

Evidence for immature perception in adolescents: Adults process reduced speech better and faster than 16-year olds

Karin Wanrooij^a and Maartje E.J. Raijmakers^b

^aLeiden University; ^bVrije Universiteit Amsterdam

ABSTRACT

Previous work suggests that adolescents are still refining acoustic-phonetic cue use in clear-speech perception. This study shows adolescents' immature perception of reduced speech, in which speech sounds are naturally deleted and merged within and across words. German adults and 16-year-olds listened to either German reduced or unreduced (few or full cues) part- and full phrases (without and with context) in a phrase-intelligibility task. As expected, adolescents had lower scores when adequate perception required flexible acoustic-phonetic cue use most, i.e., when hearing reduced speech without context. Participants also listened to reduced and unreduced words and pseudowords (no context) in a lexical decision task. Here, 16-year-olds had poorer and slower responses than adults overall and particularly when hearing pseudowords. Explanations for the age effects are discussed. We conclude that experience continues to refine linguistic representations, at least until adulthood.

ARTICLE HISTORY

Received 16 January 2019

Accepted 20 April 2020

1. Introduction

Natural speech is inherently different from written texts: Words are shortened or even deleted, instead of being spelled out, and words are merged with surrounding words, rather than being separated by pauses. For instance, an English speaker may say “nahpmeam” [nɑ:pmi:m]¹ for *Do you know what I mean* (Field 2003), and speakers of German, the target language in this article, may say “haspmmomentsait” [haspmmomentsait] for *Hast du einen Moment Zeit* [has^h du: ʔainən mo:mənt^h tsait^h] ‘Do you have a moment?’ (Kohler 1998). Such “reductions” are intrinsic and abundant in natural speech across languages (e.g., German: Kohler 1990, 1998; Dutch: Ernestus 2000; English: Shockey 2003; Johnson 2004; Japanese: Maekawa & Kikuchi 2005).

The term *reductions* sometimes refers to (partial) deletions of speech sounds only (Brown & Kondo-Brown 2006). For the purpose of this study, we use it in a wider sense and include mergers or “assimilations” between speech sounds. The German example demonstrates both: the words *du* ‘you’ and *einen* ‘a’ are shortened and merged with the next word *Moment* ‘moment.’ Specifically, the article *einen* [ʔainən], which contains two alveolar nasals [n] in a full pronunciation, is reduced to [m]. This [m] has inherited the nasal feature but is articulated at a bilabial rather than alveolar place of articulation as a preparation for the subsequent bilabial [m] in *Moment*. The word *du* [du], which starts with an alveolar plosive [d] in a full pronunciation, is reduced to the plosive [p]. This [p] has inherited the plosive feature of [d] but is also articulated at a bilabial rather than alveolar place of articulation, thus expressing the rounding of the lips for [u] and preparing for the bilabiality of the

CONTACT Karin Wanrooij  karin.wanrooij@hotmail.com  Educational Sciences, Universiteit Leiden, Leiden, AK 2333, Netherlands.

This article was originally published with errors introduced during the publication process, which have now been corrected in the online version.

¹We follow the convention of using square brackets [] to denote phonetic representations (realizations of tokens) and slashes // to denote phonological representations.

 Supplemental data for this article can be accessed [publisher's website](#).

© 2020 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

upcoming [m]. Reductions affect both function words, which are often highly reduced (such as the pronoun *du* in the example) and content words (such as the nouns *Moment* and *Zeit* in the example; Ernestus & Warner 2011; Kohler 1998).

It is useful to realize that words or phrases can be reduced to different extents. For instance, the [pm] in “haspmmomentsait” can be reduced further to just a glottalized [m] (Kohler 1998). Another example is the 50 different variants of the German word *eigentlich* [ʔaɪgəntliç] ‘actually’ (which include *eingklich* [aɪŋkliç], *eintlch* [aɪntlɪç], and *eni* [aɪni]) in the German Kiel Corpus of Spontaneous Speech, which we also use in this study (IPDS 1996, 1997, 1995; Niebuhr & Kohler 2011). Reductions are therefore a major source of variation in the speech signal, alongside other sources of variation that listeners have to deal with, such as variation due to different dialects or individual speakers.

Reductions are based partly on perceptual or articulatory constraints that hold across languages (Mitterer 2008; Mitterer et al. 2006) and partly on language-specific phonological rules (Darcy et al. 2009; Mitterer & Tuinman 2012; Torreira & Ernestus 2011). Children must acquire this system to understand reduced speech in their native language. Our knowledge about how they do this and how long they need for this is limited.

For instance, in theoretical terms, it is not evident what children’s initial phonological representations look like and how they develop. Evidence from the past decades shows that these representations probably emerge under the influence of experience, as we will also argue in the course of this Introduction. From the start of their lives, infants experience both unreduced (“full”) and reduced variants. Buckler, Goy & Johnson (2018) show that even in infant-directed speech, around 75% of the tokens is reduced. If experience influences children’s representations, this suggests that initial representations are relatively strongly influenced by full pronunciations (still 25% of the tokens in Buckler, Goy & Johnson 2018) and that early exposure to much variation in “reducedness” is needed for learning how to deal with this variation. Children’s early productions reflect this pattern: For instance, they contain relatively few assimilations between words initially (Thompson & Howard 2007). However, it is too simple to say that initial representations will always reflect the full pronunciation mainly, since the frequency distributions of variants differ between words. Indeed, early pronunciations tend to be reduced when the reduced variant occurs frequently (Newton & Wells 2002; Zukowski & Larsen 2011). Hence, production studies show that early representations are probably influenced by full and reduced variants, and children need time to acquire reduction patterns as general regularities.

Hardly any studies examine children’s reduced-speech perception. Two studies show that toddlers already perceive simple assimilations adequately (e.g., when *pen* is realized as *pem* when followed by *please*, as in *Can you find the pen please?*; Skoruppa, Mani & Peperkamp 2013; Skoruppa et al. 2013). One other study (Smiljanic & Sladen 2013) reports difficulties with reduced-speech perception in 6- to 13-year-olds: The children failed to use semantic cues to predict final words in sentences masked by noise when sentences were pronounced in “conversational” and hence reduced speech as opposed to “clear speech” (with acoustic-phonetic enhancements). Recently, van der Feest, Blanco & Smiljanic (2019) presented adults with the same sentences in three speaking styles (infant-directed speech was added) and two signal-to-noise ratios (a condition without noise was added) and found similar results: Adults had more problems with using semantic cues when sentences were presented in conversational speech than in the clearer speaking styles and failed to use these cues when sentences in conversational speech were masked by noise (we will discuss perception in noise further in the following).

In itself, a clear-speech advantage is not surprising: This speaking style generally improves intelligibility for adults (see Smiljanic & Bradlow 2009). Similarly, studies on adults’ reduced-speech perception report a “full-form advantage,” i.e., pronunciations without reductions are processed faster and better than reduced counterparts (Drijvers, Mulder & Ernestus 2016; Ernestus, Baayen & Schreuder 2002; Ranbom & Connine 2007), and at a more refined level, the more reduced a word pronunciation is, the less accurate participants’ recognition (Ernestus, Baayen & Schreuder 2002). The availability of more cues and more enhanced cues thus facilitates perception. Relatedly, studies on adults’ reduced-speech perception report a beneficial effect of different types of cues in the context, including phonetic and phonological cues (Mitterer 2011), semantic cues (van de Ven, Tucker &

Ernestus 2011), and also syntactic cues (Ernestus, Baayen & Schreuder 2002; Janse & Ernestus 2011). Particularly acoustic-phonetic cues in the reduced forms and their contexts appear to be important for reduced-speech perception (Janse & Ernestus 2011).

Taken together, studies on reduced-speech perception suggest that children learn to deal with at least simple reductions early in life; that word recognition in conversational (reduced) speech is more difficult than in clearer speaking styles both for children and adults; and that acoustic-phonetic cues play a special role in reduced-speech perception. Unfortunately, none of the studies compared reduced-speech perception across ages, so that possible developmental patterns remain concealed. Additionally, the reduced-speech perception experiments with children only featured simple reductions, so that it also remains uncertain how children perceive more complex reductions.

Possible development in reduced-speech perception can also be analyzed in terms of how children learn competences needed for reduced-speech perception. At least two competences seem necessary. First, listeners must be sensitive to acoustic-phonetic detail (Janse & Ernestus 2011). For instance, when hearing [hasmomentsart] (mentioned previously; Kohler 1998), it benefits listeners if they are able to pick up the glottalization of the [m], which is the only remnant of the plosive phoneme in the word *du*. Second, listeners must be able to flexibly use the cues at hand. Irrespective of the absence or specific realization of phonemes and words (e.g., the absence of /d/ or its realization as [d], [p], or a glottalization in the example), the listener must be able to use and integrate the available cues to understand what was said. These “available cues” can be small-scale acoustic-phonetic cues (such as the glottalization) but also larger-scale “contextual cues” of different kinds (such as syntactic and semantic cues). Evidence shows that precisely these two competences (i.e., sensitivity to acoustic-phonetic detail and flexibility in using available cues) require time to develop (see also Werner 2007), as we will now discuss per competence.

As for the first competence (sensitivity to acoustic-phonetic detail), 1.5- to 2-year old infants are already sensitive to slight mispronunciations caused by single-feature alterations in familiar words, such as *vaby* versus *baby*, pronounced in slow, infant-directed speech (Swingley & Aslin 2000). Hence, infants can pick up acoustic-phonetic detail when speech is slow and clear enough and the words are familiar enough. Other evidence shows that this sensitivity develops throughout childhood and adolescence. A consistent observation is that children identify speech sounds as certain phonemes with increasingly sharper and more adult-like boundaries; this development extends well beyond infancy, at least until 12 years in Hazan & Barrett (2000), 15 years in Rigler et al. (2015), 17 years in Flege & Eefting (1986), and 18 years in McMurray et al. (2018).

It is commonly assumed that sharper phoneme boundaries reflect more mature representations of phonemes in the brain. Until recently, it was believed that such increasingly mature representations were also characterized by a diminishing sensitivity to within-category differences (Werker & Tees 1984; Kuhl et al. 1992; Tsushima et al. 1994; Best et al. 1995). However, McMurray et al. (2018), who zoomed in on real-time processing by using eye tracking, observe that increasingly sharper phoneme boundaries are coupled with a *rising* sensitivity to within-category differences. They argue convincingly that sharper boundaries probably do not reflect the structure of phonemic representations but rather an increasing ability to deal with variability in the speech signal. As children grow more sensitive to acoustic-phonetic detail, they become more able to deal with such variability. Hence, a rising sensitivity to acoustic-phonetic detail, which in this view reflects the true structure of phonemic representations, enables the sharp identifications.

A rising sensitivity to acoustic-phonetic details matches another observation, namely children’s increasing attention to smaller units of speech. Initially, children focus on larger units than single speech sounds (Hirsh-Pasek et al. 1987; Jusczyk et al. 1992; Myers et al. 1996). In doing so, they pay particular attention to cues defining the syllable rather than the separate phonemes (Jusczyk & Derrah 1987; Nittrouer & Miller 1997; Nittrouer et al. 2000). For instance, Nittrouer et al. (2000) presented 5-year-olds with syllables consisting of a fricative and a vowel (FV) and syllables with the same acoustic material presented in the reverse order (VF). While adults weighted the acoustic cues differently depending on the order, no such difference was observed in the children. Hence, young

children may perceive syllables more as wholes than as a series of separate speech sounds. Of course, children will need to start differentiating between FV and VF syllables and also between equivalent syllables that represent the beginnings of different words. As the lexicon grows, the child will need to focus more on cues defining individual phonemes (Jusczyk et al. 1992; Metsala & Walley 1998; Walley, Metsala & Garlock 2003).

The growing lexicon not only stimulates the sensitivity to acoustic-phonetic details defining phonemes (or in theoretical terms “richer” representations) but also increases efficiency in word recognition (at the level of word representations) already at 2 years of age (Fernald et al. 1998; Hurtado, Marchman & Fernald 2008; Marchman, Fernald & Hurtado 2010). This efficiency reflects an increasingly efficient lexical competition process, which continues throughout adolescence (McMurray et al. 2018; see also Sekerina & Brooks 2007; Rigler et al. 2015). Lexical competition occurs while listening to ongoing speech. The brain supposedly activates all word representations that are possible at a certain time; these word candidates “compete” with one another and are considered as long as the acoustic information matches the listener’s (word) knowledge, until one candidate remains (Marslen-Wilson 1990; Marslen-Wilson & Welsh 1978; McClelland & Elman 1986). As children accumulate word knowledge, the number of competitors increases, thereby raising the need for making the competition process more efficient. A rising sensitivity to acoustic-phonetic detail probably facilitates this development, since it coincides with increased lexical efficiency (McMurray et al. 2018). Considering these findings, it is not surprising that adults are highly sensitive to acoustic-phonetic detail and that this sensitivity benefits their perception—for instance, through allowing them to predict upcoming speech (Andruski, Blumstein & Burton 1994; Kapnoula et al. 2017; McMurray, Tanenhaus & Aslin 2002). To summarize, children grow increasingly sensitive to acoustic-phonetic detail and use this ability to process phonemes and words more and more efficiently. In linguistic theory, these developments imply increasingly more mature phonemic and word representations in the brain.

We now turn to the second competence necessary for adequate reduced-speech perception (flexibility in cue use). Adults’ extensive (non-)linguistic knowledge allows them to recognize various cues (e.g., acoustic-phonetic, syntactic, and semantic cues) in the speech stream and use these flexibly to fill in missing parts and predict upcoming speech. This flexibility is particularly visible in adults’ adaptations of cue weightings to the clearness of speech. In experiments, such clearness is manipulated by changing the signal-to-noise ratio (Mattys, White & Melhorn 2005), or the speaking style (Bradlow & Alexander 2007), or both (van der Feest, Blanco & Smiljanic 2019). When changing the speaking style, acoustic-phonetic cues are either enhanced (as in “clear speech” or infant-directed speech; Bradlow & Alexander 2007; van der Feest, Blanco & Smiljanic 2019) or reduced (as in “conversational” speech; van der Feest, Blanco & Smiljanic 2019). Interestingly, manipulating the signal-to-noise ratio yields similar cue weighting adaptations in adults to manipulating the speaking style. In general, the clearer the speech is, the larger the weighting of larger-scale contextual cues (such as semantic cues; Bradlow & Alexander 2007) and conversely, the less clear it is (as in noisy or reduced speech), the larger the weighting of smaller-scale acoustic-phonetic cues (Mattys, White & Melhorn 2005; van der Feest, Blanco & Smiljanic 2019).

Children can also fill in missing information and predict upcoming words already at a young age (Cole & Perfetti 1980; Fernald et al. 1998; Fernald, Swingle & Pinto 2001). For instance, 18-month-olds were found to look at the correct picture (rather than another picture) before the depicted word was pronounced in full (Fernald et al. 2001). However, children are less flexible in adapting their cue weightings when the clearness of speech is reduced. In studies with children, this was done primarily through noise addition. When speech is presented in noise, children of different ages struggle more with perception than older children and adults, irrespective of whether the noise replaces or masks speech and irrespective of whether it masks (or replaces) a phoneme, word, or sentence. For instance, 5- to 6-year-olds have more problems recognizing familiar single words in which a phoneme has been replaced by noise than adults, who generally experience a “phoneme restoration effect” (Warren 1970) in this case, i.e., an illusion that the word is complete (Newman 2004; Walley 1988); 11- to 13-year-olds are less competent than 15- to 17-year-olds in recognizing a final predictable word in a sentence presented in noise (Elliott 1979; similar results in Smiljanic & Sladen 2013); and children continue

having trouble with correctly repeating nonsense words that are degraded with reverberation and noise at least until the age of 15 (Johnson 2000; similar results in Taitelbaum-Swead & Fostick 2016).

Hence, children can use word and semantic knowledge to recover and predict words (even at a young age) but are less competent in this than adults when tested in noise (even at an older age). Although it has been hypothesized that this incompetence reflects predominantly lesser word and semantic knowledge, resulting in a smaller ability to use larger-scale contextual cues (which is definitely a possibility; Newman 2006; Nittrouer & Boothroyd 1990), adults' behavior in speech-in-noise experiments highlight an additional possibility, namely an insufficient flexibility to switch attention to smaller-scale acoustic-phonetic cues (aforementioned competence 2). Relatedly, children may not be able to pick up these cues in all detail (competence 1), which would also hamper such flexibility.

In sum, two competences—a sensitivity to acoustic-phonetic cues (“rich” phonemic representations) and an ability to use such cues flexibly (“stable” representations that facilitate phoneme identification and word recognition)—grow increasingly better throughout childhood and adolescence as language experience accumulates. Since the competences are vital for adequate reduced-speech perception, it is likely that reduced-speech perception also continues developing in this time and that hence reduced-speech perception is still not as proficient in adolescents as in adults. The present study is the first to examine this hypothesis. Specifically, we examine whether reduced-speech perception is still immature in 16-year-olds as compared to young adults whose native language is German. We target 16-year-olds rather than younger ages because we expect still immature perception at this age in late adolescence, when children continue learning new vocabulary and refining their language skills. Additionally, this study is part of a larger project that includes second-language learners in high school (in our case, Dutch adolescents learning German), and even simple reduced speech appeared too difficult for younger learners.

Given the aforementioned literature, we expect the adolescents to get in trouble particularly when the need for relying on subtle acoustic-phonetic cues is acute. For examining this, we use two manipulations: a manipulation of the speaking style (reduced and unreduced) and a manipulation of the amount of context surrounding the target stimuli (without and with). The need for proficient acoustic-phonetic cue use is most acute when participants hear reduced targets without context: In reduced speech, the cues are limited and subtle and vary with the specific pronunciation; in fragments without context, listeners must rely on whatever of these cues remains.

2. Method

2.1. Participants

Participants were 36 German adults and 38 German high school pupils. As mentioned in the Introduction, this study is part of a larger project on adolescents' perception of German reduced speech, which includes nonnative adolescents (Wanrooij & Raijmakers, [Under review](#)). The German adolescents represent the control group in this larger project.

The adults were German students recruited at Leiden University (the Netherlands). The adolescents were recruited in the region of Hannover in northern Germany, which is a region where people tend to speak the most common variant of German. They attended a school type that prepares pupils for university; this was chosen to ensure comparable educational backgrounds. All participants were raised monolingually. The adolescents had not lived abroad at all, and the adults had not lived in a country where German is not the main language longer than 6 months before the age of 18. The length of residence in the Netherlands of the adults varied (range = 1 to 54 months) but was mostly short (median = 6 months, modus = 1 month) and did not differ between the two conditions, which are explained in the following (Bayes Factor expressing the odds of “no difference” versus “a difference” = 3.1). All participants signed informed consent forms, as did adolescents' parents. The procedure was approved by the Ethical Committee of the Department of Educational Sciences of Leiden University.

Each participant was randomly assigned to either the “Reduced” or “Unreduced” Condition (where they listened to either reduced or unreduced speech respectively). Table 1 lists their ages and sex distributions per Age Group and Condition. For the Condition, a between-subject design was preferred over a within-subject design to avoid nuisance factors that could influence intelligibility: Presenting the same participant with the same stimulus twice (once reduced and once unreduced) would elicit unwanted priming effects, and using dissimilar reduced and unreduced stimuli would evoke the risk of differences in intelligibility between reduced and unreduced stimuli such as more or less known words or shorter and longer fragments).

2.2. Phrase-intelligibility task

The main task used for examining an age effect in reduced-speech perception was a phrase-intelligibility task, in which participants were presented with speech fragments and then wrote down what they heard (Brown & Hilferty 1986; Ernestus, Baayen & Schreuder 2002; Ernestus, Dikmans & Giezenaar 2017; Henrichsen 1984; Norris 1995; Ten Bosch et al. 2016; Wong et al. 2015). This format is well suited for studying reduced-speech perception because it is possible to present more than a single word and thus to include reductions that involve multiple words, which occurs frequently in natural speech.

The first goal was to confirm effects previously reported for adults, namely a full-form advantage (unreduced speech should be easier to perceive than reduced speech; Drijvers, Mulder & Ernestus 2016; Ernestus, Baayen & Schreuder 2002; Ranbom & Connine 2007) and a context effect (context should facilitate perception; Ernestus, Baayen & Schreuder 2002; Janse & Ernestus 2011; Mitterer 2011; van de Ven, Tucker & Ernestus 2011). For examining the influence of context, the task was divided in two subtasks (Table 2). In subtask 1, the first presentation was always a full phrase (e.g., *schönen guten Tag*) and the second presentation a simple twofold repetition (e.g., *schönen guten Tag—schönen guten Tag*). In subtask 2, the first presentation was limited to only a part-phrase (e.g., *zusammen*, without any further “context”), and the second presentation consisted of the full phrase (i.e., with more “context”) plus the part-phrase again (e.g., *zusammen Kaffee trinken—zusammen*). (The repetition of the part-phrase was added to remind participants that they only had to write down the part-phrase. To make the procedure in subtask 1 identical to that in subtask 2, a third repetition was added in this subtask too).

The second and primary goal was to examine an effect of experience with the native language (which we label “age effect”). As explained in the Introduction, an age effect was expected chiefly when listeners had to rely most on whatever subtle acoustic-phonetic cues were available. Generally, the reliance on acoustic-phonetic cues was expected to grow more urgent and an age effect more likely in going from unreduced (full cues) to reduced (partial cues) phrases and in going from presentation 1B (full phrase, second presentation) to 1A (full phrase, first presentation) and 2B (full phrase, first presentation; after hearing part-phrase) to 2A (part-phrase, first presentation). Hence, an age effect was expected chiefly in presentation 2A, where new information was presented “without context” in reduced speech; and least in presentation 1B, where old information was presented “with context” in unreduced speech.

Since we expected an age effect mainly for the part-phrases, one may wonder why we included subtask 1 at all, particularly because all full phrases were probably simple for native listeners (they had to be usable for nonnative adolescent listeners too; Wanrooij & Raijmakers, *Under review*, as

Table 1. Number (*N*) of Adolescents and Adults in each Condition, their sex (F = female; M = male), and ages (years;months).

Age Group	Condition	<i>N</i> (sex)	Age	
			Mean	Range
Adolescents	Reduced	19 (14 F; 5 M)	16;02	15;05–17;03
	Unreduced	19 (11 F; 8 M)	16;02	15;09–16;08
Adults	Reduced	18 (13 F; 5 M)	24;01	20;05–32;02
	Unreduced	18 (14 F; 4 M)	23;09	18;03–28;09

Table 2. Trial examples in the two subtasks in the phrase-intelligibility task. In subtask 1, a full phrase was repeated three times. In subtask 2, a part-phrase was presented first, then the full phrase, and then the part-phrase again.^a

Presentation	Subtask 1	Subtask 2
A	1. <i>schönen guten Tag</i>	1. <i>zusammen</i>
B	2. <i>schönen guten Tag</i>	2. <i>zusammen Kaffee trinken</i>
	3. <i>schönen guten Tag</i>	3. <i>zusammen</i>

^aA reduced variant of each fragment was presented to participants in the Reduced condition, e.g., [ʃø:ŋguntə:k] for *schönen guten Tag*, and an unreduced variant to those in the Unreduced condition, e.g. [ʃø:nəngu:təntə:k].

mentioned in the Introduction). We did not limit the task to subtask 2, to be able to (i) test for the full-form advantage irrespective of limited context, as is done in subtask 1 (this had not been attested before in adolescents); (ii) examine the context effect in two ways rather one way (Section 3.1.2); and (iii) compare the results to those of nonnative listeners, for whom subtask 1 was already considered challenging, later (Wanrooij & Raijmakers, [Under review](#)).

2.2.1. Stimuli

The stimuli were based on spoken fragments selected from the Kiel Corpus of Spontaneous Speech (IPDS (Institut für Phonetik und digitale Sprachverarbeitung der Christian-Albrechts-Universität zu Kiel) 1995, IPDS 1996, IPDS 1997). All selected fragments were reduced short phrases that were comprehensible as stand-alone units (e.g., *schönen guten Tag*). They had been produced by different male and female speakers. To prevent differences in intelligibility due to these speakers' different voices rather than to differences in speaking style, the fragments were reproduced by a single male native speaker of German in a sound-proof booth. The speaker came from the region of Hannover and spoke with a "Standard German" accent. He was not a linguist and had experience as an actor. He was not informed about the phenomenon of reduced speech before the recordings. The speaker listened to several short fragments of German reduced speech and repeated each fragment immediately after hearing it. First, he was instructed to repeat the fragments as closely as possible to the original versions (yielding the "Reduced" stimuli). Then he was asked to repeat them as clearly as possible, as in formal speech ("Unreduced" stimuli). He produced almost all reduced versions with the intended reductions at first trial. Many unreduced versions required some additional recordings.

The duration of unreduced speech is usually longer than that of corresponding reduced speech, thus providing listeners not only with more cues but also with more processing and decision time. To clarify the influence of reductions separate from durational differences, the durations of the reduced and unreduced versions were equated by means of the PSOLA-procedure in the computer program Praat (Boersma & Weenink 1998–2018). In each pair of reduced and unreduced stimuli, the longer version was shortened and the shorter version lengthened (mean shortening factor = 0.94, standard deviation = 0.05). The adaptations did not result in strange sounding or unnatural stimuli: Two native listeners of German were asked to identify such stimuli, and no stimulus had to be removed. The shortness of the stimuli allowed us to avoid visual support (such as a written text presenting nontargets) and concomitant possible orthographic or other visual nuisance effects. Such support becomes necessary when presenting longer stretches of speech, which are more demanding to memorize (as in, e.g., Ernestus, Dikmans & Giezenaar 2017; Ten Bosch et al. 2016).

Appendices A and B list the 48 stimuli with their phonetic transcriptions in each version (reduced and unreduced), for subtask 1 (24 stimuli) and 2 (24 stimuli) respectively. Each stimulus contained one to four targets, i.e., words that were reduced in the Reduced version and unreduced in the Unreduced version. The stimuli contained different types and degrees of reductions. The main type that occurred in almost every reduced stimulus (except in stimulus numbers 25, 35, 37, 41, 43, 44, and 46 in Appendix B) were reductions of unstressed consonant-schwa-nasal (CəN) combinations, which are frequent in German (e.g., *brauchen* [brauxən] 'to need,' with CəN = [xən]). Among these, the most frequent type were reductions of unstressed plosive-schwa-/n/ sequences. The typical pattern that can

be discerned when arranging tokens with such sequences from least to most reduced is as follows: (0) there is no reduction, e.g., *Gelegenheit* [gələ:ɡənhart] ‘occasion,’ where CəN = [gən]; (1) the schwa drops [gələ:ɡnhart]; (2) if the consonant’s place of articulation is bilabial or velar, i.e., differs from that of the nasal /n/, the nasal’s place of articulation shifts to that of the consonant [gələ:ɡnhart]; (3) this consonant becomes a nasal resulting in a geminate [gələ:ŋnhart]; (4) the geminate reduces to a single nasal [gələ:nhart] (Kohler 2001). The phonetic transcriptions in [Appendices A](#) and [B](#) demonstrate that all these degrees featured in the stimuli.

The stimuli also contained reductions in the form of further assimilations with surrounding speech sounds within and across words, beyond the four-step pattern just described. Particularly highly frequent words and phrases are prone to such extremely reduced patterns. For instance, the reduced variant of *eigentlich* ([ʔaɪɡəntliç] ‘actually’ (with CəN = [gən]) in the stimuli is not simply [ʔaɪntliç] (which would be the end product of the four-step pattern and which also occurs as a token in the corpus) but [amɪ] (see Niebuhr & Kohler 2011 for a description of variants of *eigentlich* in the corpus). And the reduced version of *haben wir* ([ha:bən vi:v] (CəN = [bən]) is one step beyond [ha:m vɐ] (which contains the end product of the four-step pattern for the verb), namely [ha:mɐ]. Another reduction type in the stimuli was the nasalization of plosives in voiced nasal-plosive combinations where nasal and plosive have the same place of articulation (Kohler 2001). An example is *Kalenner* [khalənnɐ], a reduced variant of *Kalender* ‘calendar,’ in which /nd/ is realized as [nn]. We did not use the precise degree of reduction as a factor in the analyses but simply compared answers to reduced and unreduced speech. This was because most phrases were highly reduced (i.e., step 4), and the scoring of the precise “amount of reduction” was difficult due to the fact that tokens were natural (not synthetically controlled) and included complex reductions across words.

As visible in the [Appendices \(A and B\)](#), reduced targets included both function words (such as auxiliary verbs) and content words (such as adjectives, nouns, and verbs). The order of the stimuli was randomized for each participant separately, with one condition—if a word appeared more than once, it appeared first as a target and later appearances were not targets. For example, the word *haben* ‘have’ appeared first as a target ([Appendix B](#), stimulus number 40) and later as a nontarget (stimuli 32 and 41). Stimuli were presented with the software package E-prime (Version 2.0, Psychology Software Tools, Inc., 2012).

2.2.2. Procedure

In each trial, all participants wrote down their answers twice: once after the first presentation of a fragment (answer A) and once after the second plus third presentations (answer B; [Table 2](#)). Between writing down answers A and B, participants judged the intelligibility of the fragment on a 5-point scale. This intelligibility judgment was added to verify whether accuracies would reflect experienced difficulty. The task was self-paced.

Participants wore headphones. Before the task, the sound level was adjusted to each participant’s preferred level. Before each subtask, an explanation appeared on the screen, and participants performed a practice trial, which proceeded in precisely the same way as the subtask after it. Participants also received the explanation on paper and could ask questions before and after practising. They were informed how to write German umlauts on the laptops (i.e., the letters ä, ö, and ü): They could leave the dots out or write the basic letter plus an e (yielding ae, oe, and ue). They were also informed that they would not have to write capital letters and that the phrases that they were about to hear were taken from dialogs in which people make appointments with one another.

Testing took place in a quiet room, at a university (adults) and in a school (adolescents; in a classroom in a separate part of the school, not adjacent to other classrooms). The adolescents were tested during school hours. Participants were tested individually or in small groups up to four persons (adults) or seven persons (adolescents). When testing was done in small groups, the explanations were done simultaneously and individuals were seated behind nonadjacent desks.

2.3. Auditory lexical decision task

After the phrase-intelligibility task, participants did an auditory lexical decision task (Holley-Wilcox 1977; McCusker, Holley-Wilcox & Hillinger 1979). This task is common in reduced-speech perception research because it allows for measuring processing speed, in addition to accuracies (e.g., Connine, Ranbom & Patterson 2008; Morano, Ernestus & Ten Bosch 2015; Pitt, Dilley & Tat 2011; Poelmans 2003; Tucker 2011). Participants are presented with a list of words and pseudowords (or nonwords), and they indicate as quickly and accurately as possible if the just-presented item is a word or not. The reaction times are usually faster for words than pseudowords. The commonly given explanation is that the reaction times reflect how long listeners spend searching for a stimulus in the mental lexicon: Decisions for words can be made as soon as the word is found; for pseudowords, however, extra time elapses before listeners decide that their search is in vain. We will test for this “item effect” as one way of checking the validity of the task. We used this task so that we could measure processing speed (in addition to accuracies). Also, the task provided another “difficult listening situation,” in which sensitivity to subtle acoustic-phonetic cues and the ability to use such cues flexibly were required and hence where we expected an age effect. This is because the stimuli in this task are always isolated items (“without” context), which are presented under time and accuracy pressure. A reliance on subtle acoustic-phonetic cues was anticipated to be more urgent and consequently an age effect more likely for reduced items (only partial cues) than unreduced items (full cues) and also for pseudowords than words. This is because for pseudoword perception, cues cannot be mapped, and should also not be mapped accidentally, onto word representations. In other words, listeners have to weigh acoustic-phonetic cues more heavily than “larger-scale” cues (in this case word knowledge), an ability that children seem to struggle with (see Section 1).

2.3.1. Stimuli

Appendix C lists the stimuli in the lexical decision task. Stimulus creation and presentation were the same as in Task 1. There were 40 words and 40 pseudowords (22 two-syllable items and 18 three-syllable items for each of these item types) in each version. All words were German verbs, and all pseudowords were German pseudoverbs. All onsets (e.g., “bl” in *bleiben* ‘to stay’ and “ver” in *versuchen* ‘to try’; see Appendix C), and final syllables of the words were reused to create the pseudowords (on paper only; we did not resynthesize the sounds). The middle portion of the pseudowords was chosen as much as possible in such a way that the onset of the last syllable in each stimulus provided the cue to whether it was a real word or not. Since the stimuli were produced by a native speaker and not (re-)synthesized, the preceding speech sounds already probably contained cues hinting at the nature of these onsets. Listeners who can pick up these subtle cues well are thus at an advantage. In all unreduced versions, the final syllable consisted of a consonant–schwa–[n] sequence. In the reduced versions, this sequence was reduced according to the pattern described previously as the main reduction pattern used in the phrase-intelligibility task (Section 2.2.1). The degree of reduction differed per stimulus but was similar in corresponding words and pseudowords. Stimuli were presented with the software package E-prime (Version 2.0, Psychology Software Tools, Inc., 2012). Chronos button boxes (Model PST-100430, Psychological Software Tools) were used to collect responses.

2.3.2. Procedure

The task was performed after the phrase-intelligibility task, in the same session. Before the task, the sound level was adjusted to each participant’s preferred level again, and participants received instructions in the same way as in the previous task.

In each trial, participants first saw a star on the screen (1,000 ms), and then they heard either a German verb or a pseudoverb, while the screen showed *NEIN* ‘no’ in the left corner and *JA* ‘yes’ in the right corner. Their task was to indicate as quickly and as accurately as possible whether they considered the just-heard stimulus a German word or not by clicking on either the left button on the button box, which was labeled “N” for *NEIN*, or on the right button labeled “J” for *JA*. Participants

were informed that all words were German verbs and all pseudowords were German pseudoverbs. A next trial appeared after a participant clicked a button or after 3,000 ms after stimulus onset.

Before the task, participants did two practice exercises. The aim of the first exercise (20 practice trials) was to practice that the “no-button” was on the left and the “yes-button” on the right of the button box. For this, either *NEIN* or *JA* appeared in blue in the middle of the screen (500 ms), and participants had to click the correct button as quickly as possible (i.e., the left button after *NEIN* and the right button after *JA*). During the decision time, the screen showed *NEIN* ‘no’ in the left corner and *JA* ‘yes’ in the right corner, just as in the later real task. There was no sound in this exercise, and the next trial would only appear after the participant had clicked a button. The eight trials in the second exercise proceeded in exactly the same fashion as the trials in the subsequent task.

3. Analyses and results

We used nonparametric tests for ordinal measures and whenever distributions of ratio measures were nonnormal and could not be transformed. Bonferroni corrections for multiple testing were applied; the resulting alpha is reported per analysis. Whenever the *p* value is reported to be one-sided, we test for a specific direction of an effect.

3.1. Phrase-intelligibility task

Two dependent variables were measured: (i) accuracies and (ii) intelligibility judgments (“judgments”). Accuracies were the primary dependent variable. For the accuracies, participants’ answers were analyzed with a script (written in the computer program Praat; Boersma & Weenink 1998–2018), which disregarded small spelling mistakes (e.g., full points were given for the target *hervorragend* ‘excellent’ when written with only one *r* and for the target *noch mal* ‘again’ when written without a space as “nochmal”). Points were given only for targets. The maximum accuracy for each *stimulus* (which consisted of one to four targets; see [Appendices A and B](#)) was 1 point. Hence, a participant could obtain a maximum accuracy of 48 points. (The maximum accuracy for each *target* was calculated as 1 divided by the number of targets in the stimulus). The judgments were ordinal measures on a 5-point scale ([Section 2.2.2](#)). In the following, the numbers 1 and 2 refer to subtasks 1 and 2 respectively; the letters A and B to the first and second answers respectively.

3.1.1. Accuracies per stimulus

To check the validity of the stimuli, we first inspected how many participants had each stimulus correct (the “stimulus accuracy”). Since the stimuli had been developed to be also usable for nonnative listeners, it was expected that participants would find it relatively easy to give the right answers, at least for the second presentation of the stimulus. Consequently, it was relevant to reconsider any stimulus with a “low” stimulus accuracy, which was operationalized as follows: If less than half of the participants gave the correct answer for a stimulus, the assignment of points for this stimulus would have to be adapted (if possible) or the stimulus would have to be removed. The former option was necessary for only one stimulus (*ganz hervorragend passen*; [Appendix A](#), stimulus number 12): Many participants in the Reduced condition (i.e., for A: 4 Adolescents and 4 Adults; for B: 6 Adolescents and 3 Adults) responded “fassen” instead of *passen*. This is not surprising in view of the bilabiality of both /f/ and /p/. Additionally, some participants (i.e., for A: 5 Adolescents and 1 Adult in the Unreduced condition and 2 Adolescents and 1 Adult in the Reduced condition; for B: 1 Adolescent and 1 Adult in the Unreduced condition and 3 Adults in the Reduced condition) responded “passend” instead of *passen*. This is not surprising either: /n/ and /d/ are both alveolar (stop) consonants. In view of these considerations, “fassen” and “passend” were also counted as correct answers in both conditions. After these adaptations, not a single stimulus had to be removed.

Overall, the stimuli seemed easy to comprehend for the German participants. This held in particular for the unreduced versions, where minimally 34 out of 37 participants gave the correct

answer to the second stimulus presentation (answer B). A Wilcoxon signed-rank test for paired samples (each pair consisting of a reduced and the corresponding unreduced version) confirmed that the reduced versions (median across subtasks = 35.7) were harder to comprehend than the unreduced versions (median = 36.60; $T = 54.50$, $z = 4.3$, $p = 1.027e-9$, $r = .44$; the effect size r is calculated as: $r = z/\sqrt{N} = 4.3/\sqrt{96}$; Rosenthal 1991).

3.1.2. Accuracies per participant

The distributions of the accuracies as measured per participant were nonnormal, due to the fact that participants, particularly those in the Unreduced condition, tended to score at ceiling (which confirms that the stimuli were relatively easy for the native listeners; see previous section). Therefore, we report medians and ranges per Age Group and Condition (Table 3) and show the accuracies per participant (Figure 1).

First, we tested whether the data would support earlier findings of the “context effect”, i.e., that stimuli are easier to process with than without context. Since all participants heard both phrases with and phrases without context, the effect was analyzed within subjects, and it was analyzed in two ways: by comparing the accuracies of answers 2A (part-phrases, i.e., “without context”) versus 2B (corresponding full phrases, i.e., “with context”) and versus 1A (other full phrases, “with context”; Table 2). Two Wilcoxon signed-rank tests were performed per Age Group ($\alpha = 0.05/4$ tests = 0.0125; one-sided ps): one testing for Context1 (2A vs. 2B) and one for Context2 (2a vs. 1A). Participants’ accuracies were indeed significantly lower for 2A than 2B and for 2A than 1A (Adults 2A vs. 2B: median 2A = 22.33, median 2B = 24.00, $T = 3.00$, $z = -4.221$, $p = 1.334e-05$, $r = -.70$; Adolescents 2A vs. 2B: median 2A = 21.33, median 2B = 23.00, $T = 21.00$, $z = -4.154$, $p = 1.641e-5$, $r = -.67$; Adults 2A vs. 1A: median 1A = 23.25, $T = 52.00$, $z = -2.805$, $p = .003$, $r = -.47$; Adolescents 2A vs. 1A: median 1A = 23.08, $T = 61.50$, $z = -3.788$, $p = 7.894e-5$, $r = -.61$). Thus, the context effect is replicated for adults and now also established for adolescents.

Second, we tested whether the data would support earlier findings of the “full-form advantage”. Since participants heard either reduced or unreduced phrases, the effect was analyzed between subjects. Eight Mann-Whitney tests ($\alpha = 0.05/8 = 0.00625$; one-sided ps) for independent samples (Reduced vs. Unreduced) were conducted, one for each Age Group, subtask (1 and 2) and answer (A and B). Reduced accuracies were indeed lower than Unreduced accuracies (Adolescents 1A: $U = 26.5$, $z = -4.533$, $p = 2.91e-6$, $r = -.74$; Adolescents 1B: $U = 44.0$, $z = -4.004$, $p = 3.11e-5$, $r = -.65$; Adolescents 2A: $U = 6.5$, $z = -5.118$, $p = 1.54e-7$, $r = -.83$; Adolescents 2B: $U = 79.5$, $z = -3.022$, $p = .001$, $r = -.49$;

Table 3. Median accuracies and ranges for adolescents and adults per condition (Reduced and Unreduced), subtask (1 and 2), and answer (A and B) in the phrase-intelligibility task.

Age group	Condition	Median	Range	(Range in %)
Answer 1A				
Adolescents	Reduced	21.75	18.00 ~ 23.50	(75.00 ~ 97.92)
	Unreduced	23.75	21.92 ~ 24.00	(91.32 ~ 100)
Adults	Reduced	22.00	20.50 ~ 23.50	(85.42 ~ 97.92)
	Unreduced	24.00	21.50 ~ 24.00	(89.58 ~ 100)
Answer 1B				
Adolescents	Reduced	21.50	16.50 ~ 24.00	(68.75 ~ 100)
	Unreduced	23.25	22.67 ~ 24.00	(94.44 ~ 100)
Adults	Reduced	22.33	20.33 ~ 23.25	(84.72 ~ 96.88)
	Unreduced	24.00	23.50 ~ 24.00	(97.92 ~ 100)
Answer 2A				
Adolescents	Reduced	20.00	16.33 ~ 22.00	(68.06 ~ 91.67)
	Unreduced	23.00	20.33 ~ 24.00	(84.72 ~ 100)
Adults	Reduced	21.83	16.67 ~ 23.00	(69.44 ~ 95.83)
	Unreduced	24.00	22.00 ~ 24.00	(91.67 ~ 100)
Answer 2B				
Adolescents	Reduced	22.00	17.00 ~ 24.00	(70.83 ~ 100)
	Unreduced	24.00	20.33 ~ 24.00	(84.72 ~ 100)
Adults	Reduced	23.83	22.17 ~ 24.00	(92.36 ~ 100)
	Unreduced	24.00	23.00 ~ 24.00	(95.83 ~ 100)

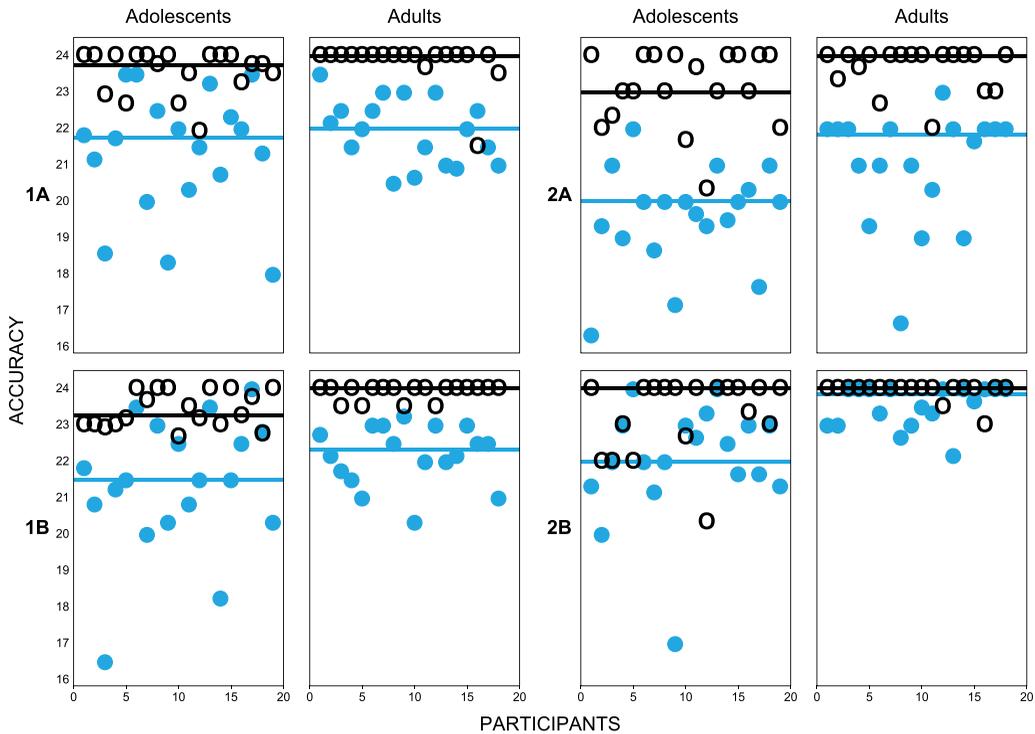


Figure 1. The accuracies in subtasks 1 (left) and 2 (right) in the phrase-intelligibility task, for answers A (top) and B (bottom), per Adolescent and Adult (see top titles) in the Reduced (blue discs) and Unreduced (black circles) conditions. Horizontal lines represent the medians.

Adults 1A: $U = 12.0$, $z = -4.934$, $p = 4.03e-7$, $r = -.82$; Adults 1B: $U < 0.001$, $z = -5.292$, $p = 6.05e-8$, $r = -.88$; Adults 2A: $U = 7.0$, $z = -5.040$, $p = 2.33e-7$, $r = -.84$; Adults 2B: $U = 97.0$, $z = -2.524$, $p = .006$, $r = -.42$). This replicates the full-form advantage for adults and establishes it for adolescents.

Subsequently, the main research question was addressed: We tested whether adolescents score worse than adults, particularly when listening to reduced part-phrases (answers 2A). Eight Mann-Whitney tests (Adolescents vs. Adults) were conducted ($\alpha = 0.05/8 = 0.00625$), one for each Condition (Reduced and Unreduced), subtask (1 and 2), and answer (A and B).

As expected, Adolescents listening to reduced speech scored significantly lower than Adults listening to reduced speech for the part-phrase answer 2A ($U = 77.0$, $z = -2.888$, $p = .003$, $r = -.47$). Additionally, they scored significantly lower for the subsequent full-phrase answer 2B ($U = 49.5$, $z = -3.756$, $p = 1.73e-4$, $r = -.62$). Hence, adolescents seem more confused than adults by reduced forms presented out of context and continue having trouble with these forms even after being presented subsequently with the forms in context.

As expected, accuracies did not differ significantly between the Age Groups listening to unreduced speech with one exception: An age effect appeared when unreduced full phrases were repeated, i.e., precisely in the most unexpected case when acoustic-phonetic cue use was considered easiest (1A: $U = 111.5$, $z = -2.122$, $p = .034$; 1B: $U = 76.0$, $z = -3.152$, $p = .001$, $r = -.52$; 2A: $U = 118.5$, $z = -1.746$, $p = .083$; 2B: $U = 112.0$, $z = -2.294$, $p = .021$). However, this age effect should be disregarded: When inspecting Adolescents' answers, we noticed many perfect (i.e., ceiling) answers to the first presentation. Therefore, we suspect that they considered it unnecessary to give the same answer to a simple stimulus for a second time.

Nonparametric tests do not yield interactions. Therefore, we inspected a possible interaction between Age Group and Condition differently: We calculated the accuracy per stimulus for each Age Group

x Condition (i.e., Adolescents Reduced, Adolescents Unreduced, Adults Reduced, Adults Unreduced). Then we computed the difference between the accuracy in the Unreduced condition and the accuracy in the Reduced condition per Age Group. One Wilcoxon signed-ranks test per answer (A and B) showed a larger difference for Adolescents (A: median = 1.00, range = -3.00 to 18.00; B: median = 0.50, range = -2.00 to 9.00) than Adults (A: median = 0.00, range = -1.00 to 17.00, $T = 199.00$, $z = -2.317$, $p = .019$, $r = -.33$; B: median = 0.00, range = -1.00 to 8.50, $T = 204.50$, $z = -2.236$, $p = .024$, $r = -.32$). This reflects that adolescents have more trouble with reduced speech in particular than adults.

3.1.3. Intelligibility judgments

Judgments were first inspected for consistency. Judgments of two Adolescents in the Unreduced condition were removed because a comparison with their own accuracies and those of other Adolescents revealed that they most likely judged the stimuli reversely (i.e., the most difficult stimulus as the most intelligible and vice versa; for instance, they had higher accuracies in subtask 1 than 2, as expected, but judged the stimuli in subtask 1 as more difficult). Table 4 displays participants' median intelligibility judgments (disregarding these two participants).

Two Wilcoxon signed-rank tests ($\alpha = 0.05/2 = 0.025$; one-sided p s), one per Age Group, showed that Adolescents and Adults judged isolated targets (subtask 2) to be less intelligible than those embedded in a phrase (subtask 1) (Adults 2 vs. 1: median 1 = 4.25, median 2 = 4.00, $T = 17.00$, $z = -2.680$, $p = .003$, $r = -.45$; Adolescents 2 vs. 1: median 1 = 4.00, median 2 = 4.00, $T = 10.00$, $z = -3.255$, $p = 6.187e-4$, $r = -.54$). Hence, the context effect is not only reflected in participants' accuracies (see previous) but also in their intelligibility judgments.

Participants in the Reduced condition judged the stimuli to be less intelligible than participants in the Unreduced condition in both Age Groups in subtask 2 and in the Adolescent group in subtask 1; Adults' judgments in subtask 1 approached ceiling judgments in both conditions (four Mann-Whitney tests, one per Age Group and subtask, $\alpha = 0.05/4 = 0.0125$; Adolescents 1: $U = 60.50$, $z = -3.363$, $p = 7.71e-4$, $r = -.56$; Adolescents 2: $U = 51.00$, $z = -3.644$, $p = 2.68e-4$, $r = -.61$; Adults 1: $U = 118.50$, $z = -1.536$, $p = .167$; Adults 2: $U = 54.00$, $z = -3.622$, $p = 2.92e-4$, $r = -.60$). Hence, the full-form advantage is not only reflected in participants' accuracies (see previous) but also in their intelligibility judgments.

Four Mann-Whitney tests ($\alpha = 0.05/4 = 0.0125$) checked whether Adolescents and Adults in each Condition differed significantly in their intelligibility judgments. This was the case only for the Reduced condition in subtask 1 (Reduced 1: $U = 86.00$, $z = -2.724$, $p = .006$, $r = -.45$; Reduced 2: $U = 121.50$, $z = -1.562$, $p = .123$; Unreduced 1: $U = 151.50$, $z = -0.057$, $p = 1.0$; Unreduced 2: $U = 128.00$, $z = -0.923$, $p = .328$). Hence, adolescents judge reduced full phrases to be less intelligible than adults.

3.2. Auditory lexical decision task

Left-handed participants were removed from the data set (1 from Adolescents Reduced; 3 from Adolescents Unreduced; 2 from Adults Reduced, and 2 from Adults Unreduced). All analyses were

Table 4. Intelligibility judgments: medians and ranges for Adolescents and Adults per Condition (Reduced and Unreduced) and subtask (1 and 2) in the phrase-intelligibilitytask.^a

Condition	Adolescents		Adults	
	Median	Range	Median	Range
Subtask 1				
Reduced	3.5	3 ~ 5	4	4 ~ 5
Unreduced	5	4 ~ 5	5	4 ~ 5
Subtask 2				
Reduced	3	1 ~ 4	3.75	2 ~ 5
Unreduced	4	3 ~ 5	5	4 ~ 5

^aJudgments are those without the two participants who probably judged the stimuli reversely (see text).

performed on the log-transformed (base 10) values of the reaction times, which had been measured from stimulus onset. Undefined answers, which occurred when a participant did not click the button for an item, were removed from the data set, as were excessively short or long reaction times, i.e., items with log-transformed reaction times that lay outside the range of plus or minus 2 standard deviations from the mean of each Group's answers to each Item Type (following e.g., Poelmans 2003; Item Types are: Words and Pseudowords; Groups are: Adults Reduced, Adults Unreduced, Adolescents Reduced, and Adolescents Unreduced). Inaccurate answers were also removed from the data set.

Two mixed-design ANOVAs (i.e., analyses of variance with between-subjects and within-subjects independent variables; $\alpha = 0.05/2 = 0.025$) were performed, one with the accuracies (i.e., the number of correct responses) and the other with the log-transformed reaction times as the dependent variable. Both had Age Group (Adolescents vs. Adults) and Condition (Reduced vs. Unreduced) as between-subjects variables and Item Type (Words vs. Pseudowords) as within-subjects variable. Both included all possible interaction effects and used Type-III Sum of Squares.

3.2.1. Accuracies

The top row of Figure 2 shows the accuracies in response to Words and Pseudowords for Adolescents and Adults in each Condition. The top half of Table 5 presents the corresponding means and 95% confidence intervals (henceforth CIs).

The ANOVA showed a grand mean of 35.10 with a CI (+34.58 ~ +35.62) that did not include 20 (the chance level). This supports the validity of the test: It shows that German participants, irrespective of Age Group or Condition, score significantly above chance level in this task. The ANOVA also showed main effects of Age Group (mean difference Adolescents–Adults = -1.42 , CI = $-2.45 \sim -0.38$, $F[1,64] = 7.472$, $p = .008$), Condition (mean difference Reduced–Unreduced = -1.89 , CI = $-2.93 \sim -0.86$, $F[1,64] = 13.35$, $p = 5.236e-4$), and Item Type (mean difference Words–Pseudowords = 4.13 , CI = $+3.22 \sim +5.04$, $F[1,64] = 82.54$, $p = 4.022e-13$). As for Age Group, Adults (mean = 35.81, CI = $+35.08 \sim +36.54$) had more items correct than Adolescents (mean = 34.39, CI = $+33.66 \sim +35.13$), thus showing an age effect. As for Condition, participants in the Unreduced condition (mean = 36.05, CI = $+35.31 \sim +36.79$) had more items correct than participants in the Reduced condition (mean = 34.16, CI = $+33.44 \sim +34.88$), thus replicating the full-form advantage for isolated items. As for Item Type, participants had more correct responses to Words (mean = 37.16, CI = $+36.69 \sim +37.64$) than Pseudowords (mean = 32.97, CI = $+31.99 \sim +33.95$), thus demonstrating an item effect.

Additionally, the ANOVA yielded significant interactions between Age Group and Item Type ($F[1,64] = 5.87$, $p = .018$) and between Condition and Item Type ($F[1,64] = 10.15$, $p = .002$). Hence, differences in accuracy between responses to words and responses to pseudowords in a lexical decision task depend on age and speaking style. For the former interaction (Age Group x Items Type; Figure 3, left), a post hoc independent-samples (Adults vs. Adolescents) *t*-test showed that the accuracy difference between responses to Words and responses to Pseudowords (= dependent variable) was smaller for Adults than Adolescents (mean difference Adolescents – Adults = 2.32 , CI = $+0.38 \sim +4.27$, $t[66] = 2.39$, $p = .020$). Since Adults and Adolescents had very similar accuracies for Words (Figure 2), adolescents seem to have more problems in particular with pseudowords than adults. For the second interaction (Condition x Item Type; Figure 3, right), a post hoc independent-samples (Reduced vs. Unreduced) *t*-test showed that the accuracy difference between responses to Words and responses to Pseudowords (= dependent variable) was larger for the Reduced condition than for the Unreduced condition (mean difference Reduced–Unreduced = 2.96 , CI = $+1.07 \sim +4.85$, $t[66] = 3.12$, $p = .003$). This implies that listeners are more eager to recognize a word in a pseudoword when the item is reduced (and presented in isolation). The other interactions were not significant (Age Group x Condition: $F[1,64] = 0.97$, $p = .330$; Age Group x Condition x Item Type: $F[1,64] = 1.72$, $p = .195$).

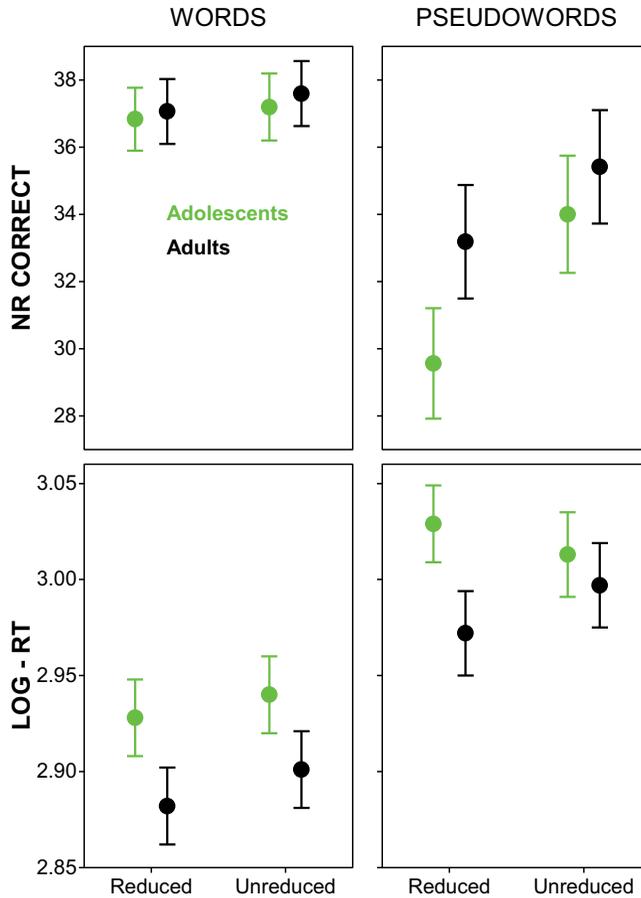


Figure 2. The mean number of correct answers (top row) and log-transformed reaction times (bottom row) in response to Words (left) and Pseudowords (right) in the lexical decision task, for Adolescents (green) and Adults (black) in each condition (see bottom titles). Bars show 95% confidence intervals.

Table 5. Accuracies (i.e., the number of correct answers; top) and log-transformed reaction times (with nontransformed equivalents between brackets; bottom; in milliseconds) in response to Words and Pseudowords in the lexical decision task, for Adolescents and Adults per condition (Reduced and Unreduced): means and 95% confidence intervals (CIs).

Age Group	Condition	Words		Pseudowords	
		Mean	CI	Mean	CI
Accuracies					
Adolescents	Reduced	36.83	+35.90 ~ +37.77	29.56	+27.92 ~ +31.20
	Unreduced	37.19	+36.19 ~ +38.18	34.00	+32.26 ~ +35.74
Adults	Reduced	37.06	+36.09 ~ +38.02	33.18	+31.49 ~ +34.87
	Unreduced	37.59	+36.62 ~ +38.55	35.41	+33.72 ~ +37.10
Reaction times					
		Mean	CI	Mean	CI
Adolescents	Reduced	2.93 (847.23)	+2.91 ~ +2.95	3.03 (1069.05)	+3.01 ~ +3.05
	Unreduced	2.94 (870.96)	+2.92 ~ +2.96	3.01 (1030.39)	+2.99 ~ +3.04
Adults	Reduced	2.88 (762.08)	+2.86 ~ +2.90	2.97 (937.56)	+2.95 ~ +2.99
	Unreduced	2.90 (796.16)	+2.88 ~ +2.92	3.00 (993.12)	+2.98 ~ +3.02

3.2.2. Reaction times

The bottom of Figure 2 shows the mean log-transformed reaction times that Adolescents and Adults had when responding correctly to Words (left) and Pseudowords (right). Table 5 (bottom) lists the corresponding means and CIs for each Age Group in each Condition.

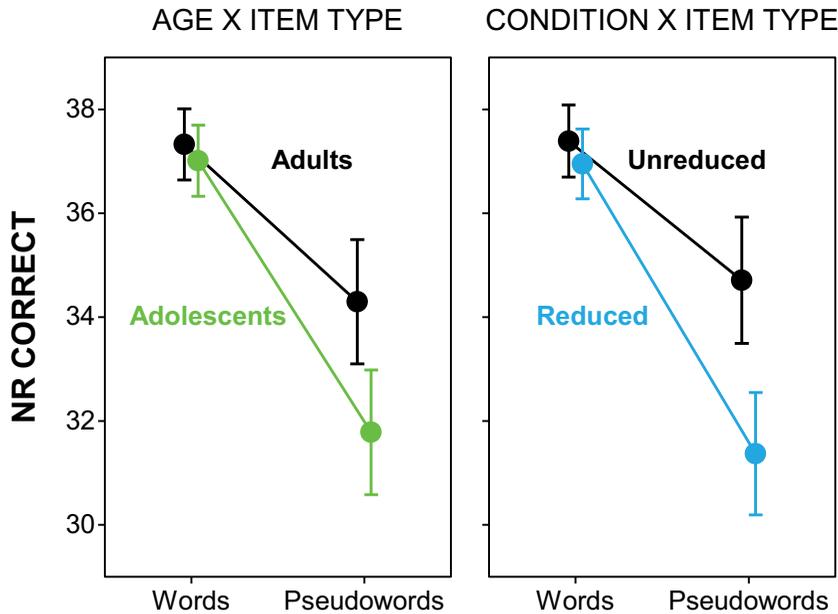


Figure 3. The significant interactions with Item Type (Words vs. Pseudowords) of Age Group (Adult vs. Adolescent; left) and Condition (Reduced vs. Unreduced; right), for the number of correct answers (dependent variable) in the lexical decision task.

The ANOVA (see Section 3.2) produced main effects of Age Group (mean difference Adolescents–Adults = 0.040, CI = +0.021 ~ +0.059, $F[1,64] = 17.37$, $p = 9.414e-5$) and Item Type (mean difference Words–Pseudowords = -0.090, CI = -0.097 ~ -0.082, $F[1,64] = 541.57$, $p = 6.153e-33$). As for Age Group, Adolescents (mean = 2.978, CI = +2.96 ~ +2.99; in nontransformed reaction times: mean = 951 ms, CI = +912 ~ +977 ms) had longer log-transformed reaction times than Adults (mean = 2.938, CI = +2.92 ~ +2.95; in nontransformed reaction times: mean = 867 ms; CI = +832 ~ +891 ms), thus demonstrating an age effect in reaction times. As for Item Type, participants had shorter log-transformed reaction times to Words (mean = 2.91, CI = +2.90 ~ +2.92; in nontransformed reaction times: mean = 813 ms; CI = +794 ~ +832 ms) than Pseudowords (mean = 3.00, CI = +2.99 ~ +3.01; in nontransformed reaction times: mean = 1,000 ms; CI = +977 ~ +1,023 ms). Hence, the reaction times also reflect the item effect. The other effects in the ANOVA were not significant (Condition: $F[1,64] = 1.096$, $p = .299$; Age Group x Condition: $F[1,64] = 1.590$, $p = .212$; Age Group x Item Type: $F[1,64] = 0.703$, $p = .405$; Condition x Item Type: $F[1,64] = 2.206$, $p = .142$; Age Group x Condition x Item Type: $F[1,64] = 4.894$, $p = .031$).

4. Discussion

The present study is the first to examine reduced-speech perception in adolescents. It demonstrates less proficient speech perception in German 16-year-olds than young adults (in their twenties) in general and particularly when reduced speech is presented “without context.” In that case, successful perception forces listeners to detect and flexibly use remaining acoustic-phonetic cues. Reduced speech is difficult in this respect, due to deletions of speech sounds (fewer cues), shortenings of speech sounds (less salient cues), and assimilations of speech sounds within and across words (cues end up in the context and thus become unavailable when the context is removed). Overall, both age groups perceived unreduced speech better than reduced speech and speech with context better than without context, which is in line with previous work (e.g., Ernestus, Baayen & Schreuder 2002; Section 1). The

“age effects,” which we interpret as reflecting differences in experience with the native language, appeared in a phrase-intelligibility task and an auditory lexical decision task, as follows.

In the phrase-intelligibility task, we presented listeners with part-phrases (without context) and full phrases (with context) in either reduced (limited subtle cues) or unreduced (full clear cues) speech. Overall, accuracies were high. This was not surprising, since the phrases had to be usable for nonnative listeners too (Wanrooij & Raijmakers, [Under review](#)) and thus were simple for our native listeners. Although we had to use nonparametric tests for dealing with ceiling scores (particularly for the adults), the outcomes matched the predictions: Adolescents had consistently lower accuracies than adults precisely when proficient acoustic-phonetic cue use was most required, namely when listening to part-phrases in reduced speech. Subsequently, their accuracies remained lower than those of adults when hearing the corresponding reduced full phrase. This situation can still be hard in terms of acoustic-phonetic cue use because cues picked up in the part-phrase must now be integrated with those in the full phrase. Adolescents seem to struggle more with this than adults. Additionally, an analysis based on accuracies per stimulus (which we used as a way to inspect the interaction between age group and speaking style, in the absence of a direct possibility to examine this interaction in nonparametric tests; [Section 3.1.2](#)) confirmed that adolescents have more problems than adults specifically with reduced speech (more than with unreduced speech).

In the auditory lexical decision task, we presented listeners with words and pseudowords (both “without” context) in either reduced or unreduced speech. A uniform age effect emerged: Adolescents had consistently lower accuracies and slower reaction times than adults generally. Also, they tended to hear words in pseudowords more than adults and thus seemed to make their word knowledge prevail over acoustic-phonetic cues signaling the pseudoword status more. This finding resembles earlier observations that nonnative listeners struggle with increasing the weight of acoustic-phonetic cues relative to larger-scale semantic cues when this is required in noise (Bradlow & Alexander 2007).

The lexical decision task did not provide evidence that adolescents had more trouble than adults specifically with reduced speech because it did not yield significant interactions involving the age groups (adolescents versus adults) and the speaking styles (reduced versus unreduced). This could mean different things. First, it could be that in fact there was an interaction, but the task did not reveal it. This possibility is suggested by the graphical display of accuracies and reaction times ([Figure 2](#)), where an age effect is visible for reduced and not for unreduced pseudowords (if tested separately, the former will be significant, the latter will not). A second possibility is that there were no interactions involving age groups and speaking styles because all items in this task can be considered more difficult for adolescents than adults in terms of acoustic-phonetic cue use. After all, context was always absent, and the task was performed under time and accuracy pressure. Even a slightly better ability to pick up subtle acoustic-phonetic cues that signaled the content of the final syllable before this syllable started in full (participants could only make a word or pseudoword decision when recognizing the onset of the final syllable) would give listeners a crucial advantage for all items. A third possibility is that there were no interactions involving age groups and speaking styles because the age effects were based on other factors than experience with the native language. We will now consider such other explanations.

For instance, it could be suspected that the adolescents were more tired, were concentrating less, or were less motivated to participate than the adults. Although we cannot rule out these possibilities, we do not have reasons to believe they were the case: Testing took place during school hours in a quiet classroom, and participation was voluntary. Similarly, there are no indications that the tasks were too demanding for adolescents. As mentioned, the stimuli were simple for native listeners, and scores were rather high. It could also be argued that the results were affected by the fact that the adults were tested in the Netherlands and the adolescents in Germany. However, if this were true, the age effects should have been smaller rather than larger, since adults’ exposure to and use of Dutch and English at the time

of testing more likely hampered than facilitated their performance in German, particularly as compared to the adolescents, who were immersed in a German environment at the time of testing.

An alternative explanation for the faster processing of speech in adults than adolescents in the lexical decision task is a domain-general increase in processing speed that can be observed until adulthood (Kail 1991). In adolescence, this increase in processing speed is probably related to brain maturation in the prefrontal cortex and the development of executive functions (Crone & Steinbeis 2017; Diamond 2013), which can be separated into the monitoring and updating of information in working memory, inhibitory control, and cognitive flexibility (Miyake et al. 2000). However, tasks that measure domain-general inhibitory control and cognitive flexibility largely yield adult-like levels before the age of 15 (Huizinga, Dolan & van der Molen 2006). Also, as pointed out by McMurray et al. (2018), measures of domain-general inhibitory control do not correlate well with measures of speech processing efficiency (Kapnola et al. 2017; Kong & Edwards 2016; see also Fernald, Perfors & Marchman 2006), and successful models of speech processing also implement inhibition at a local rather than a domain-general level (McClelland & Elman 1986). Hence, speech perception does probably not involve domain-general inhibitory control. Similarly, working memory does probably not exist as a domain-general mechanism in speech perception; rather, differences in processing capacity are tied to linguistic knowledge and experience (Andringa et al. 2012; MacDonald & Christiansen 2002). Hence, if inhibition and processing capacity (“working memory”) play a role in our age effect in reaction times, they more likely reflect differences in linguistic experience than a domain-general rise in processing speed. An additional counterargument against such a domain-general explanation is that this cannot clarify why apart from being slower in the lexical decision task, adolescents were also less accurate than adults in both tasks and judged reduced phrases in the phrase-intelligibility task as less comprehensible.

A direct argument in favor of an experience-based explanation is that the relative immature perception in our adolescents as compared to the adults is similar to the relative immature perception that we observed in another study (Wanrooij & Raijmakers, [Under review](#)) in nonnative as compared to native listeners. The former reflects a difference in experience related to age, the latter a difference in experience irrespective of age. Specifically, in the larger project where this study was part of, we observed increasingly better accuracies and faster reaction times in response to reduced-speech fragments in the same tasks as in this study, for listeners in the following order: nonnative adolescents (little experience), native adolescents (more experience), native adults (most experience; Wanrooij & Raijmakers, [Under review](#)). Nonnative listeners have also shown perceptual patterns similar to younger listeners when listening to speech in noise (Bradlow & Alexander 2007; Erickson & Newman 2017; Mayo, Florentine & Buus 1997).

Finally, an experience-based explanation for the age effects fits in well with diverse linguistic theories that view representations as gradually emerging with experience, e.g., the Gradual Learning Algorithm (Boersma 1998), the Native Language Magnet theory (Kuhl et al. 2008), neuroconstructivist approaches (Karmiloff-Smith 2006) and usage-based approaches (Tomasello 1992, 2009; Bybee 1999).

Based on this, we believe that the age effects in this study should not be attributed to nonlinguistic causes but to better linguistic capabilities in the adults than the adolescents, which are related to more linguistic experience. In line with the literature discussed in the Introduction (e.g., McMurray et al. 2018; [Section 1](#)), this study supports the idea that, as experience accumulates, listeners grow more sensitive to subtle acoustic-phonetic cues (implying “richer” phonemic representations) and are more able to use these cues flexibly depending on available information (“more stable” representations and more efficient lexical competition). These are two competences vital for reduced-speech perception. As also discussed in the Introduction, a major driving force behind these advances is probably the growth of the vocabulary in breadth and depth, which still continues in adolescence. As children know more words (breadth), the need for an efficient word recognition process increases. This puts pressure on the child to identify phonemes competently (Fernald et al. 1998; McMurray et al. 2018; Rigler et al. 2015; Sekerina & Brooks 2007) and thus to know the acoustic-phonetic details that define the various ways in which phonemes appear in natural speech. Experience with reduced variants (depth) probably helps them to learn these details. That children are exposed to reduced variants from the start of their

lives (Buckler, Goy & Johnson 2018) and that age effects are observed in adolescence in this study suggest that this is a protracted and dynamic process.

Acknowledgments

We thank Jos Pacilly, Malte Steinhoff and Mané van Veldhuizen for their indispensable assistance with technical matters, participant recruitment and testing.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the 'Nationaal Regieorgaan Onderwijsonderzoek' (Netherlands Initiative for Education Research) [grant 405-16-382 awarded to the first author].

References

- Andringa, Sible, Nomi Olsthoorn, Catherine van Beuningen, Rob Schoonen, and Jan Hulstijn. 2012. "Determinants of Success in Native and Non-native Listening Comprehension: An Individual Differences Approach." *Language Learning* 62 (suppl. 2):49–78. doi:10.1111/j.1467-9922.2012.00706.x.
- Andruski, J E., S E. Blumstein, and Martha Burton. 1994. "The Effect of Subphonetic Differences on Lexical Access." *Cognition* 52 (3):163–87. doi:10.1016/0010-0277(94)90042-6.
- Best, Catherine T., Gerald W. McRoberts, Rosemarie LaFleur, and Jean Silver-Isenstadt. 1995. "Divergent Developmental Patterns for Infants' Perception of Two Nonnative Consonant Contrasts." *Infant Behavior & Development* 18 (3):339–50. doi:10.1016/0163-6383(95)90022-5.
- Boersma, Paul. 1998. "Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives." Doctoral diss., University of Amsterdam: LOT Dissertation Series.
- Boersma, Paul, and David Weenink. 1998-2018. *Praat: Doing Phonetics by Computer [Computer Program]* (Version 6.0.23). Accessed January 19, 2017. <http://www.praat.org/>.
- Bradlow, Anne R., and Jennifer A. Alexander. 2007. "Semantic and Phonetic Enhancements for Speech-in-noise Recognition by Native and Non-native Listeners." *The Journal of the Acoustical Society of America* 121 (4):2339–49. doi:10.1121/1.2642103.
- Brown, James D., and Ann Hilferty. 1986. "The Effectiveness of Teaching Reduced Forms of Listening Comprehension." *RELC Journal* 17 (2):59–70. doi:10.1177/003368828601700204.
- Brown, James D., and Kimi Kondo-Brown, eds. 2006. *Perspectives on Teaching Connected Speech to Second Language Speakers*. Honolulu, HI: University of Hawai'i at Manoa, National Foreign Language Resource Center.
- Buckler, Helen, Huiwen Goy, and Elizabeth Johnson. 2018. "What Infant-directed Speech Tells Us about the Development of Compensation for Assimilation." *Journal of Phonetics* 66:45–62. doi 10.1016/j.wocn.2017.09.004.
- Bybee, Joan. 1999. "Usage-based Phonology." *Functionalism and Formalism in Linguistics* 1:211–42.
- Cole, Ronald A., and Charles A. Perfetti. 1980. "Listening for Mispronunciations in a Children's Story: The Use of Context by Children and Adults." *Journal of Verbal Learning and Verbal Behavior* 19 (3):297–315. doi:10.1016/S0022-5371(80)90239-X.
- Connine, Cynthia M., Larissa J. Ranbom, and David J. Patterson. 2008. "Processing Variant Forms in Spoken Word Recognition: The Role of Variant Frequency." *Perception & Psychophysics* 70 (3):403–11. doi:10.3758/PP.70.3.403.
- Crone, Eveline A., and Nikolaus Steinbeis. 2017. "Neural Perspectives on Cognitive Control Development during Childhood and Adolescence." *Trends in Cognitive Sciences* 21 (3):205–15. doi:10.1016/j.tics.2017.01.003.
- Darcy, Isabelle, Franck Ramus, Anne Christophe, Katherine Kinzler, and Emmanuel Dupoux. 2009. "Phonological Knowledge in Compensation for Native and Non-native Assimilation." In *Variation and Gradience in Phonetics and Phonology*, edited by Frank Kügler, Caroline Féry and Ruben van de Vijver, 265–309. Berlin, Germany: Mouton De Gruyter.
- Diamond, Adele. 2013. "Executive Functions." *Annual Review of Psychology* 64 (1):135–68. doi:10.1146/annurev-psy-113011-143750.
- Drijvers, Linda, Kimberley Mulder, and Mirjam Ernestus. 2016. "Alpha and Gamma Band Oscillations Index Differential Processing of Acoustically Reduced and Full Forms." *Brain & Language* 153-154:27–27. doi 10.1016/j.bandl.2016.01.003.

- Elliott, Lois L. 1979. "Performance of Children Aged 9 to 17 Years on a Test of Speech Intelligibility in Noise Using Sentence Material with Controlled Word Predictability." *The Journal of the Acoustical Society of America* 66 (3):651–53. doi:10.1121/1.383691.
- Erickson, Lucy C., and Rochelle S. Newman. 2017. "Influences of Background Noise on Infants and Children." *Current Directions in Psychological Science* 26 (5):451–57. doi:10.1177/0963721417709087.
- Ernestus, Mirjam. 2000. "Voice Assimilation and Segment Reduction in Casual Dutch. A Corpus-based Study of the Phonology-phonetics Interface." Doctoral diss., LOT, Utrecht.
- Ernestus, Mirjam, Harald Baayen, and Rob Schreuder. 2002. "The Recognition of Reduced Word Forms." *Brain and Language* 81 (1–3):162–73. doi:10.1006/brln.2001.2514.
- Ernestus, Mirjam, Mirte E. Dikmans, and Ghislaine Giezenaar. 2017. "Advanced Second Language Learners Experience Difficulties Processing Reduced Word Pronunciation Variants." *Dutch Journal of Applied Linguistics* 6 (1):1–20. doi:10.1075/dujal.6.1.01ern.
- Ernestus, Mirjam, and Natasha Warner. 2011. "An Introduction to Reduced Pronunciation Variants." *Journal of Phonetics* 39:253–60.
- Feest, Suzanne van der, Cynthia P. Blanco, and Rajka Smiljanic. 2019. "Influence of Speaking Style Adaptations and Semantic Context on the Time Course of Word Recognition in Quiet and in Noise." *Journal of Phonetics* 73:158–77. doi:10.1016/j.wocn.2019.01.003.
- Fernald, Anne, Amy Perfors, and Virginia A. Marchman. 2006. "Picking up Speed in Understanding: Speech Processing Efficiency and Vocabulary Growth across the 2nd Year." *Developmental Psychology* 42 (1):98–116. doi:10.1037/0012-1649.42.1.98.
- Fernald, Anne, Daniel Swingley, and John P. Pinto. 2001. "When Half a Word Is Enough: Infants Can Recognize Spoken Words Using Partial Phonetic Information." *Child Development* 72 (4):1003–15. doi:10.1111/1467-8624.00331.
- Fernald, Anne, John P. Pinto, Daniel Swingley, Amy Weinberg, and Gerald W. McRoberts. 1998. "Rapid Gains in Speed of Verbal Processing by Infants in the 2nd Year." *Psychological Science* 9 (3):228–31. doi:10.1111/1467-9280.00044.
- Field, John. 2003. "Promoting Perception: Lexical Segmentation in L2 Listening." *ELT (English Language Teaching) Journal* 57 (4):325–34. doi:10.1093/elt/57.4.325.
- Flege, James E., and Wieke Eefting. 1986. "Linguistic and Developmental Effects on the Production and Perception of Stop Consonants." *Phonetica* 43 (4):155–71. doi:10.1159/000261768.
- Hazan, Valerie, and Sarah Barrett. 2000. "The Development of Phonemic Categorization in Children Aged 6–12." *Journal of Phonetics* 28 (4):377–96. doi:10.1006/jpho.2000.0121.
- Henrichsen, Lynn E. 1984. "Sandhi Variation: A Filter of Input for Learners of ESL." *Language Learning* 34 (3):103–26. doi:10.1111/j.1467-1770.1984.tb00343.x.
- Hirsh-Pasek, Kathy, Deborah G. Kemler Nelson, Peter W. Jusczyk, Kimberley Wright Cassidy, Benjamin Druss, and Lori Kennedy. 1987. "Clauses are Perceptual Units for Young Infants." *Cognition* 26 (3):269–86. doi:10.1016/S0010-0277(87)80002-1.
- Holley-Wilcox, P. 1977. *The Effect of Homophony with Auditory Presentation of Stimuli*. Paper presented at the meeting of the Midwestern Psychological Association. Chicago, United States of America. May (cited in McCusker et al., 1981).
- Huizinga, Mariëtte, Conor V. Dolan, and Maurits W. van der Molen. 2006. "Age-related Change in Executive Function: Developmental Trends and a Latent Variable Analysis." *Neuropsychologia* 44 (11):2017–36. doi:10.1016/j.neuropsychologia.2006.01.010.
- Hurtado, Nereyda, Virginia A. Marchman, and Anne Fernald. 2008. "Does Input Influence Intake? Links between Maternal Talk, Processing Speed and Vocabulary Size in Spanish-learning Children." *Developmental Science* 11 (6):31–39. doi:10.1111/j.1467-7687.2008.00768.x.
- IPDS. 1996. "The Kiel Corpus of Spontaneous speech, Vol.2," CD-ROM#3, IPDS, Kiel.
- IPDS. 1997. "The Kiel Corpus of Spontaneous speech, Vol.3," CD-ROM#4, IPDS, Kiel.
- IPDS (Institut für Phonetik und digitale Sprachverarbeitung der Christian-Albrechts-Universität zu Kiel). 1995. "The Kiel Corpus of Spontaneous speech, Vol.1." CD-ROM#2, IPDS, Kiel.
- Janse, Esther, and Mirjam Ernestus. 2011. "The Roles of Bottom-up and Top-down Information in the Recognition of Reduced Speech: Evidence from Listeners with Normal and Impaired Hearing." *Journal of Phonetics* 39 (3):330–43. doi:10.1016/j.wocn.2011.03.005.
- Johnson, Carole E. 2000. "Children's Phoneme Identification in Reverberation and Noise." *Journal of Speech, Language, and Hearing Research* 43 (1):144–57. doi:10.1044/jslhr.4301.144.
- Johnson, Keith. 2004. "Massive Reduction in Conversational American English." In *Spontaneous Speech: Data and Analysis. Proceedings of the 1st Session of the 10th International Symposium*, edited by K. Yoneyama and K. Maekawa, 29–54. Tokyo, Japan: The National International Institute for Japanese Language.
- Jusczyk, Peter W., and Carolyn Derrah. 1987. "Representation of Speech Sounds by Young Infants." *Developmental Psychology* 23 (5):648–54. doi:10.1037/0012-1649.23.5.648.
- Jusczyk, Peter W., Kathy Hirsh-Pasek, Deborah Kemler Nelson, Lori Kennedy, Amanda Woodward, and Julie Piwow. 1992. "Perception of Acoustic Correlates of Major Phrasal Units by Young Infants." *Cognitive Psychology* 24 (2):252–93. doi:10.1016/0010-0285(92)90009-Q.

- Kail, Robert. 1991. "Developmental Change in Speed of Processing during Childhood and Adolescence." *Psychological Bulletin* 109 (3):490–501. doi:10.1037/0033-2909.109.3.490.
- Kapnoula, Efthymia C., Matthew B. Winn, Eun Jong Kong, Jan Edwards, and McMurray. Bob. 2017. "Evaluating the Sources and Functions of Gradiency in Phoneme Categorization: An Individual Differences Approach." *Journal of Experimental Psychology: Human Perception and Performance* 43 (9):1594–611. doi:10.1037/xhp0000410.
- Karmiloff-Smith, Annette. 2006. "The Tortuous Route from Genes to Behavior: A Neuroconstructivist Approach." *Cognitive, Affective, & Behavioral Neuroscience* 6 (1):9–17. doi:10.3758/CABN.6.1.9.
- Kohler, Klaus J. 1990. "Segmental Reduction in Connected Speech in German: Phonological Facts and Phonetic Explanations." In *Speech Production and Speech Modelling* (=nato ASI Series, Series D, Vol. 55), edited by W. J. Hardcastle and A. Marchal, 69–92. Dordrecht, Boston, London: Kluwer.
- Kohler, Klaus J. 1998. "The Disappearance of Words in Connected Speech." *ZAS Working Papers in Linguistics* 11:21–34.
- Kohler, Klaus J. 2001. "Articulatory Dynamics of Vowels and Consonants in Speech Communication." *Journal of the International Phonetic Association* 31 (1):1–16. doi:10.1017/S0025100301001013.
- Kong, Eun Jong, and Jan Edwards. 2016. "Individual Differences in Categorical Perception of Speech: Cue Weighting and Executive Function." *Journal of Phonetics* 59:40–57. doi 10.1016/j.wocn.2016.08.006.
- Kuhl, Patricia K., Barbara T. Conboy, Sharon Coffey-Corina, Denise Padden, Maritza Rivera-Gaxiola, and Tobey Nelson. 2008. "Phonetic Learning as a Pathway to Language: New Data and Native Language Magnet Theory Expanded (Nlm-e)." *Philosophical Transactions of the Royal Society B* 363 (1493):979–1000. doi:10.1098/rstb.2007.2154.
- Kuhl, Patricia K., Karen A. Williams, Francisco Lacerda, Kenneth N. Stevens, and Björn Lindblom. 1992. "Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age." *Science*, New Series 255 (5044): 606–08.
- MacDonald, Maryellen C., and Morten H. Christiansen. 2002. "Reassessing Working Memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996)." *Psychological Review* 109 (1):35–54. doi:10.1037/0033-295X.109.1.35.
- Maekawa, Kikuo, and Hideaki Kikuchi. 2005. "Corpus-based Analysis of Vowel Devoicing in Spontaneous Japanese. An Interim Report." In *Voicing in Japanese*, edited by Jeroen van de Weijer, Kensuke Nanjo and Tetsuo Nishihara, 205–28. Berlin: Mouton de Gruyter.
- Marchman, Virginia A., Anne Fernald, and Nereyda Hurtado. 2010. "How Vocabulary Size in Two Languages Relates to Efficiency in Spoken Word Recognition by Young Spanish-English Bilinguals." *Journal of Child Language* 37 (4):817–40. doi:10.1017/S0305000909990055.
- Marslen-Wilson, William. 1990. "Activation, Competition and Frequency in Lexical Access." In *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, edited by Gerry T.M. Altmann, 148–72. Cambridge, MA: MIT Press.
- Marslen-Wilson, William, and Alan Welsh. 1978. "Processing Interactions and Lexical Access during Word Recognition in Continuous Speech." *Cognitive Psychology* 10 (1):29–63. doi:10.1016/0010-0285(78)90018-X.
- Mattys, Sven L., Laurence White, and James F. Melhorn. 2005. "Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework." *Journal of Experimental Psychology: General* 134 (4):477–500. doi:10.1037/0096-3445.134.4.477.
- Mayo, Lynn H., Mary Florentine, and Søren Buus. 1997. "Age of Second-language Acquisition and Perception of Speech in Noise." *Journal of Speech, Language, and Hearing Research* 40 (3):686–93. doi:10.1044/jslhr.4003.686.
- McClelland, James L., and Jeffery L. Elman. 1986. "The TRACE Model of Speech Perception." *Cognitive Psychology* 18 (1):1–86. doi:10.1016/0010-0285(86)90015-0.
- McCusker, Leo X., Michael L. Hillinger, and Randolphe G. Bias. 1981. "Phonological Recoding and Reading." *Psychological Bulletin* 89 (2):217–45. doi:10.1037/0033-2909.89.2.217.
- McCusker, Leo X., P. Holley-Wilcox, and Michael L. Hillinger. 1979. *Frequency Effects in Auditory and Visual Word Recognition*. Paper presented at the meeting of the Southwestern Psychological Association. San Antonio, Texas, United States of America. April (cited in McCusker et al., 1981).
- McMurray, Bob, Ani Danelz, Hannah Rigler, and Michael Seedorff. 2018. "Speech Categorization Develops Slowly Through Adolescence." *Developmental Psychology* 54 (8):1472–91. doi:10.1037/dev0000542.
- McMurray, Bob, Michael K. Tanenhaus, and Richard N. Aslin. 2002. "Gradient Effects of Within-category Phonetic Variation on Lexical Access." *Cognition* 86 (2):B33–B42. doi:10.1016/S0010-0277(02)00157-9.
- Metsala, Jamie L., and Amanda C. Walley. 1998. "Spoken Vocabulary Growth and the Segmental Restructuring of Lexical Representations: Precursors to Phonemic Awareness and Early Reading Ability." In *Word Recognition in Beginning Literacy*, edited by Jamie L. Metsala and Linnea C. Ehri, 89–120. Mahwah, NJ: Erlbaum.
- Mitterer, Holger. 2008. "How are Words Reduced in Spontaneous Speech?" In *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics*, edited by Antonis Botinis, 165–68. Athens: University of Athens.
- Mitterer, Holger. 2011. "Recognizing Reduced Forms: Different Processing Mechanisms for Similar Reductions." *Journal of Phonetics* 39 (3):298–303. doi:10.1016/j.wocn.2010.11.009.
- Mitterer, Holger, and Annelie Tuinman. 2012. "The Role of Native-language Knowledge in the Perception of Casual Speech in a Second Language." *Frontiers in Psychology* 3 Article 249:1–13.

- Mitterer, Holger, Valéria Csépe, Ferenc Honbolygo, and Leo Blomert. 2006. "The Recognition of Phonologically Assimilated Words Does Not Depend on Specific Language Experience." *Cognitive Science* 30 (3):451–79. doi:10.1207/s15516709cog0000_57.
- Miyake, Akira, Naomi P. Friedman, Michael J. Emerson, Alexander H. Witzki, and Amy Howerter. 2000. "The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis." *Cognitive Psychology* 41 (1):49–100. doi:10.1006/cogp.1999.0734.
- Morano, Lisa, Mirjam Ernestus, and Louis Ten Bosch. 2015. "Schwa Reduction in Low-proficiency L2 Speakers: Learning and Generalization." In *Proceedings of the 18th International Congress of Phonetic Sciences (Icphs 2015)*, edited by Maria Wolters, Judy Livingstone, Bernie Beattie, Rachel Smith, Mike MacMahon, Jane Stuart-Smith and Jim Scobbie, 5 pages. Glasgow: University of Glasgow.
- Myers, James, P W. Jusczyk, Deborah G. Kemler Nelson, Jan Charles-Luce, A L. Woodward, and Kathryn Hirsh-Pasek. 1996. "Infants' Sensitivity to Word Boundaries in Fluent Speech." *Journal of Child Language* 23 (1):1–30. doi:10.1017/S0305000900010072.
- Newman, Rochelle S. 2004. "Perceptual Restoration in Children versus Adults." *Applied Psycholinguistics* 25 (4):481–93. doi:10.1017/S0142716404001237.
- Newman, Rochelle S. 2006. "Perceptual Restoration in Toddlers." *Perception & Psychophysics* 68 (4):625–42. doi:10.3758/BF03208764.
- Newton, Caroline, and Bill Wells. 2002. "Between-word Junctures in Early Multi-word Speech." *Journal of Child Language* 29 (2):275–99. doi:10.1017/S0305000902005044.
- Niebuhr, Oliver, and Klaus J. Kohler. 2011. "Perception of Phonetic Detail in the Identification of Highly Reduced Words." *Journal of Phonetics* 39 (3):319–29. doi:10.1016/j.wocn.2010.12.003.
- Nittrouer, Susan, and Arthur Boothroyd. 1990. "Context Effects in Phoneme and Word Recognition by Young Children and Older Adults." *The Journal of the Acoustical Society of America* 87 (6):2705–15. doi:10.1121/1.399061.
- Nittrouer, Susan, and Marnie E. Miller. 1997. "Predicting Developmental Shifts in Perceptual Weighting Schemes." *The Journal of the Acoustical Society of America* 101 (4):2253–66. doi:10.1121/1.418207.
- Nittrouer, Susan, Marnie E. Miller, Court S. Crowther, and Mary Jane Manhart. 2000. "The Effect of Segmental Order on Fricative Labeling by Children and Adults." *Perception & Psychophysics* 62 (2):266–84. doi:10.3758/BF03205548.
- Norris, Robert W. 1995. "Teaching Reduced Forms: Putting the Horse before the Cart." *English Teaching Forum* 33:47–50.
- Pitt, Mark A., Laura Dilley, and Michael Tat. 2011. "Exploring the Role of Exposure Frequency in Recognizing Pronunciation Variants." *Journal of Phonetics* 39 (3):304–11. doi:10.1016/j.wocn.2010.07.004.
- Poelmans, Petra. 2003. "Developing Second-language Listening Comprehension: Effects of Training Lower-order Skills Versus Higher-order Strategy." Doctoral diss., LOT, Utrecht, the Netherlands.
- Ranbom, Larissa J., and Cynthia M. Connine. 2007. "Lexical Representation of Phonological Variation in Spoken Word Recognition." *Journal of Memory and Language* 57 (2):273–98. doi:10.1016/j.jml.2007.04.001.
- Rigler, Hannah, Ashley Farris-Trimble, Lea Greiner, J. Jessica Walker, Bruce Tomblin, and McMurray. Bob. 2015. "The Slow Developmental Time Course of Real-time Spoken Word Recognition." *Developmental Psychology* 51 (12):1690–703. doi:10.1037/dev0000044.
- Rosenthal, Robert. 1991. *Meta-analytic Procedures for Social Research (Revised)*. Newbury Park, CA: Sage Publications.
- Sekerina, I A., and P J. Brooks. 2007. "Eye Movements during Spoken Word Recognition in Russian Children." *Journal of Experimental Child Psychology* 98 (1):20–45. doi:10.1016/j.jecp.2007.04.005.
- Shockey, Linda. 2003. *Sound Patterns of Spoken English*. Cambridge: Blackwell.
- Skoruppa, Katrin, Nivedita Mani, Kim Plunkett, Dominique Cabrol, and Sharon Peperkamp. 2013. "Early Word Recognition in Sentence Context: French and English 24-Month-Olds' Sensitivity to Sentence-Medial Mispronunciations and Assimilations." *Infancy* 18 (6):1007–29. doi:10.1111/inf.12020.
- Skoruppa, Katrin, Nivedita Mani, and Sharon Peperkamp. 2013. "Toddlers' Processing of Phonological Alternations: Early Compensation for Assimilation in English and French." *Child Development* 84 (1):313–30. doi:10.1111/j.1467-8624.2012.01845.x.
- Smiljanic, Rajka, and A R. Bradlow. 2009. "Speaking and Hearing Clearly: Talker and Listener Factors in Speaking Style Changes." *Language and Linguistics Compass* 3 (1):236–64. doi:10.1111/j.1749-818X.2008.00112.x.
- Smiljanic, Rajka, and Douglas Sladen. 2013. "Acoustic and Semantic Enhancements for Children with Cochlear Implants." *Journal of Speech, Language, and Hearing Research* 56 (4):1085–96. doi:10.1044/1092-4388(2012/12-0097).
- Swingle, Daniel, and Richard N. Aslin. 2000. "Spoken Word Recognition and Lexical Representation in Very Young Children." *Cognition* 76 (2):147–66. doi:10.1016/S0010-0277(00)00081-0.
- Taitelbaum-Swead, Riki, and Lea Fostick. 2016. "Auditory and Visual Information in Speech Perception: A Developmental Perspective." *Clinical Linguistics & Phonetics* 30 (7):531–45. doi:10.3109/02699206.2016.1151938.
- Ten Bosch, Louis, Ghislaine Giezenaar, Lou Boves, and Mirjam Ernestus. 2016. "Modeling Language-learners' Errors in Understanding Casual Speech." In *Errors by Humans and Machines in Multimedia, Multimodal, Multilingual Data Processing. Proceedings of Errare 2015*, edited by G. Adda, V. Barbu Mititelu, J. Mariani, D. Tufiş and I. Vasilescu, 107–21. Bucharest: Editura Academiei Române.

- Thompson, Joy, and Sara Howard. 2007. "Word Juncture Behaviours in Young Children's Spontaneous Speech Production." *Clinical Linguistic & Phonetics* 21 (11–12):895–99. doi:10.1080/02699200701600221.
- Tomasello, Michael. 1992. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.
- Tomasello, Michael. 2009. "The Usage-based Theory of Language Acquisition." In *The Cambridge Handbook of Child Language*, edited by Edith L. Bavin, 69–87. Cambridge: Cambridge University Press.
- Torreira, Francisco, and Mirjam Ernestus. 2011. "Realization of Voiceless Stops and Vowels in Conversational French and Spanish." *Laboratory Phonology* 2 (2):331–53. doi:10.1515/labphon.2011.012.
- Tsushima, Teruaki, Osamu Takizawa, Midori Sasaki, Satoshi Shiraki, Kanae Nishi, Morio Kohno, Paula Menyuk, and Catherine Best (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6–12 months: language-specific developmental changes in speech perception abilities. In International Conference on Spoken Language Processing (ICSLP), 1695–1698. Yokohama, Japan
- Tucker, B V. 2011. "The Effect of Reduction on the Processing of Flaps and /G/ in Isolated Words." *Journal of Phonetics* 39 (3):312–18. doi:10.1016/j.wocn.2010.12.001.
- van de Ven, Marco (see Ven)
- van der Feest, Suzanne (see Feest)
- Ven, Marco van de, Benjamin Tucker, and Mirjam Ernestus. 2011. "Semantic Context Effects in the Comprehension of Reduced Pronunciation Variants." *Memory and Cognition* 39 (7):1301–16.
- Walley, A C. 1988. "Spoken Word Recognition by Young Children and Adults." *Cognitive Development* 3 (2):137–65. doi:10.1016/0885-2014(88)90016-0.
- Walley, Amanda C., Jamie L. Metsala, and Victoria M. Garlock. 2003. "Spoken Vocabulary Growth: Its Role in the Development of Phoneme Awareness and Early Reading Ability." *Reading and Writing* 16 (1/2):5–20. doi:10.1023/A:1021789804977.
- Wanrooij, Karin and Maartje E.J. Raijmakers. *Under review*. "Hama"? Reduced pronunciations in non-native natural speech obstruct high-school students' comprehension severely at lower processing levels.
- Warren, R M. 1970. "Perceptual Restoration of Missing Speech Sounds." *Science* 167 (3917):392–93. doi:10.1126/science.167.3917.392.
- Werker, Janet F., and Richard C. Tees. 1984. "Cross-language Speech Perception: Evidence for Perceptual Reorganization during the First Year of Life." *Infant Behavior and Development* 7 (1):49–63. doi:10.1016/S0163-6383(84)80022-3.
- Werner, Lynne A. 2007. "Issues in Human Auditory Development." *Journal of Communication Disorders* 40 (4):275–83. doi:10.1016/j.jcomdis.2007.03.004.
- Wong, Simpson W.L., P.K. Peggy, P. P. Mok, Kevin Kien-Hoa Chung, Vina W.H. Leung, Dorothy V.M. Bishop, and Bonnie Wing-Yin Chow. 2015. "Perception of Native English Reduced Forms in Chinese Learners: Its Role in Listening Comprehension and Its Phonological Correlates." *TESOL Quarterly* 51 (1):7–31. doi:10.1002/tesq.273.
- Zukowski, Andrea, and Jaiva Larsen. 2011. "Wanna Contraction in Children: Retesting and Revising the Developmental Facts." *Language Acquisition* 18 (4):211–41. doi:10.1080/10489223.2011.605043.

Appendix A

Table A1. Overview of stimuli in subtask 1 of the phrase-intelligibility task.

Nr	Stimulus	Phonetic transcription of targets ([...] = non-targets)		Number of targets
		Reduced	Unreduced	
1	<i>damit einverstanden</i>	[...] ʔaɪnfɛʃtann(n)	[...] ʔaɪnfɛʃtandən	1
2	<i>gegen 14, 15 Uhr</i>	ge:ŋ [...]	ge:gən [...]	1
3	<i>bei dieser Gelegenheit</i>	[...] gəle:ɡŋhart	[...] gəle:gənhart	1
4	<i>wenn Ihnen das nichts ausmacht</i>	[...] i:n [...]	[...] ʔi:nən [...]	1
5	<i>zu Ihrem Büro</i>	[...] j:əm [...]	[...] ʔi:rəm [...]	1
6	<i>gleich morgens</i>	[...] mœɡŋs	[...] mœgəns	1
7	<i>im selben Monat</i>	[...] zɛlm(m) [...]	[...] zɛlbən [...]	1
8	<i>Wiederhören</i>	vi:dəhø:ən	vi:dəhø:rən	1
9	<i>kleinen Augenblick</i>	klain ʔʊŋblɪk	klainən ʔʊgənbɪk	2
10	<i>nach dem Essen</i>	[...] m ɡsn	[...] de:m ʔɛsən	2
11	<i>einen halben Tag</i>	n halm(m) [...]	ʔaɪnən halbən [...]	2
12	<i>ganz hervorragend passen</i>	[...] hɛfo:ɛra:ŋ pasn	[...] hɛfo:ɛra:gənt pasən	2
13	<i>in meinem Kalender</i>	[...] maɪm khalɛn(n)ɐ	[...] maɪnəm kalɛndɐ	2
14	<i>die letzten beiden Tage</i>	[...] lɛtstn baɪn(n) [...]	[...] lɛtstən baɪdən [...]	2
15	<i>nehmen wir das</i>	ne:m vi:ɐ [...]	ne:mən vi:ɐ [...]	2
16	<i>offenbar ganz was Spannendes</i>	ɔfma:ɐ [...] ʃpan(n)dəs	ʔɔfənba:ɐ [...] ʃpanəndəs	2
17	<i>schönen guten Tag</i>	ʃø:ŋ ɡun [...]	ʃø:nən ɡu:tən [...]	2
18	<i>wegen der Uhrzeit</i>	ve:ŋ dɐ [...]	ve:gən de:ɐ [...]	2
19	<i>an einem Wochenende</i>	[...] ʔaɪm vɔxŋɛn(n)ə	[...] ʔaɪnəm vɔxənʔɛndə	2
20	<i>mit anderen Worten</i>	[...] ʔnən vɔŋ	[...] ʔanderən vɔɛtən	2
21	<i>noch mal folgendes überlegen</i>	nɔ ma: fɔlŋəs y:bɛle:ɡŋ	nɔx ma:l fɔlgəndəs ʔy:bɛle:gən	3
22	<i>mal eben sehen</i>	ma: ɛ:bm ze:n	ma:l ʔɛ:bən ze:ən	3
23	<i>einen gemütlichen Abend verbringen</i>	ʔɪŋ ɡɛmy:tɪçŋ ʔ:mp fɛbrɪŋŋ	ʔaɪnən ɡɛmy:tɪçən ʔa:bənt fɛbrɪŋən	4
24	<i>die wichtigen noch offenen Fragen klären</i>	[...] vɪçtɪɡŋ [...] ʔf(ə)nən fra:ɡŋ khɛlɛ:ən	[...] vɪçtɪɡən [...] ʔɔfənən fra:gən klɛ:rən	4

^aThe word combination *noch mal* in the fragment *noch mal folgendes überlegen* was assigned 1 point at maximum: only 8 German natives (2 Adolescents and 6 Adults) wrote this combination correctly as two separate words; the other 66 natives considered it to be one word.

Appendix B

Table B1. Overview of stimuli in subtask 2 of the phrase-intelligibility task.

Nr	Stimulus (Context between parentheses)	Phonetic transcription of targets ([...] = non-targets)		Number of targets
		Reduced	Unreduced	
25	<i>(donnerstags sieht es allerdings (schlecht aus)</i>	gɛɪŋs	ʔaledɪŋs	1
26	<i>(immer) ausserordentlich (gut)</i>	əʊsɛʁəntliç	ʔausɛʔœdɛntliç	1
27	<i>(ich bin) eigentlich (gar nicht da)</i>	aɪnɪ	ʔaɪgɛntliç	1
28	<i>entschuldigen (Sie bitte)</i>	ʔɛnʃʊlɪŋ	ʔɛnʃʊldɪgən	1
29	<i>(zuviel Arbeit zu) erledigen</i>	ɛɛlə:dɪŋ	ʔɛələ:dɪgən	1
30	<i>(ich könnte) höchstens (nach den Feiertagen)</i>	hø:stns	hø:çstəns	1
31	<i>(Anfang November) irgendwann (oder am Montag)</i>	ʔɛgɪvən	ʔɛgɛntvən	1
32	<i>auf jeden Fall (haben wir ja die Einladung)</i>	[...] je:n [...]	[...] je:den [...]	1
33	<i>(noch genauer) verabreden</i>	fɛəpre:dən	fɛʔapre:dən	1
34	<i>(okay) vereinbaren (wir das)</i>	fɛəɪnma:ən	fɛʔaɪnba:rən	1
35	<i>vielleicht (wieder 8 Uhr)</i>	flaɪt	flaɪaɪçt	1
36	<i>ab dem sechzehnten (bis Weihnachten)</i>	[...] zɛçtse:ʔn	[...] zɛçtse:ntən	1
37	<i>(ich fände es) wirklich (toll)</i>	vɪɛkiç	vɪɛkiç	1
38	<i>zusammen (Kaffee trinken)</i>	tsuzam(m)	tsuzamen	1
39	<i>(das ist) im Grunde genommen (keine schlechte Idee)</i>	[...] grun(n)ə genomm	[...] grunde genomən	2
40	<i>dann haben wir (das auch gleich mit eingetragen)</i>	ha:me	ha:bən vi:ə	2
41	<i>(da haben Sie) natürlich nicht (ganz unrecht)</i>	naty:e(lɪ)ç nɪç	naty:ɛlɪç nɪçt	2
42	<i>(ich denke das dass) mit sieben Tagen (auch nicht hinkommen wird)</i>	[...] zi:m ta:ŋ	[...] zi:bən ta:gen	2
43	<i>zum Beispiel (vom 3ten bis zum 11ten Februar)</i>	tsu maɪʃl	tsum baɪʃpi:l	2
44	<i>da habe ich (nicht dran gedacht)</i>	da:bɪh	da: ha:bə ʔɪç	3
45	<i>das können wir (also machen)</i>	s kœm mi:ə	das kœnən vi:ə	3
46	<i>dann sind wir uns (da doch noch einig geworden)</i>	[...] zɪm v ʊns	[...] zɪnt vi:ə ʔʊns	3
47	<i>(das wird dann) wahrscheinlich ein bißchen (schwierig)</i>	vəʃaɪn(l) m bɪsçn	va:əʃaɪnliç ʔaɪn bɪsçən	3
48	<i>zwei Stunden werden wir (brauchen)</i>	[...] ʃtʊnʔn vɛ:əm vɛ	[...] ʃtʊndən vɛ:ɛdən vi:ə	3

Appendix C

Table C1. Overview of stimuli in the auditory lexical decision task.

Nr	Words	Pseudowords
	2 syllables	
1	<i>bleiben</i>	bleffen
2	<i>brauchen</i>	brufen
3	<i>fassen</i>	fagen
4	<i>fehlen</i>	fälchen
5	<i>finden</i>	fauben
6	<i>fragen</i>	frocken
7	<i>geben</i>	gachen
8	<i>gehen</i>	gelchen
9	<i>halten</i>	hehen
10	<i>helfen</i>	hinsen
11	<i>klappen</i>	kaffen
12	<i>kucken</i>	klanan
13	<i>legen</i>	lalen
14	<i>machen</i>	muchen
15	<i>melden</i>	mählen
16	<i>planen</i>	plappen
17	<i>rufen</i>	relen
18	<i>sagen</i>	saffen
19	<i>schaffen</i>	scheigen
20	<i>treffen</i>	trehen
21	<i>wollen</i>	welben
22	<i>wählen</i>	weten
	3 syllables	
23	<i>anbieten</i>	ankongen
24	<i>anfangen</i>	anschreichen
25	<i>aufschreiben</i>	aufreigen
26	<i>ausreichen</i>	ausfaggen
27	<i>ausschlafen</i>	aussuden
28	<i>beeilen</i>	bekrielen
29	<i>beenden</i>	benüten
30	<i>bekommen</i>	beschlarer
31	<i>besprechen</i>	bezichten
32	<i>eintragen</i>	einfagen
33	<i>genügen</i>	gesteden
34	<i>hinkriegen</i>	hinschlaben
35	<i>losfahren</i>	losschiegen
36	<i>verschieben</i>	verbieben
37	<i>verstehen</i>	vereigen
38	<i>versuchen</i>	verspreffen
39	<i>verzichten</i>	vertramen
40	<i>vorschlagen</i>	vorenden