Mid-Phon 18: Program Booklet

TALKS

(Organized alphabetically by last name of first author)

The Interaction of Syntactic Cues and Timing Information in Spoken Word Recognition

**Melissa Baese-Berk[1], Laura Dilley[1,3], Chris Heffner[2], J. Devin McAuley[3], Tuuli Morrill[1,3], and Mark Pitt[4]

1 - Department of Communicative Sciences and Disorders, Michigan State University; 2 – Department of Linguistics, University of Maryland; 3 – Department of Psychology, Michigan State University; 4 – Department of Psychology, The Ohio State University

Recent evidence has demonstrated the importance of timing information in speech perception and spoken word recognition (e.g., Davis, Marslen-Wilson, & Gaskell, 2002; Salverda, Dahan, & McQueen, 2003). Shannon, Zeng, Kamath, Wygonski, & Ekelid, (1995) have demonstrated that when frequency information is severely degraded the signal such that timing information remains largely intact, spoken word recognition is still remarkably robust. This suggests that temporal information plays an important role in recognizing spoken words, above and beyond spectral information alone.

Dilley and Pitt (2010) demonstrated that a specific type of timing information, distal (i.e., non-local) speech rate, influences spoken word recognition. Specifically, they examined the effect of distal speech rate on the perception of function words that were heavily coarticulated with context speech and thus acoustically ambiguous with respect to the number of words and word boundaries (e.g., *leisure or time* vs. *leisure time*). Dilley and Pitt manipulated the distal speech rate of the context speech before the critical acoustically ambiguous region, but not the duration of the region itself. They demonstrated that the distal speech rate influences whether a word boundary is heard during the critical ambiguous region. They interpreted these findings as evidence for the fact that distal duration information influences perception of local timing events in speech, thus influencing word segmentation.

The current study examined how distal speech rate information interacts with syntactic information during perception. Specifically, we examined how the distal speech rate effect (i.e., the perceptual appearance or disappearance of function words as a function of the rate of surrounding speech) is influenced by whether or not the function word is syntactically obligatory. In a first study, we examined cases in which the function word is syntactically optional vs. cases in which it is obligatory (see examples 1 and 2 below).

1) Sally might try *a* liquid detergent. (Optional)
2) John said he would obey *a* rebel leader. (Obligatory)

Listeners were asked to transcribe the sentence they heard. We then calculated whether the listeners reported hearing a function word or not. The results from this study suggest that the speech rate effect overrides information from syntax. That is, listeners were willing to report the function word as being absent, even in cases in which it is obligatory.

In Experiment 2, we designed the study to ensure that listeners were using grammatical parsing when listening to the sentences. We asked participants to first determine whether the sentence they heard was grammatical, and then asked them to transcribe the sentence. Results from this study will help determine the interplay between top-down and bottom-up information during spoken word recognition.

The Acoustics of Phonation Types in Obstruents and Sonorants: The Case of Marathi

Kelly Harper Berkson, University of Kansas

This research presents a comprehensive acoustic analysis of phonation type distinctions in Marathi, an Indo-Aryan language with numerous breathy voiced sonorants and obstruents. Few previous studies have investigated breathy voiced sonorants, perhaps due in part to their crosslinguistic rarity. As such, this work establishes important new facts about phonation type distinctions in sonorants: the data reveal that male and female speakers cue breathy phonation quite differently, that there are an abundance of trading relations, and that sonorants do not cue phonation type distinctions as well as obstruents do. The claim is made that these factors contribute to the crosslinguistic rarity of employing phonemic breathy voice in sonorants.

Previous work has established that breathy voice in vowels is often associated with increased H1-H2 values (Blankenship 2002, Esposito 2006, Huffman 1987, Khan 2012, Wayland & Jongman 2003), increased H1-A3 values (Esposito 2006, Khan 2012), and decreased Cepstral Peak Prominence (CPP) values (Blankenship 1997, Esposito 2006, Khan 2012). Breathy voiced obstruents have also been found to trigger increased H1-H2 and H1-A3 values in subsequent vowels (Dutta 2007). While H1-H2 is often measured and found to distinguish between breathy and modal phonation, however, H1-A3 and CPP are not measured as commonly and prove significant less often. Furthermore, data for breathy voiced vowels outweighs that available for consonants, and consonant data relates primarily to obstruents. Some work has investigated breathy nasals (Traill & Jackson 1988, Esposito *et al.* 2005) and laterals (Harris 2009), but to date there is no comprehensive study of breathy voice in a language that utilizes this feature across multiple manners of articulation.

The current study addresses this gap by investigating breathy voice in Marathi. Ten native speakers (five male, five female) were recorded producing Marathi words embedded in a carrier sentence. Tokens included plain and breathy voiced stops, affricates, nasals, laterals, rhotics, and approximants before the vowels [a] and [e]. Measures reported for consonants and subsequent vowels include duration, F0, CPP, and corrected H1-H2* and H1-A3* values. The findings reveal that the spectral effect of breathiness extends into the subsequent vowel for both obstruents and sonorants, and that – as expected – breathy voice is associated with decreased CPP and increased H1-H2* and H1-A3* values.

A strong gender difference is also revealed, with breathy phonation cued by low-frequency measures like H1-H2* in male speech and by CPP – a measure which provides information about the aspiration noise included in the signal – in female speech. Multiple instances of trading are also reorted: time and again, where one cue is weak or absent, another cue is strong or present. Furthermore, the cues that are present for obstruents are not necessarily mirrored by sonorants. These findings underscore the importance of including both genders and multiple vowel contexts when testing phonation type differences. In addition, the fact that the acoustic differences triggered by consonant phonation type are often weaker after sonorants than after obstruents may partially explain why breathy voiced sonorants appear so rarely in the world's languages: they simply are not cued that well.

First language versus dominant language intelligibility in "switched dominance" bilinguals

Michael Blasingame, Northwestern University
Ann R. Bradlow, Northwestern University

Many models of bilingualism assume that a talker's L1 will influence speech production in the L2 (e.g., Kroll, 2010; Flege, 1995) at the phoneme and word levels. This L1 filtering produces a foreign-accent in the L2 (Anderson-Hsieh & Koehler, 2006). However, heritage speakers (bilinguals who grew up in a home in which one of the languages spoken was not the dominant language of society) present a challenge to these models as heritage speakers dissociate L1 and language dominance. For example, in the United States, Spanish heritage speakers (SHS) are L1 Spanish speakers, yet due to English instruction in schools, they have become English dominant (Polinsky and Kagan, 2007), dissociating their L1 (Spanish) and dominant language (English). Based on studies showing that age of acquisition (AoA) is a strong predictor of performance in speech perception (Mayo et al, 1997; Shi, 2010), we may expect that that the early-acquired L1 (Spanish) of SHS will show native-like production (i.e. SHS will have native like production in both languages). However, based on limited usage and non-dominance, we may expect the Spanish of SHS to exhibit a foreign-accent and lower overall intelligibility than their dominant L2 (English). Thus, in this study we ask which language (if either) will be produced with lower intelligibility.

It has also been established that there is individual variation in L1 and L2 intelligibility at the sentence level (Bradlow et al, 1996; Munro, 1998). Work on L2 intelligibility variation has focused on L2 proficiency as the main predictor (e.g. van Wijngaarden et al, 2002). In view of the possibility of native-like production in both languages, we also ask whether SHS L1 and L2 intelligibility are correlated, demonstrating that talker-specific factors contribute to SHS speech intelligibility regardless of language status (i.e. L1 vs. L2, or dominant vs. non-dominant).

In the current study, we compared SHS speech intelligibility in Spanish and English. We recorded 10 Spanish heritage speakers' productions of 110 simple sentences taken from the Hearing-in-Noise-Test (HINT; Soli, 2008) and embedded them in speech-shaped noise at two signal to noise ratios (SNR), -4 and -8 dB. Native L1 Spanish listeners and native L1 English listeners typed written responses to SHS productions of these sentences. Responses were scored as overall percent words correctly recognized, providing a speech intelligibility score for each SHS talker in each language. Results showed a significant drop in intelligibility from the easy to the hard SNR, but there was no significant effect of task language nor any task language by SNR interaction (Figure 1), indicating equivalent intelligibility across the two languages of these SHS. Acoustic analyses were also performed; however, no acoustic correlate significantly correlated with speech intelligibility suggesting that a combination of articulatory-acoustic factors underlies variation in overall intelligibility in both languages. At the easier SNR, no significant correlation was found between L1 and L2 speech intelligibility. However, results indicate a very strong, positive correlation between Spanish and English intelligibility (r = .84, t(18) = 6.7, p < .01) at the harder SNR (-8 dB) as shown in Figure 2 below.

These results suggest that Spanish heritage speakers, who dissociate L1 and language dominance, exhibit equivalent (and very high) intelligibility in both languages. Moreover, these SHS also demonstrate a cross-language dependency in overall intelligibility between both languages. SHS with low (or high) intelligibility in one language tend also to have low (or high) intelligibility in the other language, demonstrating a pattern of cross-language consistency rather than independence.

"What is a letter?" - *Nomen, Figura, Potestas*

Anna Bosch, University of Kentucky

"Letters are the figures and colours wherewith the image of man's voice is painted." John Hart, 1569

Three broad areas of intellectual interest inform the shape of dialect study through the 19th century: first, the development of nationalism, and with it a regional focus that saw value in regional or "ethnic" distinctions; second, the development of phonetics as a science drawing from both physiology and acoustics, and including the invention of instruments such as the laryngoscope and early experiments in sound recording technology; and third, the increasing professionalism of the educational system in Europe, demonstrated by the growth of professional academies and societies, and by an explicit focus on pedagogy. The pedagogical aim of this movement is not insignificant and can be seen as an important impetus for the development of the International Phonetic Alphabet, primarily to aid in language instruction (see also the journal "Le Maitre Phonétique").

As J.R. Firth outlines in great detail in his historical peroration, "The English School of Phonetics," British (or English-language) academics boast a long tradition of debate on orthography, orthoepy, "alphabetics", and phonetics. Firth holds up a number of Elizabethan figures who were influential in academic and policy matters, from Sir Thomas Smith, Secretary to Queen Elizabeth, to John Hart, who published *"an orthographie, conteyning the due order and reason, howe to write or paint the image of mannes voice, most like to the life or nature."* Later scholars built on this intellectual tradition, and 19th c. public figures such as Isaac Pitman, Alexander Ellis, and Henry Sweet developed orthographic systems that aimed at phonetic transcription. The British tradition over several centuries, then, encompassed an interest in the written expression of human speech, in regional pronunciation, and in language pedagogy; all three of these areas of focus came together in 19th century linguistic descriptions. This essay examines the transcription practices employed by early dialect descriptions of Celtic languages, in light of contemporaneous discussions of transcription and spelling reform of the period.

In 1880 Gilliéron published his first dialect description, the *Petit atlas phonétique du Valais roman*, and in 1884 Henry Sweet published his description of Spoken North Welsh. For Breton we find dialect descriptions such as Emile Ernault's "Etude sur le dialecte breton de la presqu'ile de Batz" (1883) and Loth's "Le Dialecte de l'Ile aux Moines" (1893). Descriptive essays on all the existing Celtic languages are found in the earliest issues of Celtic-study journals such as *Revue Celtique* (established 1870) and *Transactions of the Gaelic Society of Inverness* (established 1871). These early descriptive texts have had a lasting impact on linguistic study of the Celtic languages (see Jackson 1958, Hamp 1988, Gillies 1988), serving as the foundation for early 20th century dialect atlases, and directly influencing the fieldwork research and the shape of publication of later dialect atlases such as the *Atlas linguistique de la Basse-Bretagne* (Le Roux 1924-1963; see also Jackson 1967), the *Linguistic Atlas and Survey of Irish Dialects* (Wagner 1958-69)

An Acoustic Analysis of Vowel Assimilation in Mandarin

Kelly Carden, University of Iowa

Cheng (1973), Duanmu (2000), and Lin (2007), among others, describe vowel assimilation in Mandarin, a process in which the mid vowel and/or the low vowel assimilates in backness and sometimes roundness to an adjacent glide or nasal. There is little agreement, however, on the exact quality of the underlying vowels and their alternations. According to Cheng (1973), the mid vowel phoneme, which he represents as /ɤ/, has three allophones [e, ɤ, o] and the low vowel phoneme /a/ has two allophones [a, ɑ]. Cheng accounts for the alternations with his "backness rule." According to the backness rule, both the mid and low vowels assimilate in backness to an adjacent glide. Only the low vowel assimilates to an adjacent nasal. Duanmu (2000) limits vowel assimilation to only the mid vowel, which he represents in its underlying form as /ə/. According to Duanmu, the mid vowel has four allophones [e, ə, ɤ, o]. The quality of the surface vowel is determined by the backness of a preceding glide or following nasal. Lin (2007) represents the mid vowel as /ə/ and, like Cheng, claims that mid vowel assimilation occurs only with adjacent glides and not with nasals. According to Lin, the low vowel assimilates to a following nasal, but not to a preceding glide. The relevant similarities and differences of the three analyses are compared in the following table.

Table 1: Three representations of Mandarin vowel assimilation

	Cheng (1974)	Duanmu (2000)	Lin (2007)
mid vowel in isolation	[ɤ]	[ə]	[ə]
mid vowel with glide onset	[je], [ɥe], [wo]	[je], [ɥe], [wo]	[je], [ɥe], [wo]
mid vowel with nasal coda	[ɤn], [ɤŋ] (no change)	[en], [ɤŋ]	[ən], [əŋ] (no change)
low vowel in isolation	[a]	[a]	[a]
low vowel with glide onset	[ja], [wɑ]	[ja], [wɑ] (no change)	[ja], [wɑ]
low vowel with nasal coda	[an], [ɑŋ]	[an], [aŋ] (no change)	[an], [ɑŋ]

Unfortunately, there are no acoustic studies that would allow us to resolve these differences. Howie (1976) is the most thorough acoustic study of Mandarin. Howie's measurements confirm the generalizations above for the mid vowel with a glide onset, but he does not include measurements for vowels in most of the other contexts.

This paper presents an acoustic description of Mandarin mid and low vowels in varying contexts with the goal of providing a detailed account of vowel assimilation in Mandarin. In this study, eight native speakers of Mandarin were recorded pronouncing the mid vowel and the low vowel with three different glide onsets (/j/, /w/, /ɥ/) and two different nasal codas (/n/, /ŋ/). The F1 and F2 of the vowels were then measured and compared to those of the vowels in isolation. In isolation, the formant measurements of the mid vowel point to the back vowel [ɤ], while the exact quality of the low vowel in isolation varied among speakers. If assimilation of backness and roundness in different consonantal contexts does indeed occur then statistically significant differences in F2 values are expected among the different contexts. For the mid vowel, a repeated-measures analysis of variance revealed a statistically significant two-way contrast in F2 between /Vn/ and /Vŋ/. A significant four-way contrast was found among /jV/, /ɥV/, /V/, and /wV/. Due to speaker variation, the data for the low vowel were more complex, but some statistically significant F2 effects were found. Additionally, nasal codas had a significant effect on the F1 of both mid and low vowels, which is consistent with the generalizations in Beddor et al. (1986).

Memory for prosody

Jennifer Cole, University of Illinois
Stefanie Shattuck-Hufnagel, MIT

Lacking the automatic playback of audio recording devices, humans recall a previously heard utterance on the basis of a cognitive representation of the linguistic object encoded in memory. When repeating or imitating a heard utterance, this representation serves to guide the phonology and phonetics of the spoken output. A number of recent studies show that the memory encoding of perceived speech includes (sub-phonemic) phonetic detail that reflects the individual speaker's voice and variation in the phonetic implementation of phones. The study presented in this talk asks about the memory encoding of prosody: Does the cognitive representation of the prosodic form of a heard utterance specify the abstract phonological features that encode pitch-accents and prosodic phrase boundaries? Does it include speaker-and utterance-dependent phonetic detail?

Two different production tasks are used to explore the prosodic aspect of these representations: an imitation experiment in which 10 speakers of American English heard and then imitated 32 spontaneous utterances from a Maptask corpus, and a read enactment task in which a different group of 10 speakers read the same 32 sentences aloud from text presentation. For each task, the resulting utterances were compared for similarity to the original Maptask utterance in their phonological prosodic features (prominences and boundaries), where similarity was assessed through agreement statistics over ToBI transcriptions produced by the authors. Similarity was assessed for phonetic cues to prosody in glottalization and pause location, as labeled at the word level from auditory and visual inspection of the acoustic signal, using agreement statistics. Phonetic similarity was also assessed through normalized acoustic measures of F0, intensity, and duration from the stressed vowel of each word, using linear regression. The main empirical findings from these studies are (i) imitators reproduce the phonological prosodic features more reliably than phonetic cues of glottalization and pause (Figs. 1-3), and (ii) the prosodic form of read-enacted utterances is phonologically and phonetically more variable across speakers than with imitated utterances.

These findings, considered in the context of our ongoing work on prosody perception, are discussed in relation to a complex model of prosody encoding. Prosodic form is encoded in terms of both phonetic detail and abstract prosodic features, but information at these two "levels" is accessed differentially. Abstract features are accessed more reliably and accurately than phonetic detail in later tasks involving recall for reproduction.

Too Much, Too Little, Too Late: Hybrid Opacity in Berbice Dutch Creole

Michael Dow, Indiana University

Recent revisions to opacity (e.g. Baković 2011) have, in addition to expanding the typology of opaque interactions, challenged Kiparsky's (1973) widely accepted definition of opacity. For one, the distinction between transparency and opacity is not as clear-cut as previously thought; a single interaction may display both, e.g. fed counterfeeding (Kavitskaya & Staroverov 2010). Opaque interactions may also interact with each other, resulting in doubly opaque forms, e.g. double counterbleeding in Alsatian French (Montreuil 2010). In this paper, I present a case of opacity in Berbice Dutch Creole (Kouwenberg 1994) which challenges the mutual exclusivity of underapplication and overapplication. In this new type of opacity, which I dub HYBRID OPACITY, B simultaneously obscures the motivation for A (overapplication) and creates new inputs for A to apply (to which it crucially does not, i.e. underapplication) within the same derivation. While it is still too early to make any steadfast claims about hybrid opacity's impact on larger terminological issues, that both types of effects may coexist in a single interaction may prove influential in our attempts to redefine opacity.

Berbice Dutch Creole (a Dutch-Ijo creole spoken in Guyana from the 17th century till 2005) displays frequent unstressed vowel deletion (e.g. /nimi/ → [nim]), which optionally leads to place assimilation in nasal-final roots with the anterior suffix /-tɛ/ (1a-c). The mapping of /nimi-tɛ/ → [nin-tɛ] (1a) exemplifies a transparent feeding scenario between Syncope and Assimilation. Meanwhile, underlying NC-final roots either undergo epenthesis or obstruent deletion (e.g. /maŋg/ → [maŋgi], [maŋ]), but agreement of the nasal-final roots with the anterior marker is either blocked (2a, b), e.g. [maŋ-tɛ], *[man-tɛ], or vacuously satisfied (2c).

Given the predictability of NC place agreement in monomorphs, most rule-based frameworks require a placeless nasal archiphoneme in NC clusters (e.g. /maNg-tɛ/). In addition to this, Optimality Theory must account for both heterorganic (e.g. /ma**mg**-tɛ/) and homorganic nasals (e.g. /ma**ŋg**-tɛ/) in the input. With the first two UR types, the nasal actively gains its place from the velar stop. The surface cluster [ŋt] therefore violates the nasal place agreement generalization in two ways: a) it has assimilated to an ultimately absent segment, and b) it fails to assimilate to the ultimately adjacent segment. Deletion thereby removes the motivation of Assimilation (overapplication) as well as creating additional environments for Assimilation too late (underapplication).

I show that, while Lexicon Optimization's preference for the homorganic UR type /ma**ŋg**-tɛ/ challenges the overapplication aspect of hybrid opacity, this is not a necessary stipulation (cf. Archiphonemic Prudence in Bermúdez-Otero (2003)).   Hybrid opacity is problematic for the standard definition of opacity, since traditionally underapplication and overapplication are: a) derived by separate interactions, orderings or mechanisms, and b) defined in mutually exclusive terms. The fuzziness between transparency and opacity has fueled much of the drive for providing a revised, principled definition of the phenomenon. In the same manner, hybrid opacity suggests a fuzziness between *types* of opacity, which may prove an essential wrinkle in the theory. To this effect, I offer general criteria for identifying further cases of hybrid opacity in hopes that more empirical data will clarify the situation.

Vowel Height

San Duanmu, University of Michigan

In this study I examine whether two degrees of height, i.e. using four binary features [high], [back], [tense] (or ATR), and [round], are sufficient to distinguish all basic vowels in the databases UNSID and P-base.

Most phonologists assume that, to distinguish all vowels in the world's languages, at least three degrees of heights are needed, plus the feature tense (or ATR). Based on an examination of UPSID and P-base, Duanmu (2012) argues that there are at most three degrees of vowel height, where low vowels do not contrast in tenseness. Duanmu (2012) further suggests that two degrees of vowel height seem to be sufficient, although sufficient evidence remains to be provided.

In this study I examine UPSID and P-base in order to find out which languages require both [high] and [low], besides [back], [round], and [tense]. I shall follow two principles, which I call the Principle of Contrast and Known Feature First, given in (1) and (2).

(1) The Principle of Contrast:
  a. If two sounds can contrast in any language, they must be distinguished by at least one feature.
  b. If two sounds never contrast in any language, they need not be distinguished by a feature.
(2) Known Feature First:
  Unless evidence requires otherwise, use known features or properties first before introducing to a new feature (or a new feature combination).

First, I extract the basic vowel (BV) inventory of every language in UPSID and P-base by excluding non-basic vowels (diphthongs and those that involve length, nasalization, murmur, or glottalization). Then I examine every inventory that contains thirteen or more BVs in order to see if they can fit into a sixteen-cell BV table, using four binary features [high], [back], [tense], and [round] only. For example, in P-base, twelve inventories have thirteen or more BVs: Turkana, German, Saami (Central South-Lappish), Danish, Dutch (most varieties), Dutch (southern and Belgium), Dutch, Karimojong (Karamojong), Welsh, Welsh (south), Welsh (north), and Eastern Ostyak (Khanty). The largest BV inventory is found in German, analyzed in (3).

(3) Two-height analysis of basic vowels (BV) in German, the largest BV inventory in P-base

i	y		u
ɪ	ʏ		ʊ
e	ø	ɐ (ə)	o
ɛ	œ	a	ɔ

No language is found to require three degrees of height. We shall also compare our result, which is based on contrast, with those based on natural classes (Mielke 2008).

A socio-phonetic study on the relationship between /tr/ and /tʃ/ variation in Chilean Spanish

Tanya L. Flores, Indiana University Bloomington

This project begins with an acoustic analysis of the productions of two phonetic variables in Chilean Spanish that have non-standard variant forms: (1) the /tr/ cluster produced as [tʃ], as in "otro" *'other'* pronounced [o.tʃro] or [o.tʃo], and (2) the alveopalatal /tʃ/ produced with a preceding /t/, as in "ocho" *'eight'* pronounced [ot.tʃo]. This study provides the first detailed acoustic description of these variables with spectrographic evidence of the variant productions present in this dialect. These descriptions are important in replacing outdated dialectal descriptions that were based on impressionistic data (Oroz 1966, Lenz 1940). This study then examines several linguistic and extra-linguistic factors that favor the non-standard variants of these two variables. Finally, the results of the /tr/ analysis will be compared to the results of the /tʃ/ analysis to explore the relationship between the two variation processes via their internal and external motivations, with implications for the consonantal system of this dialect as a whole.

The aforementioned accounts of these variables have claimed that variation only occurs in lower class and/or informal speech registers. The data for this study comes from approximately 40 hours of high quality digital radio programming revealing that variation has since spread into formal public speech among middle and upper class speakers. The recordings were selected from 12 radio stations based in Santiago, Chile. There are a total of 1200 tokens of each variable, /tr/ and /tʃ/, for a total of 2400 tokens. These were extracted from the speech of 52 reporters, program hosts, and invited guests, all natives of the Santiago metropolitan area. Speakers were coded for sex, age, and social class. The contexts of the speech range in formality across four main genres: newscasts, formal journalistic interviews, sport-show commentaries, and light entertainment shows.

I am using spectrographic analysis to determine the place and manner of articulation of each token. Duration, voicing, and formant energy were measured for each segment of every token, including (1) closures, (2) [t] releases, (3) rhotic segments, and (4) frication segments for [tʃ], where present. Linguistic factors such as variable position within the word, word position in sentence, phonetic context, and stress were coded. Extra-linguistic factors are sex, age, and social class of speakers and target audiences, speech genre, and word frequency. A logistic regression analysis is being used to determine which linguistic and extra-linguistic factors favor the production of the variant forms.

Results of the /tr/ analysis reveal various phonetic realizations of each variant, including a voiced variant [dr], not previously reported. Significant differences between the standard [tr] and variant [tʃ] forms were found in both the F1 and F2 values, and significant durational differences were found between the standard and non-standard forms' closure and rhotic segments. While speech genre, sex, age, and word frequency were the most influential factors on variant production, unstressed syllables and back vowels in the post-context were also found to favor the non-standard forms. The /tʃ/ analysis is currently in progress.

This study extends our knowledge of how separate variation processes are related within one consonantal system through an acoustic and sociolinguistic analysis of internal and external factors affecting the variant productions.

Task differences enhance cross-language phonetic interactions in bilingual speech

Erin Gustafson, Northwestern University
Caroline Engstler, Northwestern Univeristy
Matthew Goldrick, Northwestern University

Researchers often attribute bilingual accentedness to interaction between L1 and L2 phonetic systems (Flege, 1995). For example, French and English use voice onset time (VOT; the duration between stop burst release to vowel onset; Lisker & Abramson, 1964) to cue the contrast between voiced and voiceless stops (pre-voice/short-lag in French vs. short-lag/long-lag for English). In a bilingual, these conflicting systems interact, leading to systematic deviations in non-native productions (e.g., native English speakers produce a larger number of short-lag voiced stops in French than native French speakers). Many factors are thought to influence this interaction, such as age of acquisition and similarity between L1 and L2 phones (Piske, MacKay, & Flege, 2001). However, very little research has considered how the interaction is modulated by different processing conditions. In this study, we examine whether differences in processing across tasks influence the accuracy with which bilinguals articulate L2 sounds.

Evidence of task effects have been found in research with stutterers and individuals with Parkinson's disease, where low demand tasks (e.g., repetition) elicit more intelligible (e.g., Kempler & Van Lancker, 2002) and less errorful (e.g., Saltuklarogly, Dayalu, Kalinowsky, Stuart, & Rastatter, 2004) speech than high demand tasks (e.g., conversation). Effects of task have also been found in research with unimpaired populations. In an EMG study, Riès, Legou, Burle, Alario, and Malfait (2012) demonstrated that motor execution times (the duration between facial muscle activation to verbal onset) were longer in picture naming vs. repetition, suggesting that articulatory processing is sensitive to task demands.

To examine the influence of different processing demands on bilingual accentedness, we compare VOT elicited by picture naming and word repetition, which have been shown to impose different cognitive demands on speakers (e.g., Ferrand, 1999; Riès et al., 2012). If processing context influences the ability to produce a foreign language, a low demand task like repetition would allow speakers, even those that have no experience with that language, to focus on careful articulation (i.e., more French-like VOT, or more pre-voiced tokens). However, when bilinguals perform a high demand task like picture naming, they may not have the resources available to carefully articulate their L2 (i.e., more English-like VOT, or more short-lag tokens).

Twenty-five English(L1)-French(L2) bilinguals who recently returned from studying abroad in France, eight French(L1)-English(L2) bilinguals living in the United States, and eight monolingual English controls participated in this study. The two groups of bilinguals named 27 pictures and repeated 36 French words beginning in voiced and voiceless labial and alveolar stops, while the monolingual controls performed only the repetition task.

The figure below illustrates that English-French bilinguals produced more short-lag stops than French-English bilinguals in both tasks. Both groups of L1 English speakers performed similarly in repetition, suggesting that accurate repetition is possible regardless of language background. Crucially, we found that English-French bilinguals, but not French-English bilinguals, used short-lag voiced stops significantly more in picture naming than repetition. For voiceless stops, no reliable effects were obtained, likely due to low statistical power.

Our results indicate that a bilingual's ability to access L2 articulatory representations varies across processing contexts, and are consistent with previous findings with monolinguals, where different modes of processing elicit quantitatively different speech. These results will be discussed in the context of both exemplar theory and cascading activation.

Liquid Glide Clusters in Picard: What They Reveal about the Liquids

Ryan Hendrickson, Indiana University

One of the fundamental goals of phonological analysis is to group segments into classes based on shared features and/or behaviors (Mielke, 2008). While researchers generally agree upon certain classes due to a shared articulatory trait (e.g. nasals are [nasal]), there is more uncertainty for others, such as liquids, which do not have a clear common feature. The liquid class most often includes the rhotic and lateral consonants of a language, which are more often grouped based on behavior than an articulatory feature. While these consonants do often pattern similarly, there are just as many examples of them patterning differently (Dickey, 1997; Colantoni & Steele, 2005; among others).

Liquid-Glide clusters in Picard, a Romance language variety spoken in northern France and southern Belgium, provide numerous examples of both similarities and differences between the liquids. In (1), we see that liquid-glide onsets are permissible, however (2) shows that they are variably avoided, triggering epenthesis[1]. The examples in (3) suggest that the marginality of /LG/ clusters is due to the closeness in sonority, as an occlusive-glide onset is permitted (3a), while nasal-glide onsets variably trigger epenthesis (3b-c).

1. a. *él rouet d'Angerla* [el.rwe] 'Angerla's spinning wheel' (*Picardiries*: 138)
   b. *j' n'oc- connouos point ni d'loin* [nid.lwɛ]… 'I know neither from afar…' (*Lettes*: 476)
2. a. *o volez été Roé?* [e.t**e**r.we] 'you want to be king?' (*Piéches*: 31)
   b. *un Homme dé Loi* [d**e**l.wa] (*Du tout insanne*: 38)
3. a. *qu'j'aime bien ch'poéyis-leu* [bjɛ̃ʃ.pwe.ji] 'that I love this country' (*Picardiries*)
   b. *pour mieu* [pur.mjø] 'for better' (*Ch'coin*: 22)
   c. *a iroait pétète é-miu!* [tɛ.t**e**m.jy] 'that would go perhaps better' (*Piéches*: 18)

While (1-3) show similar behavior, a crucial difference arises in /lj/ clusters (4-5). In (4), we continue to see variable epenthesis in /rj/ onsets, while in (5), we see variable deletion in /lj/.

4. a. *éj n'ai pu rien dit* [pyr.jɛ̃.di] 'I couldn't say anything else' (*Lettes*: 164)
   b. *i n' gaingne érien* [gɛɲ.**e**r.jɛ̃] 'he wins nothing' (*Lettes*: 184)
5. a. *is racache'te chés ieuves* [ʃe.jøv] 'they lure the hares' (*I nn'est qu'd'és croére!*, CD)
   b. *aveuc eune mémouère dé liève* [d**e**l.jɛv] 'with a hare's memory'(*Ch'coin* 10)

This asymmetry suggests that /lj/ clusters are less well formed than /rj/ clusters. An analysis based solely on sonority cannot adequately account for the data, given that the rhotic is often analyzed as more sonorous than the lateral (Côté, 2004) and would predict its deletion as well. In order to account for this difference, an OCP constraint on two adjacent [coronal] consonants is invoked. This accounts for the deletion of the lateral before a /j/ but not before a /w/. This crucially assumes that /r/ is underspecified for place, despite its coronal realization – an analysis supported by Lindau (1985), among others. However, (6) shows that other coronal consonants followed by /j/ cause epenthesis rather than delete, complicating the analysis.

6. *inne é-tchiène* [et**ʃ**.jɛn] 'a dog' (*Lecat* 73:24)

I therefore propose that a highly-ranked combined OCP constraint against shared place features and close sonority in adjacent consonants is responsible for the deletion of the lateral before /j/. This analysis makes it possible to view /r/ and /l/ as members of a natural class by attributing their differences to the place feature specific to the lateral. Finally, the high ranking of this constraint is supported by variable lateral deletion in word-internal /lj/ clusters (7), showing that the process is not bound to one specific domain of application.

7. *cavailleu* [ka.va.jø], *cavalieu* [ka.val.jø] 'dance partners' (*Vasseur*)

Phonological awareness and conventionalization in sound change

José I. Hualde, University of Illinois at Urbana-Champaign

In the standard neogrammarian view a distinction is made between regular, biomechanically-induced, sound change and psychologically-based analogy. In a sense, however, all sound change has a psychological aspect, even when its origin is in biomechanics, since at some point phonological recategorization is required for sound change to take place (e.g. /p/ > /b/). In Labovian sociolinguistic research a distinction is also made between change from below and change from above related to speakers' awareness.

In this presentation I will consider the role of phonological awareness in regular sound change drawing from my recent acoustic research on intervocalic consonant lenition in a number of languages (including Spanish, Italian and Basque). I will argue that, at an initial stage, lenition applies as the neogrammarians envisioned: across morphological boundaries and without regard to lexical identity. At this initial stage the process may be below speakers' consciousness, and yet may operate as a conventionalized reductive process in the speech community, beyond biomechanical reduction. A number of factors may cause awareness of the phenomenon and its phonologization. It is at this stage that word- and morpheme-boundaries start to matter as conditioning environments and we also find lexical effects.
This research has also revealed the existence of important individual differences in phenomena such as intervocalic consonant voicing, correlated in part with the sex of the speaker. I will discuss the possible eventual conventionalization of sociolinguistic variation from biomechanic biases in lenition processes perhaps through the social construction of these individual differences in speech.

Reexamining the Nature of English Vowel Nasalization

Ho-Hsin Huang, Michigan State University
Karthik Durvasula, Michigan State University

This paper reexamines the nature of vowel nasalization in English. It specifically aims to tease apart whether the nasalization of vowels adjacent to nasal consonants in English is a controlled phonological process or is an uncontrolled/automatic phonetic process (Solé 2007). We examine the process by analyzing the nasalization patterns in acoustic tokens of English vowels in various nasal contexts. Our results broadly support Solé's claim that vowel nasalization of English is a controlled process rather than an automatic phonetic process.

Cohn (1993) observes that, in English, there is a substantial amount of nasalization in vowels adjacent to nasal consonants. It is possible to account for this nasalization in English through at least two different mechanisms. On one hand, vowel nasalization could be a phonetic/coarticulatory effect from the adjacent nasal consonant (1a). On the other hand, it could be a result of a controlled/phonological process that involves the nasal gesture being consistently co-ordinated with the vowel gesture(s) (1b). If the nasalization of the vowel is due to coarticulation (1a), a consistent *duration* of nasalization is expected irrespective of vowel duration. However, if the nasalization is due to a controlled phonological process (1b), then a consistent *ratio* of vowel nasalization to vowel duration is expected, i.e., the duration of nasalization is expected to increase with vowel duration.

In the current paper, we use Solé's experimental paradigm of varying speech rates to obtain a cross-section of vowel durations, and thereby test if the nasal gesture has a consistent durational manifestation or a consistent ratio (of total vowel duration) manifestation. However, unlike Solé, we attempt to study the nasalization patterns through purely acoustic methods. We ran a production experiment on 6 male American English speakers from Michigan. The participants were prompted to pronounce 12 test items and 24 fillers in the carrier sentence, "I will say _____ soon" at three speech rates (careful, conversational, and fast) to induce a good amount of variation in vowel durations. The syllable structure of all test items and fillers was CVN, where C=[b, d, g], V=[a, ʌ], and N=[ŋ, m], e.g. 'bomb'. The recorded tokens were analyzed using Praat (Boersma & Weenink 2013). The onset of nasalization was identified with abrupt changes in formant intensity, the appearance of anti-formants, and the appearance of the nasal pole (F_{P1}). The duration of nasalization and the whole vowel were recorded for each token.

Our data show that the onset of nasalization in the vowel is close to the onset of the vowel. The nasal portion for both the vowels ([a] and [ʌ]) is more than 50% in all the test items. An analysis of the duration of the oral and nasal portions of the vowel shows that the duration of nasalization is not the same for different vowel durations (2). Instead, there is a steady increase in nasalization with an increase in total vowel duration [$r(214)=0.985$, $p<2e-16$], thereby lending support to Solé's claim that the nasalization is controlled and phonological. However, a closer look at the data suggests that the support for the phonological mechanism is also not clean. The percentage of nasalization appears to increase with increase in vowel duration (3) [$r(214) = 0.335$, $p<8e-8$].

Although our results are slightly different from Solé's, we still argue that vowel nasalization in English is a controlled/phonological process. This is due to two reasons observed in the data: (a) The duration of nasalization varied linearly with vowel durations, (b) The nasalization of the vowel consistently started close to the onset the vowel. However, the results raise the need for a more nuanced view of how controlled/phonological processes vary with speech rate than has been previously appreciated.

Phenomenal Accent and Metrical Accent

Brett Hyde, Washington University
Shannon Howell, Washington University

In theories of musical rhythm, such as Lehrdahl and Jackendoff 1983, there is a distinction between phenomenal accents and metrical accents. Phenomenal accents are abrupt changes in pitch, duration, volume, etc. Metrical accents are the points where different metrical levels intersect. The distinction allows theories of musical rhythm to account for the perception of meter when, as is often the case, there is a mismatch between accented notes and strong metrical positions. Such mismatches occur, for example, when accented notes are a subset of strong metrical positions or when accented notes occupy weak metrical positions (syncopation).

While there a few cases where there is evidence for foot structure in the absence of acoustically accented syllables (e.g. Seminole/Creek and Cairene Arabic), theories of metrical stress typically assume that metrically strong syllables are distinguished in some way that can be measured acoustically. Recent acoustic studies challenge this assumption. While native speakers of Spanish report a pattern of secondary stress (Harris 1983, Hyde 2012), acoustic analyses have failed to detect stress on the relevant syllables (Diaz-Campos 2000). Similarly, while it has been long accepted that Polish has a pattern of secondary stress (Rubach and Booij 1985), a recent study was unable to detect secondary stress acoustically (Newlin-Lukowicz 2012).

Incorporating the distinction between phenomenal accent and metrical accent into the theory of metrical stress can help to account for the discrepancy between acoustic analyses and the perceptions of native speakers. Acoustic analyses detect only phenomenal accents, while native speakers can make judgments reflecting metrical accents.

In this talk, we address the implications that a distinction between phenomenal accent and metrical accent has for the typology of stress patterns. Drawing on empirical studies of the perception of rhythm in music (Parncutt 1994, Barnes and Jones 2000), we argue that single and dual stress systems should be reanalyzed as binary stress systems. Studies of musical rhythm show that metrical analyses appear to be both automatic and rich with the preferred ratio between metrical levels being 1 to 2. Single and dual stress systems imply impoverished metrical analyses that switch between dispreferred metrical ratios.

We also demonstrate that, given appropriate assumptions, one or two phenomenal stresses per form are sufficient to distinguish between metrical patterns. The key assumption is that clash and lapse are avoided unless required by the distribution of phenomenal accents. As a result, the reanalysis of single and dual stress systems as binary stress systems is much more straightforward than one might expect. It is typically the case that fixing the position of the primary stress is sufficient to determine the position of secondary stresses.

Social aspects of loanword adaptation: Adaptation of English word-final stops into Korean

Harim Kwon, University of Michigan – Ann Arbor

When English words with a postvocalic word-final plosive are borrowed into Korean, the vowel /ɨ/ is variably inserted after the plosive. This vowel epenthesis is variable in three ways: not all words take the epenthetic vowel (inter-item variation), some words variably take the vowel (intra-item variation), and not all speakers classify words in the epenthetic/non-epenthetic/ variable in the same manner (inter-speaker variation). Previous research has shown that the inter-item variation is conditioned by the phonological properties of the words being borrowed: vowel insertion rate is higher (i) when the final plosive is voiced than when it is voiceless (e.g. *kid* → kʰitɨ but *kit* → kʰit), (ii) when the pre-final vowel is tense than when it is lax (*beat* → pitʰɨ but *bit* → pit), and (iii) when the plosive is coronal than non-coronal (*mat* → mɛtʰɨ but *Mac* → mɛk, map → mɛp) (Kang 2003, Kwon 2012). Kwon (2012) claims that inter-speaker variation at least partially originates from individual borrowers' perception of English coda release. Listeners with different English experience perceived the word-final coda of English non-words differently, which leads to the different adaptation pattern among individuals. This study examines how Korean speakers with various English experience adapt actual English words into Korean, and suggests that the inter-speaker variation in loanword usage is influenced by the speakers' attitude toward the two languages, Korean and English.

Near-monolingual Koreans and Korean-English bilinguals were tested in a written survey using actual English words ending on a coda plosive (e.g. *seat*, *gap*, *strike*). Korean case-marker allomorphy was used to elicit the data of interest from the participants. In addition to the survey on loanword usage, the participants also completed a language survey asking (1) age of arrival if they are in the US, (2) length of residence in the US, (3) which language they use in everyday life, (4) self-reported Korean fluency, (5) attitude towards foreign-accented Korean, (6) attitude towards foreign-accented English, (7) which language they want to teach to their children, and (8) which language they use when conversing with Korean-English bilinguals. 5-point likert scale was used for questions (3)-(8).

Results: Among the questions asked, (3) the language of everyday use, (6) attitude toward foreign-accented English, and (8) the language of use when conversing with a bilingual were significant predictors of the likelihood of vowel epenthesis. Participants were more frequently insert /ɨ/ if they use more English in their everyday life [$F(1,1841) = 22.266$, $p < 0.001$], if they think it is important to speak English without foreign accent [$F(1,1841) = 7.795$, $p < 0.01$], and if they use more Korean when they converse with Korean-English bilinguals [$F(1,1841) = 3.979$, $p < 0.05$]. In conclusion, the current results suggest that loanword usage is influenced by the speakers' attitude toward the source language as well as their native language. Although speech perception influences online adaptation process of loanwords, perception in a laboratory setting do not fully reflect what happens in real world.

Speaker- and Reader-based Account of Discourse Prominence in a Free Word Order Language

Tatiana Luchkina, University of Illinois at Urbana-Champaign

Prosodic prominence realized at the level of a phonological phrase has been traditionally interpreted as one way of signaling discourse-salient words which carry information in discourse that is new and important (Selkirk 1996, Calhoun 2010). In free word order languages, word order choice is motivated by the information structure of the surrounding discourse (Downing & Noonan, 1995); as a result, a word can enter discourse (1) positioned in-situ, (2) fronted, or (3) post-posed, relative to its canonical position. Ex-situ discourse prominence has been documented for languages that allow for a relative freedom of word order (Howarth 2010, Skopeteas et al. 2010) and may even require that a constituent moves out of its canonical position in order to be interpreted as discourse-prominent (e.g., Hungarian & Mohawk). The present study analyzes perception-production cues to discourse prominence in Russian, a language in which in- and ex-situ ways of marking discourse-prominent words are possible.

Acoustic-prosodic correlates of prominence, word order, and information structure in a miniature Russian corpus are explored in two perception experiments to determine (1) how variable prosodic manifestation of discourse prominence is as a function of the sentential position of the word; (2) whether in Russian non-canonical word order is sufficient to mark discourse-prominence, or whether concurrent prosodic marking of the dislocated constituent is required.

Native speakers of Russian (N=12) were recorded reading two published narratives. The texts were annotated for information (IS) by two native Russian-speaking linguists. The three information structure categories adopted from Calhoun (2005) were Theme (given information), Rheme (new information), and Mediator (information predictable from context). Acoustic-prosodic measures of F0, intensity, and duration of the stressed vowel were taken from the stressed syllable of each IS-coded word. The relationship between these acoustic measures and IS was then tested using multinomial regressions. We hypothesized that there is a significant and reliable differences in the acoustic measures as a function of (1) IS and (2) the sentence position of the word. The results of regression analyses reveal that intensity and duration of the stressed vowel of the IS-coded words reliably predict whether the word's IS category is Theme or Rheme. Intensity, duration, and peak F0 value of the stressed vowel predict the word position as canonical vs. fronted. Perceptually salient text-based and acoustic-prosodic cues to discourse prominence were determined in the reading comprehension and perceptual (auditory) comprehension tasks in which native speakers of Russian (N=86, 26 participated in the listening task) identified and rated discourse-prominent word(s) in each sentence in the narratives on the scale from 0 to 100. Results reveal that participants interpreted words which carry new information, are located at the juncture of intonation phrases, and/or are positioned ex-situ as highly discourse-prominent; in the listening modality, increased intensity and duration of the stressed vowels are also associated with higher discourse prominence scores.

We conclude that viewing sentential position of a word and acoustic marking as independent or complementary ways of signaling discourse prominence in Russian is problematic for our corpus because in the spoken modality, ex-situ constituents are consistently marked acoustically.

We will discuss these findings in light of the Multiple Source view of discourse prominence (Watson et al. 2008 & Watson 2010) and Scrambling Complexity hypothesis (Sekerina, 1996).

Who's talking now? Infant processing of vowels with infant vocal tract parameters

Matthew Masapollo, McGill University
Linda Polka, McGill University
Lucie Ménard, University of Québec at Montréal

To develop normal speech production abilities, infants must perceptually match their own vocalizations to the speech patterns they hear produced by their adult caregivers. In this study, we explore how infants respond to infant vowels and how this impacts their ability to recognize infant vowel productions as phonetically equivalent to vowels produced by adult and child talkers. The results indicate that pre-babbling (4-to-6-month-old) infants have some ability to match infant-produced vowels to adult and child productions before they are able to produce the target vowel sounds. Infants are also very responsive to talker changes, display a strong attraction to infant speech signals, and detect phonetic differences more quickly for signals emitted by an infant talker than an adult talker. These findings help us better understand speech processing skills that support early vocal learning.

Prosodic patterns in distal speech context affect perception of prosodic constituency and word segmentation: effects of list intonation and f0 downtrend

Tuuli H. Morrill[1,2], Laura C. Dilley[1,2,3] & J. Devin McAuley[2]

[1]Department of Communicative Sciences and Disorders, Michigan State University, [2]Department of Psychology, Michigan State University, [3]Department of Linguistics and Germanic, Slavic, Asian and African Languages, Michigan State University

Perception of prosodic structure – particularly phrasal boundaries - is generally assumed to be mediated by phonetic variation local to the prosodic boundaries themselves. Phonetic variation associated with local prosodic phrasal boundaries includes articulatory strengthening and lengthening (Byrd & Saltzman, 1998; Duez, 1982; Fougeron & Keating, 1997; Keating, Cho, Fougeron, & Hsu, 2003; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). However, prosody is also described as showing repetitions or perceived regularities in tonal and stress patterns that may span multiple successive prosodic domains (Couper-Kuhlen, 1993; Lehiste, 1977; Pierrehumbert, 2000). Recently, it has been shown that prosodic regularities in the distal (non-local) context can influence the segmentation of syllables into words, and thus their perceived prosodic constituency (Dilley, Mattys, & Vinke, 2010; Dilley & McAuley, 2008). In the current study, three experiments tested which types of distal prosodic patterns elicit these effects. In particular, we tested all of the following kinds of distal prosodic patterns: (1) patterns consisting of regularity in only timing cues, only f0 cues, or both (Experiments 1-2), (2) items ending with either a low-high (Experiment 1) or a high-low (Experiment 2) tonal pattern, and (3) tonal patterns showing a progressive f0 decrease, i.e., a 'downtrend' (Experiment 3).

Experimental items were eight-syllable sequences beginning with two disyllabic, trochaic words (e.g., "magnet guilty") followed by four syllables that were ambiguous with respect to their lexical affiliation (e.g., *crisis#turnip* vs. *cry#sister#nip)*. Each experiment used two distinct distal prosodic patterns resynthesized on the initial five syllables of experimental items; these patterns favored grouping of final syllables into prosodic constituents such that listeners were predicted to hear a final disyllabic word (e.g., *turnip*) or monosyllabic word (e.g., *nip*), with the acoustic characteristics of the final three syllables identical across conditions. The proportion of final disyllabic word responses was determined for each condition. Results (presented in Figures 1-3) show that, as predicted, in each of the experiments, distal contexts influenced segmentation of the final syllables, even though sequences with identical acoustic characteristics were segmented in each condition. Disyllabic responses occurred a significantly greater proportion of the time with the disyllabic distal context than with the monosyllabic distal context. Moreover, distal prosodic effects on segmentation could be obtained using temporal cues only; the strongest effects were observed in the combined pitch and temporal cues condition (as measured with *d'* in a signal detection analysis). These results reveal that perception of prosodic structure does not depend only on local phonetic modifications (e.g., articulatory strengthening); rather, a variety of prosodic patterns in distal context can influence perceived prosodic constituency as well. These findings shed light on processes associated with perceiving prosodic structure while further elucidating a new, powerful factor that has recently been identified to influence segmentation of words from continuous speech.

A Morpho-Prosodic Typology of Compound Truncation

Michael Opper, University of Michigan

Compound truncation refers to a word-formation process by which at least one of the constituents in a compound word is truncated (as in English *sitcom* which is derived from *situational comedy*). This study introduces a typology of the mapping between syntactic constituency and prosodic domains in compound truncation focusing on compound truncations with two syntactic constituents. In each of the languages examined in this study – (Standard) Chinese, Japanese, and English – compound words which have non-truncated constituents do not violate syntactic-prosodic constituency correspondence (Selkirk 2006, Selkirk 2011); each constituent is both a syntactic word and a prosodic word. However, in compound truncation, the possible mappings between the two syntactic constituents and prosodic units are listed in (1).

Chinese utilizes mappings (1a-b) in compound truncation. A large percentage (60-70% depending on style) of the Chinese lexicon exhibits the property of "Word length flexibility" – the property in which a given word has a monosyllabic short form and a disyllabic long form (Huang & Duanmu 2012). A consequence of this property is that many compound words have semantically-equivalent variant phonological outputs. Each of the aforementioned mappings is observed in the "flexible word" compounds in (2) which are taken from Duanmu 2012.

Compound truncation is very productive in Japanese (Itô 1990, Nishihara et. al 2001). Japanese compound truncation is observed for Sino-Japanese compounds (3a) and for English-based compounds (3b). Both base compounds and compound truncations map two syntactic constituents to two minimal phonological words – mapping (1c).

English compound truncations consist of two separate truncated syntactic constituents which are each their own prosodic words with no size restrictions – mapping (1d). Lack of resyllabification in English compound truncations provides evidence for syntactic-prosodic independence; consider the compound truncation *phys-ed* which is syllabified as [fɪz.ɛd] and not [fɪ.zɛd]. Compound truncation has been mentioned, but not accounted for in the English word formation literature (Bauer 1982, Plag 2003). This study intends to contribute to our understanding of English compound truncation through data compiled from an experiment. In the experiment, 76 native English speakers were each recruited through Amazon Mechanical Turk to create 700 compound truncations of semi-random compounds. The data compiled suggest two relevant trends regarding English truncation. First, there is no set process for truncation in English. For instance, there is a substantial amount of variation with regard to the total number of syllables for the truncated modifiers in our dataset - 29.3% are monosyllabic, 61.1% are disyllabic, 9% are trisyllabic, and 0.6% are longer than three syllables. Second, the hypothesis of prosodic mutual-independence of constituents in compound truncation is enhanced by the discovery that the number of syllables in one constituent of a compound truncation does not impact the number of syllables in the other constituent (through the results of a LMM).

This study provides evidence that the mapping between prosodic units and syntactic constituents in compound truncations differs across languages. Chinese compound truncations with two syntactic constituents are contained within the domain of a single prosodic word – mappings (1a-b). In Japanese, each constituent in compound truncations is a quantity-insensitive foot (minimal word) – mapping (1c). Finally, in English, each constituent is of variable prosodic size and shape – mapping (1d). These three languages demonstrate each of the theoretically logical syntactic-prosodic mappings for compound truncation.

Contrastive features for the English vowel system

Thomas Purnell, University of Wisconsin–Madison
Eric Raimy, University of Wisconsin–Madison
Joseph Salmons, University of Wisconsin–Madison

An outstanding problem for theoretical phonology is accounting for structured heterogeneity in speech variation (Weinreich et al. 1968:149). *Hypermodularity* (Carruthers 2008, Purnell, Raimy & Salmons in preparation) posits more grammatical structure, providing more distinct niches as sources of explanation. This paper presents an analysis of Modern American English (MAE) dialects based on proposals by Avery and Idsardi (2001) and Dresher (2009). We develop abstract representations to support lexical access, to demarcate phonetic vs. phonological aspects of MAE vowels and to provide the structural basis for dialectal variation. A necessary aspect of hypermodularity is the parceling out of explanation of different phenomena to different levels of representation and/or interactions among levels of representation.

Avery & Idsardi's (2001) model of distinctive features is more articulated than other contemporary models (e.g. Halle 1995, Halle et al. 2000) with respect to both physiology and abstraction. Physiologically, it recognizes that many gestures are oppositional in nature (e.g. a segment can not be [+high] and [+low] or both [+round] and [-round]). An abstract *dimension* node is added to distinctive features, capturing the oppositional nature of gestures. Avery & Idsardi propose that (narrow) phonological specification occurs at the level of *dimensions* not *gestures.* This entails that one main function of phonology is to specify abstract dimensions with concrete gestures. We combine these principles with Dresher's (2009) *successive binary division* (SBD) algorithm to further limit the structure posited in (narrow) phonological representations. Dresher's SBD posits that phonological inventories are divided in a binary fashion with one of the resulting groups being 'marked'. We interpret 'marked' to mean being given structure (Dresher, Piggott & Rice 1994) in line with Avery & Idsardi's principles. Finally, we follow Hall's (2011) proposals on how phonological contrast is mapped to phonetic categories taking into account dispersion and enhancement (Keyser & Stevens 2001, 2006). We set forth an analysis of contemporary English vowels on these principles.

A hypermodular analysis accounts for phonological differences among major MAE shifting vowel systems (Figure 1): Canadian (Roeder, in press), Northern Cities (Labov et al. 2006) and Southern Cities (Sledd 1966) shifts. We focus here on the short vowel features for /ɪ/, /ɛ/ and /ʊ/. Specifically, the hierarchy for the standard system resembles the long vowel hierarchy, where Tongue Thrust dominates Tongue Height. If we take MAE as a basis for comparison, the NCS and Canadian systems differ by employing Labial actively. Moreover, Southern Cities and Canadian increase specificity by adding Tongue Root features. Phonetic drift (Rice 2009) and variation can be accounted for by phonological underspecification. The consequence of synthesizing the Avery & Idsardi / Dresher approaches to features with hypermodularity is that phonology has a seat at the table alongside sociophonetics in the effort to understand and model structured heterogeneity.

Characterizing the phonological context of roots and affixes

Anne Pycha, University of Wisconsin, Milwaukee

Phonology presents us with the fundamental challenge of defining contexts. Of course, we do this whenever we specify the environment in which a specific process takes place (e.g., "in word-final position"), but recurring cross-linguistic patterns suggest the need for deeper characterizations. For example, Hyman (2008) and others have pointed to the odd status of prefixes: unlike suffixes, prefixes rarely trigger changes on roots, and this asymmetry holds for a puzzlingly diverse set of processes, including local assimilation, long-distance vowel and consonant harmonies, and vowel elision. Could we offer a characterization of the "prefix context" (as well as "root" and "suffix") that would encompass this diversity and account for the recurrence of this asymmetry?

I address this question with an experimental paradigm designed to address whether listeners perceive – and remember – prefixes, roots, and suffixes differently. I investigate both perception and memory because they interact continuously during speech processing, as listeners compare incoming speech with their stored representations (McClelland & Elman 1986, Poeppel & Monahan 2010). At the same time, perception and memory behave differently: for example, although listeners recognize high-frequency words more quickly than low-frequency words, they remember them more poorly because they lack distinctiveness (Joordens & Hockley 2000). Since affixes have token higher frequencies than roots, we expect the same logic to apply; the broader point is that a complete characterization of context must take into account both on-line perception and stored representations.

Participants completed a study phase in which they listened to seventeen affixed words presented in the clear, such as *reshred* and *clownship.* They then completed a test phase in which they listened to thirty-four words partially overlaid with white background noise, and rated the loudness of the noise on a subjective scale from 1 (softest) to 5 (loudest). At test, half of the words were old, having been already presented at study, and half were new. Furthermore, half of the words had noise overlaid only the root, while half had noise only on the affix. As a control, an identical procedure used simple words with pseudo-affixes, such as *regret* and *worship.*

Study (listen and identify):   reshred clownship bewhirl bleakish …

Test (rate noise loudness):   bleakish unchalk bewhirl fudgeless reshred clownship …

Previous studies have shown that if people create a memory for a stimulus, they perceive it differently later on. In Goldinger, Kleider & Shelley (1999), participants gave lower ratings to background noise if they were listening to the same words in the same voice they had heard earlier; this is a perceptual illusion, because noise volume was equivalent across old and new. Thus, the key questions in the current study were a) whether listeners would give different noise ratings to prefixes, roots, and suffixes, which would suggest that they perceive these morphemes with different degrees of clarity, b) whether this effect was modulated for old vs. new words, which would suggest that listeners store memories for these morphemes in different ways.

110 subjects participated. Results indicate no effect of old versus new words, so we have no evidence that listeners store memories for morphemes in different ways. However, the results also indicate that participants rate noise on prefixed versus suffixed words very differently, demonstrating that prefixed words (and in particular, roots preceded by a prefix) provide more perceptual clarity overall than suffixed words do. This finding could provide an intriguing solution to the problem of the alternation asymmetry: post-prefix roots provide a "context" that is uniquely perceptually clear, therefore uniquely resistant to listener mis-perception and the development of alternations (Ohala 1990), while pre-suffix roots enjoy no such special status.

On the Developmental Path of Final Consonant Cluster Acquisition in Cairene Arabic

Marwa Ragheb, Indiana University
Stuart Davis, Indiana University)

The acquisition of consonant clusters (CC) in Cairene Arabic (CA) poses an interesting challenge. As (1a-c) show, CA permits final CCs of any two phonemes regardless of sonority profile. Furthermore, CA has final geminate consonants (1d). Given the range of final consonant clusters in CA as exemplified in (1), along with the difficulty that children acquiring English have in mastering final clusters (McLeod et al. 2002), it would be hypothesized that final clusters would be difficult for children acquiring CA to master. Moreover, if they are difficult to master, a sonority effect can be hypothesized in which clusters of falling sonority should be more easily mastered (i.e. occur earlier in acquisition) than clusters of rising sonority, given that they are much more common across languages. In this talk we present our preliminary work on the developmental path of L1 CC acquisition in CA, a topic not previously studied. We examine the CC acquisition in CA based on a fairly detailed study of two children, brother and sister, 3 years apart in age, living in a monolingual setting in Cairo, Egypt. Data from the first child, MG were elicited when he was 2 years 8 months old (2:8) and gathered over a one-month period. Data from RG, his sister, was elicited three years later over a longer period of time, starting when she was also 2:8. Data were elicited by use of a probe specifically designed to elicit word-final CCs. Spontaneous data was also recorded. Sessions ranged from 30 minutes to 1 hour, and were conducted multiple times during the observation period. By examining the data obtained from both children, we can demonstrate a developmental path for CA cluster acquisition. We focus on the two children when both were at 2:8. Representative data from MG and RG are shown in (2) and (3), respectively. While both children were at the same age, RG had more target appropriate clusters. The comparison of the two allows us to posit a developmental path of cluster acquisition in CA. First, both children had target-appropriate word-final geminates and pharyngeal-initial CCs. These were the only target appropriate final clusters observed for MG. RG also had both these cluster types in a target-like manner at 2:8; in addition she had acquired nasal-obstruent clusters and sibilant-obstruent clusters. Given our data, we posit the general developmental path of CC acquisition for CA in (4). A closer look at the data shows that acquisition does indeed follow a sonority profile, with some nuances. First, geminates seem to be acquired very early, as in many languages that have phonemic geminates (Ota 2002 for Japanese). Second, pharyngeal consonants have been shown to be phonetically high sonority glides in CA (Elgendy 2001), thus any pharyngeal-initial final CC would be a falling sonority cluster; nasal-obstruent clusters are also acquired early but not nasal-sonorant clusters. Further, the nasal-obstruent clusters always show up as assimilated in the child data even if unassimilated in the target; assimilated nasal- obstruent clusters can be considered an unmarked type of falling sonority cluster since they entail just a single place gesture. The early acquisition of sibilant-obstruent clusters is interesting in light of their sometimes early acquisition in English (Barlow 1997) and their special status in other languages (Goad 2011). RG treats sibilant-obstruent clusters as special in acquisition even though they do not have phonotactic properties that distinguish them from other clusters in CA.

Effects of autistic traits on differences in temporal reduction and perceptual compensation

Rory Turnbull, Ohio State University

It is well known that words undergo temporal reduction in contexts of high semantic predictability (Lieberman, 1963). This effect has been hypothesized to exist due to communicative pressures on language production (Lindblom, 1990). According to these accounts, temporal reduction is licensed in predictable contexts, when it is less likely for a listener to misparse the meaning of an utterance. In unpredictable contexts, the speaker employs temporal enhancement to be sure that the message is received by the listener.

Nevertheless, studies of these kinds of predictability effects have typically observed high degrees of variance between subjects in the extent and size of the effects (e.g. Clopper and Pierrehumbert, 2008). The current study attempts to explain this variance in terms of individual differences in cognitive processing style—particularly, 'autistic' traits.

The effects of autism and autistic traits on language processing have been receiving increasing attention recently. One such study is Yu's (2010) investigation into perceptual compensation and autistic traits. Perceptual compensation is a phenomenon where listeners take into account various contextual factors in their perception of speech sounds. For example, /s/ is more /S/-like (i.e. has a lower frequency) when in the context of rounded vowels. It has been shown that listeners have a tacit knowledge of this, and shift their /s/ – /S/ category boundaries in the context of rounded vowels (Mitterer, 2006). Yu found that participants with more autistic traits exhibited more perceptual compensation than participants with fewer autistic traits.

In this study, we replicate the main results of Yu's study, and provide evidence of an individual differences effect in language production. Participants were recorded reading matched pairs of high- and low-predictability sentences (Kalikow et al., 1977). The final word of each sentence is either predictable or unpredictable given the preceding context—for example, "For your birthday I baked a cake" vs. "Tom was thinking about the cake". Subsequent to the reading task, participants took part in an /s/ – /S/ identification task, designed as a complete replication of Yu (2010). Stimuli consisted of four /sV/ – /SV/ continua, with the vowels /a, u/ from a male and a female talker spliced in; the continua were constructed along a seven-step series. The task was forced-choice identification. After the two linguistic tasks, participants completed an autistic traits questionnaire (Baron-Cohen et al., 2001). The durations of sentence-final target words were measured, and participants' responses to the identification task and questionnaire scored.

Results from the identification task replicate those of Yu (2010): participants who scored higher on the questionnaire (i.e. exhibit more autistic traits) generally exhibited more perceptual compensation than the lower-scoring participants. Results from the production task suggest a similar story: participants who scored higher on the questionnaire exhibited greater temporal differences between the high- and low-predictability conditions (see Figure 1). This study demonstrates that individual differences in autistic traits can affect language production as well as perception, and provides a basis for future research on individual differences in predictability effects.

Accounting for Metathesis in Highland East Cushitic

Coral Williams, Indiana University

The phenomenon of metathesis has long posed interesting problems for phonological theory. Recent explanatory accounts of metathesis include Blevins & Garrett (2004), who propose a typology of phonetic motivations for metathesis, and Hume (2001), who proposes as motivation the interactions between the constraint LINEARITY and an array of contextual markedness constraints targeting specific segments involved in their dispreferred context. In my paper, however, I argue that an analysis of the metathesis process found in a group of three Highland East Cushitic languages, Kambaata, Alaaba, and Sidaama, reveals a flaw in both Blevins & Garrett's and Hume's accounts.

In these three HEC languages, there is a conspiracy of three processes, epenthesis, assimilation, and metathesis, which work together to avoid illicit consonant clusters resulting from verbal inflection. These languages do not allow complex syllable boundaries, and preferred consonant sequences include geminate consonant, sonorant-obstruent sequences, and glottal stop-sonorant sequences. Each process in the conspiracy targets a different type of illicit cluster: epenthesis applies to avoid complex syllable boundaries, assimilation applies to avoid non-geminate sequences which do not consist of both a sonorant and an obstruent, and metathesis applies to avoid illicit obstruent-sonorant or sonorant-glottal stop sequences. As a side effect of the affixes used in verbal inflection, all cases of metathesis are obstruent-nasal metathesis.

The overarching generalization is that these languages prefer consonant sequences which are either geminate or falling sonority sequences, and the metathesis process enforces this preference in the context of verbal inflection. This generalization, however, would be missed by both the approach of Blevins & Garrett and that of Hume. Blevins & Garrett, in fact, argue that obstruent-nasal metathesis should not exist as a synchronic process, but as a result of analogical historical change. With no evidence of motivation for such an analogical change in these HEC languages, Blevins & Garrett's claims about metathesis can neither positively account for the phenomenon nor be disproved, which makes this position difficult to support.

Hume's approach, on the other hand, could conceivably account for the metathesis observed with a contextual markedness constraint which targets obstruents preceding sonorant consonants, *OBS/__SON. This would effectively behave as a Syllable Contact constraint, and capture the dispreference for rising sonority clusters. However, Hume's family of contextual markedness constraints comprises constraints which target "all segments in all observed contexts" (Hume 2001). This creates a very powerful family which allows for the existence of phonetically unmotivated constraints, such as *C/__V which would effectively ban onsets, which do not need to exist and therefore do not need to be proposed.

As seen above, an analysis of the metathesis pattern found in a group of three Highland East Cushitic languages reveals flaws in the approaches of Blevins & Garrett (2004) and Hume (2001), two recent accounts which propose global motivation for the phenomenon of metathesis. Blevins & Garrett's approach would leave the metathesis pattern unaccounted for due to lack of evidence, and Hume's powerful family of constraints leaves too much leeway for inventing constraints. Instead, the consonant metathesis in these HEC languages is best analyzed as the emerging effect of the preference for falling sonority consonant sequences.