Nederlandse Vereniging Voor Fonetische Wetenschappen

http://www.fon.hum.uva.nl/rob/FonetischeVereniging/

Abstracts

Intonational realisation of topic and focus in child Dutch

Aoju Chen, Max Planck Institute for Psycholinguistics, Nijmegen

Languages express the topic-focus distinction in different ways (e.g. word order, particles, intonation). In this study we focus on intonation and in particular look at type of pitch accent and phrasing. There has been relatively little discussion on the use of intonation to express the topic-focus distinction in child language. Prior work is mostly concerned with the use of accentuation in expressing contrast. In this study, we examined how Dutch children use pitch accent types (including deaccentuation) and phrasing to mark topic and focus in different sentence positions and how they differ from adults. The topic and focus under investigation were non-contrastive and realised as full NPs (e.g. what did the boy draw? [The boy]$_{topic}$ drew [a castle]$_{focus}$).

A picture-matching game was used to elicit topic-focus structures as answers to WH-questions. Two variables were controlled for in the answer sentences: PRAGMATIC CONDITION (topic, focus), SENTENCE POSITION (initial, final). Data were collected from monolingual Dutch children (aged 4-5 years, 7-8 years and 10-11 years) as well as adults. The intonation patterns were transcribed following ToDI notation.

Our analysis has revealed five major findings:

1. Children of all age groups and adults employ a similar set of pitch accent types. These accent types (e.g. L*H, H*L, H*, !H*L) form the core of the inventory of pitch accents in Dutch. This finding thus shows that children as young as 4 have adult-like inventory of accent types.
2. Like adults, children of all age groups deaccent topic more frequently than focus independent of sentence position. This result indicates children's early sensitivity to the accentuation-focus and deaccentuation-topic associations, as suggested in prior work.
3. Children acquire H*L as the typical 'focus accent' at the age of 7 or 8. 4- to 5-year-olds exhibit a weak preference for H*L over other accent types in sentence-initial position and no preference for H*L in sentence-final position. Possibly, frequent use of H*L in sentence-initial topic in adult Dutch has made it difficult for young children to associate H*L primarily with focus.
4. Children also acquire deaccentuation as the typical 'topic intonation' at the age of 7 or 8. 4- to 5-year-olds realise topic similarly frequently with deaccentuation, L*H and H*L in sentence-final position.
5. Only 4- to 5-year-olds appear to use phrasing to realise topic, which forms its own intonational phrase (IP). Older children and adults utter the topic-focus structure mostly as one IP.

The prosodic marking of the contrast between restrictive and appositive clause in Dutch

Vincent van Heuven, Leiden University Centre for Linguistics

It is well known that there are prosodic differences between restrictive versus appositive relative clauses in English. In a current project we are in the process of establishing the phonetics of the prosodic marking of this contrast in Dutch, in both speech production and perception. In my talk today I will report results of a pilot experiment with a speech-technology orientation. We ask to what extent Dutch listeners are sensitive to the prosodic

marking of the difference between the two clause types, and by implication, if the quality of a text-to-speech system can be improved if the contrast is properly modelled in the prosodic component of the TTS.

We resynthesized Dutch sentences whose lexico-syntactic properties are compatible with either a restrictive or an appositive interpretation of a relative clause (but never both). We generated multiple versions of each sentence by systematically varying (i) presence versus absence of domain-final lengthening before the onset of the relative clause (ii) presence versus absence of a physical silence at the onset of the relative clause and (iii) the melodic configuration at the boundary. Four different melodies were used, which were hypothesized to mark the absence or presence of a prosodic boundary with different degrees of strength.

Listeners were ask to judge for each stimulus on a scale from 0 to 10 how well the speaker read the stimulus, where '0' stood for 'very poorly/inappropriately read' and 10 represented 'read perfectly').

Our results indicate that domain-final lengthening and the presence of a physical pause at the beginning of the relative clause are dispreferred for the restrictive type. Melody seems to contribute very little to the contrast.

Prosodic marking of information status in L1 and L2. A comparative study of Dutch and French

Laurent Rasier, Université Catholique de Louvain

In recent years quite a lot of attention has been paid to the suprasegmental features of speech, i.e. stress, accent, intonation, tone, rhythm, and speech pauses (Fox 2000). By contrast, the study of prosodic systems suffers from a considerable under-representation in the field of second language acquisition research. Indeed, most work devoted to L2 pronunciation has hitherto focussed on segmental issues (Chun 2002, Leather & James 1991, Rasier 2006).

In the first part of this talk, I discuss some recent research on interlanguage prosody. Then, I will set out to investigate L1/L2 speakers' use of prosody to signal information status in Dutch and French. The data consist in the L1 speech of 20 native speakers of Dutch and 20 native speakers of French on the one hand and in the L2 speech of 20 advanced French-speaking learners of Dutch and 20 advanced Dutch-speaking learners of French on the other (Rasier 2006). The corpus was gathered by means of an experimental accentuation test consisting in a picture description game in which the information value of target words was kept under control (see also Barlow 1998, Swerts et.al. 2002. The results show a strong transfer effect in the L2 learners' use of prosody (accentuation, deaccentuation) in order to indicate information status. It is shown that the differences between the native and non-native speakers of Dutch and French regarding the prosodic marking of information status can be explained in terms of the markedness relations (in the sense of Eckmann 1988) between Dutch and French.

Prosody - a missing link between phonetic detail and phonemic categories?

Claudia Kuzla, Max Planck Institute for Psycholinguistics, Nijmegen

The prosodic structure of an utterance influences the fine-grained phonetic details of spoken word forms. Speech sounds are articulated more strongly in prosodically prominent positions, e.g., at the onsets of higher prosodic domains.

Importantly, this prosodic strengthening involves phonetic characterisitcs of the speech signal which also cue phonological distinctions. While articulatory strengthening suggests that, e.g., a prosodically strong /b/ would become more similar to a [p] (what I will call the 'Fortition Account'), it has also been hypothesized that prosodic strengthening enhances language-specific distinctive features, which would make a prosodically strong /b/ less [p]-like. We investigated how domain-initial strengthening affects the acoustic cuees to phonemic contrasts in German obstruents.

Experiment 1 focused on the plosives /b, p, d, t, g, k/, which form pairs of phonemes differentiated by the fortis-lenis contrast. Important acoustic cues to this contrast are closure duration, glottal vibration during closure, voice onset time, and intensitiy of the release noise. Closure durations were longer At higher prosodic boundaries, closure durations were longer, and in lenis plosives, a smaller proportion of the closure was produced with glottal vibration. Voice onset time in lenis plosives was not affected by prosody. In contrast, VOT decreased at higher boundaries for the fortis plosives, as did the maximal intensity of the release. These results suggest that the effects of prosody on different phonetic cues can go into opposite directions, but are overall constrained by the need to maintain paradigmatic phonemic contrasts.

In Experiment 2, we examined how prosody constrains a sandhi process, the progressive assimilatory devoicing of the word-initial lenis fricatives /v, z/ folllowing /t/. Reduction in glottal vibration makes lenis fricatives more fortis-like (/f, s/). This devoicing was especially strong across small prosodic boundaries. However, prosodic structure affected the fricative duration, another cue to the fortis-lenis distinction, in the opposite direction. Duration was shorter at smaller boundaries, just as the closure duration of the plosives. Hence, at smaller prosodic boundaries, fricatives were more devoiced (more fortis-like), but also shorter (more lenis-like).

In conclusion, our results show that neither the Fortition Account nor the Feature Enhancement explanation are fully supported by the syntagmatic effects of prosody on individual cues. At the phonemic level, however, the paradigmatic fortis - lenis contrast remains distinctive in all prosodic contexts, irrespective of prosodic strengthening and assimilation.

Effects of dialect and context on the realization of German pitch accents

Bettina Braun, Max Planck Institute for Psycholinguistics, Nijmegen

We investigated whether alignment differences reported for Southern and Northern German speakers (Southerners align peaks in prenuclear accents later than Northerners) are carried over to the production of different functional categories such as contrast. To this end, the realization of non-contrastive theme accents is compared with those in contrastive theme-rheme pairs such as 'Sam rented a truck and Johanna rented a car.'

We found that when producing this 'double-contrast', speakers mark contrast both phonetically by delaying and rising the peak of the theme accent ('Johanna') and/or phonologically by a change in rheme accent type (from high to falling 'car').

The effect of dialect is complex: a) only in non-contrastive contexts produced with a high rheme accent Southerners align peaks later than Northerners; b) peak delay as a means to signal functional contrast is not used uniformly by the two varieties. Dialect clearly affects the realization of prenuclear accents but its effect is conditioned by the pragmatic and intonational context.

Music and lyrics: a comparative study of the tone-melody correspondence Thai and Cantonese popular songs

Vincie W.S. Ho, The University of Hongkong

Despite the close resemblance of language and music in terms of their acoustic nature, the relationship between lexical tone and melody in Thai music has received little attention. This study aims to address the issue whether Thai popular songs demonstrate close correspondence between linguistic tone and musical melody. A comparative approach will be adopted based on the findings of a previous study on the tone-tune interface in Cantonese popular songs (Ho 2006).

The Thai language possesses a rich tonal inventory which consists of five contrastive tones – high, mid, low, falling and rising. One may wonder if the lexical tones and the melody of a song need to match in order that the comprehensibility of the lyrics to be preserved. According to Ho (1998), an almost perfect tone-melody mapping was found in a traditional Thai song composed in the 60s. Not only global pitch movement of word phrases or stanzas but also localized tone contour of individual syllables is fully represented by the melody. Nonetheless, tone-melody mismatch appears to be more tolerated in recent works of Thai pop music. Saurman (1999) suggested that the degree of

correspondence is related to the genres of music and the year of production – a higher degree of tone-melody correspondence is found in classical and traditional songs, whereas contemporary pop songs especially those that are reinterpretation of western music seem to abide tone-tune divergence. In Cantonese popular songs, however, tone-tune mismatches are extremely rare regardless of their origin or the year of composition. Strict matching between musical pitch transition and tonal target transition is observed in all the Cantonese pop songs analyzed. With data from 20 Thai songs composed in the past few decades to present, this study attempts to reveal and account for the structural differences, if any, between the tone-tune correspondence of Thai songs and that of Cantopop.

Baart, J.L.G. (2004). Tone and song in Kalam Kohistani (Pakistan). Hugo Quene & Vincent van Heuven (eds.): *On Language and Speech: Studies for Sieb G. Nooteboom.* Utrecht: Netherlands Graduate Institute of Linguistics. 2004.

Chan, M. (1987). Tone and Melody in Cantonese. *Berkeley Linguistic Society, Proceedings of the Thirteenth Annual Meeting 1987.*

Ho, W.S.V. (1998). Tone and Tune: A Study on the Relationship of Text and Melody in Contemporary Vocal Music. *MA Dissertation, University College London, University of London.*

Ho, W.S.V. (2006) "The Tone-melody Interface of Popular Songs written in Tone Languages". Conference Proceedings of the 9th International Conference on Music Perception and Cognition, University of Bologna, Italy, 22-26 August, 2006.

Ho, W.S.V. (2006) "An Optimality-theoretic Analysis of the Tone-melody Mapping in Cantonese Popular Songs". *Conference Proceedings of the Second International Symposium on Tonal Aspects of Languages, University of La Rochelle, France, 27-29 April, 2006.*

Saurman, M.E. (1999). The agreement of Thai speech tones and melodic pitches. *Notes on Anthropology, 3/3, 15-24.* Dallas, TX: Summer Institute of Linguistics.

Wong, P.C.M., & Diehl, R.L. (1999). Melody-tone relation in Cantonese songs. *Journal of the Acoustical Society of America, 106: 2286.*

Prosodic information in audiovisual spoken-word recognition

Alexandra Jesse, Max Planck Institute for Psycholinguistics, Nijmegen

Prosodic information influences the spoken-word recognition process. For example, auditory lexical stress information contributes to the activation of word candidates during spoken-word recognition (Cooper, Cutler, & Wales, 2002; van Donselaar, Koster, Cutler, 2005; Soto-Faraco, Sebastián-Gallés, & Cutler, 2001). However, we typically do not only hear but also see speakers in conversations. Visual speech (i.e., information from seeing the face of a speaker) is known to contribute to the robust recognition of speech segments (e.g., for an overview, see Massaro and Jesse, in press). Segments are better recognized when presented as audiovisual than as auditory-only speech. But little is known about visual speech's ability to provide prosodic information. The project reported here will address whether visual speech informs about lexical stress and whether this information can alter lexical competition during the audiovisual spoken-word recognition process.

Dutch word pairs that overlap in their first two syllables segmentally but differ in lexical stress were selected (e.g., OCtopus vs. okTOber; capital letters marking primary stress). In an audiovisual speech version of a cross-modal repetition priming task, the first two syllables of these pairs were presented sentence-finally either as auditory-only, visual-only, or audiovisual speech (e.g., 'The password was OCto-'). On these critical trials, these primes were followed by printed presentations of either matching ('octopus') or stress-mismatching ('oktober') target words. Filler trials included nonword targets. Response times needed to indicate whether the printed items were words or nonwords were analyzed. Replicating previous results for auditory-only conditions (e.g., van Donselaar et al., 2005), matching primes should speed up and mismatching primes slow down correct target recognition compared to when unrelated primes precede target presentations (e.g., The password was machi-', where 'machi-' was taken from 'machine'). If visual speech also conveys lexical stress information and this information influences indeed lexical activation, then for audiovisual primes, target response times should be similarly modulated by overlap in lexical stress. Results are discussed within the framework of current models of auditory and audiovisual spoken-word recognition.

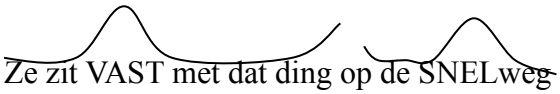
Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech, 45,* 207-228.

Donselaar, W. van, Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology, 58A*, 251-273.
Massaro, D.W., & Jesse, A. (in press). Audiovisual speech perception and word recognition. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics*. Oxford, U.K.: Oxford University Press.
Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45*, 412-432.

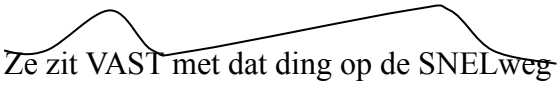
Arguments for ToDI

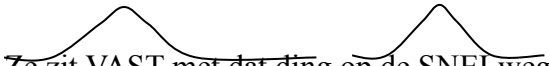
Carlos Gussenhoven, Radboud Universiteit Nijmegen

Arguments for phonological analyses can be internal or external. Among the most highly prized external arguments are experimental data on speech behaviour. Boundaries of intonational phrases need not obviously be marked by pauses. Our expectation in situations in which they are not is that they are marked by a boundary tone, such as H%, to signal the separation between the preceding and the following portion of speech. The boundary tone will cause some characteristic change in the pitch contour at the location of the boundary. An example is the peak caused by a final H% in (1).

(1) 
Ze zit VAST met dat ding op de SNELweg
'She has got stuck with that vehicle on the motorway'
or 'She must be on the motorway with the vehicle'

A closer look at data from Dutch suggests that the above generalization is often true, but that there are exceptions either way. First, a low valley immediately after an accent peak does not induce a boundary after *VAST* in (2). Second, absence of a clear pitch movement nevertheless results in a boundary after *ding* in (3).

(2) 
Ze zit VAST met dat ding op de SNELweg

(3) 
Ze zit VAST met dat ding op de SNELweg

Where (1) is ambiguous, due to the interpretation of *vast* as an adjective 'stuck' or as an adverb 'therefore probably', (2) means 'She has got stuck with that vehicle on the motorway', while (3) means 'So she must be on the motorway with that vehicle'. The adjectival meaning of *vast* is best triggered by the presence of a boundary and identity of pitch accents in the two intonational phrases. The adverbial meaning is most likely to be triggered by the absence of an intonational boundary.

We designed an experiment with artificial two-accent intonation contours having identical pitch accents in the two positions. We varied the length of the first phrase (2), the presence of the medial boundary (2), and the type of pitch accent (3). Source sentences were

(4) Ze zit vast met dat ding op de snelweg	'She has got stuck (with that vehicle) on the motorway'
Ze zit vast op de snelweg	*or* 'She must be on the motorway (with that vehicle)'
Hij zit alleen met die man in 't café	'He is alone (with that man) in the pub' *or*
Hij zit alleen in 't café	'The problem is he's (with that man) in the pub'

The presence or absence of the contours is determined on the basis of ToDI transcription. The results will be argued to support the analysis underlying that transcription system.

The importance of prosody for TRP projection

Wieneke Wesseling, University of Amsterdam

In three different Reaction Times (RT) experiments, the effect of various prosodic features - like end tone, position of prominent words and utterance length - on the projection of potential turn changes, also known as Turn Relevance Places (TRPs) was investigated. Subjects were asked to react with minimal responses to prerecorded dialogs and impoverished versions of these dialogs.

The first two experiments - with both full and impoverished versions of the dialogs, containing only intonation and pause information (*hummed* stimuli), or no periodic component at all (*whispered* stimuli) - already showed that end tone information can be used by listeners to predict upcoming utterance ends, although in natural speech the pitch signal itself might be redundant for predicting TRPs.

It also seemed that the presence of non prominent words right before an utterance end reduced the delays of both elicited and natural responses. This would suggest that the presence of a prominent, informative, word starts the projection of a possible upcoming TRP, while the availability of non prominent, predictable, speech then allows listeners to improve their predictions of the exact timing of the TRP.

However, in a third experiment, where at random, of each utterance, either one of the last four words was replaced by white noise (*masked* condition), or no word was replaced (*non masked* condition), the masking of prominent words did not affect the response times of our subjects. Only when it was the very last word of the utterance that was masked, the reaction time was delayed. It seems that effect of prominent word position that was found earlier, is in fact an effect of utterance length.

Our results suggest that predicting the relative position of the last word before the TRP is robust enough to be unaffected by missing either intonation, or individual words. The strong facilitating effect of utterance length on RTs also points to the use of global syntactic and discourse structure in predicting the relative position of the last word.