Gutachten 1

"The focus of the project, speech segmentation by children in an L2 setting, is important and, as the authors note, indeed underresearched. There are, nevertheless, several important aspects of the way in which a foreign language is instructed in Germany, that are not included in experiments of the project. In addition, there are theoretical linguistic aspects of phonotactics that are also not mentioned. Despite these reservations I recommend the project be funded in the hope that my recommendations can be incorporated in the work packages.

*We thank the reviewer for their positive assessment of our project. We have now revised the project to accommodate the recommendations of the reviewer and believe that they strengthen the project.*

A difficult aspect of learning a foreign language is speech segmentation. An important cue that children use to do so concern the transitional probabilities of segments in speech (Adriaans, 2007; Adriaans & Kager, 2010), whichh has to be acquired in tandem with learning how to segment speech. The transitional probabilities of speech segments paint a more complicated and differentiated picture of phonotactics than the present project suggests. They involve probabilities instead of the binary choice ‘exists/does not exist’. However, it is not complicated to derive these probabilities from a corpus (Hall et al., 2015). I strongly recommend that the authors do this and that TP replaces the binary ‘exists/does not exist’ category and is used as a variable in the statistical anylsis.

*We agree fully with the assessment that TPs play a role in speech segmentation abilities. For this reason, we plan to use the same procedure developed in our pilot project “Foreign speech segmentation in adults and young children” to develop the sentence stimuli in both experiments:*

“A set of 120 grammatically correct and semantically meaningful standard German sentences were prepared, each containing one target word. Unlike Gullberg and colleagues, however, we controlled for the transitional probability, altering sentences such that target words could not be preceded or followed by the same word more than once, and that target word syllables only occurred within target words. This was accomplished by transforming the sentences from orthographic letter to phonological codes using CELEX (Baayen, Piepenbrock, & Gulikers, 1995) and using the R package *markovchain* (Spedicato, 2017) to compute transitional probabilities between syllables. This demonstrated the feasibility of controlling for complex cues to segmentation even within natural language, which will be applied in the current proposal.”

“In designing the study, we were able to control for cues to word boundaries (transitional probabilities), which are often overlooked in studies with naïve learners using natural language.”

*We believe that this manipulation addresses the reviewer’s concern. As we are interested in children’s use of phonotactic cues, it is important that transitional probabilities at the word boundaries of target words are held constant across our conditions of interest, namely between target words that are either phonotactically legal or illegal. We will accomplish this by ensuring that the syllables that occur in our target words do not occur in the rest of the sentences and that target words can not be preceded or followed by the same word more than once.*

*That being said, it is possible that the German transitional probabilities of the syllables of our English target words may influence their successful segmentation. To address this issue, we will opt to control for the German transitional probability of the syllables in our English target words, holding them within the conditions legal/illegal to be as similar as possible to ensure that variation in TPs do not influence children’s performance:*

“We will choose English target words such that the transitional probability of their syllables in German child-directed speech, as determined by a corpus analysis (Behrens, 2006), is similar within the conditions of legal/illegal target words.”

Another aspect that is not discussed in the project is the way in which foreign languages are taught in Germany in primary schools. Usually the teacher is the only form of English language input. This is a very different form of English than the children will be exposed to in the experiment. In the experiment children have to segment speech of a native speaker. Irrespective of the number of years in primary school, segmenting speech of a native speaker will be a new experience for them. In my view ignoring this different form of input for the children (in school and in the experiment) sets up the project for many null results, which will then lead to execution of the contingency plan. If I understood correctly, the subject pool will then be adults or older children. This will be a rather different set of experiments.

This problem can be addressed by also testing the children on segmentation of English as spoken by a German native speaker. This speech will have phonetic and phonological properties that will be closer to their experience, and more likely to yield interpretable results. It would be most revealing to compare the result obtained in this way to the results of speech segmentation of speech spoken by native speakers of English. I am aware that this a larger change to make. However, the number of people working in the project (3 PIs and 2 PhD students) leaves room to implement these changes, and still execute the working program within 36 months.

*We thank the reviewer for this interesting and practice-relevant suggestion. Certainly, some teachers of English in Germany produce English speech with a German accent, although this varies drastically between teachers. It is possible that the acoustic-phonetic similarity of children’s German knowledge and that of German-accented English speech could be facilitatory in comparison to that of a native English speaker. Having said that, even in primary school FL teaching, teachers use plenty of audio and video media that provide native English input (e.g. songs and videos in Playway, Sally or other textbooks). More generally, we would argue that the goal of teaching children English in German schools is not so they can understand and produce German-accented English speech but instead to understand and produce English speech in general. This is also the goal of the current project. Critically, we test children BEFORE they begin foreign-language instruction, so none of the children will have had significant experience with the English language, whether native or German-accented. We include a set of comparison studies with German-accented speakers as a possible related dissertation topic for the PhD students employed in this project.*

Gutachten 2

"1. Wie beurteilen Sie die Qualität des Vorhabens, vor allem hinsichtlich Originalität und erwartetem Erkenntnisgewinn?

Der Gegenstand des Projekts ist die Fähigkeit von Kindern, fremdsprachlichen Input zu verarbeiten. Genauer eingegrenzt geht es um Fähigkeiten der phonologischen Segmentierung, der Identifikation von Wortgrenzen bei der auditiven Wahrnehmung fremdsprachlichen Materials. Motiviert wird dieser Forschungsgegenstand mit der gegenwärtigen Debatte um den besten Zeitpunkt für den Beginn des schulischen Fremdsprachenunterrichts. Hierzu liegen unterschiedliche Ergebnisse aus wissenschaftlichen Studien vor, weitere wissenschaftliche Klärung ist erforderlich. Das Projekt greift daher eine wichtige Frage auf, deren Bearbeitung im Prinzip zu Erkenntnissen führen kann, die für die schulpädagogische Praxis von Bedeutung sein sollten. Das Vorhaben ist zweifellos originell, eine vergleichbare Studie wurde noch nicht durchgeführt. Die Antragsteller planen drei Experimente, in denen die Fähigkeit von Kindern (Variablen: Alter, Sprachkenntnisse) getestet wird, englischen Input zu segmentieren. Getestet wird dabei, ob das muttersprachliche Wissen über phonotaktische Prinzipien, über lexikalische Einheiten (cognates) und über durch angrenzende Einheiten erschließbare Wortgrenzen einen Effekt auf die Verarbeitung fremdsprachlichen Materials hat. Die Experimente sind sorgfältig konzipiert und sind in erster Linie aus psycholinguistischer Sicht sehr interessant.

Ich sehe allerdings eine Reihe von Problemen, die von den Antragstellern angesprochen bzw. in der Darstellung der Experimente bei einer Überarbeitung berücksichtigt werden sollten. Meine Anmerkungen beziehen sich auf die Anlage der Untersuchung generell sowie auf das experimentelle Design im Besonderen.

*We thank the reviewer for their suggestions to improve our project. We have now revised the project to accommodate the recommendations of the reviewer and believe that they strengthen the project.*

- Die Fremdsprache, die untersucht werden soll, ist das Englische. Es ist anzunehmen, dass Kinder im Alter von 6/7 und 9/10 bereits erheblichen Kontakt mit dem Englischen haben. Dazu werden auch Wörter gehören, die nicht-deutsche Silbenphonologie aufweisen.

*We agree that it is likely that children in Germany will have had some non-systematic form of contact with the English language before they receive instruction in it in their school. This is just a fact of life, and – as the study addresses a real-life situation – we do not think that it is a problem as such. In particular, older children will have likely encountered more English than younger children, which needs to be considered when assessing the appropriate starting age of EFL instruction. We agree with the reviewer that it is important to assess this aspect. We will therefore include a receptive test of vocabulary (CLT-Litmus English) that comprises 25 high-frequency items familiar in child-directed speech and use the participants’ scores as a fixed effect in the analyses-*

- Die Testung bilingualer Kinder mit ganz unterschiedlichen Ausgangssprachen beinhaltet Variablen, die in keiner Weise kontrolliert werden. Kinder können z. B. Russisch, Polnisch, Türkisch und Arabisch als Muttersprachen haben. Es ist nicht bekannt, inwieweit unterschiedliche phonologische Komplexität (z. B. Russisch versus Türkisch) möglicherweise die phonologische Verarbeitung grundsätzlich beeinflusst und damit auch die Fähigkeit, neue sprachliche Strukturen zu segmentieren. Die Untersuchung sollte sich zunächst auf monolinguale Kinder beschränken, die dabei berücksichtigten Variablen sind bereits sehr umfassend.

*The DFG sets a priority on engaging with society in funded research projects with all of its diverse aspects, and further warns researchers that not taking these aspects into account can lead to “blind spots” that limit the quality of scientific results (*<https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/vielfaeltigkeitsdimensionen/stellungnahme.pdf>)*. Over a quarter of the population of Germany has an immigrant background (Statistisches Bundesamt, 2020, with more than 50% of all children in the relevant age range having an extended migration background in urban centres. Regardless of the scientific question being posed, we feel it is our duty to include a diverse sample of children which reflects the make-up of the German population today. This includes bilingual children.*

*Nonetheless, we appreciate the scientific concern underlying the comment by the reviewer. We agree that phonological processing could be influenced by bilingualism; in fact, this is one of the reasons we find it of scientific interest to test bilingual children, in addition to its societal importance. Based on previous results on cross-linguistic influence in child second and third language acquisition (e.g. Hopp et al., 2019), we expect the majority language, German, to have a strong influence on the processing of a novel foreign language, regardless of whether children speak a (possibly additional) first language next to German or not. We thus think that studying the influence of German on the segmentation of English makes sense for the whole of the targeted population. Nevertheless, to ensure that we do not gloss over important differences, we will opt to analyze the results of the bilingual and monolingual children separately. If the pattern of results is similar for monolinguals and bilinguals, then we will combine the monolingual and bilingual datasets.*

*Furthermore, the fact that we think studying the influence of German on the segmentation of English is a sensible undertaking for the group as a whole does not preclude that it is also sensible to ask whether in addition of strategies common to both subgroups, an influence either of specific first languages or of the bilingual experience per se on segmentation strategies in a foreign language may exist in the bilingual children when compared with the monolingual ones. To be able to detect such potential differences, one would need to focus on specific subgroups in terms of first language as well as degrees of bilingualism (or however one would call factors such as ages of onset and quantity and quality of exposure) and design materials appropriate for dissociating influences of German from influences of other first languages. To do so, one of the PhD-projects may additionally focus on bilingual children, carry out additional and/or ancillary tasks, and explore potential effects of the minority-language phonology on the bilingual children’s performance (see below).*

Ein Problem sehe ich auch bei der Bestimmung der illegalen Onsets. [sl] ist als solcher definiert. Allerdings haben deutsche Kinder häufig noch länger Schwierigkeiten den [ʃ] zu artikulieren, z.B. ist die Aussprache [slitn] für Schlitten häufig und anhaltend zu finden. Geht dieses Wissen dann verloren? (auch in wenigen Worten der Standardsprache vorhanden, wie Slalom, Slip, Slogan).

*In general, the perceptual and productive abilities of children are not identical in that children’s perceptual abilities vastly outstrip their productive abilities. Hence, the problems in production do not imply difficulties in perception, and – surely – six-to-10-year-old children understand the target pronunciation of “Schlitten”. Nonetheless, we agree that earlier or production-related difficulties with the*[ʃ] *could theoretically have an influence on the phonotactic knowledge of the children tested in our project. We have therefore removed /sl/ from the possible illegal German onsets in our stimuli.*

- Einschlägige Forschung zu sprachlichen Fähigkeiten bei Bilingualen ist ungenügend rezipiert. Studien in jüngerer Zeit haben gezeigt, dass eine Aussage wie ‚bilinguals typically show an advantage in cognitive behaviour‘ sehr viel stärker differenziert gesehen werden muss. Die unterschiedlichen exekutiven Funktionen werden z.T. positiv z.T. negativ durch einen bilingualen Spracherwerb beeinflusst. (z. B. Antoniou, M. (2019). The Advantages of Bilingualism. Annual Review of Linguistics: Vol. 5, 395-415. Miller, D., Bayram, F., Rothman, J. & Serratrice (eds.) (2018). Bilingual Cognition and Language: The State of the Science Across Its Subfields. Amsterdam: John Benjamins.)

*We have now reworded this sentence to better capture the nuance in the field regarding bilingual’s cognitive abilities: “*Some studies demonstrate differences in cognitive factors for bilingual children (Adesope, Lavin, Thompson, & Ungerleider, 2010; but see Morton & Harper, 2007)…*”. We note that there is no study to date that has shown systematic DIS-advantages of bilingual children in cognitive tasks. Hence, bilingualism is either neutral or positive wrt cognitive development, so that we believe that the premise for speaking of cognitive effects of bilingualism holds. The reviewer is perfectly right, though, to say that the linguistic competences (sprachliche Fähigkeiten) of bilinguals differ from those of monolingual children, which is reflected in our hypotheses.*

- Die genaue Vorgehensweise bei den Experimenten wird in dem Antrag nicht klar. Ist es so gedacht, dass die Kinder 192 Sätze hören und danach Wortentscheidungen treffen sollen? Dies scheint eine gewaltige Gedächtnisaufgabe zu sein, wobei sich die Frage stellt, wie die Faktoren ‚Gedächtnis‘ bzw. ‚Konzentrationsfähigkeit‘ in der Auswertung berücksichtigt werden kann.

*The procedure for the experiments proposed in our project is described in detail on page 11. In the training phase children will listen to the 192 training sentences. In the subsequent test phase, response accuracy will be measured to evaluate whether children successfully recognize words presented during the training phase.*

*Although this task certainly requires concentration, previous studies have successfully conducted similar experiment with children in a similar age range. For example, Saffran and colleagues (Saffran et al., 1997) tested 6- to 7-year-old children in an artificial language learning task using a procedure similar to that of the current project, where the training phase was 21 minutes long. In the current project, the individual sentences could range between 2 and 3 seconds, leading to a total training phase length between 6.4 and 9.6 minutes, less than half the length of the training phase used by Saffran et al. (1997).*

*We agree that memory may play a role in children’s successful performance in the current project. For this reason, we included the variables working memory and phonological short-term memory. Regarding the length of the test phase, we include trial number as a random intercept in our random effects structure, to account for whether performance changes in the course of the experiment.*

- Es sollte auch erläutert werden, wie ausgewertet wird. Werden ‚Fehler‘ unter den Bedingungen legal/illegal unterschiedlich bewertet? Wie sollen die zahlreichen Kontrollvariablen in die Auswertung eingehen? Sollen hier individuelle Profile ermittelt werden? Hier müssten die Ausführungen im Antrag genauere Angaben machen.

*The analysis plan for the experiments proposed in our project is described in detail on page 11. It has now been updated to include our approach for initially analyzing the data of monolingual and bilingual children separately. Errors are considered in the accuracy measurement, which differentiates between correct and incorrect answers. We do not treat accuracy for legal and illegal target words differently, instead analyzing the influence of phonotactic status of target words on their accuracy:*

*“*Considering that phonological processing could vary depending on the minority language of the bilingual children (but see Stimuli for our plan to control for similarity in common minority languages), the first pair of models will examine bilingual and monolingual children separately. These models will compare overall performance on target words that were either phonotactically legal or illegal.*”*

*We also described in this section how we include our numerous factors which examine individual variability in our analysis approach. They will be added as fixed effects (including interactions) in the mixed-effects regression modelsAs regards the effects of bilingualism, the revised proposal states:*

*“*If the pattern of results [in the first pair of models] is similar between bilingual and monolingual children, this data will be combined and a second model will be constructed to evaluate the role of age, monolingual/bilingual background, production skills in the majority-language, and phonological awareness, with the additional control factors minority-language production skills (bilingual children only), WM, phonological STM, SES, and non-verbal IQ. If the pattern of results [in the first pair of models] is different for bilingual and monolingual children, then the second models will be conducted separately for these two groups.*” We will also include random intercepts for different classrooms and schools in our model, to account for potential variance.*

- Es wird auch nicht klar, in welcher Form und mit welcher Instruktion die lexical decision task durchgeführt werden soll. Die Kinder sollen einzelne Wörter identifizieren, wobei Studien zeigen, dass Kinder im Alter von 6 u. U. noch kein dem erwachsenen Sprecher entsprechendes Bewusstsein der Einheit Wort haben (vgl. z. B. die Diskussion und Ergebnisse in Literacy and Word Boundaries, 2007, J. Kurvers, R. van Hout, and T. Vallen) .

*It is true that children of this age do not have completely adult-like conceptualizations of the unit “word”, in particular since they have less literacy experience. The decision task we employ does not hinge on a complete abstract understanding of the term “word”, yet operates on the basis of familiarity. To make this clearer, we have now included the following sentence to describe how children will be instructed in the test phase:*

“Children will be asked to help the experimenter identify the words that the alien spoke previously by listening to individual words and indicating by button press whether they had heard the sound previously.”

*Completion of this task does not require adult-like knowledge of the abstract entity “word”, as illustrated for example in the successful performance of 6- to 7-year-olds in a similar paradigm (Saffran et al., 1997). Hence, the term “lexical decision task” we used is in fact somewhat of a misnomer of what the children actually do – although it describes the purpose of the task.*

2. Inwiefern überzeugen Ziele und Arbeitsprogramm hinsichtlich der Klarheit der Arbeitshypothesen und einer sinnvollen Eingrenzung der Thematik? Benennen Sie bitte Stärken und Schwächen der geplanten Untersuchungen, die Angemessenheit der Methoden und des Zeitplans.

Die Ziele sind klar formuliert und ergeben sich aus einer interessanten Forschungsfrage. Wie bereits unter 1 dargestellt, ist die Untersuchung der bilingualen Gruppe aus den genannten Gründen problematisch. Unklar ist auch, welche Schlussfolgerungen sich aus den erwarteten Befunden für die schulische Praxis ergeben könnten. Es wird erwartet, dass jüngere Kinder auf Grund ihrer noch nicht vollständig stabilisierten L1 besser bei der Segmentierung abschneiden könnten, andererseits könnten ältere Kinder auf Grund ihrer kognitiven Vorteile (phonological awareness) bessere Leistungen bringen. Die Fragestellung ist für die Psycholinguistik hochinteressant, ihre Praxisrelevanz ist nicht unbedingt deutlich.

*We thank the reviewer for their support that our project is relevant for the field of psycholinguistics. We would argue, however, that it is also has practical relevance. As the reviewer states themselves, our project examines whether younger or older children may have an advantage in initial foreign speech segmentation and further seeks to identify the mechanisms driving these potential differences. As outlined in the introduction to the proposal, current decisions in educational policy about pushing back the onset of FL instruction from 1st to 3rd grade (Baden-Württemberg, NRW) illustrate the practical relevance, Importantly, these policy decisions are primarily made (or justified) against the backdrop of research suggesting that the long-term outcomes of foreign-language instruction do not differ between students experiencing an early or a late(r) start of FL instruction. As shown and discussed in, e.g., Baumert et al. (in press, Language Learning), such a reasoning is flawed since long-term outcomes are predominantly affected by later effects, i.e. teaching in secondary schooling. Hence, this project centers on the potential immediate advantages of an early start by addressing the psycholinguistic mechanisms that may confer advantages on early learners.*

5. Bitte formulieren Sie ein eindeutiges Votum für oder gegen eine Förderung. Bitte machen Sie bei einem Votum für eine Förderung einen konkret ausdifferenzierten Mittelvorschlag. Berücksichtigen Sie dabei gegebenenfalls, ob die beantragten Mittel angemessen sind.

In der vorliegenden Form kann ich das Projekt nicht befürworten. Ich empfehle eine Überarbeitung, bei der zum einen die bilingualen Probanden entweder nicht aufgenommen werden, oder sehr viel deutlicher gemacht wird, wie die Varianz der Muttersprachen tatsächlich für die Interpretation der Daten berücksichtigt werden kann. Des Weiteren sollte dazu Stellung genommen werden, ob Vorkenntnisse des Englischen tatsächlich ohne Bedeutung für das Experiment sind (no prior experience erscheint fraglich) und wie die Ergebnisse konkret für die schulische Praxis ausgewertet werden könnten. Hierbei ist auch zu berücksichtigen, dass der Fremdsprachenunterricht in der Schule mit bereits segmentiertem Material arbeitet, die Kinder lernen auf der Basis von Wörtern, nicht von unsegmentiertem auditivem Input."

As we have discussed above, we believe the exclusion of bilingual children from our project would be in direct conflict with the principles of the DFG and would render the project “praxisfern” and thus largely irrelevant beyond a lab setting. However, we find the influence of the phonological properties of the non-majority language of the bilingual children to be an important issue worth pursuing. As outlined above, we therefore now *include the option of investigating this question by itself, in a specific language combination in bilingual children, such as Turkish-German, as a possible related dissertation topic for the PhD students employed in this project.*

We hope that we have addressed the concern of the role of prior experience in our direct answer above.

Regarding the concrete practical application of our research results in schools by teachers, we note that this is not the concern of the project. This psycholinguistic project investigates the mechanisms driving foreign and second language learning. Its findings will have immediate relevance to educational policy wrt the optimal starting age for FL instruction and – more generally – by highlighting how different linguistic, cognitive and biographical factors interact in child (language) learning. Beyond these issues, the findings of the project, e.g. concerning the role of cognates in speech segmentation, hold some implications for teaching and can be integrated into potential interventions and changes in curricula. Since this is not the direct concern of the project and EFL instruction needs to take into account multiple dimensions, we refrain from overstating the teaching implications of the findings. We also lay no claim for this project to be an applied or practical project on classroom practices.

We take issue with the characterization of the speech produced by FL teachers in the classroom as consisting solely of individual words. Teaching manuals, curricular guidelines and current classroom practice bear witness to the use of communicatively relevant, coherent discourse, utterances and sentences in the target language, English. English teacher training didactics, such as those taught by our colleagues at the TU Braunschweig and TU Dortmund, provide specific instruction to trainee teachers for how to communicate with young English learners. This includes producing utterances with multiple words, requiring the listener to segment the speech stream. This is also encouraged and required in the accompanying texts available for English teachers. For example, in the teacher’s edition of the English for German learners teaching series *Come In 3* for students in grade 3, the method of “Einsprachigkeit” or teaching in one language alone and using full English phrases, is explicitly encouraged (Große-Brauckmann & Hedelberg, 2018, p. 6). From our experience visiting EFL classrooms in primary school, we are confident to say that EFL instruction from early on occurs mainly in English, with teachers using non-segmented language and certainly more than isolated words in English. We have now added this information as motivation in our section “Foreign speech segmentation” in State of the art and preliminary work.