



DFG Research Unit 2373 'Spoken Morphology'

2nd Workshop on
Spoken Morphology: Phonetics and Phonology of Complex Words

Heinrich-Heine-Universität Düsseldorf
August 23-25, 2017

Conference Venue:
Schloss Mickeln, Düsseldorf-Himmelgeist

Abstracts

ordered according
to first name of first
author

Semitic roots facilitate auditory word recognition
Adam Ussishkin, University of Arizona

Words consist of a phoneme sequence that maps onto meaning. Most prominent theories of word recognition portray the recognition process as a connection between these small units and a semantic level. However, there is a growing body of evidence in the priming literature suggesting that there is an additional, morphological level that mediates the recognition process. In morphologically linear languages like English, however, morphemes and sound sequences are frequently co-extensive, so the source of priming effects between related words could be due to phonological overlap as opposed to morphological overlap. In Semitic languages, though, the morphological structure of words reduces this confound, since the morphemes are interdigitated in a non-linear fashion. Semitic words are typically composed of a discontinuous root (made up of three consonants) embedded in a word pattern specifying the vowels and the ordering between consonants and vowels. Active-passive pairs in Maltese illustrate this relationship (the root is underlined); e.g., *fetaħ* 'open'-*miftuħ* 'opened'.

In this talk, I report on a number of experiments our lab has carried out on Hebrew and Maltese investigating the extent to which root morphemes facilitate auditory word recognition, and to what extent it is possible to tease apart the potential contribution of phonological overlap from true morphological effects. These experiments make use of auditory priming techniques that vary whether or not the prime is (Goldinger 1996) or is not (Kouider and Dupoux, 2005) masked from conscious perception. The results show consistently that not only do roots facilitate word recognition in both Hebrew and Maltese, but that these morphological effects appear to be independent of phonological overlap effects.

What affects phonetic reduction in NNN compounds? Looking at the impact of morphological structure

Annika Schebesta & Gero Kunter

This paper presents results from two different studies on the effect of morphological structure on the acoustic signal in NNN compounds in English. We examined corpus data (Boston University Radio Speech Corpus) as well as data from a production experiment, in which participants were asked to read NNN compounds in short texts. The investigation is based on the Embedded Reduction Hypothesis (ERH), which has first been tested by Kunter and Plag (2016):

Embedded Reduction Hypothesis:

In a complex word [X Y] Z,

the inner boundary between X and Y is more prone to phonetic reduction than the outer boundary between Y and Z.

They examined experimental data and found that there are length differences in the constituents for both left- and right-branching compounds, which can be explained by the ERH. However, this length effect is strongly intertwined with the bigram frequencies in the data, thus it is not clear what the effect of boundary strength and branching actually is.

In our corpus study, we used data from the BURSC to test the ERH and found slightly different results. In left-branching compounds, the free constituent, i.e. the N3 in the compound, is indeed longest among all three nouns, which is predicted by the hypothesis. This effect becomes stronger with increasing N1N2 bigram frequencies. In right-branching compounds, however, the free constituent N1 as well as the embedded constituent N3 are long. This is not predicted by the ERH, but rather speaks in favor of an overall lengthening effect on N3.

In order to test the effect of branching and boundary strength while controlling for bigram frequencies, we created NNN with low frequencies for N1N2 and N2N3, and kept the syntactic position of the compound constant in the target sentences. The NNN include a W1W2 pair of constituents (e.g. W1 = ACCOUNT, W2 = SERVICE) with a plosive (/t,d/) at the word boundary that is likely to be deleted in certain conditions. Plosive deletion serves as an additional measure of phonetic reduction to constituent durations. We tested 25 such W1W2 pairs in 4 different conditions:

left-branching 1:	[guest	ACCOUNT]	SERVICE
left-branching 2:	[ACCOUNT	SERVICE]	assistant
right-branching 1:	guest	[ACCOUNT	SERVICE]
right-branching 2:	ACCOUNT	[SERVICE	assistant]

The statistical analyses we have run so far, however, show that boundary strength and branching do not have a significant impact on phonetic reduction. In fact, the data shows a lengthening effect on both N1 and N3 constituents, similar to what we have found in the corpus data.

In a future experiment, we plan to focus on the effect of bigram frequency on phonetic reduction by, for instance, contrasting high-frequency bigrams and low-frequency bigrams. We will also try to understand the effect on N2 durations in different branching directions, which might help us understand the phonetics of NNN compounds better.

References

- Kunter, Gero and Ingo Plag. 2016. Morphological Embedding and Phonetic Reduction: The Case of Triconstituent Compounds. *Morphology* 26 (2), 201-227.
- Schebesta, Annika and Gero Kunter. in preparation. Embedded Reduction in NNN Compounds: Results from a Corpus Study.

A production study on the pronunciation of English noun-verb homophones

Arne Lohmann (*Heinrich-Heine-Universität Düsseldorf*) &
Erin Conwell (*North Dakota State University*)

In this presentation we report first results from a study on the pronunciation of English noun-verb conversion homophones, e.g., *kick(V)* vs. *kick(N)*. While nouns and verbs differ in acoustic realization due to the different positions in the sentence they occur in, the aim of this study is to investigate possible differences in pronunciation that are unrelated to syntactic/prosodic context. To that end the N/V homophones are read in controlled prosodic contexts that are maximally similar across the two syntactic categories. We test the following twelve N/V pairs: *kick, dig, tip, kiss, look, book, cook, push, pack, nap, chat, snack*. The following sentences exemplify the syntactic contexts the words occur in.

- (1a) The students had a nap in the classroom next to the gym.
- (1b) The students had to nap in the classroom next to the gym.
- (2a) The actors' kiss on the movie set for the new production annoys the director.
- (2b) The actors kiss on the movie set for the new production and this annoys the director.

The test sentences are embedded in paragraphs and are read by undergraduate students at North Dakota State University. Previous research on N/V homophones indicates that there may be two main factors impacting on their pronunciation that are independent of syntactic/prosodic context: First, Watson et al. (2006) provide initial evidence for a category-specific propensity for the placement of intonational phrase (IP) boundaries, i.e., for a greater likelihood of placing IP boundaries after nouns rather than verbs. Second, previous corpus-based studies indicate an effect of lemma frequency differentiating the two words' durations (see Conwell 2016 on child-directed speech and Lohmann 2017 on the regular speech of adults). We test both of these factors on the data acquired through the production study.

References

- Conwell, Erin. 2016. Prosodic disambiguation of noun/verb homophones in child-directed speech. *Journal of Child Language*, 1–18.
- Lohmann, Arne. 2017. Lemma frequency affects the duration of homographic noun/verb conversion 'homophones'. Presentation at the *173rd Meeting of the Acoustical Society of America*.
- Watson, D., Breen, M., and Gibson, E. (2006). The role of syntactic obligatoriness in the production of intonational boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1045–1056.

Using morphology to investigate the interface between speech perception and speech production

Ben Tucker (University of Alberta, Edmonton)

Investigation of the lexical representation often involves data from speech production and speech perception. Hypotheses about the structure of the lexicon rely on data from these two modalities of language, however our understanding the interaction between these two modalities is weak. One model of the lexicon could posit that both production and perception use a single lexicon and that similar effects should be found in both production and perception. In this presentation, I will discuss two domains which explore characteristics of the lexicon using speech production data. I will describe several studies that use production data to investigate the lexicon and discuss two studies that explore these same lexical characteristics using perception studies. The first perception study explores the perception of the durational differences in homophones found by Gahl (2007). The second perception study investigates how manipulating the durational characteristics of the morphology bearing vowel in irregular verbs affects perception. This talk will conclude with a discussion of the relationship between perception and production and why knowledge of this relationship is essential to understanding the nature of lexical representation.

Perception of spoken morphology

Erin Conwell

Although homophones are defined as words that sound the same, research is increasingly suggesting that homophones are, in fact, differentiated during production. In particular, the durations of homophones appear to be influenced by meaning frequency and by grammatical category of use. As evidence mounts for the existence of differential productions of apparent homophones, the question is raised of whether listeners use these differences to facilitate processing. I will present data from three studies describing what effects acoustic differences in homophones have on processing of these words by adult listeners. Two of these studies examine the processing of isolated homophone tokens, one using a behavioral method and one using an ERP technique, while the third will examine how these words are processed in fluent sentences. Some limitations of each approach will also be considered, as well as the challenges associated with expanding these methods to language learners.

Perception of Semitic tri-lateral verbs in Maltese: Effect of reduction and morphological composition

Holger Mitterer
University of Malta

Maltese makes use of tri-consonantal roots such k-t-b (Engl. to write). These pose at least two challenges for the listener. First, some forms of these verbs have all three consonants in a sequence (e.g., jiktbu, Engl. 'they write'). In production, this leads to reduction and deletion of root consonants. This gives rise to the question whether such roots can still be easily recognized or require contextual support (see Ernestus, Baayen, & Schreuder, 2002). This was tested with an eye-tracking paradigm, in which looks to thematically related targets were compared for reduced and unreduced verb forms. For example, for jiktbu, the visual display contained a letter and two other objects not associated with writing. The results show that listeners do apparently not recognize reduced forms immediately because anticipatory looks to the thematically related target were only observed when the full form of the verb was presented but not if the reduced form was presented.

A second challenge for the perceiver is to deal with the massive number of word forms in which such roots occur. A single verb form may occur in several hundred forms just due to inflectional morphology. These forms may look as different as jiktbu and kitbet (Engl., 'he wrote'). While there seems to be a consensus that auditory word recognition involves both whole-word access and morphological decomposition (e.g., Wurm, 1997), it is unclear what the relative time course of these routes is. Given the problem of word segmentation in the auditory modality, it may be assumed that whole-word access might be faster than a route via decomposition (analogue to claims about the dual-route model of visual-word recognition). We tested this in an eye-tracking paradigm that allowed us to assess when listeners infer the gender of the agent. This was possible because person/gender information is prefixed in the present tense but suffixed in the past tense. Results are in line with early morphological decomposition: listeners can infer the gender of the agent earlier in the present tense when the information is carried by a prefix than in the past tense when the information is carried by a suffix. This shows that the recognition of verbs in a morphologically rich language is based on early morphological decomposition.

The Prosody of Derived Words in English - New Data and Theoretical Challenges

Javier Sanz & Sabine Arndt-Lappe

In descriptions of English morphophonology, traditionally a distinction is made between stress-shifting and stress-preserving affixes. This distinction has figured prominently in the theoretical literature, providing the basis for far-reaching claims about the nature of phonology-morphology interaction (most notably in Lexical Phonology: cf. Kiparsky 1982 et seq., Giegerich 1999; cf. also e.g. Pater 2000, Zamma 2012, Bermúdez-Otero 2012, Stanton & Steriade 2014).

In our project, we have looked at variability in stress assignment in complex words. Three phenomena have served as test cases:

- preserving and non-preserving primary stresses in long words suffixed with *-ory* (as in, e.g., *célébratory* vs. *celebratory* vs. *celébratory*, derived from *célébrate*),
- preserving and non-preserving primary stresses in long words suffixed with *-able* (as in, e.g., *ánalysable* vs. *analýsable*, derived from *analyse*), and
- variability of secondary stress assignment reflecting preservation or non-preservation of basal primary stress (as in, e.g., *implausibility* vs. *implàusibíity*, derived from *impláusable*).

In two reading studies, productions of pertinent words were elicited from some 60 speakers of British English. For all three phenomena under investigation we found robust evidence for stress variability, to the extent that none of the speakers recorded uses the same stress pattern throughout all words in the test set. Still, we find that the variation is highly systematic. For all three phenomena, chances of stress preservation seem to be probabilistically determined by phonological markedness effects, lexical effects (such as frequencies), and speaker-specific effects. With respect to the latter we find that there are distinct groups of speaker that differ in how they weigh the relevant factors.

On a theoretical level, our findings provide a challenge for both classical stratum-based approaches to the phonology-morphology interface that categorically affiliate every affix to a stratum, and for lexicon-driven approaches that do not incorporate representations on a level of abstraction that allows us to model phonological markedness effects (such as those stemming from syllable weight). Furthermore, the frequency and speaker effects found in all derivational categories investigated seems to us to suggest an account that establishes a more direct and constrained connection between formal markedness, transparency, and frequency, taking them as correlates of psycholinguistic parseability (cf. Hay 2002 et seq). One promising avenue that we are currently exploring to model our findings is to assume a stratal approach that is grounded in such psycholinguistic principles (cf. Bermúdez-Otero 2012, 2017), into which we incorporate a theory about what the prosodic representation of an affix looks like (Sanz 2017). The observed effects (of markedness, frequencies, and speaker groups) then emerge from an interaction of stratal affiliation (which reflects the parseability of the form and may differ for different speakers) and affix representation (which is different for different groups of speaker) with the markedness ranking of the language.

NDL Modeling of Maltese Plurals and Intuitions of Native Speakers

Jessica Nieder, Ruben van de Vijver

Our project investigates the plural formation in Maltese, a Semitic language with a very complex and variable plural system (see Azzopardi-Alexander and Borg (1997), Mayer et al. (2013) or Schembri (2012) for an overview). This raises the question as to what the basis is for native speakers to set up appropriate singular-plural mappings. We hypothesize that the phonotactics of the singular determines the shape of the plural. We conducted a corpus study and a production experiment and used several *NDL* (Baayen et al., 2011) modelings to understand better the basis of the generalizations set up by native speakers.

As a first step we conducted a corpus study where we extracted 2369 Maltese nouns from two different sources (Schembri (2012) and MLRS Corpus Malti v. 2.0), annotated the CV structure and added the corpus frequency numbers. The corpus was then used as a basis for a production experiment in which native speakers were asked to produce plural forms for existing singular nouns and phonotactically legal nonce singular nouns. Nonce forms were constructed on the basis of words in our corpus by changing either the consonants or the vowels or both systematically. The nouns that have been used as a base to create nonce words had either a broken plural form, a sound plural form or both plural forms. The results of the experiment support the hypothesis that the phonotactics of the singular noun determines the shape of the plural form: Participants produced significantly more sound plural forms for nonces that had an existing sound plural word as a base and significantly more broken plurals for nonces that had a broken plural form as a base.

We then modeled our experimental results with the *Naive Discriminative Learner*, a cognitive learning algorithm based on associative learning described in Baayen et al. (2011). The *NDL* model was trained on our corpus of 2369 Maltese nouns and tested with our nonces. As *Cues* we used singular nonces coded as sets of bigrams and *Outcomes* were the two possible plural types ‘broken’ or ‘sound’, as deduced from the answers of our participants. *NDL* classified 65,3 % of the nonces in the same way as our participants did.

Our next steps are to test different variants of the *NDL* model (with trigrams, CV

structures and syllable structure as cues) to see if the performance on our stimuli changes. In our talk we will present and discuss the results of the NDL modeling.

References

- Azzopardi-Alexander, M. and Borg, A. (1997). *Maltese*. Routledge.
- Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., and Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118(3):438–481.
- Mayer, T., Spagnol, M., and Schönhuber, F. (2013). Fixing the broken plural in maltese. In *Perspectives on Maltese Linguistics*, pages 129–158. De Gruyter.
- Schembri, T. (2012). *The Broken Plural in Maltese: A Description*. Il-Lingwa Tagħna. Univ.-Verl. Brockmeyer.

Storage and computation in word-form encoding

Joana Cholin

Ruhr-Universität Bochum

What are the relevant production units during post-lexical encoding processes in spoken language production? In my talk, I will argue for the functionally important role of syllables during morpho-phonological and phonetic encoding levels during speech planning.

There is agreement that syllabic structure contributes to processes such as phonological word formation and stress assignment. There is, however, disagreement over whether syllabic structure is retrieved as part of the lexicalized word form, or whether it only emerges in online-syllabification processes that follow lexical access. In either case, the output of word form encoding is considered to be a phonological word that contains a string of syllables specified for segmental and metrical properties.

A further question concerns the issue of whether or not these syllables activate their corresponding motor programs in a separate store that supplies speakers with precompiled motor programs to facilitate the final steps of speech planning.

In the first part of my talk, I will review data stemming from different experimental paradigms and different languages that provide evidence for and against the assumption that syllable structure is a lexical property. I will argue against stored syllabic structure at a lexical level and for a context-dependent, post-lexical online computation of syllables.

In the second part of my talk, I will review findings and present data that strengthen the view that syllabic motor programs are retrieved from a separate store. In particular, I will present data from fluent as well as disfluent speakers to explore how storage and computation of syllables might interact between the levels of morpho-phonological and phonetic encoding and articulation. Finally, I will address the question of how the incremental execution of subsequent syllables depends on factors such as frequency and the size of the phonological word.

Homophones and their representation in the mental lexicon

Peter Indefrey, Katharina Sternke-Hoffmann and Frauke Hellwig

Jescheniak, Meyer & Levelt (2003) observed that low frequent (LF) homophones are named faster than frequency-matched non-homophonous words in a translation task. They interpreted this finding as being due to shared lexical word form representations enabling the low frequency (LF) homophone to ‘inherit’ the higher frequency of its twin. Gahl (2008), on the other hand, found a systematic difference in pronunciation length between LF and HF heterographic homophonous word forms when analysing a corpus of American English natural speech, with HF homophones being pronounced faster than their LF twins. She argues that this finding is best compatible with non-shared lexical representations of homophones including fine phonetic detail.

To shed some more light on this issue we tried to replicate Gahl’s findings experimentally, embedding homophonous words in sentences and asking the participants to read the sentences first silently, then aloud and to repeat them from memory after the visually presented sentences disappeared from the screen. We used (a) the homophone pairs of Jescheniak et al. (2003) and (b) other heterographic and homographic noun homophones. We replicated Gahl’s (2008) findings for heterographic homophones, including those of Jescheniak et al. (2003), HF twins being pronounced about 50 ms faster. There was no reliable length difference for homographic homophones. Overall, and also for homographic noun homophones alone, word length differences showed medium-sized correlations ($r = 0.4$ to 0.5) with log lemma frequency differences between HF and LF homophones.

Considering that there seem to be reliable length differences, listeners might exploit these to disambiguate homophone readings. To test this option, we performed a lexical decision task, in which visually presented targets were preceded by an auditorily presented prime. The primes were homophones or non-homophonous control words preceded by a neutral sentence frame (“the following word is...”) that did (‘related’) or did not (‘unrelated’) have a semantic relationship to the target word. The homophone primes had been elicited from several speakers in disambiguating contexts. We chose homophone tokens with clear length differences between HF and LF readings.

The results showed that, irrespective of length, only the dominant (HF) reading was activated. Hence listeners did not exploit the length difference for disambiguation.

We will discuss the implications of our findings for the question of shared or separate lexical word forms of homophones and report on our next steps towards a resolution of this issue.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84, 474-496.

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 432-38.

Phonetic detail and morphological structure: affix durations and gemination

Sonia Ben Hedia, Julia Zimmermann & Ingo Plag

Recent research on the acoustic properties of complex words has yielded interesting and puzzling results on how morphological structure may co-determine certain aspects of the speech signal (duration, formant structure etc.). Such findings are a challenge for existing theories of the phonology-morphology interface and of widely accepted speech production models.

In this paper we will review our own work on phonetic effects of morphological structure. We will first show how morphological segmentability may influence the duration of derivational affixes in English (*un-*, *in-*, *dis-* and *-ly*), using conversational speech data from the Switchboard corpus. We will then take a look at the gemination and degemination behavior of the same set of affixes. Corpus data and experimental data show that morphological informativeness influences gemination. Finally, we will present corpus studies that demonstrate that word-final S in English varies by morphological status and morphological category, and that a morphological boundary preceding the S blocks voicing-dependent allophonic lengthening of the segment preceding the boundary. In sum, we find very robust morphological effects on phonetic detail. The theoretical implications of these findings will be discussed.

Does this make me look parsable? Spelling variation and morphological structure

Susanne Gahl & Ingo Plag

In recent work it has been shown that English spelling reflects important aspects of morphological structure (e.g. Berg 2013, Aronoff & Berg 2017). It has also been shown that the writing process itself may reveal interesting aspects of language processing and language production (see, for example, the work by Kandel and colleagues on handwriting).

In this paper we will present a study at the intersection of these two areas of research, the spelling variation (aka spelling mistakes) in morphologically complex words. We wanted to test the hypothesis that the spelling alternations reflect morphemic status and morphological structure. To do so, we looked at two different cases, using data from Twitter. First, the competing spellings of what is generally taken to be the same suffix, <ible> and <able>, and, second, the competition between the suffix string <ment> and a non-suffixal string <mint>. For <able/ible> we find that morphological segmentation to higher error rates, due to competition between the two suffixes. For <ment/mint> segmentation protects from errors.

The effects of morphology on spelling in our data are analogous to those in pronunciation variation (Plag & colleagues, Gahl & colleagues), hand movement (Dukhovny & Gahl 2014) and typing speed (Badecker, et al. 1990, 1996, Will et al. 2003a, 2003b, Nottbusch et al. 1998 etc.).