Language in Society (CL2.203), IIIT Hyderabad

Assignment 1

Abhinav S Menon

1 Question 1

1.1 Lyons (1970)

All the people who use a given language (or dialect). (Lyons, John ed. New horizons in linguistics. No. 410 N4. (1970)) p. 326

This is a very simple and uncomplicated definition on the face of it, but closer inspection reveals a major ambiguity – what exactly is meant by *language* and *dialect*? The possibility of distinguishing a speech community by this definition hinges on a clear definition of languages and distinction between them.

For example, it is not clear whether Hindi and Urdu are different languages. Their day-to-day variants are mutually intelligible (sometimes called Hindustani), but their higher literary styles are frequently each incomprehensible to speakers of the other (due to unfamiliar vocabulary and constructions)[1].
If they are considered separate languages, they have distinct speech communities; that would then mean that speakers belonging to different speech communities can communicate with each other, which seems undesirable. Conversely, if they are considered to have the same speech community, speakers belonging to a single speech community might be unintelligible to each other, which also appears to be a problem.

If a language is considered to be a different entity from a dialect, another ambiguity arises from the use of the word *or*; are speech communities to be decided by language or dialect? This becomes especially tricky in places where the languages form dialect continua, like South India (specifically, Kerala). The colloquial Trivandrum dialect of Malayalam (the southern part of Kerala) differs widely from the colloquial Kannur dialect (the northern part of Kerala) and the two are not fully mutually intelligible. Further, dialects spoken in districts bordering Tamil Nadu (like Palakkad) or Karnataka (like Kasaragod) bear similarities to the languages of those states as well as to other dialects of Malayalam itself. This situation leads to much the same difficulties as the Hindi/Urdu controversy, only with more variants.

[1] King, Robert D. "The poisonous potency of script: Hindi and Urdu." (2001): 43-59.

This statement appears, therefore, to be more of a rough characterisation of the notion of a speech community than a precise definition (in the rigorous sense of the word). In my opinion, to make use of this definition, one must bring in the notion of mutual intelligibility rather than use vague terms like "language" and "dialect".

1.2 Hockett (1958)

Each language defines a speech community: the whole set of people who communicate with each other, either directly or indirectly, via the common language. (Hockett, Charles F. "A course in modern linguistics." (1958)) p. 8

Here, too, we find the ambiguity induced by the word "language"; but Hockett acknowledges this difficulty[2] and the resulting uncertainty in the extent of an actual speech community. However, another logical issue presents itself in an analysis of this definition.

Hockett appears to consider the phrase *via the common language* as determined by the speaker's idea of the language they are speaking. For example, he says that a document written in the Old English of 700 AD is still English (because that was what it was called even then), even though no modern speaker of English can make sense of it. If we pursue this course, we find that Hindi and Urdu unambiguously form two speech communities, as their speakers mostly identify with one or the other.
Similarly, in Kerala, all speakers who characterise their language as "Malayalam" must fall in the same speech community, regardless of whether they can understand each other or not; and speakers who call what they speak "Tamil" or "Kannada" form distinct speech communities, even though they may be understood by Malayalam speakers from Palakkad or Kasargod.

This definition takes us a little closer to a clear understanding of speech communities, but there are still some edge cases. It appears that mutual intelligibility is less of a criterion here than language *identity*.

1.3 Bloomfield (1933)

A speech community is a group of people who interact by means of speech. (Bloomfield, Leonard. "Language. 1933." New York: Holt (1933)) p. 42

This is the most simple and unambiguous of all six definitions, as it avoids the use of unclear terms like *language*. However, it is not free from issues – since it makes no reference to language, it could refer to absolutely any group of people who speak with each other. If we assume (reasonably) that *interact* means that all utterances are understood by the addressee, here again mutual intelligibility becomes the main criterion. However, since the language is not specified, the speech community can potentially contain speakers of any number of languages

[2]Hockett, Charles F. "A course in modern linguistics." (1958)

(since only a common speaker is needed to ensure that people in the group still "interact by means of speech").

For example, Hindi and Urdu may form separate speech communities under this definition, or a single one – the group of people who speak both Hindi and Urdu interact by means of speech with people who know only one or the other. Thus this definition allows us some freedom in determining the speech community.

It is to be noted, however, that in the same work[3] (p. 29), Bloomfield offers another description of a speech community – *a group of people who use the same system of speech-signals*. He goes on to say that all members must "speak intelligibly and understand what others say". Here we find confirmation that he intended mutual intelligibility to be the basis of his definition, but also intended it to only cover a single language. The phrase *system of speech-signals*, presumably, is intended to refer to a language (or possibly only its vocabulary), but leads to issues much like those in the previous definitions. No two speakers can be using *exactly* the same set of speech-signals, so to what extent do two systems of speech-signals have to coincide for them to be the "same"? If two speakers use exactly the same function words and inflections, but one says प्रधान मंत्री and the other वज़ीरे-आज़म (and other differences in vocabulary), are they using the same set of speech-signals?

In this definition, then, the author leans towards mutual intelligibility over language identity. He also (p. 45) acknowledges the infeasibility of determining the exact boundaries of speech communities.

1.4 Gumperz (1962)

Any human aggregate characterised by regular and frequent interaction by means of a shared body of verbal signs and set off from similar aggregates by significant differences in language use. (Gumperz, John J. "Linguistic and Social Interaction in Two Communities 1." American anthropologist 66.6_PART2 (1964): 137-153) p.137

This definition shares one characteristic with Bloomfield's – the term *verbal signs*, or what Bloomfield called *speech-signals*. There is also the extra stipulation of *regular and frequent interaction* – this might indicate that a more restricted idea is in Gumperz's mind than that of a set including all speakers of a given language. He also specifies how to distinguish two speech communities, viz., by *significant differences in language use*. This is a term that has not occurred in any of the previous definitions, and the word *significant* is extremely subjective. It is not clear exactly how different the usage of two communities must be to make them two different speech-communities.

The same work[4] contains an analysis of the speech of Khalapur, a village north

[3] Bloomfield, Leonard. "Language. 1933." New York: Holt (1933)

[4] Gumperz, John J. "Linguistic and Social Interaction in Two Communities 1." American anthropologist 66.6_PART2 (1964): 137-153

of Delhi. Gumperz appears to consider the people of this village as constituting a different speech community from speakers of Hindi, and outlines some differences between the speech of the village and Hindi. This is suggestive, since the differences are not very significant – they mostly consist of slightly different pronunciations of function words, with some exceptions. Using this as a yardstick, we have some idea of where to draw the line for a certain speech community.
For example, it is immediately clear that all of Kerala cannot come under a single speech community, following this definition (since some groups of dialects differ even more significantly from each other than Hindi from the speech of Khalapur). Similarly, Hindi and Urdu almost certainly define different speech communities as well.

Thus, this definition clears the matter up further, by means of the study the author refers to in the same work. We have a less uncertain view of the boundaries of speech communities now.

1.5 Labov (1972)

The Speech community is not defined by any marked agreement in the use of language elements, so much as participation in a set of shared norms; these norms may be observed in overt types of evaluative behaviour and by uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage. (Labov, William. Sociolinguistic patterns. No. 4. University of Pennsylvania press, 1972.) p.120

This is the first definition that explicitly excludes *use of language elements* – Labov instead resorts to defining the community in terms of a *set of shared norms*.
He refers to variation in the usage of various forms at certain levels. Almost all linguistic elements have some form of variation – this is not spatial or regional variation, but variation across different contexts of usage or situations. These situations are what Labov calls *levels of usage*. Thus, a community is determined by all its members following the same *pattern of variation* across all its *levels of usage*.

The reference to levels of usage make this definition unique in unambiguously assigning Hindi and Urdu to different speech communities. This is because it ascribes to one speech community an *abstract pattern of variation* rather than a single speech system. Thus, Hindi's is the speech community that speaks Hindustani at an informal or day-to-day level, but makes use of vocabulary derived principally from Sanskrit in its "highbrow"[5] variant, spoken on television and radio broadcasts. Urdu has the same (or a very similar) system at its colloquial level, but derives its upper-class and literary vocabulary principally from Farsi and Arabic (cf. the previously noted distinction between प्रधान मन्त्री and वज़ीरे-आज़म). Thus, the two speech communities have different *patterns of variation*, even though at one level they coincide.

[5]King, Robert D. "The poisonous potency of script: Hindi and Urdu." (2001): 43-59.

Similarly, most (but not all) dialects of Malayalam in Kerala come under one speech community, as they observe similar variations across different contexts. For example, several dialects express respect by avoiding use of the second-person pronoun and repeating the addressee's designation (one might, for instance, ask one's mother `amma kalYYicco?` *did Mother eat?* without using any pronouns). Dialects which differ from others in such variations would then constitute a distinct speech community.

Therefore, this definition is less ambiguous than most of the others, avoiding as it does reference to a single communication-system (defined by many factors other than language), but instead making use of a composite, abstract entity – variation across levels of usage of a language. This helps us to more clearly mark off speech communities.

1.6 Le Page (1968)

Each individual creates the systems for his verbal behaviour so that they shall resemble those of the group or groups with which from time to time he may wish to be identified, to the extent that

- *He can identify the groups*
- *He has both opportunity and ability to observe and analyse their behavioural systems*
- *His motivation is sufficiently strong to impel him to choose and to adapt his behaviour accordingly*
- *He is still able to adapt his behaviour* (Robert Le Page, 1968)

Here again, the explicit reference to language or linguistic systems is avoided, but the term *verbal behaviour* is mentioned, which is similar to Gumperz's *verbal signals* and Bloomfield's *speech-signals*. However, Le Page describes the members of a speech community as having systems that *resemble* each other – but may not be the same.

Importantly, he also mentions *groups with which … [the speaker] may wish to be identified [and which] he can identify*. This seems to indicate that he considers the individual's description of their language – their linguistic identity – as being of significance in the demarcation of the individual's speech community. This is not, however, to the exclusion of actual, verbal similarity in speech; both of these are factors in the decision.
Le Page's next point requires that the speaker have *both opportunity and ability to observe and analyse [the groups'] behavioural systems*. This can be taken to mean that the user is physically present and can therefore observe the (verbal) behaviour of the groups referred to. This is reminiscent of Gumperz's condition of *regular and frequent interaction*, which only these two definitions share.
The third point he lists is a restriction on the individual's *motivation* to imitate the behaviour of said groups. The purpose of this point appears to be to reinforce the identity-based aspect of the community – when one feels the need to identify with a group or community, the motivation to adapt one's behaviour to emulate

that group is stronger. In fact, in a later work[6], Le Page clarifies that the social groups with which the speaker wishes to identify give cues to the speaker that indicate to them the chance that they can be integrated – this feedback affects the speaker's motivation to identify with the groups.
The fourth and final point given by Le Page concerns the speaker's *ability to adapt his behaviour*. This point appears redundant in contrast with the other three – no reason immediately comes to mind why an individual who identifies with a group, interacts with it frequently, and is motivated to integrate with it, should face difficulty in doing so. However, it can be taken as a reiteration on the constraint that a speaker must behave in a way consistent with the general behaviour of the rest of the members of the group.

An important quality of Le Page's definition is that it is centred around a single individual, and does not in fact list the characteristics of the group as an entity in itself. As a result of this, it is more explicit in allowing speech communities to have overlaps – the very specific use of the phrases *group or groups … from time to time* make this clear.

This definition's clear emphasis on linguistic identity unambiguously separates Hindi and Urdu, like Labov's (although for a different reason). It is clear that Hindi and Urdu speakers have distinct linguistic identities from one another[7]. On the other hand, all dialects of Malayalam come under a single speech community according to this definition, since the systems of verbal behaviour all over the state can be said to *resemble* each other (albeit in a very broad sense), and (perhaps more importantly) because all speakers identify their language as Malayalam and not as Trivandrum Malayalam, Ernakulam Malayalam, etc.

This definition, then, brings together in some respects Gumperz's, Bloomfield's and Hockett's characterisations of the notion of a speech community. It takes into account similarity as well as the identity of speakers themselves.

2 Question 2

All the definitions explained above have their own specifics and focus on different sets of aspects of the notion of a speech community. It is, however, immediately clear from the analysis of all of them that a speech community is an extremely broad and slightly imprecise notion, and therefore difficult to pin down exactly. We notice that there are some recurring themes in the discussion of factors that isolate a speech community. These are language identity, mutual intelligibility, language/dialect and community integration. Each definition emphasises some of these more than others.
Language identity is a language user's own idea, or characterisation, of the language they speak. It is independent of the ground reality (and sometimes at

[6]Le Page, R. B. "Sociolinguistics and the Problem of 'Competence'." Language Teaching & Linguistics: Abstracts, vol. 8, no. 3, 1975, pp. 137–156., doi:10.1017/S0261444800002743.
[7]King, Robert D. "The poisonous potency of script: Hindi and Urdu." (2001): 43-59.

odds with it, as in the case of Hindi vs. Punjabi).
Mutual intelligibility is simply a measure of how comprehensible the speech systems of two speakers are to each other; it is a valuable notion especially when discussing dialect continua.
Language/dialect are two very vague and inconsistently defined notions – they are considered separately here as they are given verbatim as a condition in the definitions, but do not overlap exactly with any of the others. Authors who use these terms do not clarify their sense either.
Community integration is the only non-linguistic criterion in the definitions; it describes the level of assimilation (in terms of frequency of interaction and similarity in behaviour) between a single speaker and a wider community.

Lyons' definition only refers to language and dialect, which, as we have seen, are not satisfactorily precise. His definition allows us enough leeway to make use of either the mutual intelligibility or the language identity criterion when deciding our speech communities.

Hockett's definition also refers to language, but he ackowledges the trickiness brought in by this word. He also emphasises (later in the text) the role of language identity, and this can be taken to be his main criterion in the definition.

Bloomfield's two definitions make no reference to language, but (like Hockett's) focus on the role of intelligibility, making, in fact, no mention of language identity as playing a role. He uses terms like *system of speech-signals* circumvent the ambiguities caused by the word *language*.

Gumperz makes use of the term *body of verbal signs* for the same purpose as Bloomfield does, but still refers to *language use* in his definition. His is also the first definition to take into account community integration. It also, uniquely, mentions how to tell when two speech communities are *not* the same, as well as when they are.

Labov's definition is special in an explicit exclusion of the vague criterion of "language". However, he also does not make use of the other notions of mutual intelligibility, language identity or community integration – rather, he develops the abstract notion of *invariance in patterns of variation across levels of usage*, which, in a way, takes the place of the notion of language in his definition. This notion is his criterion for grouping people of a speech community together.

Le Page's is the most intricate and comprehensive definition, in which the ideas of language identity, community integration and mutual intelligibility are all involved to constrain membership in a speech community. Like Bloomfield and Gumperz, he has a term (*verbal behaviour*) that allows him to avoid referring to *language*. In a sense, Le Page's definition (apart from the concern of *resemblance of systems of verbal behaviour*) is almost psychological in nature, reiterating, as it does, the social integration of the speaker into a group.

Thus, we have seen how each of the six definitions given above are important for the definition of a speech community, and the specialties of each one of them.

We have also seen how they apply to the speech communities of Hindi, Urdu and Malayalam.