What Linguists Know About Heritage Speakers

Maria Polinsky
University of California San Diego
polinsky@ling.ucsd.edu

Basic ideas

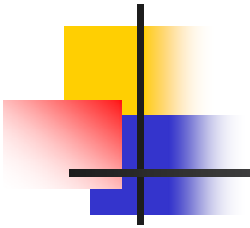
- The language of heritage speakers X is not the same as the language spoken by monolingual speakers X
- Heritage languages share a number of common features
- Understanding what heritage languages have in common and how they are different from full languages is an important step in developing efficient ways of teaching heritage speakers

Basic terms and definitions

- Heritage language/incompletely acquired language (cf. the proficiency definition, Valdés 2002)
- Full language/completely acquired language
- Baseline: full language that is used for comparison with a given heritage language
  - Baseline ≠ standard (Polinsky 1999)
- Dominant language

Linguistic questions about HL

- Do heritage speakers differ from full speakers in their linguistic competence (mental representation) or just in performance?
- Does a heritage language constitute a rule-based system or is it a collection of randomly acquired and retained “chunks”?



- If a system, what rules apply in a heritage language and do they differ from the rules of the corresponding full language?
- Do any aspects of heritage language represent universal properties of human language?
- Which aspects of heritage language are due to the interference from the dominant language?

Outline of the talk

- Heritage language as a system
- Recurrent structural features of heritage language and their possible explanation
- Conclusions
- Questions for discussion

A system or a collection of "chunks"?

Predictions of the System Model

- Heritage speaker *variation* should be lower than/similar to the variation found in full languages
- Item frequency plays a subsidiary role, similar to the role played by frequency in complete acquisition

Predictions of the Chunks Model

- Heritage speaker *variation* should be more random than variation found in full languages
- Item frequency plays a crucial role; the higher the frequency the more likely an item is to occur in the heritage language

Variation among heritage speakers

Case study I: Phonetics of Heritage Armenian (Godson 2001, 2002)

- Background

Armenian: Indo-European, two main varieties, Eastern (Armenia, Iran) and Western ("Diaspora Armenian")

Main dialects within Western Armenian: Turkish WA, Lebanese WA, Syrian WA, Egyptian WA

Vowel system: 5 basic vowels (*a, e, i, u, o*); epenthetic schwa (ə), very frequent

Experimental setup

- *Subjects*: 10 heritage speakers of Western Armenian; 10 Armenian-English bilinguals (Armenian-dominant), 1 baseline monolingual Western Armenian speaker (all females)
- *Procedure*: Reading task
- *Measurements*: Spectral representation of vowels

How similar are two speakers?

- Compare vowel spectra
- *Formal measure*: Standard Deviation of the Differences Distribution, or SDDD index (Harmegnies 1988)

  $$SDDD = [\,[\Sigma_{i=1\text{-}K} (S_i - S'_i - M_d)^2]/K\,]^{0.5}$$

  S and S' are the two spectra compared, each having an S_i and S'_i value for each of the K frequency components;

  M_d is the mean of the sum of the $S_i - S'_I$ differences.

  A standard deviation calculation on the differences between the set of amplitude measurements by frequency for the two spectra being compared. When the spectral shapes are widely different, the SDDD value increases; when the shapes are similar the SDDD decreases.

Positive control measurements

- To test the suitability of the SDDD index for the data, an initial test was done with subjects known to have different articulatory settings:
  - an English monolingual from the Midwest United States
  - an English monolingual born and raised in Southern California
  - an Armenian-dominant bilingual

SDDD range

- Midwest English monolingual vs. herself (2 speech samples): 2.77
- Midwest English monolingual vs. S. CA English monolingual: 6.12
- Midwest English monolingual vs. Armenian bilingual : 13.65
- S. CA English monolingual vs. Armenian bilingual: 13.13

The more different two dialects/languages, the higher the SDDD

0-3: same speaker
3-6: weak dialectal variation
6-9: strong dialectal variation
9+: different languages

SDDD measures for Armenian speakers

Armenian-dominant speakers by dialect

	Syria	Lebanon	Egypt	Jordan	Turkey
Syria	4.28	8.37	3.64	4.14	7.64
Lebanon		6.21	3.13	5.98	8.96
Egypt			-	7.23	7.34
Jordan				-	6.07
Turkey					-
N subjects	4	3	1	1	1

Heritage speakers by parents' dialect

	Syria	Lebanon	Egypt	Jordan	Turkey
Syria	4.54	5.72	5.92	-	7.12
Lebanon		5.86	5.90	-	4.45
Egypt			5.29	-	5.92
Jordan				-	-
Turkey					4.03
N subjects	2	4	2	0	2

Results

- Heritage speakers show a smaller range of variation in phonetic setting than full speakers
- This variation does not seem random

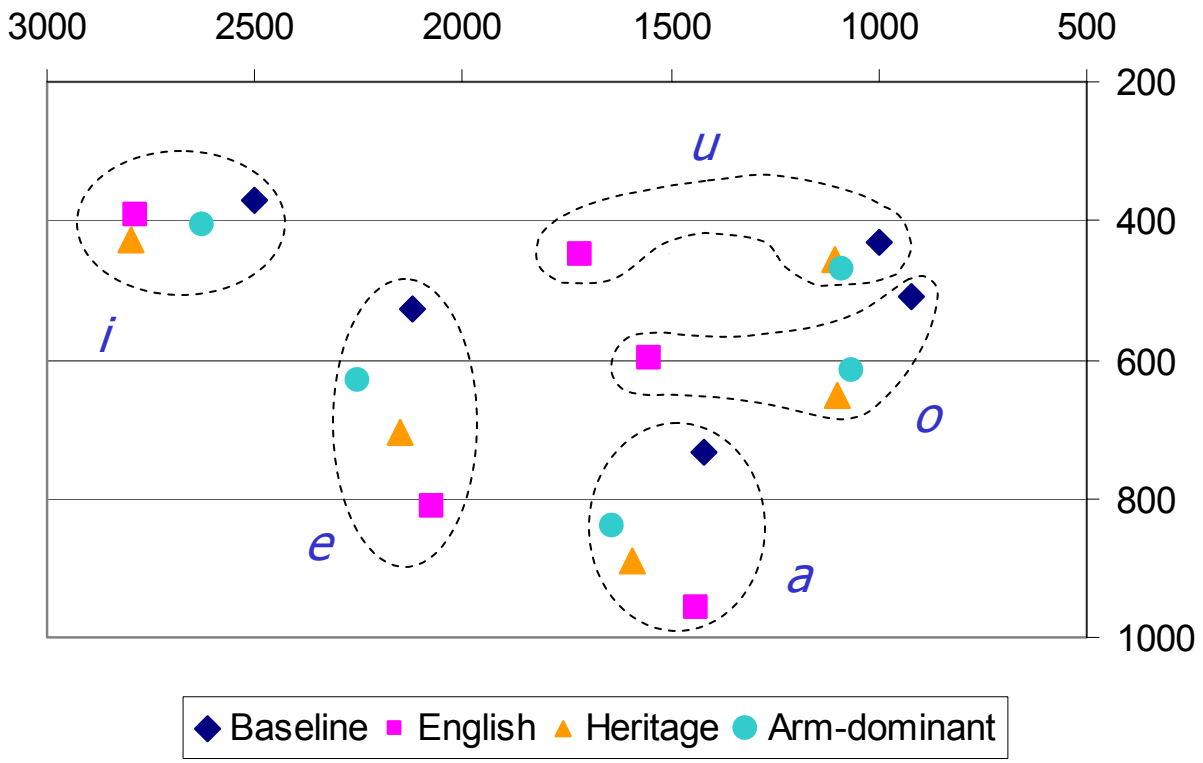
Questions

- Is the limited range of variation in Heritage W. Armenian due to the development of a new "communal standard" shared by all the heritage speakers surveyed?
  - No, because the heritage speakers in the study come from different areas and backgrounds and generally do not speak Armenian/speak it on a very limited basis
- Is the limited range of variation in Heritage W. Armenian due to the interference from English?
  - Further experiments were necessary to answer this question

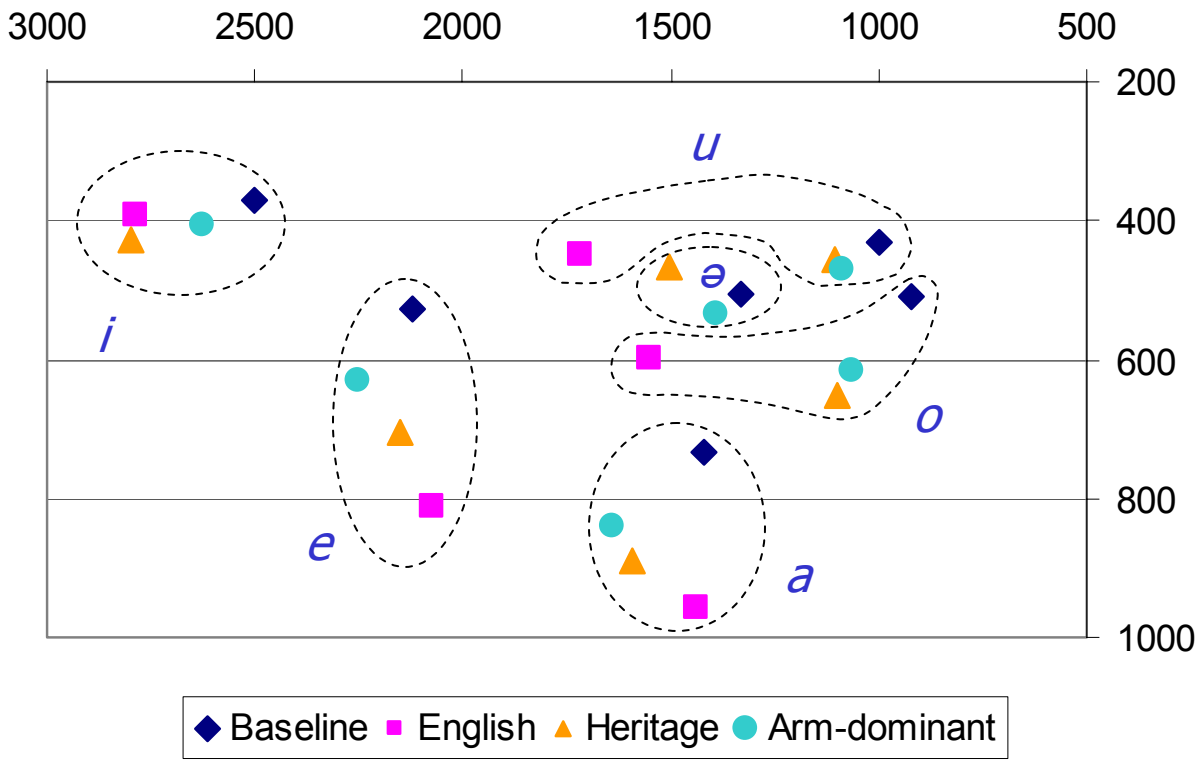
Vowel systems in Full Armenian, Heritage Armenian, and English

- There are perceptible differences in the pronunciation of heritage speakers and full speakers (the so-called 'heritage accent')
- The pronunciation of heritage speakers is judged as different from the pronunciation of L2 learners of Armenian
- Can the 'heritage accent' be reduced to the English accent?
- Godson 2002: study of vowel space in Full Western Armenian, Heritage Western Armenian, English
  - Choice of vowels: most perceptible, better measurements
  - Absent from the study: L2 learners of Armenian; heritage speakers whose dominant language is other than English

Vowel space: English, Heritage W. Armenian, and Full W. Armenian



Vowel space: English, Heritage W. Armenian, and Full W. Armenian (with *schwa*)



Study I: Conclusions

- Subjective pronunciation differences between Heritage W. Armenian speakers and Armenian-dominant speakers are corroborated by phonetic measurements
- The vowel system of Heritage W. Armenian cannot be accounted for by simple transfer from English
- Overall results are more compatible with the System Model than with the Chunks Model because they reveal systematicity within each group of speakers

The role of frequency: system or “chunks”?

Case study II: Frequency and parts of speech in Heritage Russian

- Background

Russian: Indo-European, large number of speakers in the USA, many heritage speakers, in particular among the children of the Third and Fourth Wave of immigration (Andrews 1999)

Relatively well studied grammar of the full language, good frequency data (Šteinfeldt 1965, Brown 1996)

- What factors determine lexical knowledge among heritage speakers?
  - Frequency alone
  - Frequency and some other way of categorizing lexical items
  - Factors other than frequency

Testable hypothesis

- If frequency only, there should be no difference in heritage speakers' knowledge of different lexical categories (parts of speech) controlled for frequency level

  Nouns, adjectives, verbs

  dog, junior, fit (KF75)

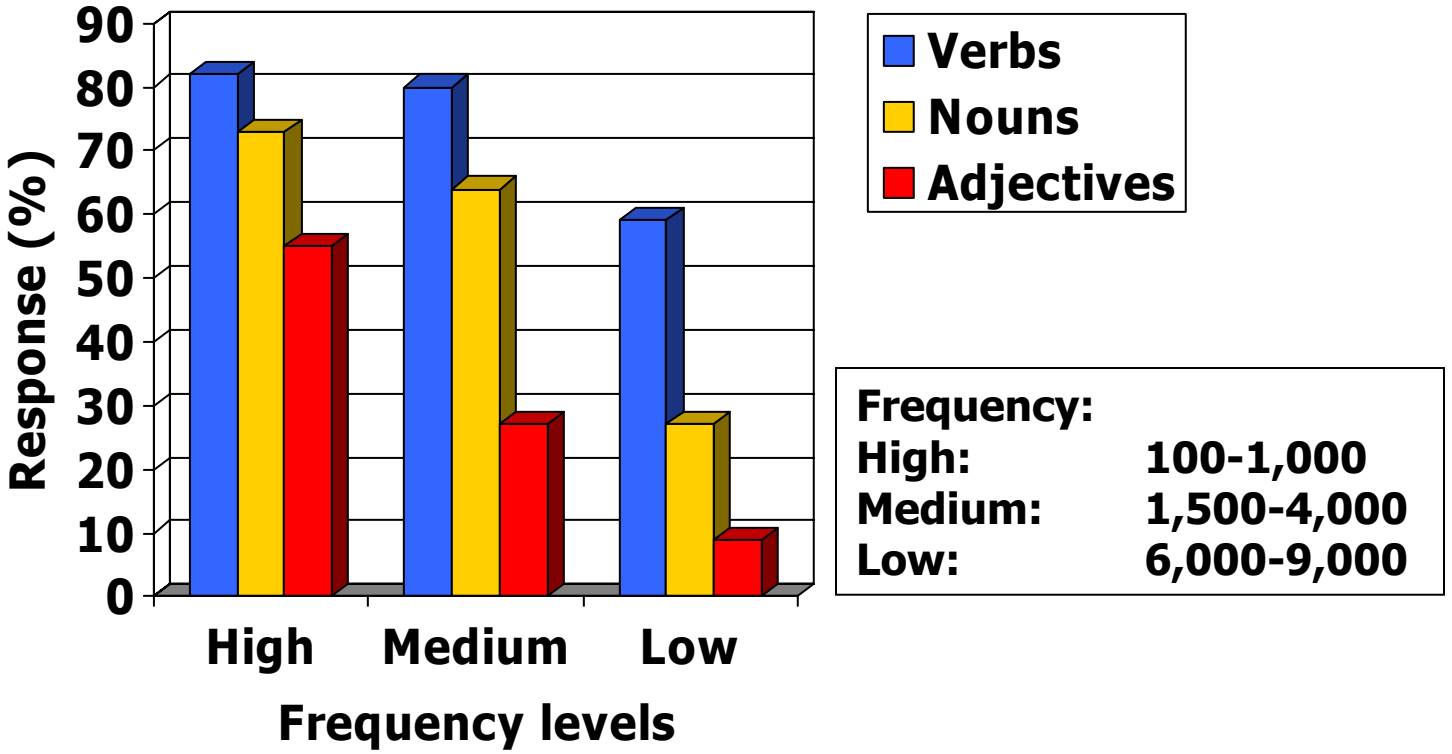
  acquisition, aggressive, abandon (KF17)

Experimental design

- *Subjects*: 8 heritage speakers of Russian, adults (average age 24). Started acquiring Russian as their first and only language, self reporting as English-dominant at 3-4 (N=4), 5 (N=2), 8 (N=1). No reading ability in Cyrillic.
- *Procedure*: Auditory presentation, 99 randomized lexical items (3x3: nouns, adjectives, verbs; high, mid, low frequency); Latinate words and recent borrowings excluded.
- *Task*: translate into English.
- *Measurements*: accuracy of translation; reaction time

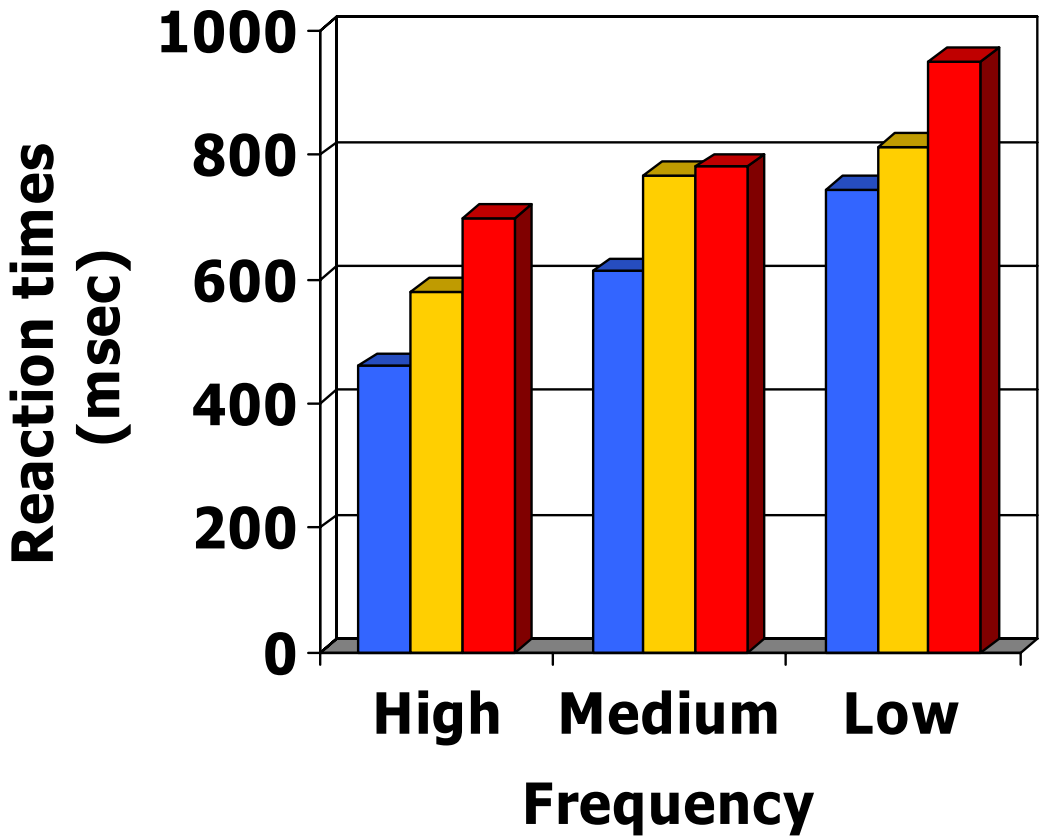
Results: Response accuracy

Averaged over 7 subjects



Results: Reaction times

Averaged over 7 subjects



Frequency:

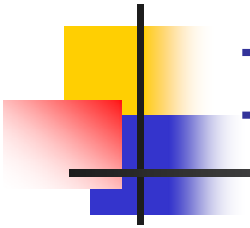
High:	**100-1,000**
Medium:	**1,500-4,000**
Low:	**6,000-9,000**

Results

- Verbs, nouns, and adjectives controlled for frequency show different levels of maintenance in Heritage Russian:

  Verbs > nouns > adjectives

- The maintenance of lexical items *within* a lexical category is sensitive to frequency



Interpreting the results

- Verbs > nouns > adjectives
  - Heads > modifiers, accounts for the poor performance on adjectives
  - Smaller class > larger class, accounts for V > N
  - Preliminary statistics: verbs (12%), nouns (48%)
  - Closed class > open class?

Closed vs. open class

- How do verbs compare with function words (prepositions, conjunctions) which constitute a clear closed class?
- Preliminary results (1 speaker tested): no difference in response to high frequency verbs and to function words of similar frequency

Why do verbs outrank nouns/adjectives?

- Closed class
  - Verbs are less of a closed class than function words; the lexical knowledge of verbs and the lexical knowledge of function words appear similar

☞ Closed class ≠ explanation

- Verbs and nouns differ in the order of acquisition
  - Conflicting views: verbs before nouns (Choi & Gopnik 1995) or nouns before verbs (Gentner 1981; Wolff & Gentner 1997)

☞ Order of acquisition ≠ explanation

Conceptual structure

- Representational differences between verbs and nouns (Gentner 1981;Nagy & Gentner 1990)
  - Difference between relational (predicative) concepts and (concrete) object concepts
  - Relational concepts (predicates) set up a frame providing links to other concepts (objects)
  - The absence of a relational concept is confounded by the absence of its frame and links; the absence of an object concept is less "costly"

Relational concepts > Object Concepts

- *Pending experiments:*

  Testing heritage speakers' knowledge of relational concepts regardless of lexical category

  - Same lexical category
  - Same frequency level

Study II: Conclusions

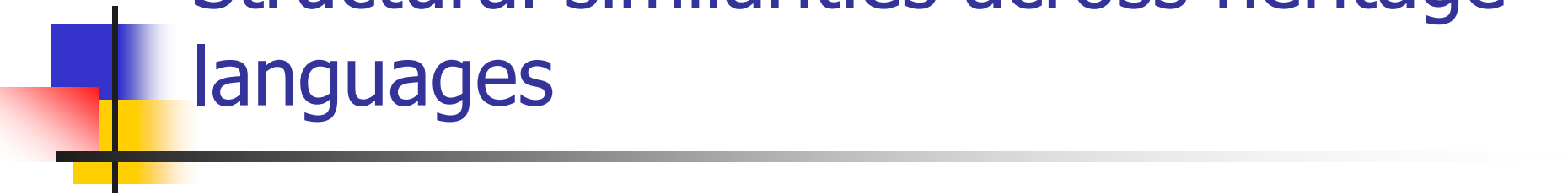
- The competence of a heritage speaker includes knowledge of lexical categories (parts of speech), possibly based on conceptual underpinnings of these classes
- The knowledge of abstract grammatical features is associated with structural heads
- Frequency alone cannot account for the results of the Heritage Russian study
- The overall results are compatible with the System Model and argue against the Chunks Model

General hypothesis

- Heritage language grammar may rely more heavily on conceptual representation
- Conceptual representation may be mapped more directly into heritage grammar

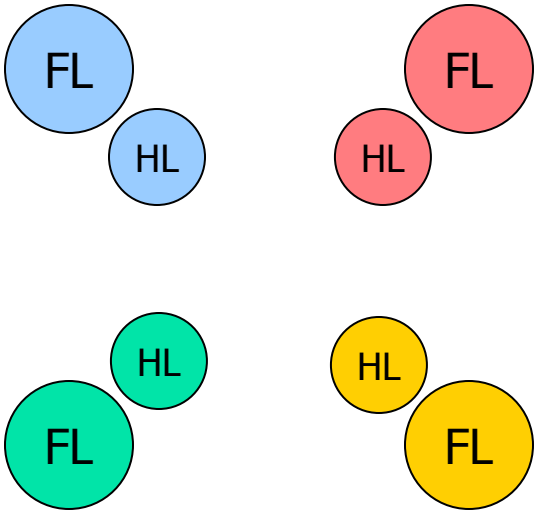
How much do heritage languages resemble each other?

Structural similarities across heritage languages

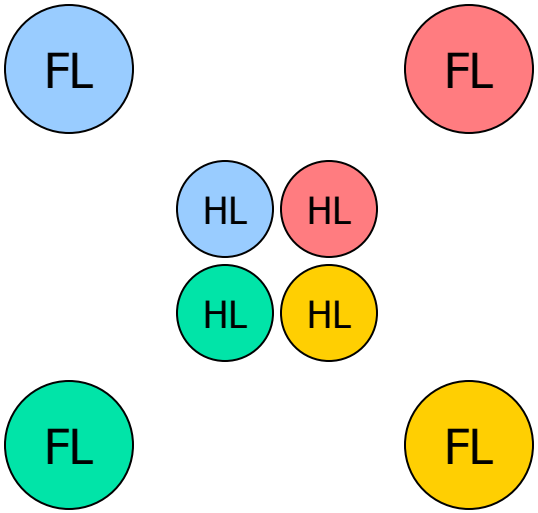


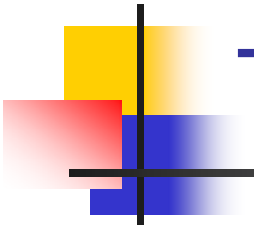
Grammars of heritage languages

- The grammar of an individual heritage language is similar to the grammar of the baseline full language



- Different heritage language grammars resemble each other more than each individual heritage language resembles the baseline





Three test cases:

- Aspect
- Agreement
- Pro-drop

Aspect in Russian and Spanish

- *Aspect*: internal composition of a situation
  - *Usual denotation*: complete or incomplete (ongoing) event
  - Aspect is a particularly interesting phenomenon because the *natural or inherent* verb semantics combines with and modifies the interpretations of the grammatical ways of expressing aspect
- *Tense*: temporal relations between situations (such as past, present and future)

Inherent verb semantics: Vendler types

Basic aspectual verb types (Vendler 1967)

- *State*: eventuality in which there is no perceptible change (*be boring, see, love*)
- *Activity:* event without a goal, homogenous process (*walk, snore, swim*)
- Achievement: event with a goal but no duration, instantaneous (*arrive*, *destroy*)
- Accomplishment: event with a goal and duration *(wash, tell)*

Verb types and their expression

- Different Vendler types receive different grammatical expression across full languages (Van Valin & LaPolla 1997)
- How do heritage languages express different verb types?
  - Variation similar to that found in full languages
  - Each verb type receives a similar expression across different heritage languages

Russian

- *Deti kormili sobaku*

'The children were feeding (IMPF) the dog.'

- *Deti nakormili sobaku*

'Children have fed (PERF) the dog.'

- *Ja vse zabyvala*

'I kept forgetting (IMPF) everything.'

- *Ja vse zabyla*

'I have forgotten (PERF) everything.'

Same tense, different aspectual forms, formally marked by prefixes/suffixes

All verb types participate in the alternation

Spanish

- *Comías bien en este restorante* 'You ate (IMPF >Imperf) well in this restaurant.' [generic]
- *Comiste bien en este restorante* 'You ate (PRET >Perf) well in this restaurant.' [specific]
- *El auto me costaba $20.000* 'The car cost (IMPF>Imperf) me 20K.'
- *El auto me costó $20.000* 'The car cost (PRET>Perf) me 20K.'

Imperfect tense > imperfective aspect

Preterite tense > perfective aspect

All verb types participate in the alternation

Aspect in Full Russian and Full Spanish

- Full Russian
  - Imperfective/perfective: inflectional and derivational morphology of aspect and the case of the main argument
  - All event types are compatible with the imperfective and perfective aspect, with weak lexical preferences
  - Perfective has many sub-meanings and is more complex than imperfective

- Full Spanish
  - Imperfective/perfective: inflectional morphology of the Imperfect and Preterite tenses
  - All event types are compatible with Imperfect and Preterite, however, lexical preferences are found
  - Imperfective has many sub-meanings and is more complex than perfective

Aspect in Heritage Russian and Heritage Spanish

- Heritage Russian (Polinsky 1994, 2002)
  - Aspectual pairs are lost; typically, one aspectual form is maintained
  - States and activities (verbs denoting homogenous events): imperfective lexicalization
  - Achievements and accomplishments (verbs denoting goal-oriented events): perfective lexicalization
- Heritage Spanish (Montrul 2001; Slabakova & Montrul 2002)
  - Preterite and Imperfect are no longer possible with all verb types
  - States (verbs denoting homogenous events) avoid Preterite: imperfective preference
  - Achievements (verbs denoting goal-oriented events) avoid Imperfect: perfective preference

Aspect restructuring in heritage language

- The mapping of event types into aspectual distinctions is lexicalized
  - Multiple mappings are disprefered (Heritage Spanish) or eliminated (Heritage Russian)
- Aspect is reanalyzed to reflect the semantic distinction between goal-oriented (telic) events and events without a goal (atelic)
  - Telic: accomplishments, achievements
  - Atelic: states, activities

Results: Aspect

- Heritage Russian and Heritage Spanish show close parallels in the mapping of verb types into aspectual distinctions
- The aspectual systems of Heritage Russian and Heritage Spanish differ from the aspectual systems of the respective full languages

Results: Aspect

- How do heritage languages represent different verb types?
  - Heritage languages show variation in the expression of verb types, just as full languages do
  - Each verb type receives a similar expression across different heritage languages

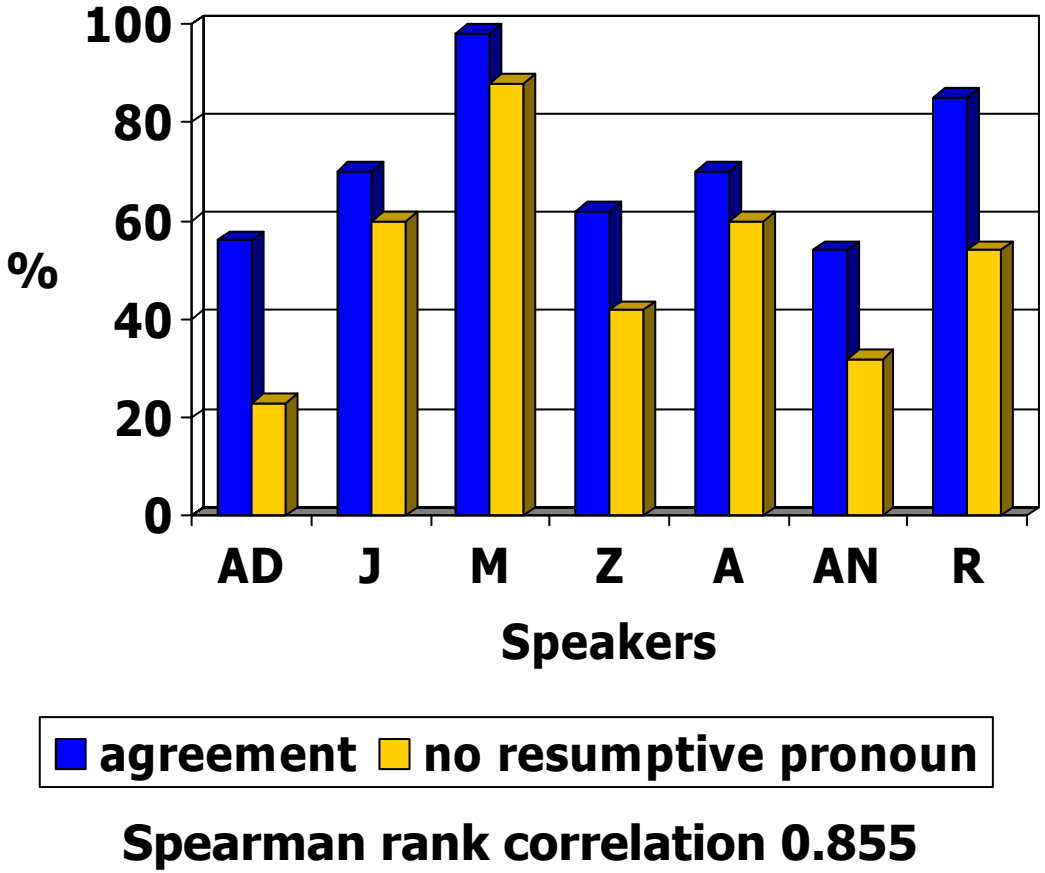
Agreement in heritage languages

- Impoverished morphology
  - Decline in agreement morphology
- Use of resumptive pronouns to compensate for the absence of agreement
  - *The candle it burn*
  - *The candles it/they burn*

Agreement and resumptive pronouns: Polish

- *Subjects*: 7 Heritage Polish speakers, mean age 27
- *Procedure*: spontaneous narrative and picture description
- *Measurements:* percentage of incorrect agreement forms; percentage of resumptive pronouns

Agreement and resumptive pronouns: Heritage Polish (Polinsky 1995)



Spearman rank correlation 0.855

Results: Agreement

- Heritage languages do not maintain the morphological agreement found in some full languages
- The absence of agreement is gradient, not categorical
- The absence of agreement is compensated for by the use of a resumptive pronoun

Pro-drop (null subjects)

- *Pro-drop*: omission of non-emphatic subject pronouns, e.g. in Spanish:

  ∅ *compraron* *un vestido*

  'They bought a dress.' [neutral statement]

  Ellas *compraron* *un vestido*

  'THEY bought a dress.' [contrast implied]

- *Full languages with pro-drop*: Spanish, Polish, Tamil, Japanese, etc.

Pro-drop in heritage languages

Predictions:

- Absence of agreement > agreement based pro-drop is no longer sustainable

- Pro-drop is not a categorical feature in full languages, therefore, heritage languages should show decline, not loss, of pro-drop

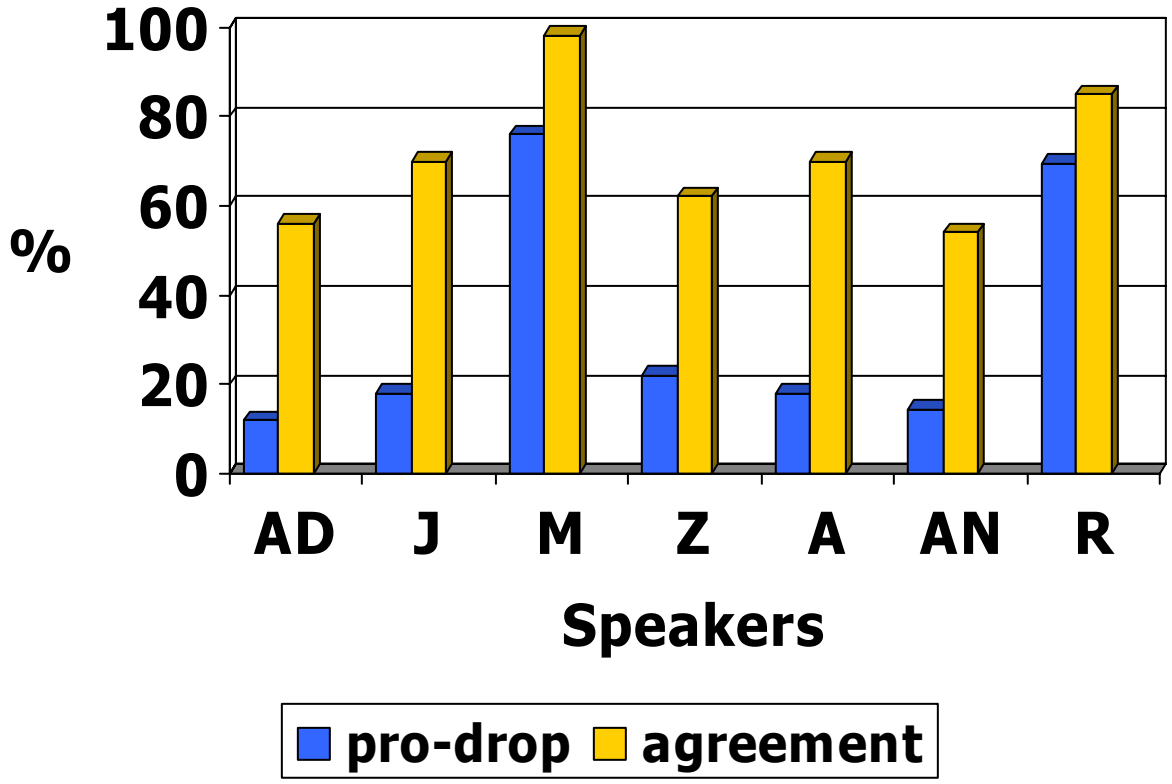
Pro-drop in Full and Heritage Spanish

- Full Spanish: pro-drop occurring in 60%-80% utterances (Bentivoglio 1992; Ávila-Shah 2000; Toribio 2001; Montrul 2001)
- Heritage Spanish: pro-drop between 50%-70% (Silva-Corvalan 1994; Lipski 1993; Ávila-Shah 2000; Toribio 2001; Montrul 2001)
  - The level of proficiency of the heritage speakers in individual studies unclear; more data needed on both Full and Heritage Spanish, with emphasis on the dialectal baseline differences in full language

Pro-drop in Full and Heritage Polish

- Full Polish: Pro-drop at about 70% (statistics based on a narrative and a play)
- Heritage Polish: Pro-drop at 32.8%, averaged over 7 subjects

Heritage Polish: Pro-drop and agreement (Polinsky 1995)



Spearman rank correlation 0.855

Results: Pro-drop

- No pro-drop in heritage language
- The absence of pro-drop is gradient, not categorical
  - Rate of pro-drop may correlate with an individual speaker's degree of proficiency
- Absence of pro-drop and absence of agreement are correlated

Results

- Heritage languages show a more restricted range of variation than full languages
  - This may be an artifact of the small sample
- Heritage languages share a number of morphosyntactic features (absence of agreement, resumptive pronouns, absence of pro-drop)
- These recurrent morphosyntactic features are correlated

Other structural features (Polinsky 1995)

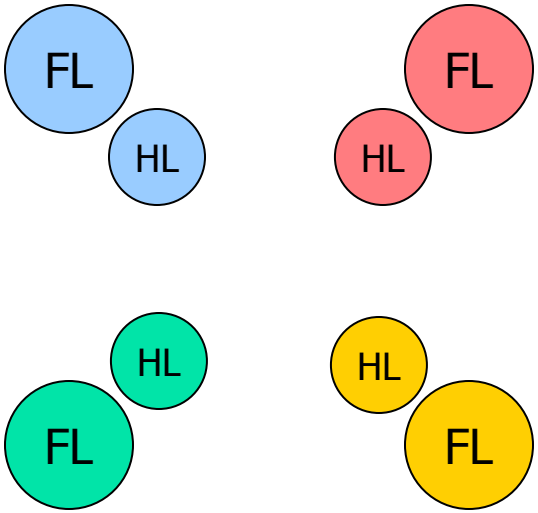
- Decline in case morphology
- Unmarked plural
- Unmarked possessive constructions: *dog Tom*
- No overt contrast between root and embedded clauses
- Coordination > subordination
- Restructuring of gender categories (Leisio 2000) and classifier systems

Conclusions

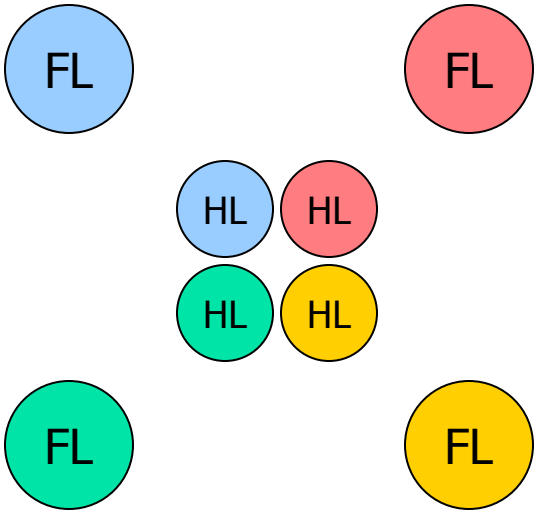
- Structural similarities across heritage languages seem greater than structural similarities between a heritage language and its corresponding full language

Grammars of heritage languages

- The grammar of an individual heritage language is similar to the grammar of the baseline full language



- Different heritage language grammars resemble each other more than each individual heritage language resembles the baseline



In search of explanation

- How can we account for the recurrence of structural features across heritage languages?

Possible explanations for the recurrent structural properties in HL

- Absence of uninterpretable features (Sorace 2000, Montrul 2001)
  - Information that has nothing to do with the lexical item is lost/suppressed
- Shortest distance in dependencies
  - Processing effects proper
  - Grammaticization of processing effects
- Choice of the most unmarked way of speaking (Bickerton's "bioprogram")

Overall conclusions

- A heritage language grammar differs from the grammar of its respective full language in a systematic, rather than random way
- Heritage languages share a number of common structural features
- At this point, there are several hypotheses as to how to account for the systematicity within and across heritage languages:
  - Rule-governed system
  - Patterns of preference
  - "Bioprogram"

Overall conclusions

- Much needed: More descriptive generalizations on the structure of heritage languages
  - These generalizations can be used to design further studies
  - These generalizations can be used to test what aspects of heritage languages are due to the interference from the dominant language
  - These generalizations can be used to inform pedagogical strategies in HL teaching

Questions for discussion

- What recurrent features have you observed in the language of heritage speakers (cf. OK's handout)
  - Low proficiency heritage speakers
  - High(er) proficiency heritage speakers
- What are these features attributable to?
  - Interference from English
  - Interference from a baseline dialect different from what is used in the classroom
  - Anything else
- What pedagogical tools can be used to modify the recurrent features if they do not fit full language?

HLI Questionnaire (Olga Kagan)

- Please compare linguistic proficiencies of non-heritage and heritage learners in the language you teach. Give specific examples where applicable
- Linguistic areas
  - Phonology
  - Vocabulary
  - Morphology and syntax
  - Register

References

- Andrews, David R. 1999. *Sociocultural perspectives on language change in diaspora: Soviet immigrants in the United States*. Amsterdam, Philadelphia: John Benjamins.
- Ávila-Shah, Barbara. 2000. Discourse connectedness in Caribbean Spanish. In Ana Roca (ed.). *Research on Spanish in the United States*, 238-251. Somerville, MA: Cascadilla.
- Bentivoglio, Paola. 1992. Linguistic correlations between subjects of one-argument verbs and subjects of more-than-one-argument verbs in spoken Spanish. In Paul Hirschbuhler and Konrad Koerner (eds.). *Romance languages and modern linguistic theory*, 11-24. Amsterdam: John Benjamins.
- Brown, Nicholas J. 1996. *Russian learners' dictionary: 10,000 words in frequency order*. London, New York: Routledge.
- Choi, Soonja, and Alison Gopnik. 1995. Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language* 22. 497-529.
- Gentner, Dedre. 1981. Some interesting differences between verbs and nouns. *Cognition and Brain Theory* 4. 161-178.
- Godson, Linda. 2001. *Articulatory setting in Western Armenian*. Ms. UCSD.
- Godson, Linda. 2002. *The phonetic aspect of incomplete acquisition: Heritage Western Armenian*. Ph. D. Dissertation in progress.
- Harmegnies, Bernard. 1988. SDDD: A new dissimilarity index for the comparison of speech spectra. *Pattern Recognition Letters* 8. 153-158.
- Lipski, John. 1993. Creoloid phenomena in the Spanish of transitional bilinguals. In Ana Roca and John Lipski (eds.). *Spanish in the United States*, 155-173. Berlin: Mouton.
- Montrul, Silvina. 2001. Incomplete acquisition and attrition of Spanish tense/aspect in adult bilinguals. *Bilingualism: Language and Cognition* 5. 39-68.
- Nagy, William, and Dedre Gentner. 1990. Semantic constraints on lexical categories. *Language and Cognitive Processes* 5. 169-201.
- Polinsky, Maria. 1994. *Loss of aspectual features.* Paper presented at the Symposium on aspect, UCLA, May 1994.
- Polinsky, Maria. 1995. Cross-linguistic parallels in language loss. *Southwest Journal of Linguistics* 14. 87-124.
- Polinsky, Maria. 1997. American Russian: Language loss meets language acquisition. In Wayles Browne et al. (eds.). *Annual Workshop on Formal Approaches to Slavic Linguistics: The Cornell Meeting (1995),* 370-406. Ann Arbor, MI: Michigan Slavic Publishers.
- Polinsky, Maria. 2000. A composite linguistic profile of a speaker of Russian in the USA. In Olga Kagan and Benjamin Rifkin, eds. *The learning and teaching of Slavic languages and cultures*. Bloomington, IN: Slavica.
- Polinsky, Maria. 2002. *American Russian as a rule-governed system*. Ms. UCSD
- http://ling.ucsd.edu/~polinsky/pubs/american-russian.pdf
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Slabakova, Roumyana, and Silvina Montrul. 2002. Aspectual tenses in L2 Spanish: A UG perspective. In Rafael M. Salaberry and Yasuhiro Shirai (eds.). *Tense-aspect morphology in L2 acquisition*. Amsterdam: John Benjamins.
- Sorace, Antonella. 2000. Differential aspects of attrition in the L1 syntax of near-native L2 speakers. In Catherine Howell, Sarah Fish, and Lucas Keith (eds.). *Proceedings of the 24th Annual Boston University Conference on Language Development, I-II*, 719-725. Somerville, MA: Cascadilla.
- Steinfeldt, Evi. 1965. *Russian word count*. Moscow: Progress.
- Toribio, Almeida Jacqueline. 2001. Spanish language decline. *Proceedings of the 25th Annual Boston University Conference on Language Development*, I-II, 768-779. Somerville, MA: Cascadilla.
- Van Valin, Robert, Jr., and Randy J. La Polla 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.
- Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca: Cornell University Press.
- Wolff, Philip, and Dedre Gentner. 1997. On the lexicalization of causal events. In Hughes, Elizabeth, Hughes, Mary, and Annabel Greenhill (eds.). *Proceedings of the 21st Annual Boston University Conference on Language Development, I-II*, 707-718. Somerville, MA: Cascadilla.