker and a new topic. Language choice does not seem to have this function at the beginning of the excerpt; however, by speaker C’s switch into English in line 5, the contrast between English and Arabic has come to symbolize (for the moment) the contrast between competing speakers and their topics.

A third way in which switching can organize discourse is by serving as a resource to set off constructed discourse. In excerpt 7 speaker A is describing an encounter with a customer service representative.

Excerpt 7: Constructed discourse

070114a, 04:08

1. A: then- SOME_ONE upgrade. ﺑﻮھ وأوھ upgrade, yes.
   Then at that moment I think you get an upgrade, yes.
2. A: هل سألوني قالولي ماذا تود تجربة ترقية مع T-Mobile؟
   They asked me, “What do you qualify for an upgrade with T-Mobile?”

3. A: لا يمكنني شراء ترقية لا يوجد ترقية.
   I said no, because T-Mobile doesn’t offer an upgrade, you have to buy it.

In line 2, speaker A sets off the constructed quote from the representative by switching into English. As noted above, switching in constructed discourse has not been found reliably to reproduce the original language of discourse; indeed, in line 3 speaker A quotes himself or herself replying in Arabic, even though the exchange almost certainly occurred in English. Instead, switching serves the more abstract purpose of leveraging the contrast between languages to metaphorically construct multiple characters in a single narrative.

PRELIMINARY CONCLUSIONS

In excerpts 2-6 we examined different ways in which the speakers in the Somali/Arabic group use the languages they speak (including English) as contextualization cues, including cues that organize and structure discourse. It is important to underscore that the phenomena we have looked at are conversational universals that monolingual speakers achieve effortlessly using cues other than mixing, such as intonation, discourse marking, or metalinguistic commentary. However, multilingual speakers have access to a broader selection of resources; moreover, they have the option to combine resources in novel ways – through what we have called “metaphorical switching” or “indexical creativity” – and leverage the novelty of the combination to achieve deeper, or parallel, contextualizations.

At the same time, the effects that the speakers in this group achieve are almost purely local to the context of discourse. When in excerpt 6, for example, speaker A uses a switch between English and Arabic to recover the floor, the meaning of that configuration of languages is dependent solely on speaker C’s use of English to interrupt seconds before, and it probably vanished from speakers’ minds seconds later. It would be unjustified – absurd, in fact – to impose an interpretation based on a country or society-level description of the social meanings of English and Arabic, say, based on the history of British imperialism in the Middle East. A bottom-up approach ensures that interpretations are justified by empirical linguistic data and relevant to the particular groups under analysis.

In terms of the significance of multilingual practice for evaluating plans and intentions, we hesitate to characterize this group based on our preliminary analysis of only a small amount of talk. We anticipate that an analyst would need to engage intensively with a group over a period of months to reach reliable conclusions.

Nevertheless, based on the prevalence of switching even in passages like excerpt 2 that are generally characterized by mixing, we feel confident that the members of the group are aware enough of the difference among the languages in their repertoire to reliably leverage the contrast among them. This is consistent with what we know about the group – that is, this is a small family living, working, and studying in a foreign environment.

Generalizing beyond this particular case, we thus expect that an analyst trained in our approach to multilingual practice will be able to contrast instances of switching and mixing in other language sets, and be able to characterize groups of language users in terms of one or the other. Our framework predicts that groups characterized by switching will orient toward boundaries between groups, while groups characterized by mixing will orient toward the in-group. The first scenario might, for example, describe a mobile group of expatriates working in a foreign environment; the second might describe a settled group whose bilingualism is nevertheless marginal to the society in which they live.
To draw cross-group conclusions about the kinds of significance associated with code-switching would require examples from a larger sample of multilingual groups of speakers and from a wider range of sociocultural contexts. A wider empirical base of multilingual events will provide a larger set of conversation types that in turn may provide clues as to whether the participants are presuming a group identity or are underscoring differences. Ultimately, this information can be fed back into a code-centered phase of analysis, as patterns of multilingual practice may prove diagnostic of an individual’s membership in a group of particular interest.

**IMPLICATIONS FOR TRAINING AND OPERATIONS**

As we indicated at the outset, our goal is to provide analysts with an exemplified approach that will allow them reliably to identify patterns of switching and mixing, use these patterns to characterize conversations in a target group, then use these characterizations to draw reliable inferences about group members’ relationships with each other and their orientation to the group.

Let us return to the example of correction in excerpt 4. Disagreement between speakers B and C about the color and pattern of the iPod creates what speaker A, at least, perceives as trouble in the conversation. The speaker resolves it with the statement ‘it was silver back’ that, in terms of content, bridges competing positions. This conversational move might have been equally effective in any language; by using English, however, speaker A leverages the formal difference between Arabic and English to “cap” the conversation, seizing the floor to shift the topic in a direction speaker A finds more interesting or productive.

Analysts can say several things about the speakers in excerpt 4. At the most basic level, because they use Arabic, Somali, and English, analysts can infer that the speakers all have some familiarity with those languages. By analyzing the structure of the conversation, however, they can also demonstrate that speaker A is asserting control. If this phenomenon re-occurs across conversations, they might hypothesize that speaker A plays a leadership role in the group.

Based on the way that speaker A asserts control by switching between Somali and Arabic, analysts can also demonstrate that the speakers are attending to the differences between the languages, and infer that they are attending primarily to boundaries between groups. If this pattern of switching holds across conversations, analysts might hypothesize that the speakers constitute a mobile group of expatriates.

The example of mixing in excerpt 2 would also allow analysts to infer that the speakers are familiar with Arabic, Somali, and English. By contrast, however, since in this excerpt language users do not leverage the differences between those languages, analysts should not make the unwarranted inference that the speakers attend to them. If this pattern of mixing holds across conversations, analysts might hypothesize that the speakers constitute a settled minority group for which mixing is emblematic of group identity. They could then start looking for evidence that other speakers who use the same pattern belong to the same group.

We believe that the best language analysts already have an intuitive sense of how patterns of multilingual practice distinguish individuals and groups. CASL research provides a framework for analysts to articulate and replicate their intuitions in a way that can support intelligence conclusions, and for managers to leverage their insights through training and professional development.

With respect to operations, the approach promotes adoption of in-group perspectives, and enables rational inference of the social characteristics of groups, grounded in the empirical details of spoken language. With respect to training, the model can be taught both in a language-specific way, as part of existing language courses, and in a language-general way – even with respect to monolingual English conversation – as part of more general training in conversational structure and intelligence analysis.

Looking beyond this study, the corpora that we have developed for Somali/Arabic and that we plan to develop for Persian/Arabic have other potential applications. With respect to operations, the
corpora can be adapted in order to train and test models for automatic language identification at the word or phrase level. And with respect to training, they can easily be used to demonstrate the realities of switching and mixing in “real life” language use.
REFERENCES


_Corresponding Author and Reprints:_ Brook Hefright, PhD, University of Maryland Center for Advanced Study of Language, (301) 226-9148, bhefright@casl.umd.edu, www.casl.umd.edu.

_Acknowledgements:_ We acknowledge the scholarly contributions of Dr. Jennifer Hill Boutz, Dr. Corey Miller, Dr. Paul Rodrigues, and Mark Vinson.

_Funding/Support:_ This material is based upon work supported, in whole or in part, with funding from the United States Government. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the University of Maryland, College Park and/or any agency or entity of the United States Government. Nothing in this report is intended to be and shall not be treated or construed as an endorsement or recommendation by the University of Maryland, United States Government, or the authors of the product, process, or service that is the subject of this report. No one may use any information contained or based on this report in advertisements or promotional materials related to any company product, process, or service or in support of other commercial purposes. This report is not Releasable to the Defense Technical Information Center per DoD Directive 3200.12. The Contracting Officer's Representative for this project is John Walker, Government Technical Director at CASL, (301) 226-8900, USG_TD@casl.umd.edu.

© 2014 University of Maryland. All rights reserved.
## APPENDIX A. ANNOTATION SCHEMA

<table>
<thead>
<tr>
<th>Potential scope</th>
<th>Phenomenon</th>
<th>Code</th>
<th>Diagnostics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linguistic form within turn</td>
<td>Disfluency</td>
<td>D</td>
<td>See transcription for truncated forms</td>
</tr>
<tr>
<td></td>
<td>Discourse markers</td>
<td>X</td>
<td>See translation</td>
</tr>
<tr>
<td></td>
<td>Correction (of self or other)</td>
<td>CS, CO</td>
<td>See translation and speaker field</td>
</tr>
<tr>
<td></td>
<td>Repetition (of self or other)</td>
<td>RS, RO</td>
<td>See transcription and translation for repeated forms and content, compare with speaker field</td>
</tr>
<tr>
<td></td>
<td>Metalinguistic comment (on self or other)</td>
<td>MS, MO</td>
<td>See translation and speaker field</td>
</tr>
<tr>
<td>Discourse structure across turns</td>
<td>Topic shift (by same, by other)</td>
<td>TS, TO</td>
<td>See translation and speaker field</td>
</tr>
<tr>
<td></td>
<td>Floor negotiation</td>
<td>F</td>
<td>Compare language and speaker field</td>
</tr>
<tr>
<td></td>
<td>Language negotiation</td>
<td>L</td>
<td>Compare language and speaker field</td>
</tr>
</tbody>
</table>
APPENDIX B. SAMPLE ELAN TRANSCRIPTIONS

<table>
<thead>
<tr>
<th>Grid</th>
<th>Text</th>
<th>Subtitles</th>
<th>Lexicon</th>
<th>Audio Recognizer</th>
<th>Metadata</th>
<th>Controls</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker A</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:28</td>
<td>baashiga weeye maalaynaa, baa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker A</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td>it's just the-the-probably from the hunger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker A</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td>ara</td>
<td>som</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker B</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker B</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td>no no. but—but</td>
<td>probably the</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker B</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td>ara</td>
<td></td>
<td>ara</td>
<td></td>
<td>ara</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker C</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td>so you slept before me and woke up all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Speaker C</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td>ara</td>
<td></td>
<td></td>
<td></td>
<td>ara</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Discourse markers</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Correction</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Repetition</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td>RS</td>
<td></td>
<td>RO</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Clarification</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Metalinguistic cont.</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Topic shifting</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Floor negotiation</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Code</th>
<th>Language negotiated</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>00:09:36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C. TRAINING MATERIALS FOR TRANSCRIPTION OF NATURAL CONVERSATION

In order to do analysis of the conversation recordings, we’ll need to have written transcriptions of what people are saying. For everything that is said, we need a transcript of what was said (in its original language) as well as a translation into English. Because it takes so long to transcribe natural conversation, we won’t be transcribing everything that we recorded; instead we’ll work with you to select portions of the conversations to transcribe.

This document will walk you through the basics of transcribing natural conversation and how to use a software program (ELAN) to input your transcriptions and align them with the audio recordings. Hopefully this document will give you the information you need to get started, but we are also happy to walk you through the process over the phone. If at any point along the way you’re not sure what to do or how to do it, please feel free to email or call so that we can help you get back on track. Don’t waste a lot of time trying to troubleshoot. We may already know how to fix the problem and can probably save you some time.

Initially, we would like you to send us approximately 2 minutes of transcribed data so that we can look it over provide you with feedback before you do the remainder of the transcription. If you’ve never transcribed natural conversation before, you may find that it is more difficult than you expected. Unlike movie or television scripts, in natural conversation, people may speak at the same time, use incomplete sentences, or be difficult to understand. Therefore it may be necessary for you to listen to a stretch of speech several times in order to transcribe it.

When you are transcribing, we would like you to break the speech down into what we call intonation units—these are natural breaking points in speech (that sometimes correspond to sentences) where there might be a pause or a change in pitch. Don’t worry too much about exactly where to put the break, but chunking speech in this way will make it easier for you to transcribe.

It is very important that we protect the privacy of the people participating in the conversations. Therefore when you are listening to the recordings, make a careful note of the timestamp anytime you hear someone use a personal name. That will allow us to go through and “bleep” out the name in order to protect the privacy of the participants. In the transcript itself, you should refer to participants as “Speaker A,” “Speaker B,” etc. Please do not use people’s real names in the transcript.

You should use the following conventions when you transcribe. (Note that the examples below use English to illustrate what we mean; of course, you’ll be transcribing in a different language):

- Use standard spelling conventions, where possible. Please let us know if you would like us to provide you with a copy of a dictionary. It’s more important, however, that you be consistent in the way you spell words.
- For the Persian-Arabic corpus, all transcription should be done using Arabic script.
- For the Somali-Arabic corpus, transcribe the Somali using standard Somali writing (for example, “af-Somaali”). You can transcribe the Arabic portions in whichever way you feel most comfortable (using either Arabic script or a romanized Arabic alphabet).
- Each intonation unit will correspond with an “annotation” in the software program ELAN. (See below for instructions on installing and using ELAN).
- If two speakers are speaking at the same time, you can transcribe what each speaker says in a different line in ELAN.
- If someone stops speaking in the middle of a word or a phrase, indicate that with a –
  
  o Maybe he’s –
• I was running down the street

• Please be sure to transcribe words such as “um,” “oh,” “uh-huh,” etc. Please also try to use a consistent spelling with these words.

• You can transcribe relevant non-linguistic sounds in angled brackets < >

  o < laughter >
  o < cough >
  o etc.

• If someone uses a word you don’t know, indicate that in angled brackets as well. This may be particularly likely if they use an Arabic word you don’t know:

  o He was going to < Arabic word >

• If you don’t understand what someone was saying, indicate that in angled brackets:

  o He was going to < unintelligible >

• For personal names, please do not write the name out. Instead, simply write < personal name >. This will help us protect the privacy of our participants.

  o I saw < personal name > yesterday.

**Language Codes**

One of the tasks we’ll ask you to do is identify the languages that are being used throughout the conversation. Please use the following sets of ISO codes. (If a language is being used that’s not on this list, please let us know and we’ll let you know the code to use!)

<table>
<thead>
<tr>
<th>Language</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>ara</td>
</tr>
<tr>
<td>English</td>
<td>eng</td>
</tr>
<tr>
<td>Persian</td>
<td>pes</td>
</tr>
<tr>
<td>Somali</td>
<td>som</td>
</tr>
<tr>
<td>Swahili</td>
<td>swa</td>
</tr>
</tbody>
</table>

If a particular item is ambiguous for which language it belongs to, indicate that by writing “ambig” followed by the two codes, for example

• ambig ara-som
• ambig ara-pes
Using ELAN

ELAN is a computer program that is specifically designed for time-aligning transcriptions with recordings. We would prefer that you use ELAN to do your transcriptions. It takes a bit of time to learn, but it makes the transcription process much easier and is worth the time it takes to learn.

A. Download and install ELAN

ELAN is a free software developed by the Max Planck Institute. To download ELAN, go to the following link: http://tla.mpi.nl/tools/tla-tools/elan/.

Click “Download the latest version,” open the installer, and follow the instructions to download ELAN on your computer.

B. Starting a new file:

We will email you a template that will make starting a new file in ELAN much easier. There will be two different templates, depending on how many speakers are participating in your conversation:

- Code-SwitchingTemplate_2Speakers.etf
- Code-SwitchingTemplate_3Speakers.etf
Save whichever template file you need (or both, if some of your conversations have two participants and some have three) to your computer. Also make sure that the audio file you will be transcribing is saved to your computer. I recommend creating a folder specifically for the Code-Switching project work, and saving all the relevant files (ELAN template, audio files, and transcript files) there.

Once you have both the template file and the audio file saved to your computer, use the following steps to start a new ELAN transcript:

1. Open ELAN.
2. If a pop-up window opens, click “OK.”
3. To start a new file, go to File > New.
4. Select “Media” on the radio button in the center of the window (A). Then, in the menu on the left, navigate to your audio file—it should end in .wav (B). Click the double arrow (>>) to add your file to the menu on the right (C). Make sure that the file shows up in the “selected files” box on the right (D). Don’t click OK yet!
5. Now, select “Template” on the radio button in the center of the window (A). Then, in the menu on the left, navigate to the appropriate template—it should end in .etf (B). Click the double arrow (>>) to add the template to the menu on the right (C). Click OK (E).

6. You should see something that looks like this:
7. Go ahead and save your file right away. Go to File > Save. Please use a file name that is identical to the name of the audio file (but they will have different extensions). So, for example, if your audio file is named CS_SOM_061614a.wav, then your ELAN file should be saved as CS_SOM_061614a.eaf.

8. I also recommend that you turn on Automatic Backup. This will save your file automatically at a time interval that you set. That way if you forget to save and the program crashes or something, you don’t lose much work. Set the backup frequency to whatever you’d like—I usually choose 5 minutes.
C. What’s on the screen

Once you’ve started your new file, spend some time clicking around to orient yourself with the basic set-up of ELAN. Don’t worry too much about what’s in most of the menus. This guide will introduce you to the features which will be relevant for you.

ELAN has a number of different “modes” that you can use for transcription. Essentially, each mode just changes the way you’re looking at a file, and provides you with different options. ELAN defaults to **Annotation mode**, and we recommend that you stay in Annotation mode the entire time you’re working with ELAN. If you accidentally end up in a different mode, simply go to Options > Annotation Mode to return.

In Annotation mode, your screen should look something like this:
At the top of your screen, you should see some bars that allow you to control the **volume** (A) and **rate** (B) of the audio file. It can be helpful to slow the rate down sometimes when you’re doing transcription, but be careful about slowing it down too far, as it might distort the speech and make it even more difficult to understand.

Below the volume and rate bars is a set of buttons that allows you to pause and play the audio file. There are a lot of different options that allow you to do things like skipping back and forward by only a second, or to navigate from one annotation to another, etc. But I wouldn’t worry too much about all the options. I primarily use the **black play/pause button (A)**, which will play starting from wherever your cursor is (or pause, if the audio is already playing), and the **blue selection play/pause button (B)** which will play only what you have currently selected.
Below this is a bar that shows the waveform of your file. (Note that if you’re transcribing from an MP3, there won’t be a waveform.) The red line shows what part of the audio file you’re in. Using the waveform can be really helpful in transcription (more on that later). For now, just practice clicking on different parts of the waveform and playing and pausing the file.

Below this are a number of tiers. This is where you will type your transcription.

And here is an example with some annotations added so that you can see what your transcript will eventually look like.
If you look carefully at the above example, you’ll notice several things. First, ELAN will allow you to type in either a roman alphabet (like English and Somali both use) or Arabic script (like Arabic and Persian use). You’ll also notice that some of the speech is transcribed in Speaker A’s tier and some in Speaker B’s tier. This allows ELAN to distinguish between the different people talking. Look at the labels on the lefthand side. For each speaker, you should see four different tiers. For example, for Speaker A, you should see:

- **Speaker A**: This is where you will transcribe everything that Speaker A says.
- **Speaker A Free Translation**: This is where you will type an English translation of what Speaker A says.
- **Speaker A ISO Codes**: This is where you will type language codes for the language(s) Speaker A is using.
- **Speaker A Notes**: You can use this tier for anything you want. For example, you might want to write some notes for yourself if you’re having trouble, or you might want to remember to ask us some questions. You don’t have to use this tier if you don’t want to, but we are providing it just in case you find it helpful.

There is a duplicate set of tiers for Speaker B (and for Speaker C, if your conversation has three speakers).

**D. Create an Annotation**

The next step is to begin creating an annotation in ELAN. To do this:

1. Click on the waveform and select a chunk of speech. (Hint: you should be able to see pauses in the speech in the waveform—use these as annotation breaks.)
2. Double-click in the part of Speaker A’s tier where you want to type your text. A white text input box with a blinking cursor should appear.

3. Click either the black play button or the blue “play selection” button to listen to the speech. Once you know what is being said, type that in the white box. When you’re done hit ctrl + enter. **Important: Be sure to hit ctrl+enter; if you only hit enter, your annotation will not stick!!** This is one of the weird quirks of ELAN.

4. To add an English translation, click on your annotation in the first tier. You should see a purple column highlighting your selection. Then simply double-click in the “Free Translation” tier. Another white box will appear. Type in the free translation, and hit ctrl+enter to make it stick.
5. Follow the same steps to add the language ISO code in the next tier. (And if you have any notes to add, you can do that in the fourth tier.)

6. To check your annotation, click on the annotation to select it, then use the “play selection” button to play it back while you read what you’ve written. 

7. Note: see Section H below for instructions on how to add language codes if a speaker code-switches (changes languages) in the middle of a chunk of speech.
8. When you’re satisfied with your annotation, move on and continue creating annotations in the same manner.

9. Some people (including me!) prefer to go through and chunk up a long stretch of audio files into annotation chunks before actually adding any text. If this helps, feel free to do this. Sometimes looking at the breaks between waveforms can make setting up your annotation chunks easier. If you decide to chunk up your file before adding text, it will look something like this:

E. Editing an annotation you’ve already created

Sometimes it may be necessary to go back and edit an existing annotation. Perhaps you misspelled something or maybe you need to change the boundaries of the annotation. Either way, it’s easy to edit something that’s already there.
In the below example, I have an annotation, but the right edge of it is in the wrong place and chops off part of what the speaker is saying.

The easiest way to change the boundary of an annotation is to hold down ALT while you select the side of the annotation that needs to be changed. While holding down ALT, drag the boundary of the annotation to where you want it to be. As you drag, you should see the boundary lines turn bright green.

If you need to change what you’ve typed into the annotation, simply double click on the relevant annotation. The box will turn white and will allow you to edit the text that is there.

F. Using different scripts (writing systems)

ELAN allows you to type in both a Latin alphabet (which is used in English and Somali) and in Arabic script, as well as in several other writing systems that we won’t be using in this project. Switching to Arabic script is easy. Double-click within the annotation so that the input box turns white, then right-click and select your desired Arabic input method:
To switch back to a Latin alphabet, follow the same steps, but select “English” instead.

G. Transcribing with multiple speakers

Each of the conversations that you transcribe will have two or three speakers. Assign each speaker a letter (A, B, or C). Then transcribe on the appropriate line. Everything Speaker A says should be transcribed on Speaker A’s line, everything Speaker B says goes on Speaker B’s line, etc. Sometimes you may notice that people talk at the same time; when this happens, your annotations may overlap, as in the following example:

Notice that even though the annotations overlap, the translations and language codes should always line up with the relevant speaker line. So, for example, Speaker A’s translations and language codes line up with what Speaker A says.
H. Adding language codes if one chunk of speech contains multiple languages

On the “ISO language codes” tier, we would like you to add an annotation for the language that is being used by the speaker. (Remember! The language codes we would like you to use are listed on page 2.)

To add a language code, begin with a stretch of speech that already has an annotation in the Speaker tier. Select that annotation, and double-click on the “Speaker ISO Code” tier. A white text input box should appear. Enter the relevant ISO code, then hit ctrl + enter to save your annotation.

In some instances, a speaker may switch between languages within a single annotation. This is a little trickier to annotate. First, follow the same steps as above: select your annotation (the line in the middle of the annotation should turn blue when it’s selected) and in the “ISO Code” tier type the ISO code for the first language in the annotation. In this case, it was Somali (som). Hit ctrl+enter to make your annotation stick.

Now, right-click on the ISO code annotation you just made. In the menu that pops up, select “new annotation after.”
This should split your annotation into two pieces in the ISO tier. In the white text-input box, type in the language code for the second language that is being used and hit ctrl + enter.

Now you have two language codes, but you’ll notice that the annotation is split exactly in half. This means that the switch in the annotation probably doesn’t line up with the point in the speech where the speaker changes languages. In order to make the boundary of the language code annotation line up correctly, you’ll need to change the boundary between the two language codes. Follow the same steps as...
described above for editing annotation boundaries. (While holding down alt, click on the boundary between the two language codes and drag it to the proper place. The line in the language code tier should turn green while you’re moving the boundary.)

When you’ve done this, the language code annotations should be in their proper place and you will be able to hear and see where the speaker switches languages. In this example, the speaker is using Somali (som) and English (eng):

I. Save and send us your transcription
J. Hints and Tricks

ELAN takes a little bit of time to learn, and gradually you’ll get used to the program and how it works. But in the meantime, here’s a summary of some things I learned along the way that I wish I had realized at the beginning.

- ELAN shades the active tier salmon (and simultaneously bolds the name of the tier, underlines it, and turns it red). In this example, the “Speaker A” tier is active:

- If you’re trying to create a new annotation and ELAN won’t let you, first check to make sure that the proper tier is “active.” If it’s not, double-click on the name of the tier and that should activate it.

- If you select an annotation, the line in the middle of the tier should turn blue. Each of the tiers will also be shaded blue for the stretch of audio that that annotation corresponds with. If you’re having trouble editing an annotation, check to make sure that you’ve selected it (is the line blue?) This example has two annotations; the one on the left is the one that is selected.
• Any time you’re moving the boundaries of an annotation, the line down the middle will turn bright green. If it doesn’t turn bright green, you can’t move it. Maybe you forgot to hold down ALT or maybe you haven’t selected it properly.

• There’s a little bar between the audio playback buttons and the waveform that has a bunch of little hatchmarks on it. Each hatchmark represents an annotation. The little red I bar shows where you are at in the audio file, and you can drag it around to quickly navigate through the audio file. If you look at this example, you’ll notice that I’ve made a bunch of annotations in the first half of the audio file (there are lots of hatchmarks) but I haven’t made any in the second half of the file (no hatchmarks).

• When you’re transcribing, you can choose to use “loop mode.” This will play the chunk of audio connected with a given annotation over and over and over and over again until you hit “pause.” This can be useful because you don’t have to keep hitting “play” but it can also be incredibly annoying!
These arrows allow you to navigate quickly through your annotations. The right arrow will navigate to the next annotation on the same tier; the left one will go to the previous annotation on the same tier; the up arrow will go to the next tier above that has an annotation; and the down arrow will go to the next tier below with an annotation. You can completely ignore these arrows and just navigate around with the mouse, but a lot of people find navigating with the arrows to be useful.