Linguistic Areas

Lyle Campbell
University of Hawai‘i Manoa
Areal linguistics is about the diffusion of structural features across the languages of a geographical area.

The term linguistic area refers to a geographical area in which, due to borrowing and language contact, languages of a region have come to share structural features –

Terms: *Sprachbund*, diffusion area, convergence area, adstratum.
The Balkans -- the best known linguistic area. Languages: Greek, Albanian, Serbo-Croatian, Bulgarian, Macedonian, and Romanian (some add Romani and Turkish).
Balkans areal traits:

1. **Central vowel** /ɨ/ (or /ə/) (not in Greek or Macedonian); 
2. Syncretism of dative and genitive cases (merged in form and function). Romanian *fetei* ‘to the girl’ or ‘girl’s’: *am data o carte fetei* ‘I gave the letter to the girl’ and *frate fetei* ‘the girl’s brother’;
3. **Postposed articles** (not in Greek). Bulgarian *məʒat* ‘the man’, *məʒ* ‘man’.
4. Periphrastic **future** (with auxiliary corresponding to ‘want’ or ‘have’, not in Bulgarian or Macedonian). Romanian *voi fuma* ‘I will smoke’ (literally ‘I want (to) smoke’) and *am a cínta* ‘I will sing’ (literally ‘I have sing’);
5. Periphrastic perfect (with auxiliary corresponding to ‘have’);
6. **Absence of infinitives** (constructions equivalent to ‘I want that I go’ for ‘I want to go’);
South Asia (Indian sub-continent) Linguistic Area
Languages: Indo-Aryan (branch of Indo-European), Dravidian, Munda, Tibeto-Burman (branch of Sino-Tibetan) families.
South Asian Areal traits:

1. Retroflex C's, particularly retroflex stops;
2. Absence of prefixes (accept in Munda);
3. Dative-subject construction (dative-experiencer. Hindi mujhe 'I knew it' [mujhe 'to me' + know + Past]);
4. Subject-Object-Verb (SOV) word order, postpositions;
5. Absence of a verb 'to have';
6. 'Conjunctive or absolutive participles' (subordinate with non-finite verbs (participles), preposed; e.g. relative clauses precede the nouns the modify;
7. Morphological causatives (verbal suffix).
South Asian Areal traits:

1. Retroflex C’s, particularly retroflex stops;
2. Absence of prefixes (accept in Munda);
3. Dative-subject construction (dative-experiencer. Hindi mujhe maaluum thaa ‘I knew it’ [mujhe ‘to me’ + know + Past]);
4. Subject-Object-Verb (SOV) word order, postpositions;
5. Absence of a verb ‘to have’;
6. ‘Conjunctive or absolutive participles’ (subordinate with non-finite verbs (participles), preposed; e.g. relative clauses precede the nouns they modify;
7. Morphological causatives (verbal suffix);
South Asia (the Indian sub-continent) linguistic area (cont.)

8. ‘Explicator compound verbs’ (auxiliary from a limited set completes the sense of the preceding main verb, the two verbs together refer to a single event. Hindi le jaanaa ‘to take (away)’ [‘take’ + ‘go’]);

9. Sound symbolic forms based on reduplication, often with $k$ suffixed (Kota [Dravidian]: kad-kadk ‘[heart ] beats fast with guilt or worry’; a:nk-a:nk ‘to be very strong [of man, bullock], very beautiful [of woman]’).

Some of these features are not limited to the Indian sub-continent, are also in neighboring languages (e.g. SOV word order is in much of Eurasia and N Africa). Some traits are not independent of one another, e.g. languages with SOV word order tend to have non-finite (participial) subordinate clauses, and not to have prefixes.
Mesoamerica

Languages: Nahua (branch of Uto-Aztecan), Totonacan, Otomanguean, Mixe-Zoquean, Mayan, Xinkan, Tarascan, Cuitlatec, Tequistlatecan, Huave.
**Mesoamerican areal traits:**

Five areal traits are shared by nearly all Mesoamerican languages, but not by neighboring languages beyond this area, and these are considered particularly diagnostic of the linguistic area:

1. Nominal possession: his-dog the man ‘the man’s dog’. Pipil (Uto-Aztecan): *i-pe:lu ne ta:kat*, lit. ‘his-dog the man’;

2. Relational nouns (locative expressions composed of noun roots +possessive pronominal affixes), my-head for ‘on me’. Tz’utujil (Mayan): *č-r-i:x* ‘behind it, in back of it’: *č*- ‘at, in’, *r*- ‘his/her/its’, and *i:x* ‘back’; *č-w-i:x* ‘behind me’, lit. ‘at-my-back’;

3. Vigesimal numeral systems. Chol (Mayan): *hun-k’al* ‘20’ (1x20), *čaʔ-k’al* ‘40’ (2x20), *uš-k’al* ‘60’ (3x20), *hoʔ-k’al* ‘100’ (5x20), *hun-bahk’* ‘400’ (1-bahk’), *čaʔ-bahk’* ‘800’ (2x400);

4. Non-verb-final word order (no SOV) – Mesoamerica is surrounded SOV languages;
Mesoamerica (cont.)


Since these five traits are shared almost unanimously throughout Mesoamerican languages but are generally not found in the languages just beyond the borders of Mesoamerica, these traits are strong evidence supporting the Mesoamerica as a linguistic area. Four of these five traits have essentially the same distribution, clustering at the borders of Mesoamerica. Such bundling is uncommon in linguistic areas.

A large number of other features are shared among several Mesoamerican languages, but some are not found in all the languages of the area, while others are found also in languages beyond the borders of the area (see Campbell et al. 1986.)
The Northwest Coast Linguistic Area

**Languages**: Tlingit, Eyak, the Athapaskan languages of the region, Haida, Tsimshian, Wakashan, Chimakuan, Salishan, Alsea, Coosan, Kalapuyan, Takelma, Lower Chinook.
NW Coast areal traits:

Elaborate systems of C’s: series of glottalized stops and affricates, labiovelars, multiple laterals, uvular stops; typically few labials (completely lacking in Tlingit and Tillamook, very limited in Eyak and most Athabaskan languages); several uvular consonants. Few V’s, only 3 (i, a, o, or i, a, u) in several languages, 4 in others. Several have pharyngeals (ħ, ꧙), most have glottalized resonants and continuants.

Shared morphological traits: suffixes, near absence of prefixes; reduplication (of several sorts, signaling iteration, continuative, progressive, plural, collective, distribution, repetition, diminutive, etc.); numeral classifiers; alienable/inalienable oppositions in nouns; evidential markers in verbs; verbal locative-directional markers.
Masculine/feminine gender (shown in demonstratives and articles); visibility/invisibility opposition in demonstratives. Aspect is more important than tense. All except Tlingit have passive-like constructions. The negative appears as the first element in a clause regardless of the usual word order. Overt marking of nominal plurals is absent or limited, as in many American Indian languages. Northwest Coast languages also have lexically paired suppletive singular and plural verb stems (that is, a lexical root may be required with a plural subject which is entirely different from the root used with a singular subject).

Some other traits are shared by a smaller number of Northwest Coast languages.
Criteria that have been used:

1. The number of traits shared by languages in an area,
2. Bundling of the traits in some significant way (e.g. at the same geographical boundaries),
3. The weight or complexity of different areal traits (some are accorded more significance for determining areal affiliation on the assumption that they are more difficult to acquire by diffusion than others).

To establish a LA, the more shared features, the better. LAs in which many diffused traits are shared among the languages are considered better established.

Difference of opinion about how many traits it takes. Some believe even one shared trait is enough to define a weak linguistic area. In the best known linguistic areas, the number and kind of shared traits vary a lot.
The idea that greater weight should be attributed to some traits for defining linguistic areas – example: borrowed word order in **Ethiopian linguistic area.**

Ethiopian Semitic languages interconnected word-order patterns, borrowed from neighboring Cushitic languages. Several of these traits reflect the diffusion of the SOV (Subject-Object-Verb) basic word order typology of Cushitic languages into the formerly VSO Ethiopian Semitic languages. Typologically the orders Noun-Postposition, Verb-Auxiliary, Relative Clause-Head Noun, and Adjective-Noun are all correlated and tend to co-occur with SOV order cross-linguistically. How to count: several shared traits?, or one general typological trait (SOV typology)? If one, it ranks high for defining a LA, since it is much more difficult for a language to change so much of its basic structure under areal influence than it is to acquire less complex traits.
Bundling of areal traits, clustering at the boundaries of a linguistic area, might be required for defining linguistic areas, though this has proven a poor criterion.

Linguistic areas are similar to traditional dialects, where often one trait spreads across more territory than another trait, so that their boundaries (or territories) do not coincide (do not ‘bundle’). Typically the geographical extent of individual traits may vary considerably. However, in the rare situation where the traits do coincide at a clear boundary, the definition of a linguistic area matching their boundary is relatively secure (e.g. Mesoamerica).

**Problem:** Lack of bundling raises serious questions for the definition of linguistic areas and thus for areal linguistics generally.
Kinds of “Linguistic Areas”  What has been called LAs include entities of widely divergent character and historical backgrounds, differing by social, cultural, political, geographical, attitudinal, and other factors.

“LAs arise in any of several ways – through social networks established by such interactions as trade and exogamy, through shifts by indigenous peoples of a region to the language(s) of invaders, through repeated instances of movement by small groups to different places within the area” (Thomason 2001:104). In the literature:

**Incipient LA:** only beginning to form though with a few shared traits.

**Moribund, decaying:** due to changes after the area was actively formed, fewer traits are currently recognizable.

**Overlapping:** overlap from different times or different reasons.

**Multilateral:** traits spread from various languages of the region.

**Unilateral:** most shared traits throughout the region are due to some particularly influential language.
**Linguistic Areas** due to rapid conquest, to population spread, migrations (traits moving with speakers);

LAs through home-grown, stay-in-place language contact (movement of traits but not peoples);

Disrupted LAs with “latecomers, earlier drop-outs, or temporary passers-by” (Stolz 2002:265).

The notion LA does not refer to a uniform phenomenon.

**Question**: is the notion LA warranted? Do all these varied things qualify as LAs, given how different they are in their composition and the in the circumstances of their birth (and decay)?
Linguistic Area vs. Borrowing generally?

It is generally acknowledged that linguistic areas are “notoriously messy,” “notoriously fuzzy” things, and that “what we understand about LAs is depressingly meager” (Thomason 2001: 99). A common perception is that the term ‘linguistic area’ is difficult to define. Thomason (2001:99): “linguistics has struggled to define the concept [LA] ever since [Trubetzkoy 1928], mainly because it isn’t always easy to decide whether a particular region constitutes a linguistic area or not.” Stolz’s (2002: 259): “the search for clearcut definitions [of “Sprachbund, linguistics area, and areal type”] has been largely futile and will probably never come to a really satisfying conclusion.”
Linguistic Area vs. Borrowing generally? (cont.)

In spite of long efforts to define ‘LA’, there is no generally agreement on its definition; for the best accepted LAs, e.g. Balkans, scholars do not agree on which languages belong to it or what traits characterize it, what its geographical extent is.

The problem: lack of a clear distinction between areal phenomena and borrowing generally.
List of proposed Linguistic Areas:
Africa as a linguistic area
Ethiopia, or the Ethiopian highlands
Benin/Surinam
Southern Africa
Macro-Sudan Belt
West African Sahel Sprachbund
The Cape Linguistic Area
Yobe State (Nigeria) Linguistic Area
Benin-Surinam Transatlantic Sprachbund
Trans-Atlantic Sprachbund
European Sprachbund (Standard Average European [SAE] Linguistic Area
Ireland
Scotland
Basque and its neighbors
Balkans
Baltic
Karelian Sprachbund
Uralic-Siberian Linguistic Area
Uralo-Altaic Sprachbund
Mediterranean
Sumero-Akkadian (Sumero-Akkadian), Ancient Near East
Anatolian area
Carpathian (Danube) Linguistic Area
Onogur Sprachbund
Rossic Sprachbund
Siberian Linguistic Area
Upper Yenisei Sprachbund (Yenisei-Kirgiz area)
Volga-Kama Sprachbund
Ostyak (Ob-Yeniseic) Sprachbund
Volga-Oka Sprachbund?
Yeniseic Sprachbund
Ob-Ugric Sprachbund
Core Uralic (Central Uralic) Sprachbund
Peripheral Uralic (Lateral Uralic) Sprachbund
Eastern Uralic Sprachbund
Caucasus
South Asian Linguistic Area (Indian subcontinent)
Sri Lanka Sprachbund
Northern Asia
Amdo Sprachbund (Qinghai-Gansu Sprachbund)
Northwest China Sprachbund
Mainland Southeast Asia
South China / Indochina
Sepik River Basin
Bird’s Head Sprachbund
East Nusantara area (E Indonesian Linguistic Area)
Timoric (Timor Leste)
North Vanuatu Sprachbund
The Pacific Linguistic Area
Australia as a linguistic area
Northwest Coast of North America
Northern Northwest Coast Linguistic Area
Northern California
Clear Lake Area (California)
South Coast Range Area (California)
California Linguistic Area
Southern California-Western Arizona Area
Plateau Linguistic Area
Great Basin Area
Pueblo Linguistic Area
Plains Area
Southeast Linguistic Area
Northeast Linguistic Area
Mesoamerica
Lowland Maya Diffusion Area
Hueheutenango Sprachbund (in Guatemala)
Areal linguistics and language classification. Often areal similarities among languages have been mistaken for evidence for remote relationships.

“Altaic” hypothesis: Turkic, Mongolian, and Tungusic, sometimes Korean, Japanese. Most specialists do not accept “Altaic.” Much of the original motivation for joining these languages was traits shared areally: vowel harmony, simple inventories of sounds, agglutination, suffixing, SOV, participial subordinate clauses – areal traits, shared by languages in surrounding regions. Failure to recognize the areal traits led to a questionable proposal of genetic relationship.
The **Mosan** hypothesis: Salishan, Wakashan, Chimakuan. The “evidence” is areal traits in the NW Coast linguistic area. Now abandoned. Swadesh’s (1953) 16 shared structural similarities to support Mosan are NW Coast areal features or typologically commonplace, e.g.:

1. “Extensive use of suffixes.”
2. “Nearly complete absence of functioning prefixes.”
3. “Extensive use of stem reduplication.”
4. “Aspect, including at least momentaneous and durative.”
5. “Tense is an optional category.”
6. “Plural is an optional category.”
7. “Non-feminine vs feminine gender (demonstratives and articles).”
8. “Numeral classifier notions, shown by suffixes.”
9. “Two alternate stems for number” (lexically paired distinct sg and pl verb stems).
10. “Lexical suffixes referring to body parts, other space references.”
11. “Predicative use of nouns.”
12. Demonstrative distinctions e.g. present vs absent, visible vs invisible.
Mahalo nui loa
(Thanks)!