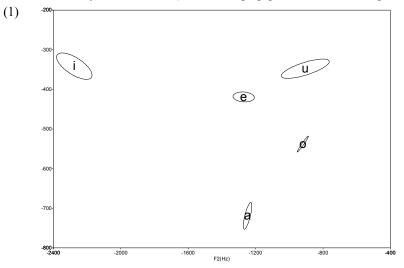
Dispersion in the Vowel System of Pima* Eric Jackson, UCLA LASSO Conference, 18 October 2003

Goals

- Instrumentally describe the vowel space of Pima, including allophonic variation due to syllable type (open *vs.* closed) and stress (two degrees)¹
 Discuss the Pima system in the context of theories of vowel distribution
- 1. Introduction: Vowel Dispersion Theories
- Cross-linguistic tendencies in segmental inventories, including vowel systems, invite explanation in terms of phonetic universals
- Simple theories of vowel dispersion (*e.g.* Liljencrants and Lindblom 1972) explain the cross-linguistic tendencies of vowel inventories by predicting a set of vowels which is maximally dispersed in the available formant space
 - Vowels are predicted to be maximally dispersed (ideally in a perceptual space), assuming that vowels are best perceived when their prototypes are maximally distinct.
 - Such models do not predict languages with relatively uneven distributions or unfilled perimeters of the vowel space; the fact that such languages exist shows that other considerations besides maximal distinctiveness are responsible for the distribution of vowels
- More complex theories of vowel dispersion (Lindblom and Maddieson 1988, Schwartz *et al* 1997*b*) respond to this by including both dispersing forces (such as maintaining maximal distinction) and attracting forces (such as articulatory simplicity or auditory salience, for example from the convergence of formants)
 - Although such theories allow for multiple vowel systems with the same number of vowels—which simpler theories did not do—large unfilled areas of vowel space are still predicted to be relatively rare
- If the canonical vowels of a language are unevenly distributed in the vowel space, a theory involving vowel dispersion might still predict that processes of allophonic variation would make use of the gaps in the vowel space

¹ This study considers only two degrees of stress—main stress *vs.* unstressed—though Fitzgerald (1999) has argued for an intermediate degree of stress in Tohono O'odham.

- Solution This is an empirical question, which would need to be posed in any language which was an apparent counterexample to a theory of vowel dispersion (*cf.* Keating and Huffman 1984)
- 2. O'odham languages: Counterexamples to Dispersion Theories?
- Pima is a Tepiman language of Southern Uto-Aztecan, spoken near Phoenix, Arizona. Pima and the closely related language Tohono O'odham (also called Papago) have been described as dialects of a single O'odham language
- The vowel system of these languages has been described by many authors (*e.g.*, Hale 1959, Saxton 1963, Hale 1965, Saxton and Saxton 1969, Hale and Alvarez 1970, Mathiot 1973, Saxton 1982, Saxton, Saxton, and Enos 1983)
 - \clubsuit These languages employ five modal vowels with two degrees of length²
 - These vowels include a high front unrounded /i/, a low central unrounded /a/, a mid back rounded /o/, a high back rounded /u/, and a high back unrounded vowel /ui/ (orthographically <e>)³
- The canonical vowels of Pima (*i.e.*, in sentence-initial stressed syllables surrounded by either /h/ or /?/) show a large gap in the mid-front region⁴



 \checkmark The symbol $<\!\!e \!\!>$ is used to represent /ui/, following the orthography

^{*} All Pima data in this presentation has been provided through the patience and good humor of Mr. Virgil Lewis, a native speaker of Pima from the Gila River Indian Community. I would like to thank him very much for all his help and insightful comments. Thanks are also due to Patricia Keating, Peter Ladefoged, Pamela Munro, Marcus Smith, Rebecca Brown, Heriberto Avelino, several audiences at UCLA, and the other members of the 2000-01 UCLA Field Methods course for their insights and annotated notes: Jill Gilkerson, Sahyang Kim, Brook Lillehaugen, Haiyong Liu, Suzanne (Lyon) Riggle, Shannon Madsen, Jason Riggle, Shabnam Shademan, and Melissa Tai. All errors remain my own. Portions of this work have been funded by the UCLA Department of Linguistics, the UCLA Institute of American Cultures, and the Phillips Fund of the American Philosophical Society.

² Some authors, *e.g.* Zepeda (1983), refer to three degrees of length in Tohono O'odham long, unmarked, and short. Tohono O'odham vowels with this third degree of length may correspond to devoiced off-glides observed in Pima on certain final consonants.

³ The authors cited above vary in describing this fifth vowel as high back or high central unrounded; speakers of Pima have described this vowel as identical to /u/ except in lip position, indicating that at least for some speakers of Pima it is phonetically [ttt].

⁴ This and all subsequent plots have ellipses drawn at a diameter of two standard deviations along the principle axes of variation (to include 86% of the data)

- This vowel system is also seen in five other Uto-Aztecan languages, and Langacker (1970) also reconstructs this system for Proto Uto-Aztecan Shoshoni (Central Numic): Miller (1972)
 - P
 - Southern Paiute (Southern Numic): Harms (1966) P
 - P Northern Paiute (Western Numic): Snapp, Anderson, and Anderson (1982)
 - P Northern Tepehuan (Tepiman): Bascom (1982)
 - P Pima Bajo (Tepiman): Shaul (1986)
- The vowel system of O'odham is a departure in two ways from the most common five vowel systems—/i, e, a, o, u/ or /i, ε , a, o, u/—seen in the world's languages (Maddieson 1984, Schwartz et al 1997a)
 - O'odham has a gap in the mid front region of the vowel space P
 - So o'odham also has a concentration of vowels—/ut, u, o/—in the mid to high back region of the vowel space, which is the typical location of a gap if the language is missing a corner vowel
- Even the more refined theories of vowel dispersion predict that such vowel systems should be cross-linguistically rare or non-existent⁵
 - This system is in fact quite rare: although the languages mentioned above have this system, they are all related
 - P The only language with this system in the UPSID database (Maddieson 1984) is Papago (Tohono O'odham)
 - That this system has persisted among Uto-Aztecan languages is a powerful P indication that factors besides perceptual distinctiveness and focalization influence the distribution of vowels
- If maintaining maximal distinctiveness is still a consideration in O'odham, however, processes of vowel allophony might plausibly take advantage of the unused vowel space, or act to more evenly fill the available vowel space
 - Vowels near the gap (front vowels or mid vowels) might vary more
 - ♦ Vowels near the crowded region (high back) might vary less

Methods of the current study 3.

- This study represents the beginning stages of an attempt to characterize the • process of vowel allophony in an O'odham language, namely Pima
- Data to date come from a single native speaker of Pima reading sentences at a ٠ self-determined pace; research with other speakers is in progress
- Measurements of vowel formants F1 through F4 were made using the Praat • software package; only F1 and F2 have been analyzed, however, using the StatView statistics package
- The measurements were coded for: •

- \checkmark Vowel quality and length (i, a, o, u, u; each long vs. short)
- \clubsuit Occurrence in a syllable with primary stress or without primary stress⁶
- \checkmark Occurrence in an open vs. closed syllable
- Left and right vocalic and consonantal environment P
- For this study only short vowels were analyzed, as long vowels are not found in all of these environments, and few were recorded overall
- A two-way factorial ANOVA (analysis of variance) was performed for F1 and F2 ٠ for each vowel phoneme
 - Stress and syllable type were used as independent categorical variables
 - Segmental context was generalized over by measuring vowels in a variety of P different segmental contexts for each condition of stress and syllable type; the number of environments and vowel tokens for each condition are shown in : the format is # environments(# tokens)

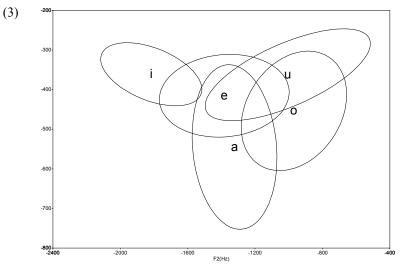
(2)	Phoneme	Open syllables		Closed syllables	
		Stressed	Unstressed	Stressed	Unstressed
	/i/	9(23)	14(35)	13(38)	12(35)
	/ ɯ /	8(28)	10(24)	11(29)	8(104)
	/a/	7(20)	12(36)	11(31)	10(150)
	/o/	5(21)	25(72)	6(21)	7(21)
	/u/	6(21)	8(28)	7(19)	7(24)

⁵ The model of Schwartz *et al* (1997*b*) comes very close to O'odham by predicting the systems /i, y, u, 'o', a/ and /i, u, u, ε , a/. These critically differ from the O'odham system, however, in that the non-peripheral vowels /y/ or /ut/ in these systems share the backness feature of the missing mid vowel (in a sense compensating for its absence), while in O'odham languages the /ui/ differs in backness from the missing mid vowel.

⁶ This refers to primary stress at the word level; the position of vowels within prosodic constituents higher than words was not controlled for.

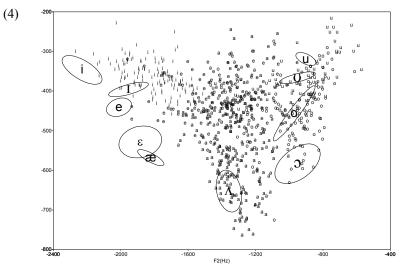
4. Vowel distribution

• The vowel space becomes appreciably filled in compared to, when vowel tokens are included from a variety of prosodic and segmental environments



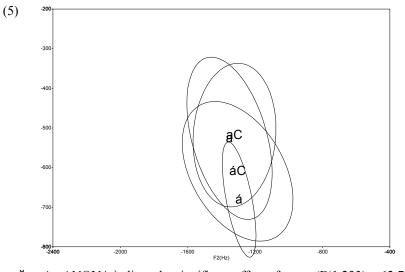
- One early aim of this study was to determine whether the Pima /tu/ is central or back, since earlier authors disagreed on this point
 - Scaling the prototypical values for [i] and [ttt] from Schwartz *et al* (1997*b*) for this particular speaker, the F2 of /ttt/ in the prototypical context is closer to the predicted F2 of a high back vowel
 - The F2 of /ttt/ measured in a variety of contexts, however, cannot be statistically distinguished from the high back or the high central vowel, since it includes such a wide range of values of F2
 - This is consistent with the target for this vowel being high back, rather than central, though the target may not be met in most contexts

• For comparison, the English vowels (in the frame [h_d], not in sentences) for this speaker are shown in , along with a scatter plot of the Pima vowel data



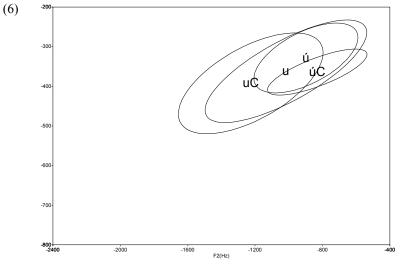
- Interestingly, none of the Pima vowels is as front as the canonical English vowels; in particular, though the Pima /i/ and /a/ do vary appreciably, they do not fill the mid front vowel space.
- 5. Vowel allophones: syllable type and stress
- Syllable type (open or closed) and stress (primary or otherwise) were examined for their effect on vowel F1 and F2
- Because ten ANOVAs were carried out (one per formant per vowel), a very conservative level of statistical certainty was used p < 0.005 (= .05/10)
- In the following figures, codes are used to indicate the condition corresponding to each set of data, exemplified for the phoneme /a/ below:
 - ✤ á: open syllable, primary stress
 - ✤ a: open syllable, unstressed
 - ♦ áC: closed syllable, primary stress
 - \Rightarrow **aC**: closed syllable, unstressed
- Significant effects of stress or syllable type were found only for the phonemes /a/, /u/, and /uu/.

The phoneme /a/



- An ANOVA indicated a significant effect of stress (F(1,233) = 62.778, p < . 0001) on F1, though no significant effects on F2.
- Solution This variation, taken to be a non-neutralizing reduction in unstressed syllables, fills in a portion of the mid front space and also results in partial overlap with the range of the /ttt/ phoneme (*cf.* and)

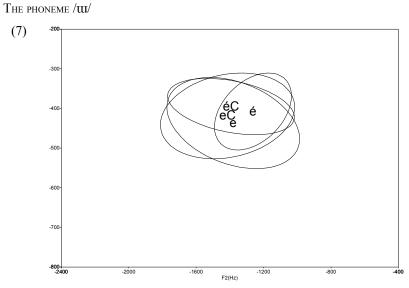
The phoneme /u/



☆ An ANOVA indicated a barely significant effect of syllable type on F1 (F (1,88) = 8.323, p = .0049), a significant effect of stress on F2 (F(1,88) =

37.604, p < .0001), and a second order effect of stress and syllable type on F2 (F(1,88)=10.845, p = .0014)

As with /a/, the reduction of /u/ in unstressed syllables results in partial overlap with /ttt/



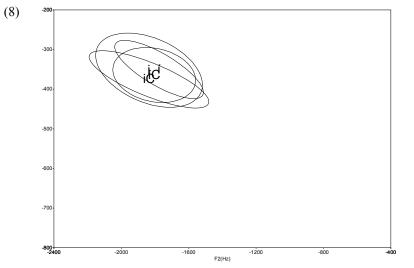
- \Rightarrow An ANOVA indicated a significant effect of stress (F(1,181) = 8.630, p = . 0037) on F1, though no significant effects on F2
- If maintaining distinctness from /u/ was a consideration, this is the direction of variation that would be expected; it may be more likely, however, that this is reduction in a mid or mid-central direction, as seen for /a/ and /u/

SUMMARY OF SIGNIFICANT EFFECTS

- Most of the significant effects on these three vowels are due to stress, and result in each of these vowels becoming more mid or more central:
 - 4 /a/ tends to be higher and /ttt/ tends to be lower in unstressed syllables
 - $\sqrt[4]{}$ /u/ tends to be more central in unstressed syllables
- For /u/, there is an interaction effect of both stress and syllable type:
 ¹/₂ /u/ tends to be more central in unstressed closed syllables than in unstressed open syllables
- The phoneme /u/ also showed the only effect due to syllable type alone:
 - \sqrt{u} /u/ tends to be lower in closed syllables
 - Since the variances of the F1 values of /u/ in the four experimental conditions varied considerably (between 29.2 and 64.8), and since this result was barely significant, re-evaluation of this result based on further data may be desirable

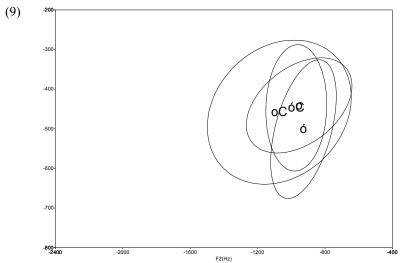
The phonemes /i/ and /o/

• The variation observed for the Pima /i/ vowel is shown below



- Although this vowel does vary, the variation does not depend significantly on stress or syllable type
- Segmental context may therefore be more responsible for this variation

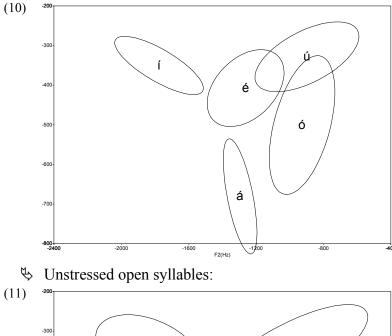
• The variation observed for the Pima /o/ vowel is shown below

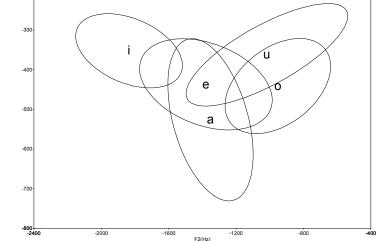


- Solution This vowel shows a great deal of variation within each of these contexts. As with /i/, segmental factors may condition a portion of the allophony seen here
- Particularly in unstressed closed syllables, /o/ and /u/ appear to overlap considerably; /u/ in this context appears to be nearly a subset of /o/

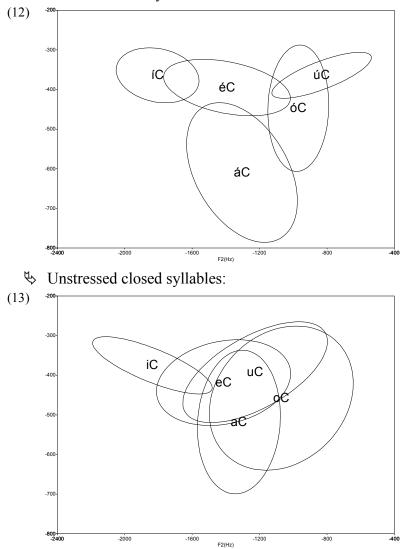
The vowels by condition

- All five vowel phonemes are shown below for each condition; in each case, the gap in the mid-front region is retained
 - Stressed open syllables:





 \Leftrightarrow Stressed closed syllables:



- Although several of the vowels in the above figures overlap appreciably, this vowel reduction is not completely neutralizing; even non-native speakers of Pima are able to distinguish most of these vowels in most of these conditions
- Several caveats are therefore necessary regarding the preceding figures:
 - Solution They indicate the range of F1 and F2 values for these phonemes in a range of segmental contexts. In a single context, they may not overlap
 - \clubsuit The vowels may also be distinguished in the higher formants or in duration,

factors which were not analyzed here.

6. Discussion

PHONETIC OVERLAP AND VOWEL REDUCTION

- These results are slightly different from a previous description of vowel reduction for Tohono O'odham
 - Crosswhite (2001) cites Hill and Zepeda (1992) as claiming that that /ttt/, /o/, and /a/ reduce to [ə] in Tohono O'odham
 - Hill and Zepeda actually claim only that /ttt/, /o/, and /a/ are realized as a centralized release (written /^h/) when they are demoraicized; they do not appear to consider this reduction *per se* (1992:367)
 - From the measurements reported here, it is /ttt/, /u/, and /a/ which reduce towards [ə] in Pima (though these vowels are not completely neutralized), and /o/ partially overlaps with [u]
- A fact quite relevant to the overlap of /o/ and /u/ is that relatively few words in Pima appear to include /u/ in an unstressed syllable
 - Smith (2002) reports that /o/ and /u/ were the two least frequent short vowels in a set of Pima texts, accounting for 3.6% and 2.7% of the phonemes, respectively; /a/ was the most frequent phoneme overall (16.4%), with /i/ ranked third (8.1%) and /tu/ ranked fifth (6.5%)
 - This fact is reflected even in the data for this study: although tokens of the phoneme /a/ were easily found in the four conditions examined here, a more thorough dictionary search was required to find words containing /o/ and /u/ in all four conditions, and these vowels remain the least frequent, as seen in the table in
- The relative rarity of /o/ and /u/ could be a motivation for the overlap seen for these two vowels:
 - Vowel reduction can act to maximize distinctness—*i.e.*, by neutralizing or blurring some vowel distinctions, the remaining distinctions may be easier to perceive in certain contexts
 - So The reduction of /o/ to [u], making these vowels less distinct, would actually increase perceptual distinctness between these and other vowels
 - Since few words include unstressed /u/, this would introduce minimal lexical confusion while increasing overall perceptual distinctness
- An alternative explanation for the rarity of /u/—namely that /u/ underwent merger with /o/ in the past—is unlikely, since such a merger has not been proposed within Proto Uto-Aztecan
 - The discussions of Proto Uto-Aztecan vowels in Langacker (1970), Miller (1967), and Voeglin, Voeglin, and Hale (1962) all have O'odham languages (exemplified by Papago) behaving conservatively, preserving the *o and *u phonemes of the proto language
 - In Voeglin, Voeglin and Hale's list of 171 cognates, *o and *u are still the least frequent short vowels

Solution by Moreover, these vowels are not lost more frequently than other vowels in Tohono O'odham: the frequency of deletion of vowels reflects the frequency of the reconstructed vowel phonemes

PATTERNS OF REDUCTION AND DISPERSION THEORY

- The vowel /a/ in reduced contexts and to some extent the vowel /i/ partially occupy the empty region in the mid front area of the vowel space, but do not completely make up for a missing mid-front vowel; this is particularly evident in the plot in
 - Solution This indicates that maintaining perceptual distinctiveness may not be sufficient to motivate /a/ or /i/ to vary allophonically with [ɛ]
- However, the pattern of vowel reduction seen with the back vowels—/u/ and /tu/ reducing towards [ə], while /o/ reduces to include [u]—may itself be motivated by considerations of maximizing distinctness
- As pointed out by Crosswhite (2001), preceptual distinctness is not the only factor which influences vowel reduction—quantal characteristics of vowels and focalization of formants also play a role—so dispersion does not need to bear the entire explanatory burden for this distribution.
- It is a challenge for theories of vowel distribution, however, that a vowel system which is predicted to be rare and potentially unstable over time has persisted in Uto-Aztecan languages to such an extent
- Interestingly, the vowel systems of Uto-Aztecan languages which do differ from Proto Uto-Aztecan generally differ so as to more evenly fill the available vowel space:
 - As mentioned in §, the O'odham vowel system is shared by at least five other Northern and Southern Uto-Aztecan languages: Shoshoni, Southern Paiute, Northern Paiute, Northern Tepehuan, and Pima Bajo
 - At least six languages—Comanche (Central Numic), Panamint (Central Numic), Kawaiisu (Southern Numic), Mono (Western Numic), Eudeve (Cahita, Sonoran), and Tübatulabal (Northern Uto-Aztecan)—differ from the proto system only by the addition of a mid front vowel (Charney 1993; Dayley 1989; Zigmond, Booth, and Munro 1991; Lamb 1958; Lionnet 1986; Voegelin 1935)
 - As many as five others—Luiseño (Cupan, Takic), Pipil (Aztecan), Mayo (Cahita, Sonoran), Yaqui (Cahita, Sonoran), and possibly Tarahumara (Tarahumaran, Sonoran)—have developed an evenly-distributed five vowel system /i, e, a, o, u/ (Hyde 1971; Campbell 1985; Freeze 1990; Lindenfeld 1973; Thord-Gray 1955)
- This change may even be beginning in O'odham languages: Miyashita (forthcoming) reports that the diphthong /ai/ in some dialects of Tohono O'odham is phonetically [e] in unstressed syllables

7. Conclusion

- Although O'odham languages do not utilize the entire formant space even under processes of vowel allophony, maintaining perceptual distinctness is still involved in the process of vowel reduction
 - Vowels near the gap do vary appreciably, though not to fill the entire gap
 - ♥ Vowels in the more crowded mid to high back region also appear to vary, though in a way which may maximize distinctness relative to reduction
- The changes seen in the vowel systems of other descendants of Proto Uto-Aztecan also show a tendency to more evenly fill the available vowel space
- Nevertheless, the persistence of this type of system is a challenge to theories which predict it to be rare or diachronically unstable
- Data from additional speakers is awaited to increase the confidence in these conclusions

References:

Bascom, Burton. 1982. Northern Tepehuan. Studies in Uto-Aztecan Grammar, vol. 3, Uto Aztecan Grammatical Sketches, ed. by Ronald Langacker. 267-393. Arlington, TX: Summer Institute of Linguistics

Campbell, Lyle. 1985. The Pipil Language of El Salvador. New York: Mouton.

Charney, Jean. 1993. A grammar of Comanche. Lincoln: University of Nebraska Press.

Crosswhite, Katherine. 2001. Vowel Reduction in Optimality Theory. New York: Routledge

Dayley, Jon. 1989. Tümpisa (Panamint) Shoshone Grammar. Berkeley: UC Press.

Fitzgerald, Colleen. 1999. Loanwords and Stress in Tohono O'odham. Anthropological Linguistics 41.193-208.

Freeze, Ray. 1990. Mayo de Los Capomos, Sinaloa. México, D.F.: Colegio de México.

Hale, Kenneth. 1959. A Papago Grammar. Bloomington: Indiana University dissertation.

Hale, Kenneth. 1965. Some preliminary observations on Papago morphophonemics. International Journal of American Linguistics 31.295-305.

Hale, Kenneth and Albert Alvarez. 1970. Toward a manual of Papago grammar: Some phonological terms. International Journal of American Linguistics 36.83-97.

- Harms, Robert. 1966. Stress, Voice, and Length in Southern Paiute. International Journal of American Linguistics 32.228-235.
- Hill, Jane and Ofelia Zepeda. 1992. Derived Words in Tohono O'odham. International Journal of American Linguistics 58.355-404.

Hyde, Villiana. 1971. An Introduction to the Luiseño Language. Banning, CA: Malki Museum Press.

Keating, Patricia and Marie Huffman. 1984. Vowel Variation in Japanese. Phonetica 41.191-207.

Lamb, Sydney. 1958. Mono Grammar. Berkeley: UC dissertation.

Langacker, Ronald. 1970. The Vowels of Proto Uto-Aztecan. International Journal of American Linguistics 36.169-180.

Liljencrants, Johan and Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. Language 48.839-862.

Lindblom, Björn and Ian Maddieson. 1988. Phonetic Universals in Consonant Systems. Language, Speech, and Mind, ed. by Larry Hyman and Charles Li, 62-78. London: Routlege.

Lindenfeld, Jacqueine. 1973. Yaqui Syntax. UC Publications in Linguistics #76. Berkeley: UC Press.

Lionnet, Andrés. 1986. Un Idioma Extinto de Sonora: el Eudeve. México, D.F.: Instituto de

Investigaciones Antropológicas, UNAM.

Maddieson, Ian. 1984. Patterns of sounds. Cambridge: Cambridge University Press.

Mathiot, Madeleine. 1973. A Dictionary of Papago Usage. Bloomington: Indiana University.

Miller, Wick. 1967. Uto-Aztecan Cognate Sets. UC Publications in Linguistics #48. Berkeley: UC Press.

Miller, Wick. 1972. Newe Natekwinappen: Shoshoni Stories and Dictionary. University of Utah Anthropological Papers 94. Salt Lake City: University of Utah Press.

Miyashita, Mizuki. forthcoming. Asymmetry to Symmetry: Gap-Filling Vowel Shift in Tohono O'odham. Paper to be presented at the 2003-04 Annual Meeting of the Society for the Study of Indigenous Languages of the Americas.

Saxton, Dean. 1963. Papago phonemes. International Journal of American Linguistics 29.29-35.

Saxton, Dean. 1982. Papago. Studies in Uto-Aztecan Grammar, Volume 3: Uto-Aztecan Grammatical Sketches, ed. by Ronald Langacker, 93-267. Arlington, TX: Summer Institute of Linguistics.

Saxton, Dean and Lucille Saxton. 1969. Papago & Pima to English, English to Papago & Pima Dictionary. Tucson: University of Arizona Press.

Saxton, Dean , Lucille Saxton and Susie Enos. 1983. Tohono 'O'odham/Pima to English, English to Tohono 'O'odham/Pima Dictionary. Second edition, ed. by R.L. Cherry. Tucson: University of Arizona Press.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée, and Christian Abry. 1997a. Major trends in vowel systems inventories. Journal of Phonetics 25.233-253.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée, and Christian Abry. 1997b. The Dispersion-Focalization Theory of vowel systems. Journal of Phonetics 25.255-286.

Shaul, David. 1986. Topics in Nevome Syntax. UC Publications in Linguistics #109. Berkeley: UCPress.

Smith, Marcus. 2002. Cross-Linguistic Text Level Letter Frequencies: Pima, Chickasaw, Menomini, and Moby Dick. Talk given at the UCLA American Indian Seminar, 16 October 2002.

Snapp, Allen, John Anderson, and Joy Anderson. 1982. Northern Paiute. Studies in Uto-Aztecan Grammar, vol 3, Uto Aztecan Grammatical Sketches, ed. by Ronald Langacker, 1-92. Arlington, TX: Summer Institute of Linguistics.

Thord-Gray, Ivor. 1955. Tarahumara-English; English-Tarahumara Dictionary, and an Introduction to Tarahumara Grammar. Coral Gables, FL: University of Miami Press.

Voegelin, Charles. 1935. Tübatulabal Grammar. UC Publications in American Archaeology and Ethnology 34(2), 55-190.

Voeglin, Charles, F. Voeglin, and Kenneth Hale. 1962. Typological and Comparative Grammar of Uto-Aztecan: I (Phonology). International Journal of American Linguistics, Memoir 17. Baltimore: Waverly Press.

Zepeda, Ofelia. 1983. A Papago Grammar. Tucson: University of Arizona Press.

Zigmond, Maurice, Curtis Booth, and Pamela Munro. 1991. Kawaiisu: A grammar and dictionary with texts. UC Publications in Linguistics #119. Berkeley: UC Press.

Eric Jackson UCLA Department of Linguistics Los Angeles, CA 90095 ejacks@humnet.ucla.edu