The Distribution and Prosodic Realization of Verb Forms in German Infant-Directed Speech

Bettina Braun, Katharina Zahner

Department of Linguistics, University of Konstanz, Germany
PO Box 186, 78467 Konstanz, Germany
E-mail: bettina.braun@uni-konstanz.de, katharina.zahner@uni-konstanz.de

Abstract

Infant-directed speech is often seen as a predictor for infants' speech processing abilities, for instance speech segmentation or word learning. In this paper, we examine the syntactic distribution (position), accentuation and prosodic phrasing of German verb forms and discuss that many verb forms are prime candidates for early segmentation: they frequently appear at the start or end of prosodic phrases; if they are not phrase-initial, they are often preceded by closed-class word forms and they are frequently accented (imperative verb forms: 72% of the cases, infinitive verb forms: 82% of the cases). It thus appears that German infants ought to be able to extract verbs as early as nouns, given appropriate stimulus materials.

Keywords: corpus, infant-directed speech, prosody, verbs, German

1. Introduction

Infant-directed speech (henceforth, IDS) is often seen as a predictor for infants' speech processing abilities, for instance speech segmentation or early word learning (e.g., Johnson et al., 2014; Laing, 2017; Laing et al., 2016). Regarding segmentation, infants have been shown to segment nouns from fluent speech earlier than verbs. The ability to extract nouns emerges between 7 and 9 months of age (e.g., Altvater-Mackensen & Mani, 2013; Bartels et al., 2009; Jusczyk & Aslin, 1995; Jusczyk et al., 1999; Kuijpers et al., 1998; Männel & Friederici, 2013; Schreiner & Mani, 2017; Zahner, Schönhuber, & Braun, 2016). When nouns occur after or before a prosodic boundary, segmentation succeeds even at 6 month of age (Johnson et al., 2014; Seidl & Johnson, 2006). Verbs, on the other hand, are segmented only several months later, by 11 months in French (Marquis & Shi, 2008) or 13.5 months in American English (Nazzi et al., 2005) – this asymmetry is observed even when verbs in the experiments have the same number of syllables and stress patterns as the noun test words. There are two frequent arguments in the literature for why nouns are segmented earlier than verbs (syntactic framing and prosodic marking). In the Introduction, we will evaluate these two arguments for the case of German drawing on German infant studies, where available, and with regard to the grammatical structure of German. Then we derive predictions for the prosodic realization of German verb forms in IDS. In the main part of the paper, we analyze these forms in the KIDS Corpus (Zahner, Schönhuber, Grijzenhout, et al., 2016).

Argument 1. Nouns are easier to recognize than verbs, since nouns occur in more restricted syntactic frames than verbs (see Höhle et al., 2004, for German; Willits et al., 2014, for English). For instance, Höhle et al. (2004) analyzed a corpus that comprises recordings of 16 mothers talking to their children aged 18 to 27 months, in total 15.000 words. They reported that indefinite articles were often followed by nouns (70% of the cases), while the personal pronoun *sie* and a verb co-occurred in only around 30% of the cases. Children have been shown to rely on such distributional information for syntactic categorization from around 12 to 16 months (Höhle et al., 2004; Mintz, 2006). Specifically, German 16 month-olds use determiners to correctly categorize words as nouns even when the kind of determiner was different during familiarization and test (Höhle et al., 2004), making it necessary to generalize. Thus, syntactic frames might assist syntactic categorization (Mintz, 2003) and segmentation (Willits et al., 2014, for verbs in frequent ing-constructions in English).

Syntactically, German is a V2 language, with the finite verb in second position. In German main clauses and *wh*-questions, the finite verb appears in second position (V2, 1a), in polar questions and imperatives in initial position (V1, 1b) and in subordinate clauses in sentence-final position (1c).

(1a) **Declarative main clause, with/without modal**
Peter **isst** eine Banane.
Peter eats a banana.
'Peter is eating a banana.'
Peter **will** seine Banane essen.
Peter wants his banana eat.
'Peter wants to eat his banana.'
***wh*-question**
What **isst** Peter?
What eats Peter?
'What does Peter eat?'

(1b) **Polar question**
Isst Peter eine Banane?
Eats Peter a banana?
'Does Peter eat a banana.'
Imperative
Iss deine Banane!
Eat your banana.
'Eat your banana!'

(1c) **Main clause, subordinate clause**
Ich bin froh, dass Peter seine Banane **isst.**
I am happy that Peter his banana eats.
'I am happy that Peter is eating his banana.'

Note that German verbs are conjugated depending on person and number (e.g., infinitival form: *essen* 'eat', personal verb forms: 1sg *esse*, 2sg *isst*, 3sg *isst*, 1pl *essen*, 2pl *esst*, 3pl *essen*).

Argument 2. Nouns are easier to recognize than verbs because nouns are prosodically more salient than verbs (accentuation and prosodic phrasing). When comparing the experimental materials used in their verb segmentation study to the materials in an analogous noun segmentation study (Jusczyk et al., 1999), Nazzi et al. (2005) found that the nouns were more often followed by a phrasal boundary than the verbs and were more often preceded by a pitch-accented syllable. The authors argue that this leads to a clearer "perceptual demarcation" for nouns compared to verbs (Nazzi et al., 2005, p. 293). These arguments may well be true for the stimuli with which infants are tested in speech segmentation tasks in the laboratory and also for read speech in English and German (but see also Conwell, 2017, for a recent corpus study on noun/verb homophones in English child-directed speech that showed that nouns are more salient than verbs in phrase-medial position). However, we know very little about the prosodic realization of nouns and verbs in IDS in German. In the adult grammar, verbs are only accented under certain conditions (Féry, 2011; Féry & Herbst, 2004; Gussenhoven, 1983, 1984; Selkirk, 2011), such as in corrective or contrastive contexts where the lexical content of the verb is contrasted (e.g., 'Peter SOLD the book; he didn't buy it'; capitals indicate accentuation), for intransitive verbs when the subject is topical ('Truman DIED'), for transitive verbs when the objects are contextually given and/or pronominalized (e.g., 'Peter BOUGHT it'). In colloquial speech of every-day conversations, verbs are expected to be more often accented, since during the course of the conversation, referents and locations become known to the interlocutors and hence tend to be pronominalized or unaccented (Arnold et al., 2013). While we do not have any analyses of adult conversational speech to back up this claim at the moment, we can test the accentuation in natural infant-directed speech.

Hypothesis. Our hypothesis is that certain verb forms are accented very frequently because conversations with young children are predominantly situated in the *here* and *now* and involve daily routines, objects and events (e.g., Dominey & Dodane, 2004; Papoušek et al., 1991; Werker & McLeod, 1989), which are in turn visually given and consequently unaccented. The verb is then the only word left that can receive the pitch accent.

For the analysis, we use the KIDS Corpus (Zahner, Schönhuber, Grijzenhout, et al., 2016), a multi-layered corpus that consists of 524 intonation phrases (IPs) directed to infants younger than one year, (see Zahner, Schönhuber, Grijzenhout, et al., 2016, for a more detailed description of the corpus; note that the TextGrids of the corpus are publically available on http://ling.uni-konstanz.de/pages/home/braun/KIDS.html - sounds are available on request).

2.2 Annotation

2.2.1 Lexical information and word class

In the KIDS Corpus, all data were labeled using Praat (Boersma & Weenink, 2014) in ten tiers (see Zahner, Schönhuber, Grijzenhout, et al., 2016, for a detailed description of annotation layers). Figure 1 shows the annotation layers that are relevant for the current distribution analysis of verb forms and the prosodic investigation of phrasing an accentuation. Tier 1 shows the orthographic transcription of the utterances. The word class of all words in the corpus was tagged, see tiers 2 and 3. Tier 2 presents broad labels for the syntactic categories, e.g., "adj" (adjective), "adv" (adverb) "conj" (conjunction), "det" (determiner), "int" (interjection) "kop" (copula), "name" (name), "nonw" (nonce word), "noun" and "part" (particle). On tier 3, we used the labels from the Stuttgart-Tübingen-TagSet (STTS, Schiller et al., 1999). For verbs, these are:

- VAFIN: finite auxiliary
- VAPP: past participle of an auxiliary
- VMFIN: finite modal auxiliary
- VMINF: infinite modal auxiliary
- VVFIN: finite lexical verb
- VVIMP: imperative of a lexical verb
- VVINF: infinitive of a lexical verb
- VVPP: past participle of a lexical verb

As described in the Introduction, finite German verbs can occur in different sentence positions. Therefore, we additionally coded the syntactic position of finite verbs as follows: sentence-initial "V1", "V2" and "sentence-final" (or "other" in case the syntactic structure was unclear due to elisions, noise, etc.). Infinitival forms typically occur in clause-final position but may be followed by adverbials, interjections, subordinate clauses, etc. The constituents were also coded, if present.
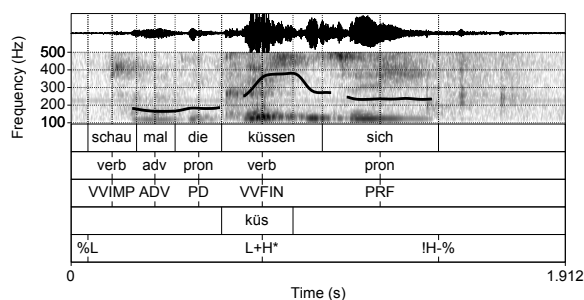

Figure 1: Example utterance from KIDS with smoothed f0 contour and the relevant annotation layers

2.2.2 Prosody

The intonational annotation in KIDS (accented syllable and tonal events, see Figure 1, tiers 4 and 5) was provided by two trained annotators (authors of this paper). Pitch accents and boundary tones were annotated according to GToBI (German Tone and Break Indices, Grice et al., 2005), a widely used annotation system within the framework of autosegmental-metrical

phonology (see, Ladd, 2008, for an overview). The pitch accent and boundary tone inventory as well as specific labeling conventions are described in more detail in Zahner et al. (2015).

For the distributional prosodic analyses of verb forms, we furthermore extracted the position of the verb forms in the respective prosodic phrase. For infinitival forms, which typically occurred in sentence-final position, we coded whether they were followed by a prosodic phrase break (intermediate phrase, ip, or full intonational phrase, IP). For imperatives, we coded their position in the intonational phrase (initial, medial, final).

3. Results

3.1 Syntactic distribution of verb forms

Verbs are the most frequent word class in the KIDS Corpus (verbs: 23%, pronouns: 19%, adverbs: 18%, nouns: 12%). There are 395 verbs, which distribute over the verb classes as follows: 153 finite verbs, 56 finite modal verbs, 77 infinitives, 76 imperatives, 18 participle forms and 13 finite copula verbs.

Of the 153 **finite verbs forms**, there are 69 different verb forms (from 37 lemmas). The most frequent lemmas are *sein* 'to be' (18 times), *haben* 'to have' (17 times) and *tun* 'to do' (12 times). Position-wise, finite verbs mostly occur in V2 (102 times), followed in frequency by V1 (38 times), the final position is rare (11 times). When the finite verb is in second position (V2), the first position is often filled by closed-class elements: most often by a demonstrative pronoun (28 cases, in 7 cases preceded by an interjection, conjunction, or particle). The next frequent categories in the preverbal field were question words (19 times) and adverbs (19 times). Noun phrases were rare (6 times). The syntactic positions of **finite modal verbs** (17 times *können* 'can', 16 times *wollen* 'want', 10 times *müssen* 'must', 9 times *sollen* 'should' and 4 times *dürfen* 'may') are similar to full finite verbs.

Infinitival verb forms occur mostly in utterance-final position (63 times), in 7 cases, the infinite verb is followed by a one-word particle interjection or a name. In the remaining 7 cases, they are followed by phrases such as *ich weiß* 'I know', by syntactic right-locations *das vielleicht haben, die Rassel* (lit: 'it maybe have, the rattle') or *das schoen zusammendruecken, das Buch* (lit: 'it well press, the book'), by a subordinate clause *muss mal aufpassen, dass ....* ('have to pay attention that...') or by an adverbial *musst du mal krabbeln lernen bis dahin* ('you have to learn to crawl till then').

The **imperative verb forms** are mostly *guck* or *schau* 'look' (28 and 22 times, respectively) and *komm* 'come' (10 times). Syntactically, they typically occur in utterance-initial position (52 times, 12 times preceded by an interjection, a name, or an object noun), but there are also 24 epenthetic occurrences and 10 occurrences in sentence-final position.

The **participle verb forms** typically occur in phrase-final position (11 times), but they are also followed by particles, adverbial phrases, dislocations, imperatives (*schau* 'look') and other parts of a verbal

complex (*bekommen haben* 'have gotten', *rausgesucht hast* 'have found out').

3.2 Prosodic Analysis

Figure 2 shows the absolute number of accented and unaccented verb forms in KIDS (for categories with N > 5); for comparison, nouns in KIDS are accented in 90% of the cases (215 accented, 15 unaccented).
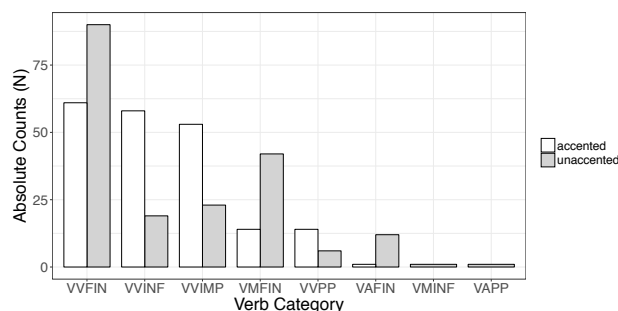

Figure 2: Distribution of accented and unaccented verb types in the KIDS Corpus

Figure 2 shows that **finite verb forms** (full lexical verbs, modal verbs and auxiliaries, VVFIN, VMFIN, VAFIN, respectively) are more often unaccented than accented, while all other verb forms are more often accented than unaccented. A x^2-square test (excluding the infrequent verb classes VAFIN, VMINF, and VAPP) shows that the accented and unaccented verb forms are not distributed equally across verb types ($x^2 = 53.7$, df = 4, p < 0.0001). The distribution of GToBI accents for these verb forms is shown in Table 1. Note that the group of finite lexical verbs (VVFIN) also includes the verbs *haben* 'to have' and *sein* 'to be', which, in general, are only accented in forms with polarity contrast (context: He does not have a car. target: He HAS a car, cf. Turco et al. (2014)). When we exclude these two lemmas, the ratio of accented to unaccented finite verbs is more balanced (47% vs. 35%; cf. last row in Table 1).

Accent type	VVFIN	VVINF	VVIMP
unaccented	**59.5% (91)**	**24.7% (19)**	**30.2% (23)**
H*	**10.5% (16)**	**22.1% (17)**	**36.9% (28)**
!H*	5.9% (9)	14.3% (1)	2.6% (2)
L+H*	7.2% (11)	14.3% (11)	13.2% (10)
L*	7.2% (11)	10.4% (8)	5.3% (4)
L* (before H%)	2.0% (3)	5.2% (4)	5.3% (4)
ratio acc:unacc	35% (w/o *have/be*: 47%)	68%	68%

Table 1: Frequent pitch accents across verb forms (bold face indicates most frequent realizations).

Since the finite forms are most variable in their syntactic distribution, we plotted the accentuation of these forms across syntactic position (V1, V2, utterance-final, and other) in Figure 3. This allows us to test whether accented and unaccented finite verbs are distributed

equally across syntactic position. This is clearly not the case ($x^2 = 12.4$, df = 2, p < 0.006, Yates' corrected and excluding the category 'other').
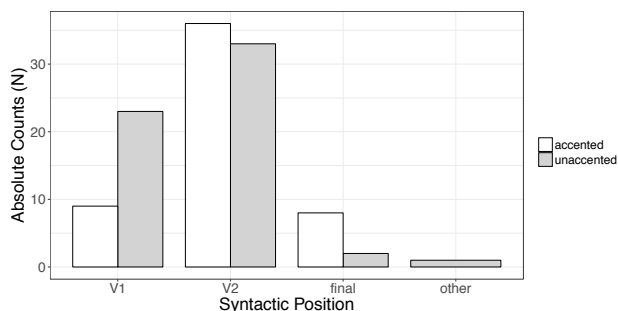

Figure 3: Distribution of accented and unaccented finite forms across syntactic position (without *haben*, *sein*)

Figure 3 shows that finite verbs are often accented in final position (a position that is rare in IDS, as it is restricted to subordinate clauses). In V1 position, the proportion of unaccented finite forms by far outnumbers accented forms, while the distribution is more balanced (or even slightly reversed) for V2 position. Figure 1 and Figure 4 show example realizations of accented finite verbs in V2 position (and an additional accented participle in Figure 4).
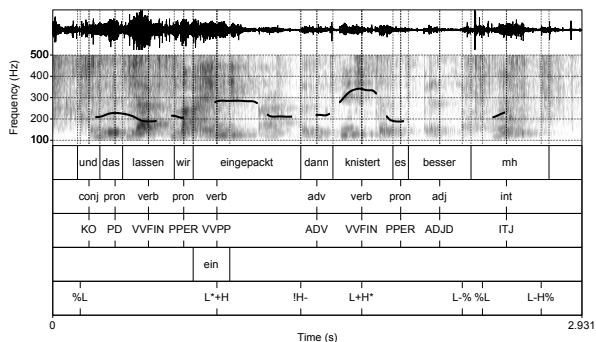

Figure 4: Example realization of an accented finite verb in V2 position

The **infinite verb forms** are twice as often accented than unaccented (Tab. 1) and typically occur at utterance ends, hence being followed by a prosodic IP phrase break (N=63, Tab. 2). In rare cases, infinitival verbs are followed by adverbials, *bitte* 'please', before a dislocation or a name, or without a prosodic phrase break (N=10).

Position of infinitival form	ip-break	IP-break	no phrase break
before adverbial	0	0	3
before bitte 'please'	0	0	1
before dislocation	0	0	1
before interjection	1	0	0
before (pet)name	0	0	4
before negation	0	1	0
before clause	1	0	1
at utterance end	0	**63**	0

Table 2: Position of infinitival verb form and prosodic phrasing.

Imperative verb forms are also more than twice as often accented than unaccented. They are almost exclusively separated from surrounding material by an intermediate phrase (75 out of 76 times), in 5 instances they constitute a separate intonation phrase (IP).

4. Discussion and Conclusion

In infant-directed speech, certain grammatical structures of a language are overrepresented, others underrepresented, mainly due to shorter utterances and the communication about events in the *here* and *now* (Dominey & Dodane, 2004; Papoušek et al., 1991; Werker & McLeod, 1989). Furthermore, the prosodic realization of word categories may be shifted because of frequent pronominalization, increased prosodic variability, and an increased number of prosodic phrases. Previous research has shown that IDS is beneficial for word segmentation (Floccia et al., 2016; Schreiner & Mani, 2017), but the driving factors have to be determined. Here, we analyzed two factors that may affect processing, the syntactic distribution and prosodic properties of German verb forms. Our distributional analyses show a high proportion of sentences in which the verb appears in second position, which is not unexpected for a V2 language. Interestingly, about half of these verb forms are accented (when we exclude the lemmas *have* and *be,* which are typically unaccented in finite forms). Verbs in V2 position are frequently preceded by closed-class forms, making their occurrence more predictable. Moreover, there are also a considerable number of imperatives and infinitival forms in IDS, which are more often accented than unaccented and, furthermore, occur mostly in phrase-initial and phrase-final position, respectively. In both positions, they are demarcated by a prosodic phrase boundary. Recent research in infant speech segmentation has shown that words are segmented better when accented (Männel & Friederici, 2013) and when occurring at phrase-initial and phrase-final positions (Johnson et al., 2014).

More fine-grained prosodic analyses show that, if accented, imperatives and infinitives are frequently realized with high-toned pitch accents (mostly H* or L+H*), which are perceived as prominent by German adult listeners (Baumann & Röhr, 2015) and are more easily segmented by German infants (Männel & Friederici, 2013; Zahner, Schönhuber, & Braun, 2016). Although these accented verb forms are not the most frequent forms that German infants encounter (finite verbs appear more often), they may qualify as candidates for early speech segmentation and thus provide German infants with a head start for the acquisition of verbs in general.

6. References

Altvater‑Mackensen, N. & Mani, N. (2013). Word‑form familiarity bootstraps infant speech segmentation. *Developmental Science,* 16(6), 980-990.

Arnold, J.E., Kaiser, E., Kahn, J.M. & Kyoungsook, K.L. (2013). Information structure: Linguistic, cognitive and processing approaches. *Wiley Interdisciplinary Reviews. Cognitive Science,* 4(4), 403-413.

Bartels, S., Darcy, I. & Höhle, B. (2009). Schwa syllables facilitate word segmentation for 9-month-old German-learning infants. *Proceedings of the 33rd Annual Boston University Conference on Language Development*, Somerville, M.A.: Cascadilla Press, 73-84.

Baumann, S. & Röhr, C. (2015). The perceptual prominence of pitch accent types in German. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK.

Boersma, P. & Weenink, D. (2014). Praat: Doing phonetics by computer. Version 5.3.84 [Computer program].

Conwell, E. (2017). Prosodic disambiguation of noun/verb homophones in child-directed speech. *Journal of Child Language,* 44(3), 734.

Dominey, P.F. & Dodane, C. (2004). Indeterminacy in language acquisition: The role of child directed speech and joint attention. *Journal of Neurolinguistics,* 17(2), 121-145.

Féry, C. (2011). German sentence accents and embedded prosodic phrases. *Lingua,* 121, 1906-1922.

Féry, C. & Herbst, L. (2004). German sentence accent revisited. In S. Ishihara, M. Schmitz & A. Schwarz (Eds.), *Interdisciplinary studies on information structure 1* (pp. 43–75).

Floccia, C., Keren-Portnoy, T., DePaolis, R., Duffy, H., Delle Luche, C., Durrant, S., et al. (2016). British English infants segment words only with exaggerated infant-directed speech stimuli. *Cognition,* 148, 1-9.

Grice, M., Baumann, S. & Benzmüller, R. (2005). German intonation in autosegmental-metrical phonology. In J. Sun-Ah (Ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing* (pp. 55-83). Oxford: Oxford University Press.

Gussenhoven, C. (1983). Focus, Mode and the Nucleus. *Journal of Linguistics,* 19(2), 377-417.

Gussenhoven, C. (1984). *On the Grammar and Semantics of Sentence Accents*. Dordrecht, Holland; Cinnaminson, N.J., U.S.A.: Foris Publications.

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A. & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy,* 5(3), 341-353.

Johnson, E.K., Seidl, A. & Tyler, M.D. (2014). The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds. *PLoS ONE 9(1):* 9(1), 1-14.

Jusczyk, P.W. & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology,* 29(1), 1-23.

Jusczyk, P.W., Houston, D.M. & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology,* 39(3), 159-207.

Kuijpers, C.T., Coolen, R., Houston, D.M. & Cutler, A. (1998). Using the head-turning technique to explore cross-linguistic performance differences. In C. Rovee-Collier, L. Lipsitt & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12). Stamford: Ablex.

Ladd, D.R. (2008). *Intonational phonology* (Vol. 119). Cambridge [u.a.]: Cambridge University Press.

Laing, C.E. (2017). A perceptual advantage for onomatopoeia in early word learning: Evidence from eye-tracking. *Journal of Experimental Child Psychology,* 161, 32.

Laing, C.E., Vihman, M. & Keren-Portnoy, T. (2016). How salient are onomatopoeia in the early input? A prosodic analysis of infant-directed speech. *Journal of Child Language,* 44(5), 1-23.

Männel, C. & Friederici, A.D. (2013). Accentuate or repeat? Brain signatures of developmental periods in infant word recognition. *Cortex,* 49(10), 2788-2798.

Marquis, A. & Shi, R. (2008). Segmentation of verb forms in preverbal infants. *Journal of the Acoustical Society of America,* 123(4), EL105-EL110.

Mintz, T.H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition,* 90, 91–117.

Mintz, T.H. (2006). Finding the verbs: Distributional cues to categories available to young learner. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Oxford: Oxford University Press.

Nazzi, T., Dilley, L.C., Jusczyk, A.M., Shattuck-Hufnagel, S. & Jusczyk, P.W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech,* 48(Pt 3), 279-298.

Papoušek, M., Papoušek, H. & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development,* 14(4), 415-440.

Schiller, A., Teufel, S., Stöckert, C. & Thielen, C. (1999). Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und großes Tagset) [Technical Report]. Retrieved from. Universitäten Stuttgart und Tübingen [Universities of Stuttart und Tübingen].

Schreiner, M.S. & Mani, N. (2017). Listen up! Developmental differences in the impact of IDS on speech segmentation. *Cognition,* 160, 98-102.

Seidl, A. & Johnson, E.K. (2006). Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science,* 9(6), 565-573.

Selkirk, E.O. (2011). The syntax-phonology interface. In J. A. Goldsmith, J. Riggle & A. Yu (Eds.), *The Handbook of Phonological Theory* (2nd edition). Oxford: Blackwell.

Turco, G., Braun, B. & Dimroth, C. (2014). When contrasting polarity, the Dutch use particles, Germans intonation. *Journal of Pragmatics,* 62(94-106).

Werker, J.F. & McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology,* 43(2), 230.

Willits, J.A., Seidenberg, M.S. & Saffran, J.R. (2014). Distributional structure in language: contributions to noun-verb difficulty differences in infant word recognition. *Cognition,* 132(3), 429.

Zahner, K., Pohl, M. & Braun, B. (2015). Pitch accent distribution in German infant-directed speech. *Proceedings of Interspeech*, Dresden, 46-50.

Zahner, K., Schönhuber, M. & Braun, B. (2016). The limits of metrical segmentation: Intonation modulates infants' extraction of embedded trochees. *Journal of Child Language,* 43(6), 1338-1364.

Zahner, K., Schönhuber, M., Grijzenhout, J. & Braun, B. (2016). Konstanz prosodically annotated infant-directed speech corpus (KIDS corpus). *Proceedings of the 8th International Conference on Speech Prosody*, Boston, USA, 562-566.